

**Heterogenität.
Wert.Schätzen**

Abstractband

**3. Jahrestagung der
Gesellschaft für Empirische
Bildungsforschung (GEBF)**

Bochum

11. – 13. März 2015

Stand: 20. Februar 2015



Dieser Band enthält die Zusammenfassungen aller Beiträge der 3. Jahrestagung der Gesellschaft für Empirische Bildungsforschung (GEBF), ausgerichtet von der Professional School of Education der Ruhr-Universität Bochum vom 11. bis 13. März 2015.

In diesen Band aufgenommen wurden alle Zusammenfassungen mit Stand vom 20. Februar 2015.

Für etwaige Änderungen sowie für weitere Details der Beiträge wie z.B. Literaturangaben siehe <http://www.conftool.com/geb2015/>.

Lipp, Robert

Einführung in die Arbeit mit den Daten des Nationalen Bildungspanel (NEPS)

Das Nationale Bildungspanel (NEPS) stellt der wissenschaftlichen Gemeinschaft umfangreiche Daten zu Kompetenzentwicklungen, Bildungsprozessen, Bildungsentscheidungen und Bildungsrenditen in formalen, nicht-formalen und informellen Kontexten als Scientific-Use-Files zur Verfügung. Die Studie umfasst insgesamt sechs im Längsschnitt verfolgte Stichproben (Neugeborene, Kindergartenkinder, Fünftklässler, Neuntklässler, Studierende, Erwachsene), die jeweils in mindestens jährlichem Abstand wiederholt befragt bzw. getestet werden. Ergänzt wird der Datenbestand durch Informationen, die von den jeweiligen Eltern, Erzieherinnen und Erziehern, Lehrerinnen und Lehrern sowie Leiterinnen und Leitern der Bildungseinrichtungen erhoben wurden. Damit bietet das NEPS eine ergiebige Quelle für die Bearbeitung bildungsbezogener Fragestellungen aus den verschiedensten Disziplinen. Für weitere Details siehe auch: www.neps-data.de.

Ziel des Workshops ist es, die Teilnehmerinnen und Teilnehmern in die Arbeit mit den NEPS-Daten einzuführen und deren vielfältige Analysemöglichkeiten zu veranschaulichen. Zu diesem Zweck werden das komplexe Design und die wesentlichen Inhalte der Studie vorgestellt, die verschiedenen Möglichkeiten des Datenzugangs demonstriert sowie ein Einblick in die verfügbaren Dokumentationsmaterialien gegeben.

Mittwoch
11. März

9:00 Uhr

GC 03/142

Trixa, Jessica

Einführung in das Forschungsdatenmanagement

Die Einhaltung der Grundsätze guter wissenschaftlicher Praxis setzt einen verantwortungsvollen Umgang mit Forschungsdaten voraus. Die angemessene Aufbereitung, Dokumentation und langfristige Sicherung der in Forschungsprozessen generierten Daten ist für die Nachvollziehbarkeit der Forschung, die Publikation wissenschaftlicher Forschungsergebnisse und die potenzielle Nachnutzung der erhobenen Daten unentbehrlich. Nicht zuletzt haben sowohl nationale als auch internationale Förderinstitutionen das Bewusstsein für Forschungsdatenmanagement sensibilisiert. Sie verknüpfen ihre Förderung damit, dass in Förderanträgen dargelegt werden soll, wie der Umgang mit generierten Forschungsdaten geplant ist und ob die Daten nach Projektabschluss für eine Nachnutzung zur Verfügung stehen werden.

Der Vortrag widmet sich den zentralen Fragen, mit denen sich Forschende im Hinblick auf gutes Forschungsdatenmanagement konfrontiert sehen und es wird erklärt, inwiefern ein Datenmanagementplan bereits in der Planungsphase eines Projekts hilfreich sein kann.

Mittwoch
11. März

9:00 Uhr

GC 03/149

Sumfleth, Elke

Forschung.Fordert.Interdisziplinarität

Aktuelle Forschungsfelder sind heute stark interdisziplinär geprägt. Da macht die Bildungsforschung keine Ausnahme. Damit ist zwangsläufig eine hohe Heterogenität verbunden. Unter dem Tagungsmotto Heterogenität.Wert.Schätzen wünscht sich die diesjährige Tagungsleitung, dass „durch die Vorstellung der wissenschaftlichen Ergebnisse Themen der Bildungsforschung aus verschiedenen disziplinären Perspektiven beleuchtet werden und aus dem Zusammenführen der verschiedenen Blickwinkel ein umfassenderes Bild zu Fragen der Bildung entsteht.“ Dies ist sicherlich ein Ziel, dass die Arbeit bei uns in Essen im Forschungsverbund von nwu in den letzten 10 Jahren geprägt hat.

Im Vortrag werde ich zum einen – teilweise auch anekdotenhaft – auf die Mühsal dieses Kooperationsprozesses eingehen und aber auch den zu erzielenden Gewinn darstellen. Zum anderen stellt sich aber auch die Frage nach dem Ausmaß an Heterogenität. Auf der Tagungshomepage wiederum kann man lesen „Beiträge aus der Bildungsökonomie, der Erziehungswissenschaft, den Didaktiken, der Psychologie und der Soziologie, die sich bildungsrelevanten Themen mit empirischen Ansätzen nähern, sind gleichermaßen erwünscht.“ Hier scheinen die „Didaktiken“ nahezu ein homogenes Fach zu sein. Abgesehen von der Spannbreite von Literaturdidaktik bis Mathematikdidaktik werden selbst in der sogenannten „Naturwissenschaftdidaktik“ mindestens drei verschiedene Sprachen gesprochen und Forschungsergebnisse sind nicht einfach von Chemieunterricht auf Biologie- oder Physikunterricht zu übertragen. An verschiedenen Projekten und ihren Ergebnissen werde ich die Vielfalt erläutern. Auch hier ist die Welt sehr bunt.

Mittwoch
11. März

12:00 Uhr

HGB 10

Karst, Karina; Frey, Andreas; Ertelt, Bernd-Joachim; Dickhäuser, Oliver

Studienorientierung durch Online Self-Assessments: Veränderung von Einstellungen zum Studienfach durch einen internetbasierten Selbsttest

Online Self-Assessments (OSA) werden als wichtiges Instrument der Studienorientierung gesehen (Hasenberg & Schmidt-Atzert, 2014). Dabei handelt es sich um Selbsttests für Studieninteressierte deren inhärenter Bestandteil eine unmittelbare Rückmeldung im Anschluss an die Bearbeitung ist. Auf dieser Grundlage kann der Studieninteressierte seine Passung zum Studienfach überprüfen und seine Entscheidung abwägen. Somit sollen OSA Selbstselektion ermöglichen und Beratungsangebote zur Studienorientierung ergänzen.

OSA können diese Funktionen nur erfüllen, wenn sich in Folge der Bearbeitung und in Abhängigkeit der erhaltenen Rückmeldung die Einstellung in Bezug auf das abgestrebte Studienfach verändert oder bewusster wird. Dieser zentralen Frage nach der Qualität von OSA wird im geplanten Vortrag nachgegangen.

Theoretisch fokussiert die Studie dabei auf die Erwartungs-Wert-Theorie sensu Wigfield und Eccles (2000), mit deren Hilfe Wahlverhalten von Personen erklärt werden kann. Im Kern besagt diese Theorie, dass die Wahl von Handlungsalternativen bestimmt wird durch erwartungsbezogene Variablen und den Wert, der dieser Handlung/Tätigkeit beigemessen wird. Danach wählt eine Person am ehesten dasjenige Studienfach, von dem sie glaubt, ihre wertbezogenen Einstellungen erfolgreich realisieren zu können. Fehlende Passung zwischen Person und Umwelt liegt dann vor, wenn eine Person ein Studium auswählt, obwohl ihre Fähigkeiten ein erfolgreiches Absolvieren dieses Studiums wenig wahrscheinlich machen oder die Gelegenheitsstrukturen des Studiums nicht diejenigen Wertkomponenten bedienen, die der Person wichtig sind. Wird eine Person im Rahmen eines OSA über eine solche Diskrepanz informiert, so sollte dies mit einer reduzierten Studienwahlintention für das zunächst angestrebte Fach einhergehen und zwar umso stärker, je stärker das Ausmaß der Diskrepanz.

An einer quasi-experimentellen Online-Studie zur Überprüfung dieser Fragestellung haben 73 Studieninteressierte mit einem Altersdurchschnitt von 19,5 Jahren (SD= 2;6) für das Fach Soziologie (n= 50) und Politikwissenschaften (n= 23) teilgenommen. Dabei wurde das Mannheimer Informationssystem für Sozialwissenschaften - MISS (Vent & Erdfelder, 2009) zur Überprüfung der Erwartungen an die Studienfächer verwendet. Die Befragung gliederte sich in drei Teile. Zunächst (t1) machten die Befragten soziodemographische Angaben und beantworteten Fragen zur Entschiedenheit und Sicherheit, das Studium aufzugreifen (je 1 Item), zu ihrer Überzeugung, erfolgreich zu studieren (Erfolgserwartung: 3 Items; 5-stufig) und zu ihren wertbezogenen Einstellungen zum Studienfach (Intrinsischer Wert, Nützlichkeit, Wichtigkeit und Kosten: je 3 Items; 5-stufig; vgl. Steinmayr und Spinath, 2010). Im zweiten Teil bearbeiteten die Befragten das OSA und bekamen unmittelbar Rückmeldung in Form eines Profilvergleichs, inwiefern ihre Erwartungen den (u.a. von Universitätsprofessoren eingeschätzten) tatsächlichen Inhalten des Studiums entsprachen. Im dritten Teil (t2) wurden die Konstrukte des ersten Teils erneut erhoben.

Zunächst wurde in einem ersten Schritt (MANOVA mit Messwiederholung) analysiert, ob sich die Einstellungen durch die Bearbeitung des Self-Assessments verändern. Während die Entschiedenheit und die Sicherheit für die Studienwahl über die Zeit signifikant stiegen, fielen die wert- und erwartungsbezogenen Einstellungen nach Bearbeitung des MISS signifikant ab. In einem nächsten Schritt wurde regressionsanalytisch geprüft, ob die rückgemeldete Diskrepanz die Einstellung zu t2 unter Kontrolle von t1 prädiziert. Dabei wurde die Diskrepanz als absolute Differenz zwischen den Erwartungen der Studieninteressierten und den Expertenaussagen berechnet. Bei drei Einstellungsvariablen hatte die Höhe der rückgemeldeten Diskrepanz einen signifikant negativen Effekt auf die mittlere Ausprägung ebendieser Variable zu t2. Je größer die Diskrepanz, desto niedriger wurde der erwartete intrinsische Wert und die Nützlichkeit des Studiums eingestuft. Außerdem reduzierte sich die Erfolgserwartung für das Studium signifikant unter Kontrolle der Diskrepanz.

Durch die Bearbeitung eines OSA lassen sich ergo die Einstellungen bezogen auf die Wahl eines Studienfachs verändern. Dabei scheint nicht nur die bloße Teilnahme an einem OSA ausschlaggebend, sondern überdies auch die Höhe der rückgemeldeten Diskrepanz. Inwiefern die Richtung der Diskrepanz im Sinne von Über-/Unterschätzung der Erwartungen differentielle Wirkung hat, muss offen bleiben und soll diskutiert werden. Zudem ist mit einer Einstellungsänderung noch nicht die sich daran anschließende Verhaltenskonsequenz sicher vorherzusagen.

Mittwoch
11. März

14:00 Uhr

GA 03/49

Chair(s)

Karst, Karina

Leichner, Nikolas; Mayer, Anne-Kathrin; Peter, Johannes; Krampen, Günter

Effizienzsteigerung eines Trainings der Informationskompetenz durch adaptierbare Online-Lehrmaterialien

Universitäre Lehrveranstaltungen zur Vermittlung von überfachlichen Schlüsselqualifikationen werden oftmals mit der Begründung außercurricular angeboten, dass sie sich inhaltlich oder zeitlich nicht in bestehende Studienpläne integrieren lassen (vgl. Heierle, 2006). Angesichts der ohnehin hohen wahrgenommenen Belastung der Studierenden (Bargel, Ramm, & Multrus, 2012) sollten außercurriculare Veranstaltungen möglichst zeitökonomisch gestaltet werden. Dies kann bspw. erreicht werden, indem die Lehrmaterialien auf das individuelle Vorwissen und die Bedürfnisse der Lernenden abgestimmt werden. Verglichen mit adaptiven Lernformen, bei denen Lernsysteme aufgrund von Prätests festlegen, welche Inhalte präsentiert werden, können bei adaptierbaren Lernformen die Lernenden selbst Einfluss auf Umfang und Inhalt der Lernmaterialien nehmen (Leutner, 2002). Dies sollte nicht allein eine höhere Lernökonomie bewirken, sondern auch eine Steigerung der Teilnehmermotivation, die ihrerseits den Lernerfolg begünstigen sollte: Zum einen werden jeweils primär individuell relevante Inhalte bearbeitet (vgl. Brophy, 2010). Zum anderen erhöht die Entscheidungsfreiheit über die Lernmaterialien das Autonomieerleben (vgl. Black & Deci, 2000; Garcia & Pintrich, 1996). Allerdings wird auch die Gefahr gesehen, dass die Lernenden nicht angemessen auswählen (Brophy, 2010) oder wichtige Lernmaterialien ignorieren (Reeve, 2009). Daher wird empfohlen, den Lernenden Empfehlungen zu geben (Brophy, 2010).

Die vorliegende Studie geht am Beispiel eines außercurricularen Blended Learning-Trainings der fachlichen Informationskompetenz zwei Fragestellungen nach: (1) Inwieweit machen Teilnehmer von einem Feedback über ihr einschlägiges Vorwissen und den damit verbundenen Empfehlungen für die Auswahl von Lerninhalten Gebrauch? (2) Lässt sich ein positiver Zusammenhang zwischen der Befolgung der Empfehlungen und a) dem objektiven Lerngewinn sowie b) den auf Informationsrecherchen bezogenen Selbstwirksamkeitsüberzeugungen nachweisen?

Methode: Die Stichprobe besteht aus N = 59 fortgeschrittenen Studierenden des BSc-Studiengangs sowie des MSc-Studiengangs Psychologie an der Universität Trier (88 % weiblich) mit einem mittleren Alter von M = 24.85 (SD = 3.69) Jahren.

Vor Trainingsbeginn wurde zur Erfassung der Selbstwirksamkeit ein Fragebogen mit zehn Items verwendet. Zur Ermittlung des einschlägigen Vorwissens wurde ein aus k = 50 Multiple-Choice-Items bestehender Wissenstest (vgl. Leichner, Peter, Mayer, & Krampen, 2013) eingesetzt. Jedem der acht Kapitel der Online-Lehrmaterialien waren mindestens fünf Items zugeordnet; wurden weniger als 66 Prozent der möglichen Punkte erreicht, wurde das Kapitel zur Bearbeitung empfohlen. Die Teilnehmer konnten jedoch frei entscheiden, inwieweit sie diesen Empfehlungen folgten. Die Lehrmaterialien wurden über die Lernplattform Moodle bereitgestellt, sodass die Befolgung der Empfehlungen anhand der Logdateien überprüft werden konnte. Der individualisierten Bearbeitung des Online-Teils folgte ein ca. 180-minütiges Präsenzseminar mit praktischen Übungen und Gruppendiskussionen. Nach Trainingsende wurden erneut der Wissenstest sowie der Selbstwirksamkeitsfragebogen bearbeitet.

Ergebnisse: Das Training erwies sich als effektiv zur Förderung des Recherchewissens ($t[58] = 17.83, p < .01$). Die Überprüfung der Logdateien ergab, dass 11 Teilnehmer (18%) weniger Kapitel bearbeiteten als empfohlen (Gruppe A); 21 Teilnehmer (36%) folgten exakt den Empfehlungen (Gruppe B) und 27 (46%) bearbeiteten mehr Kapitel als empfohlen (Gruppe C). Gruppe A erreichte unter Kontrolle des Ausgangsniveaus einen geringeren Punktezuwachs als Gruppe B ($t[30] = 2.08, p < .05$), welche sich nicht von Gruppe C unterschied ($t[46] = 0.87, ns$). Auch die auf Informationsrecherchen bezogenen Selbstwirksamkeitserwartungen nahmen im Verlauf des Trainings zu ($t[58] = 9.73, p < .01$); zwischen den Gruppen zeigte sich dasselbe Muster wie für den Wissenstest. Zwischen der Veränderung der Selbstwirksamkeitsüberzeugungen und dem Lerngewinn ergab sich ein Zusammenhang von $r = .30 (p < .05)$.

Die Ergebnisse zeigen, dass der überwiegende Teil der Studierenden den Empfehlungen folgte oder sogar mehr als die empfohlenen Kapitel bearbeitete. Zudem führte die adaptierbare Trainingsgestaltung zu einer Steigerung der Trainingseffizienz: Teilnehmer, die weniger Lernstoff bearbeitet hatten als empfohlen, profitierten objektiv wie auch subjektiv weniger von der Trainingsteilnahme, während die Bearbeitung von nicht empfohlenen Kapiteln keinen zusätzlichen Lerngewinn bewirkte.

Mittwoch
11. März

14:00 Uhr

GA 03/49

Chair(s)
Karst, Karina

Bürger, Kathrin; Wosnitza, Marold; Nolden, Philipp; Mihalca, Loredana

Indikatoren von Satisficing-Verhalten in Online Self-Assessments und deren Zusammenhang mit Leistung

Die Datenqualität von Online-Studien rückt immer mehr in den Fokus der Wissenschaft und spielt auch in Bezug auf Daten von Online Self-Assessments (OSA) eine Rolle. OSAs stellen web-basierte und kostenlose Angebote zur Studienorientierung dar.

Die anonyme und unkontrollierte Testsituation im Internet kann sogenanntes Satisficing-Verhalten fördern. Unter Satisficing (Krosnick, 1991) werden Verhaltensweisen von Befragten verstanden, die dazu dienen, den Aufwand beim Beantworten von Fragebögen möglichst gering zu halten. Dies kann sich beispielsweise in beschleunigtem („speeding“) oder geradlinigem („straightlining“) Antwortverhalten widerspiegeln. Satisficing wird durch Personenmerkmale (z.B. kognitive Leistungsfähigkeit, Interesse), Aufgabenmerkmale (z.B. Aufgabenschwierigkeit, Anzahl der Aufgaben) und Situationseffekte (z.B. Umwelteinflüsse) beeinflusst (z.B. Galesic & Bosnjak, 2009; Yan & Tourangeau, 2008).

Datenqualität ist ein Thema für Online-Studien im Allgemeinen und für OSAs im Besonderen. Die prognostische Validität von OSA-Daten hinsichtlich Studienerfolg und -zufriedenheit wurde nachgewiesen (z.B. Hasenberg & Schmidt-Atzert, 2014; Hornke, Wosnitza, & Bürger, 2013). Voraussetzung für belastbare und verlässliche Aussagen über die Eignung einer Person für einen Studiengang sowie für die Vorhersage des Studienerfolgs auf der Basis ihrer Leistung in einem OSA ist jedoch die seriöse, d.h. die ernsthafte und konzentrierte Aufgabebearbeitung.

Fraglich ist, ob und wie seriöse von unseriösen Antworten unterschieden werden können. In der vorliegenden Studie wurden drei Satisficing-Indikatoren untersucht: ‚speeding‘ (Bearbeitungszeit 1.5 Standardabweichungen unter dem Mittelwert), ‚straightlining‘ (Null-Varianz im Antwortverhalten) und Nicht-Ernsthaftigkeit (Beantwortung der Frage ‚Wie ernsthaft hast du die vorherigen Aufgaben bearbeitet?‘ mit ‚nicht ernsthaft‘). In früheren Studien konnte Satisficing anhand dieser Indikatoren nachgewiesen werden (z.B. Barge & Gehlbach, 2012; Greszki, Meyer, & Schoen, 2014). Zudem wurde gezeigt, dass die einzelnen Indikatoren miteinander zusammenhängen.

Hypothesen: Es wurde angenommen, dass alle drei Satisficing-Indikatoren in OSA-Daten beobachtet werden können. Dabei sollten sie moderat zusammenhängen aber dennoch eigenständige Varianzanteile aufweisen.

Zudem sollte Satisficing negativ mit Leistungsergebnissen zusammenhängen. Auch mit Antworten in Persönlichkeitsfragebögen (Motivation, Interesse) wurde ein negativer Zusammenhang vermutet, da diese üblicherweise durch Deckeneffekte rechtsschief verteilt sind.

Methode.

Teilnehmer: Die vorliegende Studie basiert auf OSA-Daten der RWTH Aachen University. Sie umfasst N = 6861 Personen (20% weiblich), die am Studienfeld-OSA Elektrotechnik zwischen November 2009 und April 2013 teilgenommen haben.

Untersuchungsvariablen: Die OSA-Teilnehmer bearbeiteten fünf Module: 1) drei kognitive Module: Matrizen, Prämissen und Mathematik; und b) zwei nicht-kognitive Module: Selbstregulation und Studienmotivation. Letzteres umfasst Skalen zur Selbstwirksamkeit, zur Wettbewerbs- und Erfolgsorientierung, und zum Fachinteresse/Mastery Goal. Die internen Konsistenzen der Skalen liegen zwischen .67 und .90.

Alle drei Satisficing-Indikatoren („speeding“, „straightlining“ und Nicht-Ernsthaftigkeit) wurden modulweise erstellt. Zudem wurde ein Gesamt-Satisficing-Index aus der Summe der 15 Einzel-Indikatoren gebildet (M = 1.15, SD = 2.12, α = .84, Range 0-15).

Ergebnisse: Das ‚straightlining‘ ist der schwächste Satisficing-Indikator mit einer Häufigkeit von unter 3% in allen Modulen. Demgegenüber kann ‚speeding‘ bei 4-17% bzw. Nicht-Ernsthaftigkeit bei 4-27% der Teilnehmer modulweise beobachtet werden. Die modulweisen Satisficing-Indikatoren korrelieren divergent (d.h. unterschiedlichen Indikatoren im selben Modul) schwach bis stark positiv. Die konvergenten Korrelationen (d.h. gleiche Indikatoren in unterschiedlichen Modulen) fallen zum Teil stark aus. Satisficing hängt signifikant negativ mit der Leistung in den kognitiven Modulen bzw. dem Score in den Persönlichkeitsfragebögen zusammen ($-0.20 < r < -0.43$). Der Gesamt-Satisficing-Index weist hinsichtlich der Leistung bzw. der Ergebnisse in allen Modulen eine hohe Validität auf.

Diskussion: Der Großteil der Self-Assessment-Daten kann auf Basis der Satisficing-Indikatoren als seriös beurteilt werden. Die Satisficing-Indikatoren weisen konvergente sowie divergente Validität auf. Der Gesamt-Satisficing-Index ist reliabel und kovariiert signifikant mit Leistung und Persönlichkeitsmerkmalen. Satisficing hat dementsprechend erheblichen Einfluss auf die Skalenqualität. Insofern erscheint es ratsam, Satisficing-Indikatoren in tiefergehenden Analysen weiter zu untersuchen. Hierzu zählt, dass es in zukünftigen Studien zu überprüfen wäre, ob die Häufigkeiten der modulweisen Satisficing-Indikatoren abhängig von der Position der Module innerhalb des OSAs ist. Dies wird aktuell untersucht. Die Ergebnisse werden auf der Tagung präsentiert.

Mittwoch
11. März

14:00 Uhr

GA 03/49

Chair(s)
Karst, Karina

Lernplattformen

Gerholz, Karl-Heinz; Liszt, Verena; Klingsieck, Katrin

Service Learning in der Hochschulbildung – Welchen Beitrag leistet das Lernformat für die Persönlichkeitsentwicklung und was sind didaktische Gestaltungselemente?

Theoretischer Hintergrund: Ziel von Hochschulbildung ist die Förderung einer wissenschaftlich basierten Handlungskompetenz, was neben der Vorbereitung auf zukünftige berufliche Handlungsfelder auch die akademische Persönlichkeitsentwicklung und die Teilhabe an der Zivilgesellschaft mit einschließt (vgl. Gerholz & Sloane 2011, Spoun & Wunderlich 2005, Dewey 1998). Die Lernform des Service Learning, bei welcher problemorientiert fachwissenschaftliche Inhalte des Studiums mit gemeinnützigem Engagement verbunden werden, kann hier Potentiale aufzeigen. Empirische Studien zeigen auf, dass in Service Learning-Arrangements Studierende eine höhere Selbstwirksamkeit und Lernzuwachs erleben sowie über persönliche Einsichten und ihre Engagementeinstellung stärker nachdenken (vgl. u. a. Yorio & Ye 2012, Prentice & Robinson 2010, Reinders 2010). Die hauptsächlich aus dem US-amerikanischen Kontext stammenden Studien untersuchen aber seltener, welche didaktischen Gestaltungselemente hierzu im Zusammenhang stehen. In einem wirtschaftswissenschaftlichen Studiengang wurde das Lernformat Service Learning als ein Methodenmodul für alle Bachelorstudierenden integriert. Die Studierenden bearbeiten ökonomische Problemstellungen bei gemeinnützigen Organisationen. Die Problembearbeitung wird von Inputveranstaltungen, um ein Orientierungswissen zu Methoden aufzubauen, und Beratungsphasen, u. a. zur Reflexion der Service-Erfahrungen, begleitet.

Fragestellung: In der Untersuchung des Moduls waren zwei Fragestellungen leitend: (1) Welchen Einfluss hat das Lernformat Service Learning auf die personale Kompetenzentwicklung der Studierenden und (2) welche didaktischen Gestaltungselemente korrespondieren mit der Kompetenzentwicklung?

Methodik: Das Untersuchungsdesign zum vorgestellten Modul wurde als ein Mixed Method-Ansatz („convergent parallel design“) konstruiert. Die Studierenden des Service Learning-Modul (n = 42) wurden zu Beginn und am Ende der Veranstaltung mit einem standardisierten Fragebogen zu deren Engagementbereitschaft, Selbstwirksamkeit, subjektiven Lernerfolg und Selbstbild befragt (u. a. Mabry 1998); der Fragebogen wurde auch bei einer Kontrollgruppe in einer traditionellen Veranstaltungsformat im Prä-Post-Design eingesetzt (n = 36). Weiterhin wurden problemzentrierte Interviews mit Studierendengruppen im Service Learning-Modul geführt zu Beginn und am Ende geführt. Die Daten wurden varianz- und inhaltsanalytisch ausgewertet.

Ergebnisse: In Bezug auf die Frage des Veranstaltungsformat auf die Förderung der personalen Kompetenzentwicklung zeigen sich beim Engagementbewusstseins, der Selbstwirksamkeit und dem wahrgenommenen subjektiven Lernerfolg signifikante Effekte über die Zeit, allerdings keine Interaktionseffekte hinsichtlich des Veranstaltungstyps. Weiterhin zeigt sich ein signifikanter Trend, dass der subjektive Lernerfolg von Service Learning-Studierenden höher eingeschätzt wird als von Nicht-Service Learning-Studierenden. Hinsichtlich der didaktischen Gestaltung zeigen sich im quantitativen Datenmaterial signifikante Effekte, dass die Förderung des Lernerfolgs, des Selbstbildes und der Selbstwirksamkeit im Zusammenhang mit dem zu bearbeitenden Service Learning-Projekt liegen und auf Basis der qualitativen Daten ergeben sich Hinweise, dass die didaktische Aufbereitung der Problemstellung, der Support seitens der gemeinnützigen Organisation und die Qualität der Beratung zur Veränderung des Selbstbildes und des wahrgenommenen Lernerfolgs korrespondieren. Die Untersuchung zeigt kein experimentelles Design auf, weshalb keine Kausalitäten begründet werden können, allerdings ergeben sich empirische Hinweise, dass die didaktische Ausgestaltung von Service Learning ein elementares Kriterium darstellt, um einen Beitrag zu personaler Kompetenzentwicklung der Studierenden leisten zu können.

Mittwoch
11. März

14:00 Uhr

GA 03/49

Chair(s)
Karst, Karina

Holzberger, Doris; Praetorius, Anna-Katharina; Kunter, Mareike

Unterrichtsqualität: Profile von Lehrkräften und deren Zusammenhang mit der Kompetenzentwicklung von Schülerinnen und Schülern.

Theoretischer Hintergrund: Unterrichtsqualität gilt als eine der zentralen veränderbaren Determinanten von Schulleistung (Hattie, 2009). Zur Beschreibung der Qualität von Unterricht hat sich in den letzten Jahren auf internationaler Ebene ein Modell mit den drei Dimensionen kognitive Aktivierung, Schülerorientierung und Klassenführung etabliert (Baumert et al., 2010; Lipowsky et al., 2009; Pianta & Hamre, 2009). Klieme et al. (2001) weisen darauf hin, dass Lehrkräfte in ihrem Unterricht unterschiedliche Schwerpunkte in Bezug auf die Dimensionen setzen können. Empirische Untersuchungen dazu, wie Lehrkräfte diese drei Dimensionen in ihrem Unterricht verbinden, und ob möglicherweise typische Muster im Sinne individueller Unterrichtsstile (oder Profile) existieren, liegen bislang kaum vor. Analysen im Rahmen von TALIS zeigen, dass Lehrkräfte im Vergleich zwischen den untersuchten Ländern deutlich unterschiedliche Profile aufweisen, sich innerhalb der untersuchten Länder jedoch vornehmlich in der Höhe der Ausprägung der Unterrichtsdimensionen unterscheiden und kaum in Bezug auf die Art ihrer Profile (Vieluf et al., 2012). Aufgrund der fehlenden Teilnahme von Deutschland ist bislang ungeklärt, welche Profile Lehrkräfte in Deutschland aufweisen. Darüber hinaus wurden in TALIS die Unterrichtsdimensionen ausschließlich im Selbstbericht der Lehrkräfte erfasst. Insbesondere ist zum derzeitigen Zeitpunkt unklar, ob und inwiefern sich unterschiedliche Unterrichtsprofile auf die Kompetenzentwicklung von Schülern auswirken.

Fragestellungen: Der vorliegende Beitrag beschäftigt sich daher mit den folgenden zwei Forschungsfragen: (1) Welche Profile lassen sich in Bezug auf die drei Dimensionen der Unterrichtsqualität bei deutschen Mathematik-Lehrkräften aus der Sicht der Schüler(innen) identifizieren? (2) Inwieweit können die identifizierten Unterrichtsprofile die Kompetenzentwicklung der Schüler(innen) vorhersagen?

Methode: Zur Beantwortung der Fragestellung wurden die COACTIV-Daten re-analysiert. Insgesamt nahmen 3483 Schüler(innen) an zwei Messzeitpunkten an der Befragung teil (Ende Klasse 9 und Ende Klasse 10). Als Leistungsmaß der Schüler(innen) wurde ein Mathematiktest durchgeführt. Die Freude an Mathematik der Schüler(innen) wurde mittels sechs Items im Schüler-Fragebogen erfasst. Ungefähr die Hälfte der Schüler(innen) schätzten in einem Fragebogen das Unterrichtshandeln ihrer Mathematik-Lehrkräfte (N=155) auf 11 Subskalen in Bezug auf die drei Unterrichtsdimensionen (drei Skalen für kognitive Aktivierung, sechs für Schülerorientierung, zwei für Klassenführung) ein. Die Schülerleistung, operationalisiert durch PISA-2003-Testwerte, und die Freude an Mathematik zum ersten Messzeitpunkt wurde jeweils als Kontrollvariable herangezogen.

Ergebnisse: Latente Profilanalysen auf der Basis der aggregierten Schülerdaten der Unterrichts-Skalenwerte zum ersten Messzeitpunkt identifizierten drei Lehrerprofile, die sich v. a. in der Höhe der Ausprägung der Skalen des Unterrichtshandelns unterschieden. Der p-Wert des Lo-Mendell-Rubin adjustierten LRT-Tests indizierte die Lösung mit drei Lehrerprofilen (im Vergleich zu einer Lösung mit zwei bzw. vier Profilen). Gruppe 1 (ca. 18 % der Lehrkräfte) wies unterdurchschnittliche Werte auf allen Skalen auf, Gruppe 2 (ca. 42 %) durchschnittliche Werte und Gruppe 3 (ca. 41%) überdurchschnittliche Werte auf. Neben der Höhe der Ausprägungen bestanden die größten Unterschiede in den Profilen im Ausmaß der Skalen zur Schülerorientierung. Hier unterschieden sich die Gruppen in ihren Skalenwerten durchschnittlich ca. um 0.5 (Skala von 1-4). Post-hoc Analysen ergaben signifikante Unterschiede in der Verteilung über die drei Gruppen in Abhängigkeit der Schulform. An nicht-gymnasialen Schulformen (Mittel-, Real- und Gesamtschulen) ist Gruppe 1 (unterdurchschnittliche Werte) häufiger vertreten als an Gymnasien. In den noch ausstehenden Berechnungen wird mittels Mehrebenenanalysen untersucht werden, inwieweit die identifizierten Unterrichtsprofile die Kompetenzentwicklung der Schüler(innen) (Leistung und Freude) vorhersagen können.

Diskussion und Ausblick: Die vorliegenden Profilanalysen ermöglichten zu überprüfen, inwieweit aus der Sicht der Schüler(innen) Lehrkräfte unterschiedliche Schwerpunkte in ihrer Unterrichtsgestaltung setzen (Klieme et al., 2001). Die gefundenen Unterrichtsprofile sprechen jedoch gegen eine differenzierte Profilbildung und stimmen daher mit den Ergebnissen von Vieluf et al. überein, bei denen sich die Unterrichtsprofile innerhalb eines Landes ausschließlich in der Höhe der Ausprägungen unterschieden. Zentral wird schließlich sein, ob die Profile prädiktiv für die Kompetenzentwicklung der Schüler(innen) sein wird.

Mittwoch
11. März

14:00 Uhr

GA 03/142

Chair(s)
Holzberger, Doris

Unterrichtsqualität und ihre Auswirkungen auf Schülerinnen und Schüler

Wurster, Sebastian

Nutzen Lehrkräfte Vergleichsarbeiten, zentrale Abschlussprüfungen sowie interne Evaluation für datenbasierte Unterrichtsentwicklung? Veränderungen im Unterricht und Zusammenhänge zur Schülerleistung.

Theoretischer Hintergrund: Datenbasierten Entscheidungen wird eine wichtige Bedeutung für die Weiterentwicklung schulischer Qualität zugeschrieben. In der Bundesrepublik stehen für Lehrkräfte mit den Ergebnissen aus Vergleichsarbeiten und zentralen Abschlussprüfungen verschiedene routinemäßig erhobene Daten zur Leistung von Schülerinnen und Schülern zur Verfügung. Weiterhin sind Schulen angehalten regelmäßig interne Evaluationen durchzuführen, die eine Bandbreite verschiedenster Informationen liefern können. Visscher und Coe (2003) beschreiben in ihrem Modell zu School Performance Feedback Systems, dass idealtypisch auf die Datenauswertung die Planung und Umsetzung von Entwicklungsmaßnahmen folgt, die wiederum eine Verbesserung der Qualität schulischer Prozesse und des Unterrichts nach sich ziehen sollte, um Einfluss auf die Schülerleistung entfalten zu können. Sie beschreiben als mögliche Einflussfaktoren auf den Nutzungsprozess neben schulischen Merkmalen auch die Ziele und Ausgestaltung der Evaluationsverfahren. In Deutschland sollen Vergleichsarbeiten und interne Evaluationen primär der Schul- und Unterrichtsentwicklung dienen, während zentrale Prüfungen die Hauptfunktion der Zertifizierung von Schülerleistungen erfüllen (Maag Merki, 2009). Diese verschiedenen Ziele und damit verbundenen Informationen können zu Unterschieden in der Nutzung der Datenquellen führen. Der Datennutzung wird in der Forschung zu datenbasierten Entscheidungen (Lai & Schildkamp, 2013) und der Schuleffektivitätsforschung, wie im Dynamic Model of Educational Effectiveness (Creemers & Kyriakides, 2008), ein positiver Zusammenhang zur Schülerleistung zugeschrieben. Erfolgreiche Datennutzung geht mit der Verbesserung von Schul- und Unterrichtsqualität und ausgehend davon mit einer Steigerung von Schülerleistungen einher. Schulinterner Evaluation wird im Modell von Creemers und Kyriakides (2008) eine Schlüsselrolle zugeschrieben, wobei darunter in diesem Fall auch die Rezeption von Leistungsdaten aus Vergleichsarbeiten oder zentralen Prüfungen subsumiert werden kann. In der deutschsprachigen Forschung liegen bereits einige Studien zur Verarbeitung und Nutzung von Vergleichsarbeitsergebnissen vor. Forschung zur Datennutzung aus zentralen Abschlussprüfungen und interner Evaluation sowie zum Zusammenhang zwischen datenbasierter Unterrichtsentwicklung und der Leistung von Schülerinnen und Schülern stellt in Deutschland ein Desiderat dar.

Fragestellung: Ausgehend vom Stand der Forschung und den unterschiedlichen Funktionen der drei Evaluationsverfahren werden in diesem Beitrag folgende Forschungsfragen untersucht: (1) Werden Ergebnisse aus Vergleichsarbeiten, zentralen Abschlussprüfungen und interner Evaluation von Lehrkräften als Grundlage für datenbasierte Unterrichtsentwicklung verwendet? Sind diesbezüglich Unterschiede zwischen den verschiedenen Evaluationsverfahren festzustellen? (2) Gibt es einen Zusammenhang zwischen der Datennutzung und der Leistung der Schülerinnen und Schüler?

Methode: Als Datengrundlage dient der IQB-Ländervergleich 2012, der naturwissenschaftliche und mathematische Kompetenzen von 44584 Schülerinnen und Schülern der 9. Jahrgangsstufe aus 1326 Schulen beschreibt (Pant et al., 2013). Grundlage der hier präsentierten Analysen sind N = 3099 Fragebögen von Lehrkräften aus 1137 Schulen, die bereits Erfahrung mit mindestens einem der drei Evaluationsverfahren haben und zur Datennutzung befragt wurden. Für die Analyse der Zusammenhänge zur Schülerleistung konnten jeweils ca. der Hälfte der befragten Lehrkräfte Ergebnisse aus Leistungstests in den Fächern Mathematik, Biologie, Chemie und Physik zugeordnet werden. Zur Beantwortung der Frage nach dem Zusammenhang zwischen Datennutzung und Schülerleistung werden Multilevelanalysen durchgeführt.

Ergebnisse: Daten aus Vergleichsarbeiten, zentralen Abschlussprüfungen und interner Evaluation werden von einer substanziellen Anzahl an Lehrkräften (26% bis 52%) zur datengestützten Unterrichtsentwicklung in Form der Einführung neuer Methoden, der Entwicklung von Maßnahmen zur Einzelförderung, Fortbildungsteilnahme, Arbeit am schulinternen Curriculum und der Verbesserung kollegialer Kommunikation herangezogen. Jedoch berichtet etwas mehr als die Hälfte der Befragten, dass die erfragten Aktivitäten datenbasierte Unterrichtsentwicklung eher gering ausgeprägt sind bzw. gar nicht stattfinden. ANOVAs zeigen signifikante Verfahrensunterschiede. Die Befragten berichten im Vergleich von einer häufigeren Nutzung von Daten aus interner Evaluation und zentralen Abschlussprüfungen bei der Fortbildungsteilnahme während Arbeit am schulinternen Curriculum eher auf Informationen aus interner Evaluation beruht. Hinsichtlich der anderen Aktivitäten gibt es keine bedeutsamen Differenzen. Multilevelanalysen zeigen, dass nach Kontrolle des sozioökonomischen Hintergrunds, Geschlechts, kognitiver Grundfertigkeiten und Herkunft auf Ebene der Schüler sowie der darauf beruhenden Klassenkomposition in keinem der betrachteten Fächer signifikante Zusammenhänge zwischen datenbasierten Unterrichtsentwicklungsaktivitäten und der Schülerleistung bestehen.

Fauth, Benjamin; Decristan, Jasmin; Rieser, Svenja; Klieme, Eckhard; Büttner, Gerhard

Popularität von Lehrpersonen in der Grundschule: Zusammenhänge mit Lehrermerkmalen und dem Lernerfolg der Schülerinnen und Schüler

Theoretischer Hintergrund: Der Beitrag untersucht Bedingungen und Konsequenzen der Popularität von Lehrpersonen, erhoben durch die Urteile von Schülerinnen und Schülern (SuS) in der Grundschule. Dazu wird geprüft, ob die Popularität von Lehrpersonen eine Erklärungskraft für die Leistungsentwicklung von SuS hat. In einem zweiten Schritt wird untersucht, wie die Popularität mit Merkmalen der professionellen Kompetenz und der Persönlichkeit von Lehrpersonen zusammenhängt.

Eine einfache, aber überzeugende Operationalisierung der Popularität von Lehrpersonen ist das Item „Ich mag meine Lehrerin“ (vgl. z.B. Wagner, 2008). Beim Einsatz solcher Items im Klassenkontext werden Mehrebenenproblematiken relevant (Lüdtke et al., 2009). Während auf der Individualebene ein Aspekt der je individuellen Lehrer-Schüler-Beziehung erfasst wird, wird auf der Klassenebene die Popularität als Merkmal einer Lehrperson erfasst.

In einer Studie von Montalvo et al. (2007) berichten SuS über eine größere Anstrengungsbereitschaft und bessere Schulnoten, wenn sie von Lehrpersonen unterrichtet werden, die sie mögen (vgl. auch Wagner, 2008). Zur Frage, was Lehrpersonen beliebt macht, diskutiert Wagner (2008) Zusammenhänge mit der Unterrichtsqualität. Weitere Einflussgrößen für die Popularität könnten in Merkmalen der professionellen Kompetenz von Lehrpersonen (Kunter et al., 2013) und in allgemeinen Persönlichkeitseigenschaften (Klusmann et al., 2009; Mayr, 2011) liegen.

Forschungsfragen: (1) Hat die Beliebtheit einer Lehrperson prädiktive Kraft für die Leistungsentwicklung von SuS? (2) Finden sich die Effekte auf der Individual- oder auf der Klassenebene? (3) Wie hängt die Qualität des Unterrichts einer Lehrperson mit ihrer Beliebtheit zusammen? (4) Wie hängen der Enthusiasmus von Lehrpersonen, ihre konstruktivistischen Überzeugungen (als Merkmale der professionellen Kompetenz) und ihre Offenheit für neue Erfahrungen (als allgemeines Persönlichkeitsmerkmal) mit ihrer Popularität zusammen?

Methoden: Die Studie basiert auf Daten von 1070 SuS aus 54 Klassen (Durchschnittsalter 8,8 Jahre, 49% weiblich) und ihren jeweiligen Lehrpersonen (Durchschnittsalter 42.6 Jahre, 86% weiblich).

In einem längsschnittlichen Design wurden zwei Unterrichtseinheiten in den Sachunterricht der dritten Jahrgangsstufe implementiert. Das konzeptuelle Verständnis der SuS zum Thema der Unterrichtseinheiten wurde mittels standardisierter Leistungstests erfasst (Prä-Test: 13 Items; Post-Test: 16 Items). Als Kovariaten wurden zusätzlich kognitive Grundfähigkeiten (CFT 20-R) und naturwissenschaftliche Kompetenz (adaptiert nach TIMSS 2007) erfasst.

Die Schülerurteile zur Popularität der Lehrperson wurden vor den Unterrichtseinheiten mit einer Drei-Item-Skala erfasst (Cronbachs alpha = .92; ICC1 = .15; ICC2 = .74).

Der Lehrerfragebogen umfasste 6 Items zum Enthusiasmus (Kunter, et al., 2008), 6 Items zu konstruktivistischen Überzeugungen (Staub & Stern, 2002; Warwas et al., 2011) und 12 Items zu Offenheit für neue Erfahrungen (Borkenau und Ostendorf, 1991).

Die Unterrichtsqualität wurde anhand von drei Basisdimensionen (kognitive Aktivierung, unterstützendes Klima, strukturierte Klassenführung; Klieme, et al., 2009) von externen Beobachtern eingeschätzt (Beurteilerübereinstimmung (ICC): .72–.81).

Ergebnisse: Zu den Fragestellungen (1) und (2) wurden Mehrebenen-Regressionsanalysen mit dem konzeptuellen Verständnis im Post-Test als abhängiger Variable und den oben genannten Kovariaten berechnet. Die Popularität der Lehrperson zeigte sich als signifikanter Prädiktor auf der Klassenebene, nicht jedoch auf der Individualebene.

Zu Fragestellung (3) wurden Korrelationen zwischen den Unterrichtsqualitätsdimensionen und der Popularität der Lehrperson auf der Klassenebene berechnet. Es zeigten sich bedeutsame Zusammenhänge nur mit dem unterstützenden Klima.

Zu Fragestellung (4) wurden Mehrebenen-Regressionsanalysen mit den Schülerurteilen zur Popularität der Lehrperson als abhängiger Variable berechnet. Als Kovariaten auf der Individualebene wurden das Geschlecht und die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler kontrolliert. Im Bereich der Lehrermerkmale zeigten sich signifikante Effekte für Enthusiasmus und Offenheit, nicht jedoch für konstruktivistische Überzeugungen.

Für die Entstehung von Popularität scheinen eher allgemeine, fachunspezifische als fachdidaktisch spezifische Merkmale von Unterricht und von Lehrpersonen relevant zu sein. Das Konstrukt kann auf der Klassenebene reliabel erfasst werden, wie der ICC2 zeigt. Für die Leistungsentwicklung ist dieses Klassenaggregat als Eigenschaft von Lehrpersonen relevant, nicht jedoch die individuelle Schüler-Lehrer-Beziehung.

Mittwoch
11. März

14:00 Uhr

GA 03/142

Chair(s)
Holzberger, Doris

Unterrichtsqualität und ihre Auswirkungen auf Schülerinnen und Schüler

Gärtner, Holger

Stabilität von Unterrichtsqualität in der Schülerwahrnehmung

Im Zusammenhang mit den zahlreichen Reformen des letzten Jahrzehnts werden in Schule und Unterricht zunehmend (wissenschaftlich fundierte) Evaluationsinstrumente eingesetzt, um empirische Grundlagen zur Einschätzung von Schul- und Unterrichtsqualität zu gewinnen.

Beispiele für die Einschätzung von Unterrichtsqualität sind z.B. die Schulinspektion (die mittels Beobachtung von zumeist Unterrichtssequenzen Aussagen über die Unterrichtsqualität an einer Schule trifft), die Durchführung von Schülerfeedbacks durch Lehrkräfte, die Unterrichtsbeobachtung und das anschließende Feedback durch Schulleitungen oder die kollegiale Hospitation von Lehrkräften untereinander. All diese Formen der Erfassung von Unterrichtsqualität sollen zu einer datengestützten Unterrichtsentwicklung beitragen. Vor diesem Hintergrund gewinnt die Frage der Stabilität von Unterrichtshandeln an Bedeutung, d.h. ob die jeweils gemessenen Aspekte von Unterrichtsqualität stabil genug sind, so dass die gewonnenen Erkenntnisse als Grundlage längerfristiger Entscheidungsprozesse dienen können.

Diese Studie geht der Frage nach der Stabilität von Unterrichtsqualität anhand von Schülerwahrnehmungen zum Unterricht nach. Zwei Aspekte werden hierbei untersucht: Zum einen die Stabilität unterschiedlicher Aspekte von Unterrichtsqualität, zum anderen, wie ähnlich der Unterricht ein und derselben Lehrkraft in unterschiedlichen Befragungskontexten wahrgenommen wird.

Um diese Forschungsfragen zu überprüfen, wird auf Daten des Selbstevaluationsportals zurückgegriffen (Gärtner, 2014). Das Selbstevaluationsportal ist ein Online-Angebot für Lehrkräfte, mit dem sie ein Schülerfeedback organisieren können, indem sie ihre Schülerinnen und Schüler zur Wahrnehmung des Unterrichts befragen. Bis zum 1.9.2014 haben Lehrkräfte insgesamt 11.877 Befragungen mit dem Selbstevaluationsportal durchgeführt. 1.369 Lehrkräfte haben mehrere Befragungen in unterschiedlichen Kontexten (unterschiedlichen Klassen, Fächern oder Jahrgangsstufen) durchgeführt. In diesen Befragungen werden bis zu 16 Aspekte von Unterrichtsqualität erfragt (angelehnt an das QUAIT-Modell; Gärtner, 2010; Slavin, 1995; Ditton, 2002). Sowohl die interne Konsistenz als auch die ICC2 als Maß der Reliabilität der auf Klassenebene aggregierten Einschätzung der Schüler/-innen sind mehrheitlich gut (Gärtner, 2010). Als Maß für die Ähnlichkeit dienen Korrelationskoeffizienten zwischen Schülerwahrnehmungen auf Klassenebene. Als Referenzgröße werden die Ergebnisse von Ditton und Arnold (2004) herangezogen, die für Schülerwahrnehmungen derselben Lehrkräfte in denselben Klassen zu Beginn und zum Ende eines Schuljahres Zusammenhänge von .74 (Unterrichtsformen) bis .81 (Klassenmanagement) berichten.

Die insgesamt 1.369 Lehrkräfte, die zwei Schülerbefragungen durchgeführt haben, haben dies zu 59% zweimal zum gleichen Fach getan und zu 41% in dergleichen Jahrgangsstufe. Im Durchschnitt lagen 136 Tage zwischen den Befragungen. Die Höhe des Zusammenhangs zwischen Erst- und Zweitbefragung variiert in der Gesamtstichprobe zwischen den Aspekten von $r = .38$ für Leistungserwartungen und $r = .61$ für Differenzierung. In ähnlichen Kontexten sind die Zusammenhänge stärker (im gleichen Fach mittleres $r = .52$ vs. $r = .41$ bei verschiedenen Fächern; in der gleichen Jahrgangsstufe mittleres $r = .56$ vs. $r = .41$ in verschiedenen Jahrgangsstufen). Handelt es sich um zwei Befragungen derselben Klasse zum selben Fach, steigt die mittlere Korrelation auf .63. Diese Befragungskonstellation trifft jedoch nur für 25 Lehrkräfte zu. Insgesamt 279 Lehrkräfte haben vier Schülerbefragungen durchgeführt. Zwischen der ersten und der letzten Befragung lagen durchschnittlich 292 Tage. Zwischen diesen Befragungen variiert die Höhe des Zusammenhangs zwischen .29 für Angsfreie Atmosphäre und .56 für Differenzierung. Auch hier sind die Zusammenhänge zwischen Befragungen in ähnlichen Kontexten höher als in unterschiedlichen.

Die bisherigen Ergebnisse verdeutlichen, dass mit zunehmender Ähnlichkeit der Befragungskontexte die Größe des Zusammenhangs zwischen den Befragungen zunimmt. Des Weiteren zeigen sich deutliche Unterschiede in der Stabilität verschiedener Aspekte von Unterrichtsqualität. Maßnahmen der Differenzierung scheinen in unterschiedlichen Klassenkontexten ähnlich ausgeprägt wahrgenommen zu werden, während dies für die Wahrnehmung von Leistungserwartungen nicht der Fall ist. Insgesamt unterstützen die Befunde die Annahme, eine Schülerbefragung zur Wahrnehmung des Unterrichts als Entscheidungsgrundlage für Unterrichtsentwicklung in dieser Klasse zu nutzen. Die Befunde zeigen jedoch auch, dass die Stabilität vom jeweils erfragten Aspekt abhängt.

Mittwoch
11. März

14:00 Uhr

GA 03/142

Chair(s)

Holzberger, Doris

Bender, Elena; Niclas, Schaper

Relevante Überzeugungen von Experten und Novizen im Informatiklehramt

Als einer der wichtigsten Einflussfaktoren auf das Unterrichtsverhalten gelten Überzeugungen von Lehrkräften (z. B. Staub und Stern, 2002). Lehrerüberzeugungen beinhalten die Vorstellungen über schul- und unterrichtsbezogene Phänomene und Prozesse mit einer bewertenden Komponente (Kunter & Pohlmann, 2009). Sie beeinflussen den Umgang mit den Schülerinnen und Schülern sowie Zielvorstellungen und Handlungspläne im Unterricht (Köller, Baumert & Neubrand, 2000). Überzeugungen gelten als fachspezifisch (Koballa, Gräber, Coleman & Kemp, 2000). Studien aus den Fächern Mathematik und den Naturwissenschaften berücksichtigen insbesondere lehr-lerntheoretische und epistemologische Überzeugungen (z. B. Kunter et al., 2011; Riese & Reinhold, 2008). Kunter et al. (2013) konnten einen Zusammenhang zwischen konstruktivistischen Überzeugungen von Lehrkräften und Schülerleistungen nachweisen. Welche epistemologischen und lehr-lerntheoretischen Überzeugungen für das Fach Informatik relevant sind, wurde erstmalig in einer Interviewstudie mit Lehrkräften und Expertinnen und Experten der Lehramtsausbildung untersucht (Bender et al., 2014, eingereicht). Das Fach Informatik ist dadurch gekennzeichnet, dass eine Vielzahl an didaktischen Konzepten, Zielsetzungen und Ausrichtungen existiert, wodurch ein Mangel an Orientierung für die Lehrkräfte zu deren Unterrichtsplanung und -gestaltung entsteht (Mesaros & Diethelm, 2014), was vor allem eine Schwierigkeit für angehende Lehrkräfte des Fachs darstellt. Befunde zeigen, dass sich Annahmen zu relevanten Überzeugungen von Experten eines Fachs zu denen Lehramtsstudierender unterscheiden. Tendenziell haben Novizen weniger differenzierte Vorstellungen darüber, wie Wissen in einem Unterrichtsfach aufgebaut ist (Sosu & Gray, 2012). Außerdem findet der Perspektivwechsel von der Rolle des Lernenden zur Rolle des Lehrenden erst nach dem Studium statt (Jones, Brown, Hanley, McNamara 2000).

Wie sich die Annahmen zu den relevanten Überzeugungen Lehramtsstudierender in Informatik von denen erfahrener Lehrkräfte und Experten der Lehramtsausbildung unterscheiden, ist unklar. Ziel dieses Beitrags ist zu analysieren, welche Unterschiede und Gemeinsamkeiten die Annahmen der Novizen zu den relevanten fachspezifischen Überzeugungen der Experten aufweisen.

Dazu wurden auf der Basis eines theoretisch hergeleiteten Kategoriensystems zu den relevanten professionellen Überzeugungen Analysen in Form von Interviewstudien durchgeführt. In einer Interviewstudie mit 22 Expert(inne)n der Lehramtsausbildung (11 erfahrene Lehrkräfte des Fachs, 11 Expert(inne)n der Lehramtsausbildung) wurden relevante epistemologische und lehr-lerntheoretische Überzeugungen im Rahmen der Kompetenzmodellierung identifiziert und ausformuliert (Bender et al., 2014, eingereicht). Für den Vergleich der Annahmen der Expertinnen und Experten mit den Annahmen der Novizen wurde eine Interviewstudie mit 11 Studierenden im Informatiklehramt zum Ende ihres Lehramtsstudiums in NRW und Bayern durchgeführt. Wie auch in der Expertenstudie wurde die Befragung in Anlehnung an die an Critical Incident Technique vorgenommen, unter Verwendung dafür entwickelter (Problem)Szenarien zur Erfassung der relevanten Überzeugungen. Die Analyse der Interviews wurde mithilfe der strukturierenden inhaltsanalytischen Auswertungstechnik nach Mayring (2010) durchgeführt, wobei sowohl deduktive Techniken zur Strukturierung des Materials gemäß der theoriegeleiteten Kategorien, als auch induktive Techniken zur Zusammenfassung der Aussagen innerhalb der Kategorien verwendet wurden. Die Interrater-Reliabilität für das Kategoriensystem ist mit Cohen's (1960) Kappa = .86 ($p < .0.001$) für die Expertenstudie und Kappa = .62 ($p < .0.001$) für die Studierendeninterviews zufriedenstellend.

In den Ergebnissen zeigt sich, dass die Expertinnen und Experten eher eine prozesshafte Sicht auf das Fach Informatik haben und das Fach in seinen Zusammenhängen sehen, bestehend aus übergeordneten Prinzipien und Strategien. Die Lehramtsstudierenden halten tendenziell Vorstellungen von Informatik für hilfreich, die auf einen konstanten Kern des Fachs hindeuten, der den rapiden technologischen Wandel überdauert. In den lehr-lerntheoretischen Überzeugungen teilen beide Gruppen Annahmen zu einer förderlichen konstruktivistischen Sicht in Bezug auf das Lehren und Lernen in Informatik. Allerdings schildern die Expert(inne)n die Überzeugungen in differenzierterer Art und Weise und beschreiben die Zusammenhänge lehr-lerntheoretischer und epistemologischer Sichtweisen. Außerdem gibt die Studie Hinweise auf einen noch nicht vollständig vollzogenen Transitionsprozess vom Lernenden zum Lehrenden in Informatik auf Seiten der Studierenden. Die Studierenden haben mehr Schwierigkeiten, eine normative Sicht zu förderlichen Einstellungen einzunehmen.

Mittwoch
11. März

14:00 Uhr

GA 03/149

Ott, Natalie; Malone, Sarah; Vogel, Markus; Brünken, Roland

ICH SEHE WAS, WAS DU NICHT SIEHST- Expertiseabhängige Informationsverarbeitung mathematischer Notation

Zu den grundlegendsten Erkenntnissen der Forschung zum Lernen mit multiplen Repräsentationen gehört, dass die Lernleistung abhängig ist vom gewählten Repräsentationsformat, mit dem Lerninhalte präsentiert werden (Brünken et al., 2005), bei der Nutzung multipler Repräsentationen ist außerdem die domänenspezifische Expertise zentral (Seufert et al. 2007; Sweller et al., 2003). Gängige Modelle zum Lernen mit multiplen Repräsentationen beschreiben die Informationsverarbeitung textuell/bildlich präsentierter Information (Mayer, 2001; Schnotz & Bannert, 1999; Schnotz et al., 2001). Weitgehend ungeklärt ist, inwiefern diese Modelle auf mathematische Notationsformate übertragbar sind, in denen neben grafischen Darstellungen multiple symbolische Repräsentationen (sprachlich/numerisch/formal) verwendet werden.

Wir untersuchen das Lernen mit multiplen Repräsentationen in der Mathematik unter Verwendung eines Expertiseparadigmas (Expert-Performance-Approach; Ericsson & Smith, 1991). Die Expertiseforschung zeigt, dass Experten im Aufgabenmaterial ihrer spezifischen Domäne wesentliche Informationen besser von unwesentlichen Informationen unterscheiden können als Novizen, da Experten eher nach übergreifenden Mustern suchen, während Novizen Informationen eher aus einer lokal begrenzten Suche erhalten (Spillich et al., 1979). Befunde der Expertiseforschung gelten als über verschiedene Domänen generalisierbar (Chi et al., 1988).

In der vorliegenden Studie wurde geprüft, ob sich die Ergebnisse bisheriger Expertiseforschung für den Umgang mit Formeln replizieren lassen. Zentral war die Frage, ob die Expertise darin begründet liegt, dass Experten mit Repräsentationswechseln zwischen multiplen Repräsentationen (z.B. Text/Formel) besser umgehen können als Novizen. Wir nehmen dabei an, dass die Experten in ihrer Leistung weniger von der Repräsentationsform abhängig sind (Hypothese 1) und dass den Experten der Repräsentationswechsel besser gelingt als den Novizen (Hypothese 2). Als mathematischer Gegenstandsbereich diente die Aussagenlogik, in der sich mathematisch-logische Ausdrücke in unterschiedlichen, inhaltsäquivalenten symbolischen Repräsentationsformaten darstellen lassen.

An der Studie nahmen 93 Studierende der Mathematikeinführungsvorlesung (Novizen) und 110 Studierende der mathematischen Vertiefungsseminare (Experten) teil. Den Versuchspersonen wurden 47 Teilaufgaben aus der Aussagenlogik mit gängigen Formeln (Bedingung 1) und versprachlichten Formeln bzw. alltagssprachliche logische Formulierungen (Bedingung 2). Die jeweils vorgelegte Repräsentationsform (formal/sprachlich) musste in die komplementäre transformiert werden (AV1: Übersetzen); außerdem sollten die vorgelegten logischen Aussagen auf ihren mathematischen Gehalt hin überprüft werden (richtig/falsch) (AV2: Bewerten).

Die Datenanalyse erfolgt für beide AVs getrennt mittels 2-faktorieller MV-ANOVA (Between-Faktor: Expertise; Within-Faktor: Repräsentationsform). Analog zur bisherigen Forschung zeigte sich für die Repräsentationsform ein Experten-Novizen-Unterschied für den Umgang mit diesen spezifisch-nichtredundanten Repräsentationen. Die Repräsentationsform (formal/alltagssprachlich) hat erwartungsgemäß keinen Einfluss auf die Leistung der Experten, wohl aber der Novizen. Wir finden zwei Haupteffekte (Expertise: $F(1,201)=61,600$; $p \leq .001$; $\eta_p^2 = .235$; Repräsentationsformat: $F(1,201)=21,965$; $p \leq .001$; $\eta_p^2 = .099$) sowie einen Wechselwirkungseffekt (Repräsentationsformat*Expertise: $F(1,201)=18,290$; $p \leq .001$; $\eta_p^2 = .083$). Dies deutet darauf hin, dass Experten weniger abhängig von der Repräsentationsform sind als Novizen. Ein expertise reversal effect konnte nicht gefunden werden.

Für die Gesamtleistung zeigen sich ein Expertiseeffekt sowie ein Effekt des Repräsentationswechsels. Erwartungsgemäß waren Experten den Novizen generell überlegen. Weiterhin zeigen sich bessere Leistungen bei Aufgaben, die einen Repräsentationswechsel erfordern als bei Aufgaben ohne Repräsentationswechsel. Mittels einer MANCOVA (AVs: Übersetzen/Bewerten; Fester Faktor: Expertise; Kovariaten: alltagssprachlich/formalisiert) konnten wir auch hier Haupteffekte finden für Expertise ($F(1,201)=9,219$; $p \leq .001$; $\eta_p^2 = .085$) und die beiden Repräsentationsformen „alltagssprachlich“ ($F(1,201)=1359,564$; $p \leq .001$; $\eta_p^2 = 0,932$) und „formalisiert“ ($F(1,201)=8972,132$; $p \leq .001$; $\eta_p^2 = .989$). In der anschließenden ANOVA erscheinen signifikante Gruppenunterschiede (Expertise Übersetzen: $F = 18,520$, $p < 0,001$, $\eta^2 = 0,085$; Bewerten: $F = 17,301$, $p < 0,001$, $\eta^2 = 0,080$; Alltagssprachlich Übersetzen: $F = 287,537$, $p < 0,001$, $\eta^2 = 0,591$; Bewerten: $F = 5,648$, $p = 0,018$, $\eta^2 = 0,028$; Formalisiert Übersetzen: $F = 425,072$, $p < 0,001$, $\eta^2 = 0,681$; Bewerten: $F = 306,498$, $p < 0,001$, $\eta^2 = 0,606$)

In diesem Kooperationsprojekt zwischen empirischer Bildungsforschung und Mathematikdidaktik untersuchen wir als nächstes einerseits die Frage, welches die ideale Repräsentationsform für Novizen darstellt und andererseits, wie der Umgang damit erfolgreich trainiert werden kann. Dazu werden Handlungsprofile der mathematischen Experten und Novizen ermittelt und in Trainingsstudien untersucht, inwiefern die Novizen vom Vorgehen der Experten profitieren können.

Mittwoch
11. März

14:00 Uhr

GA 03/149

Thomm, Eva; Barzilai, Sarit; Bromme, Rainer

Warum Widersprechen Sich Experten? Eine Untersuchung des Verständnisses von Wissenschaftskontroversen in Biologie und Geschichte

Informieren sich Laien über wissenschaftliche Themen, stoßen sie oftmals auf widersprüchliche Expertenaussagen. Für die Bewältigung solcher Konflikte ist es wichtig zu verstehen, warum und wie Wissenschaftler zu verschiedenen Aussagen über das gleiche Thema gelangen können. Solche subjektiven Erklärungen ermöglichen es Laien, die Gültigkeit der Aussagen sowie die Vertrauenswürdigkeit der Experten zu bewerten und den Konflikt subjektiv aufzulösen. Studien zeigen, dass Laien spezifische Annahmen über die Ursachen von Expertenkontroversen haben und zwischen inhaltsbezogenen Konflikterklärungen (d.h. durch Unterschiede im Forschungsprozess und thematische Komplexität des Untersuchungsgegenstandes) sowie personenbezogenen Konflikterklärungen (d.h. durch Unterschiede in der Kompetenz und Motivation der Experten) differenzieren (Kajanne & Pirttilä-Backman, 1999; Thomm et al., eingereicht).

Konflikterklärungen wurden vornehmlich im Kontext naturwissenschaftlicher Themen untersucht. Wissenschaftsdisziplinen unterscheiden sich jedoch hinsichtlich der Struktur des Wissens, der Art und Weise der Wissensgenerierung, sowie in den epistemischen Normen, die als gültig erachtet werden. Disziplinspezifische Unterschiede in der Wahrnehmung der Konstruktion und Rechtfertigung wissenschaftlichen Wissens wurden bereits eingehend belegt (z.B. Buehl & Alexander, 2001; Hofer, 2000). Folglich ist es möglich, dass Laien in Abhängigkeit des disziplinären Kontexts unterschiedliche Erklärungen für Expertenkontroversen heranziehen. Hierbei mag die individuelle epistemische Perspektive hinsichtlich der Natur, Quelle und Rechtfertigung wissenschaftlichen Wissens ebenfalls eine Rolle spielen und die Interpretation beeinflussen. Diese Studie untersuchte daher die Frage, in welcher Weise der disziplinäre Kontext einer Kontroverse sowie die individuelle epistemische Perspektive bedeutsam dafür sind, wie Laien Kontroversen zwischen Wissenschaftlern erklären.

Die Studie folgte einem Messwiederholungsdesign. 177 Studierende lasen zwei Kontroversen (Geschichte, Biologie). Je Kontroverse wurde Konflikterklärung sowie die kontextspezifische epistemische Perspektive gemessen. Konflikterklärung wurde mittels des ExplainingConflictingScientificClaims Fragebogens (ECSC) (Thomm et al., eingereicht) erfasst. Die Probanden erhalten eine Liste unterschiedlicher Erklärungen und beurteilen, in welchem Maße sie die einzelne Erklärung als relevant für die spezifische Kontroverse erachten. Es werden vier Konflikterklärungen erfasst: Unterschiede im Forschungsprozess, thematische Komplexität, Unterschiede in der Kompetenz sowie durch Motivationen der Experten. Die interne Konsistenz der Skalen war $\alpha = .62-.91$. Epistemische Perspektive wurde mittels des Epistemic Thinking Assessments (ETA) nach Barzilai und Weinstock (eingereicht) erfasst. Die Probanden lesen Fragen bezüglich verschiedener epistemischer Dimensionen (z.B. Sicherheit von Wissen). Je Frage werden drei Antworten präsentiert, die typische absolutistische, multiplizistische und evaluativistische Sichtweisen widerspiegeln. Die Probanden beurteilen für jede Antwort, inwieweit sie dieser zustimmen. Die interne Konsistenz der Skalen (Absolutismus, Multiplizismus, Evaluativismus) war $\alpha = .77-.88$.

Da der ECSC für Kontroversen in den Naturwissenschaften konzipiert wurde, wurde durch explorative Faktorenanalysen überprüft, ob sich die gefundenen Konflikterklärungen im Geschichtskontext replizieren lassen. Die Analysen ergaben eine Drei-Faktoren-Lösung für die Geschichtskontroverse. Thematische Komplexität sowie Motivation erwiesen sich als stabile Erklärungen in beiden Kontexten. Im Kontrast hierzu bildeten Items, die Unterschiede im Forschungsprozess und Kompetenzunterschiede erfassten, einen gemeinsamen Faktor in der Geschichtskontroverse, während diese Erklärungen in der Biologiekontroverse differenziert wurden. Varianzanalysen sowie Post-Hoc-Analysen zeigten, dass in Biologie die Kontroverse insbesondere auf inhaltsbezogenen Ursachen (Forschungsprozess und Themenkomplexität) und weniger auf personenbezogene Ursachen (Motivation und Kompetenzunterschiede) attribuiert wurde, hingegen in Geschichte der Konflikt durch inhaltsbezogene Erklärungen (v.a. Themenkomplexität) und personenbezogenen Erklärungen (v.a. Motivation) begründet wurde.

Korrelationen belegten substantielle, kontextkonsistente Zusammenhänge zwischen spezifischen Konflikterklärungen und epistemischer Perspektive. Absolutismus und Konflikterklärung korrelierten nicht. Multiplizismus korrelierte positiv mit der Konflikterklärung Motivation in beiden Kontexten, sowie positiv mit der Erklärung durch Kompetenzunterschied in der Biologiekontroverse und mit dem Erklärungsfaktor Prozess- und Kompetenzunterschiede in der Geschichtskontroverse. Kontextübergreifend war Evaluativismus positiv mit thematischer Komplexität assoziiert.

Disziplinärer Kontext sowie epistemische Perspektive beeinflussen bedeutsam die Erklärung und Interpretation von Expertenkontroversen. Besonders die Rolle und die angenommene Subjektivität, die Wissenschaftlern in der Wissenskonstruktion zu geschrieben wird, werden in den Kontexten unterschiedlich interpretiert. In der Biologiekontroverse wird der Einfluss von Kompetenz und Motivation als gering erachtet, hingegen werden in der Geschichtskontroverse personenbezogene Erklärungen und somit der Einfluss der subjektiven Sichtweise hervorgehoben.

Mittwoch
11. März

14:00 Uhr

GA 03/149

Chair(s)
Bender, Elena

Beltermann, Esther; Wessels, Insa; Kollar, Ingo; Fischer, Martin R.

Identifikation expertisebezogener Unterschiede im Verstehen von Stationsvisiten zwischen Medizinstudierenden und Ärzten – eine Analyse von Skripten

Theoretischer Hintergrund: Trotz der Bedeutsamkeit von Stationsvisiten für den Arbeitsalltag von Ärzten bereiten medizinische Curricula Studierende nicht ausreichend auf diese Tätigkeit vor (Norgaard, Ringsted & Dolmanns, 2004). Somit kann erwartet werden, dass sich Medizinstudierende aufgrund ihrer geringeren Erfahrung mit Stationsvisiten von erfahrenen Ärzten hinsichtlich ihrer visitenbezogenen kognitiven Skripts (Schank, 1999) unterscheiden. Diese Annahme soll im Kontext dieses Beitrags untersucht werden.

Skripts beschreiben dynamische Gedächtnisrepräsentationen von Ereignissequenzen. Im Kontext dieses Beitrags interessieren drei Aspekte von Visitenskripten: (a) ihre Struktur (Fischer et al., 2013), d.h. die Anzahl von Szenen (Wissen über Phasen) und Scriptlets (Wissen über Aktivitäten, die während dieser Phasen gezeigt werden), die Medizinstudierende und Ärzte als typisch für Visiten erachten; (b) ihr Inhalt, d.h. inwiefern sich die Visitenskripts von Medizinstudierenden und Ärzten hinsichtlich der Häufigkeit von Aktivitäten unterscheiden, die unterschiedlichen Inhaltsaspekten zuzuordnen sind (medizinisch, sozial, administrativ, lehr-lernbezogen, unspezifisch; Walton & Steinert, 2010); und (c) ihr wahrgenommenes Potenzial zur Wissenskonstruktion, d.h. inwieweit während Visiten passive, aktive, konstruktive oder interaktive Aktivitäten erwartet werden, wobei nach Chi (2009) insbesondere interaktiven und konstruktiven Aktivitäten ein hohes Potenzial für die Wissenskonstruktion zugeschrieben wird.

Fragestellung: Inwiefern unterscheiden sich Visitenskripts von Medizinstudierenden und Ärzten unterschiedlicher Expertisestufen in Bezug auf (1) die Anzahl von Szenen und Scriptlets? Wir nehmen an, dass Experten eine höherwertige Organisation ihres Situationswissens haben, das sich in einer geringeren Anzahl von Szenen und differenzierteren Abbildung von Scriptlets zeigt (Nivelstein et al., 2008); (2) die Häufigkeit der Nennung von Aktivitäten aus unterschiedlichen inhaltlichen Themenbereichen? Es wird erwartet, dass Studierende signifikant weniger medizinische und administrative Tätigkeiten nennen als erfahrenere Ärzte; (3) in Bezug auf das diesen Aktivitäten zuschreibbare Potenzial zur Wissenskonstruktion? Hier wird erwartet, dass Studierende häufiger als Ärzte passive und aktive Aktivitäten nennen, während für konstruktive und interaktive Aktivitäten ein umgekehrtes Muster anzunehmen ist.

Methoden: Es wurde eine Interviewstudie mit N = 50 Medizinstudierenden und ÄrztInnen der Inneren Medizin am Klinikum der Universität München durchgeführt: 15 Medizinstudierende mit durchschnittlich M=3.00 Jahren klinischer Erfahrung (SD=.00), elf Medizinstudierende im Praktischen Jahr (M=6.18 Jahre, SD=.60), zwölf Assistenzärzte (M=8.53 Jahre, SD=1.12) und 12 Stations- bzw. Oberärzte (M=19.50 Jahre, SD=9.94). Die Interviews wurden auf Basis der Struktur-lege-Technik (Scheele, & Groeben, 1988) durchgeführt; während der Interviews notierte der Interviewer die von den TeilnehmerInnen als typisch genannten Szenen und Scriptlets auf farblich kodierte Karten. Anschließend wurde die Struktur mit dem Interviewer konsentiert. Die resultierenden Strukturen wurden zunächst nach Häufigkeit der Skriptkomponenten (Fischer et al., 2013) kodiert. Anschließend wurden sie auf unterschiedliche inhaltliche Themenbereiche (Walton & Steinert, 2010) und ihr Potenzial zur Wissenskonstruktion (Chi, 2009) hin analysiert. Die Interrater-Reliabilität war auf allen drei Ebenen sehr gut (Cohens Kappa: .79 < K < .99).

Ergebnisse und Diskussion: Bezüglich der Häufigkeit genannter Szenen ergaben sich keine expertisebezogenen Unterschiede Allerdings berichteten Stations- und Oberärzte sowie Medizinstudierende signifikant mehr Scriptlets als Medizinstudierende im Praktischen Jahr ($Z=-1.921$, $p=.056$) und Assistenzärzte ($Z=-2.147$, $p=.033$). Auf der inhaltlichen Ebene zeigte sich, dass Stations- und Oberärzte und Medizinstudierende die meisten lehr-lernbezogenen Tätigkeiten und Medizinstudierende zusätzlich signifikant mehr unspezifische Aktivitäten als erfahrenere Personen nannten. Die Analyse hinsichtlich des Potentials zur Wissenskonstruktion zeigte signifikante Unterschiede in der Nennung aktiver ($H(3) = 8.11$, $p = .04$) und passiver ($H(3) = 15.607$, $p = .001$) Tätigkeiten. Assistenzärzte berichteten signifikant mehr aktive Aktivitäten und Medizinstudierende mehr passive Aktivitäten, insbesondere für ihre eigene Rolle ($H(3) = 16.665$, $p = 0.001$), als alle anderen Gruppen.

Während auf struktureller Ebene also nur wenige expertisebedingte Unterschiede in Skripten vorliegen, zeigt sich auf inhaltlicher Ebene, dass Medizinstudierende häufig irrelevante Aktivitäten als visitentypisch ansehen. Zusätzlich verstehen sie insbesondere ihre eigene Rolle als sehr passiv, weswegen das Potenzial von Stationsvisiten als Lehr- bzw. Lernsetting möglicherweise nicht ausgeschöpft wird (Chi, 2009).

Mittwoch
11. März

14:00 Uhr

GA 03/149

Chair(s)

Bender, Elena

3

Wegner, Elisabeth; Kurz, David; Burkhardt, Christian; Nückles, Matthias

Hürden laufen, Fotos machen, Bäume pflanzen? Metaphern des Lernens von Schüler/innen, Studierenden und Berufstätigen im Vergleich

Überzeugungen über das Lernen beeinflussen Motivation, Lernstrategien und damit auch die Lernergebnisse von Lernenden (z.B. Gijbels et al., 2013). Überzeugungen sind jedoch schwer zu erfassen, da sie häufig nur implizit vorliegen (z.B. Kane, Sandretto & Heath, 2004). Weiterhin lassen sich Ergebnisse zu Lernüberzeugungen schlecht zwischen verschiedenen Kontexten (Schule vs. Hochschule vs. Weiterbildung) vergleichen, da die meisten Erhebungsinstrumente kontextspezifisch sind. Eine Möglichkeit, Überzeugungen kontextübergreifend sichtbar zu machen, können Metaphern sein (z.B. Saban, Kocbeker & Saban, 2007; Lakoff & Johnson, 1980). Bisher wurden Metaphern vorwiegend genutzt, um Überzeugungen von Lehrkräften zu erheben. In zwei explorativen Studien wurde daher untersucht, welche Metaphern Lernende nutzen, um Lernen zu beschreiben, und wie sich die in den Metaphern zum Ausdruck gebrachten Lernüberzeugungen zwischen verschiedenen Kontexten unterscheiden.

Studie 1: Insgesamt N=68 Studierende aus verschiedenen Studienabschnitten und N=69 Schüler/innen (Gymnasium, 8. und 11. Klasse) wurden dazu aufgefordert, ihre Metaphern des Lernens zu beschreiben („Lernen ist wie... weil... Ziel von Lernen ist...“). Die Befragung fand jeweils zu Beginn einer Unterrichts- bzw. Seminarstunde statt. Die Metaphern wurden inhaltsanalytisch nach vier induktiv entwickelten Kategorien ausgewertet und von zwei unabhängigen Ratern klassifiziert (Cohens $\kappa = .72$) (1) Lernen als Selbstregulationsaufgabe („Lernen ist wie Hürdenlaufen - man muss sich durchbeißen“) (2) Lernen als Wissenserwerb (Lernen ist wie Fotos machen – man stellt sich ein Album zusammen“) (3) Lernen als Handeln können („Lernen ist wie Kleidung kaufen – dann hat man für jede Gelegenheit etwas zum Anziehen“) (4) Lernen als Persönlichkeitsentwicklung („Lernen ist wie einer Pflanze beim Wachsen zu zuschauen – man entwickelt sich immer weiter“)

Wir stellten signifikante Unterschiede in der Häufigkeit der verschiedenen Metaphern zwischen Schule und Universität fest, $\chi^2(3,137) = 26,69, p < .000$ (Tabelle 1). Knapp 70% der Schüler/innen beschrieben in ihrer Metapher Lernen als eine Selbstregulationsaufgabe. Wissenserwerb, Handlungsfähigkeit oder Persönlichkeitsentwicklung spielten bei den Schüler/innen nur eine untergeordnete Rolle. Bei den Studierenden schilderten nur 26,5% Lernen im Sinne einer Selbstregulationsaufgabe, am häufigsten waren Wissenserwerbsmetaphern (33,8%). 22,1% der Metaphern der Studierenden ließen sich im Sinne von Lernen als Persönlichkeitsentwicklung kategorisieren, 17,6% als Handeln können. Vergleiche innerhalb der Schule bzw. der Universität zeigten keine signifikanten Unterschiede zwischen den Klassenstufen oder zwischen den verschiedenen Studienabschnitten.

Studie 2: In der zweiten Studie wurden Studierende und Berufstätige in einer Onlinebefragung zu ihren Metaphern des Lernens befragt. Insgesamt N=95 Personen (47 Studierende, 48 Berufstätige) füllten den Fragebogen komplett aus. Die Teilnehmenden erhielten zunächst ein Beispiel einer Metapher aus einem anderen Themenfeld, und wurden dann wie in Studie 1 dazu aufgefordert, ihre Metapher des Lernens zu bilden. Die Auswertung erfolgte nach den gleichen Kategorien wie in Studie 1. Die Übereinstimmung zwischen den Ratern lag in dieser Stichprobe bei $\kappa = .81$.

Es zeigte sich eine signifikante Unterschiede in der Verteilung der Kategorien, $\chi^2(3,95) = 9,14, p = .027$. Die Hälfte der Berufstätigen schilderten Lernen als Persönlichkeitsentwicklung, 22,9% im Sinne einer Selbstregulationsaufgabe, 18,8 % als Handeln & Problemlösen, und nur 8,3% als Wissenserwerb. Bei den Studierenden waren in dieser Stichprobe Selbstregulationsmetaphern mit 38,3% am häufigsten, 21,3% schilderten Lernen als Wissenserwerb. Metaphern, die Lernen als Handeln können oder als Persönlichkeitsentwicklung schilderten, kamen mit 17% bzw. mit 23,4% ähnlich häufig vor wie bei den Studierenden in Studie 1.

Zusammenfassung und Diskussion: Die Ergebnisse zeigen, dass Metaphern eine Möglichkeit sind, Lernüberzeugungen zwischen Kontexten zu vergleichen, und dass sich die in Metaphern zum Ausdruck gebrachten Überzeugungen über Lernen zwischen Personen aus verschiedenen Kontexten sehr stark unterscheiden. Während in der Schule die Aspekte der Selbstregulation im Vordergrund stehen, rücken im Studium der Wissenserwerb und im Beruf die Persönlichkeitsentwicklung ins Blickfeld. Inwiefern sich die Metaphern auch auf das Lernverhalten auswirken, soll in weiteren Studien untersucht werden.

Mittwoch
11. März

14:00 Uhr

GA 03/149

Preckel, Franzis; Baudson, Tanja Gabriele; Krolak-Schwerdt, Sabine; Glock, Sabine
Hochbegabt und verhaltensauffällig? Automatische Assoziationen und implizite Einstellungen von Lehramtsstudierenden gegenüber hochbegabten Kindern

Theoretischer Hintergrund: Nach der sogenannten Disharmoniehypothese geht eine außergewöhnliche intellektuelle Begabung mit nicht adaptiven Persönlichkeitsmerkmalen und Verhaltensauffälligkeiten einher (Persson, 1998). Obwohl durch empirische Befunde zur Persönlichkeit Hochbegabter widerlegt (Zeidner & Shani-Zinovich, 2011), fanden Studien zu impliziten Theorien von Lehrkräften zu hochbegabten Schüler/innen Vorurteile im Sinne der Disharmoniehypothese (z.B. Baudson & Preckel, 2013). Implizite Lehrertheorien sind von großer pädagogischer Bedeutung, u.a. da sie über die Erwartungen, Ziele oder das Verhalten der Lehrkräfte die Entwicklung der Schüler/innen beeinflussen können (Jussim, Eccles, & Madon, 1996).

Fragestellung: (1). Lassen sich implizite, automatische Assoziationen zwischen Hochbegabung und beiden Dimensionen der Disharmoniehypothese „hohe intellektuelle Fähigkeit“ und „Verhaltensauffälligkeit“ nachweisen? Empirische Evidenz für diese Assoziationen unterstützte die Annahme eines impliziten Hochbegabungstereotyps i.S.d. Disharmoniehypothese. (2). Unterscheiden sich implizite Einstellungen gegenüber Hochbegabten in Abhängigkeit von der Aktivierung der Konzepte „hohe intellektuelle Fähigkeit“ bzw. „Verhaltensauffälligkeit“? Wir erwarteten neutrale Einstellungen, wenn keines dieser Konzepte salient gemacht würde, da in diesem Fall zwei Konzepte unterschiedlicher Valenz (positiv vs. negativ) interagieren, so dass insgesamt neutrale Einstellungen resultieren sollten (Maio, Essess & Bell, 2000). Andererseits sollten bei Salienz des Konzepts „hohe intellektuelle Fähigkeit“ positive und bei Salienz des Konzepts „Verhaltensauffälligkeit“ negative Einstellungen resultieren.

Methode.

Stichprobe: Insgesamt nahmen 182 Lehramtsstudierende am Experiment teil (101 weiblich; Alter: M = 23.53 Jahre, SD = 3.37; Studiendauer: M = 6.6 Semester, SD = 3.74).

Material: Automatische Assoziationen erfassten wir mit zwei separaten single-target Implicit Association Tests (ST-IAT) mit den Zielkategorien „hohe intellektuelle Fähigkeit“ bzw. „Verhaltensauffälligkeit“ und den Attribut-Kategorien „hochbegabt“ vs. „normalbegabt“. Die Attribut-Kategorien wurden über Fotos von Schüler/innen repräsentiert. Die Zuweisung eines Fotos zu einer der beiden Kategorien erfolgte zufällig (die Neutralität der Fotos im Hinblick auf die Ziel- und Attributkategorien wurde vorab getestet). Die Zielkategorie „hohe intellektuelle Fähigkeit“ wurde über acht Wörter abgebildet (z.B. intelligent), die Zielkategorie „Verhaltensauffälligkeit“ über neun Wörter (z.B. schwierig). Das Wortmaterial wurde vorab über Expertenratings evaluiert. Implizite Einstellungen erfassten wir mit einer affektiven priming (AP) Aufgabe. Dabei wurden dieselben Schülerfotos wie beim ST-IAT verwendet und zudem jeweils 14 positive und negative Wörter (z.B. warm, kalt).

Design und Datenauswertung: Da sich die Lehrereinschätzungen für die hier relevanten Konstrukte je nach Schülergeschlecht unterscheiden können (positivere Einschätzung des Verhaltens von Schülerinnen, positivere Einschätzung der Fähigkeit/Begabung von Schülern; Bennett et al., 1993; Trautwein & Baeriswyl, 2007), bekamen jeweils 50% unserer Stichprobe entweder Bilder von Schülerinnen oder von Schülern (Kontrolle Einfluss Schülergeschlecht). Alle Versuchspersonen lernten zu Beginn des Experiments die Zuweisung der Schülerfotos zu beiden Kategorien hochbegabt vs. normalbegabt. Im Anschluss bearbeiteten sie entweder die AP Aufgabe oder den ST-IAT mit der Zielkategorie „hohe intellektuelle Fähigkeit“ bzw. „Verhaltensauffälligkeit“ (between subjects design). Die Reihenfolge wurde dabei ausbalanciert. Die Datenauswertung der Reaktionszeiten aus ST-IAT und AP erfolgte mittels Varianzanalysen (ST-IAT score: 2 x 2 between-subjects ANOVA mit den Faktoren ST-IAT-Variante „hohe intellektuelle Fähigkeit“ bzw. „Verhaltensauffälligkeit“ und Schülergeschlecht; AP score: 2 x 2 x 2 between-subjects ANOVA mit den Faktoren ST-IAT-Variante, Reihenfolge (AP vor vs. nach ST-IAT), und Schülergeschlecht).

Ergebnisse: Unsere Ergebnisse bestätigen die Existenz eines impliziten Hochbegabungstereotyps im Sinne der Disharmoniehypothese - allerdings nur für hochbegabte Jungen und nicht für Mädchen. Dieses scheint sich auch auf die Bewertung Hochbegabter auszuwirken: Nach Aktivierung des Konzepts „Verhaltensauffälligkeit“ (Bedingung, in der der entsprechende ST-IAT vor dem AP bearbeitet wurde), fielen die impliziten Einstellungen gegenüber hochbegabten Jungen signifikant negativer aus als gegenüber normalbegabten Jungen. Interessanterweise fielen die Einstellungen zu Hochbegabten nach Aktivierung des Konzepts „hohe intellektuelle Fähigkeit“ neutral aus. Ohne Aktivierung eines dieser Konzepte waren die Einstellungen gegenüber Hochbegabten neutral.

Schlussfolgerung: Hochbegabung scheint ein ambivalentes Konstrukt zu sein, welches sowohl mit positiven wie negativen Assoziationen besetzt ist, die situationsabhängig aktiviert werden und in Abhängigkeit davon zu unterschiedlichen Einstellungen gegenüber Hochbegabten führen können.

Mittwoch
11. März

14:00 Uhr

GB 03/49

Schiefer, Julia; Oschatz, Kerstin; Tibus, Maike; Trautwein, Ulrich

Eine Interventionsstudie zur Förderung naturwissenschaftsbezogener Überzeugungen bei begabten Grundschulkindern

Theoretischer Hintergrund und Fragestellungen: Die Entwicklung sophistizierter naturwissenschaftsbezogener Überzeugungen ist ein zentrales Ziel naturwissenschaftlicher Grundbildung (OECD, 2007) sowie ein wesentlicher Bestandteil eines adäquaten Wissenschaftsverständnisses. Naturwissenschaftsbezogene Überzeugungen lassen sich in die beiden Kerndimensionen Nature of Knowledge und Nature of Knowing (Hofer & Pintrich, 1997) differenzieren und beziehen sich auf diejenigen epistemischen Überzeugungen, die wissenschaftlichem Wissen und seiner Entwicklung inhärent sind (vgl. Lederman, 2002, 2007). Die meisten Studien zur Förderung von naturwissenschaftsbezogenen Überzeugungen wurden mit Schüler/innen der Sekundarstufe (z.B. Conley, Pintrich, Vekiri & Harrison, 2004; Elder, 2002) oder Studierenden durchgeführt. Da ein enger Zusammenhang zwischen epistemischen Überzeugungen und (meta)kognitiven Fähigkeiten besteht (vgl. Bromme, Pieschl & Stahl, 2010), ist anzunehmen, dass sich diese Überzeugungen bei begabten Kindern bereits in der Primarstufe entwickeln. In der vorliegenden Interventionsstudie wurde untersucht, ob (1) naturwissenschaftsbezogene Überzeugungen bei Teilnehmer/innen eines Förderprogramms für besonders Begabte und Hochbegabte im Grundschulalter angeregt und verbessert werden können. Da Studien darauf hinweisen, dass Mädchen sich außerschulisch weniger mit Naturwissenschaften beschäftigen und über weniger naturwissenschaftsbezogene Interessen verfügen als Jungen (Alexander, Johnson & Kelley, 2012), wurde zudem geprüft, ob (2) sich Geschlechterunterschiede in der Wirksamkeit des Förderprogramms zeigen.

Methode.

Design: Diese Forschungsfragen wurden im Rahmen einer Interventionsstudie mit 65 Dritt- und Viertklässlern in einem randomisierten Prä-Posttest-Design mit einer Kontrollgruppe untersucht (58.46% Jungen, Alter: M = 8.76, SD = 0.61, IQ: M = 117.96, SD = 16.18). Die Intervention zur Förderung der naturwissenschaftsbezogenen Überzeugungen wurde über einen Zeitraum von zwölf Wochen an vier Hector-Kinderakademien (außerschulisches Förderangebot für besonders begabte und hochbegabte Kinder in Baden-Württemberg) durchgeführt. Die Kontrollgruppe nahm an einer Intervention zur Förderung der Präsentationsleistung teil. Die Zuteilung der Kinder zu der Interventions- und der Kontrollgruppe konnte innerhalb der Akademien randomisiert erfolgen, da beide Kurse als Kursverbund ausgeschrieben waren.

Kurskonzeption: Die Intervention (betitelt Kleine Forscher – Wir arbeiten wie Wissenschaftler) fokussierte auf die explizite Anregung wissenschaftstheoretischer Reflexion. Die Kinder erhielten verschiedene Möglichkeiten wissenschaftliche Arbeitsweisen sowie Methoden und Grenzen der naturwissenschaftlichen Erkenntnisgewinnung kennenzulernen und anzuwenden (z.B. in spielerischen Forschungsprojekten und Experimenten oder dem Besuch eines Schülerlabors).

Instrumente: Die naturwissenschaftsbezogenen Überzeugungen wurden durch das von Conley et al. (2004) entwickelte Messinstrument (in der deutschen Übersetzung von Urhahne & Hopf, 2004) erfasst. Dieses beinhaltet die Dimensionen Quelle des Wissens (5 Items, α (t1/t2) = .73/.74), Rechtfertigung des Wissens (9 Items, α (t1/t2) = .58/.64), Entwicklung des Wissens (6 Items, α (t1/t2) = .58/.73) sowie Sicherheit des Wissens (6 Items, α (t1/t2) = .73/.72).

Ergebnisse: Um die Effekte der Intervention auf die Veränderung der naturwissenschaftsbezogenen Überzeugungen der Kinder zu prüfen, wurden für jede Dimension lineare multiple Regressionen mit der Differenz der beiden Messzeitpunkte als abhängige Variable berechnet. Es zeigten sich (unter Kontrolle des Ausgangsniveaus) signifikante Interventionseffekte auf die Überzeugungen der Kinder hinsichtlich Sicherheit des Wissens ($\beta = .25$, $p < .05$), Entwicklung des Wissens ($\beta = .30$, $p < .01$) sowie marginal für Rechtfertigung des Wissens ($\beta = .19$, $p < .10$). Die Untersuchung von Geschlechtereffekten – dafür wurde zusätzlich der Faktor Geschlecht mit den zugehörigen Interaktionen in das Modell mitaufgenommen – zeigte, dass Mädchen durch die Teilnahme an der Intervention in den Skalen Quelle des Wissens ($\beta = -.64$, $p < .01$) sowie Sicherheit des Wissens ($\beta = -.78$, $p < .01$) sophistiziertere naturwissenschaftsbezogene Überzeugungen entwickelten als Jungen in der Intervention.

Die Ergebnisse liefern Hinweise auf die Effektivität des Kurses, wobei Mädchen teilweise stärker profitieren als Jungen. Die an dem Förderprogramm teilnehmenden Dritt- und Viertklässler konnten durch den Besuch der Intervention reifere naturwissenschaftsbezogene Überzeugungen entwickeln. Dies deutet auf stattgefundenen Reflexions- sowie Abstraktionsprozesse im Umgang mit Naturwissenschaften und naturwissenschaftlichem Wissen hin.

Die Studienergebnisse ergänzen wird ein Ausblick auf die Implementation des Förderangebots an 20 Hector-Kinderakademien.

Mittwoch
11. März

14:00 Uhr

GB 03/49

Chair(s)

Preckel, Franzis

Herbein, Evelin; Zettler, Ingo; Tibus, Maike; Golle, Jessika; Oschatz, Kerstin; Trautwein, Ulrich

Präsentieren – Mitteilen – Verstehen: Evaluation eines Präsentationstrainings als Enrichment-Programm für hochbegabte Grundschul Kinder.

Theoretischer Hintergrund: Sprechen – vor und mit anderen – ist eine Kompetenz, deren Bedeutung unbestritten ist (Ediger, 2011; Pabst-Weinschenk, 2004), unter anderem im Zusammenhang mit Wissensaustausch und Wissenserwerb. Sprechen im Schulkontext umfasst sowohl Rede- als auch Gesprächssituationen und beinhaltet damit Kompetenzen des zusammenhängenden Sprechens vor anderen sowie den Austausch mit anderen im Gespräch. Dabei sind Kinder den Austausch in dialogischen Situationen durch ihren Alltag gewöhnt, wohingegen sie mit dem Sprechen vor anderen meist erst ab der Schule konfrontiert werden (vgl. Behrens & Eriksson, 2011). Hinzu kommt, dass die kommunikativen Aufgaben mit dem Schuleintritt komplexere und abstraktere Themenbereiche umfassen, als dies noch im Kindergarten und der Vorschule der Fall war (vgl. Ahrenholz, 2010). Um mit den Anforderungen zurechtzukommen spielen sowohl verbale Fähigkeiten, beispielsweise die Beherrschung der Bildungssprache, in Abgrenzung zu alltagssprachlichen Kompetenzen (vgl. Heppt, Dragon, Berendes, Stanat, & Weinert, 2012; Schleppegrell, 2001; Schmolzer-Eibinger, 2012), als auch kommunikative Fähigkeiten im Bereich des zusammenhängenden Sprechens vor anderen, eine bedeutsame Rolle (Eriksson, 2006).

Im Vergleich zu ihren gleichaltrigen Mitschülerinnen und Mitschülern zeigen hochbegabte Kinder häufig einen Entwicklungsvorsprung in ihren verbalen Fähigkeiten, beispielsweise einen größeren Wortschatz, präzisere Ausdrucksweise und komplexere und längere Satzkonstruktionen (vgl. Jen, Tseng, & Kuo, 2014). Beim Sprechen vor anderen, über Themenbereiche in denen sie sehr versiert sind, kann es deshalb zur Verwendung von komplexerer Sprache kommen, als Gleichaltrige dies zeigen. Eine effektive kommunikative Interaktion kann dadurch erschwert werden, indem es für beide Kommunikationsseiten schwierig ist, die Inhalte des anderen zu verstehen (Lee, Olszewski-Kubilius, & Thomson, 2012).

Im Rahmen dieser Studie wurde ein Präsentationstraining als Enrichment-Programm für besonders begabte und hochbegabte Grundschul Kinder entwickelt und evaluiert. Ziel des Trainings war es, die Präsentationsleistung zu fördern, um in Vortragssituationen die Interaktionen mit den Zuhörern effektiv gestalten zu können, und Sprechangst zu reduzieren.

Methode: Das Präsentationstraining ist Teil eines außerunterrichtlichen Förderprogramms für besonders begabte und hochbegabte Grundschul Kinder (Hector-Kinderakademien) in Baden-Württemberg und umfasst zehn 90-minütige Kurseinheiten. Um die Effektivität des Trainings hinsichtlich der Förderung der Präsentationsleistung und Reduktion von Sprechangst zu überprüfen, wurde ein randomisiertes Kontrollgruppen-Design verwendet. Dazu wurden 65 Dritt- und Viertklässler, die für die Teilnahme an einer Hector-Kinderakademie nominiert worden waren, zufällig entweder dem Präsentationstraining (N = 33) oder der Kontrollgruppe (N = 32) zugeteilt. Die Kinder der Kontrollgruppe nahmen dabei an einem Kurs zur Förderung des Wissenschaftsverständnisses teil. Die Auswertung der Daten erfolgte mittels multipler linearer Regressionen.

Ergebnisse: Das Präsentationstraining hatte verglichen mit der Kontrollgruppe, positive Effekte auf die Präsentationsleistung der Kinder. Die Ergebnisse zeigen einen Anstieg in der selbstwahrgenommenen Präsentationsleistung der Kinder (Cohen's d = 0.34) und eine Abnahme der Sprechangst (Cohen's d = 0.20).

Diskussion und Ausblick: Die Ergebnisse weisen auf die Effektivität des Präsentationstrainings hinsichtlich der Förderung von Präsentationsleistung hin. Ergänzend dazu werden die im Rahmen der Studie entstandenen Videos der Kurzpräsentationen der Kinder mittels eines Beobachtungsbogens zur Erfassung der Präsentationsleistung eingeschätzt und ausgewertet. Ziel ist es, die Selbsteinschätzungen der Kinder um die Einschätzung durch externe Rater zu ergänzen. Diesbezügliche Ergebnisse sollen ebenfalls präsentiert werden.

Mittwoch
11. März

14:00 Uhr

GB 03/49

Chair(s)

Preckel, Franzis

Golle, Jessika; Zettler, Ingo; Tibus, Maike; Nagengast, Benjamin; Hasselhorn, Marcus; Trautwein, Ulrich

Effekte eines außerschulischen Förderprogramms für besonders begabte und hochbegabte Grundschul Kinder

Hintergrund und Fragestellung: In Baden-Württemberg gibt es seit 2010 ein Programm zur Förderung besonders begabter und hochbegabter Grundschul Kinder: die Hector-Kinderakademien. Ziel des Programms ist es besonders begabte Kinder außerhalb des Unterrichts vor allem im MINT-Bereich (Mathematik, Informatik, Naturwissenschaften und Technik) intensiv zu fördern. Die Teilnahme an den Hector-Kinderakademien ist abhängig von der Nominierung der Kinder durch die Lehrkräfte. Das Programm wurde erstmals umfangreich und mit einer großen Stichprobe von Drittklässlern im ersten Schulhalbjahr 2012/2013 in Baden-Württemberg evaluiert. Das Ziel dieser Untersuchung war es herauszufinden, ob die Kurse, die an den Kinderakademien angeboten wurden, einen Effekt auf die kognitiven Fähigkeiten der Kinder, ihre Noten und Interessen hatten. Die Kinder wurden im Klassenverband getestet. In einer Klasse konnten mehrere Kinder die Hector-Kinderakademien besuchen und die Klassen waren unterschiedlich groß.

Methode: Insgesamt wurden 2751 Kinder zu Beginn und Ende eines Schulhalbjahres u.a. hinsichtlich ihrer kognitiven Fähigkeiten getestet und bezüglich ihrer Interessen am Unterricht und außerhalb der Schule befragt. Von den 2751 Kindern besuchten 423 Schüler die Hector-Kinderakademien und 2328 Kinder, die keine Akademie besuchten, dienten als Kontrollstichprobe. Dieses Design ermöglichte es die Veränderungen über ein Schulhalbjahr zwischen den Kindern, die die Akademien besuchten und den Kindern der Kontrollgruppe zu vergleichen. Damit konnte der Effekt der Hector-Kinderakademien auf die Intelligenz der Kinder, ihre Noten (Mathe und Deutsch) und Interessen (u.a. Mathe, Deutsch, Forschen) geschätzt werden. Fehlende Werte wurden mit Hilfe multipler Imputation geschätzt (Rubin, 1987). Um potentielle Unterschiede zum ersten Messzeitpunkt (selection bias) zwischen den Gruppen zu kontrollieren und den Effekt des Programms unverfälscht schätzen zu können, wurde ein propensity score matching (Rosenbaum & Rubin, 1983; Stuart, 2010) durchgeführt. Mit Hilfe dieses Verfahrens wird zu jedem der Kinder aus den Hector-Kinderakademien ein Partner mit ähnlichen Eigenschaften aus der Kontrollgruppe zugeordnet. Dieses Vorgehen ermöglicht die Schätzung des average treatment effects of the treated, das bedeutet, dass der Effekt der Akademien für diese Kinder, die typischerweise die Hector-Kinderakademien besuchen, geschätzt werden kann. Die Stichprobe nach dem matching bestand aus 423 Kindern, die die Hector-Kinderakademien besuchten und 423 Kindern der Kontrollgruppe, die mit den Akademiekindern in Intelligenz, Noten, schulischen und außerschulischen Interessen, Selbstwert, Selbstkonzept und Sozialverhalten zu Beginn des Schulhalbjahres vergleichbar waren. Aufgrund der genesteten Datenstruktur (Schüler in Klassen) wurden neben den Variablen auf Schülerebene auch Variablen auf Klassenenebene bei dem matching berücksichtigt, dazu gehörten die durchschnittliche Intelligenz der Kinder in einer Klasse und der Anteil der Kinder einer Klasse, die an den Hector-Kinderakademien teilnahmen. Die Analyse der ausgewählten Stichprobe (N = 846) erfolgte unter Berücksichtigung der Mehrebenenstruktur der Daten (Thoemmes & West, 2011). Die Teilnahme an den Hector-Kinderakademien sowie der Anteil der Kinder einer Klasse, die die Akademien besuchten, waren die unabhängigen Variablen. Intelligenz, Schulnoten und verschiedene Interessen der Kinder waren die abhängigen Variablen.

Ergebnisse und Diskussion: Die Hector-Kinderakademien hatten einen positiven Effekt auf die fluide und kristalline Intelligenz der Kinder sowie die Mathe- und Deutschnoten am Ende des Schulhalbjahres, nicht aber auf deren Interessen. Die Ergebnisse legen nahe, dass die kognitiven Fähigkeiten und schulischen Leistungen der Kinder, die typischerweise für ein Förderprogramm begabter Kinder nominiert werden, mit Hilfe eines extracurricularen Programms positiv beeinflusst werden können. Bisher erfolgte keine Differenzierung der Kurse in unterschiedliche Inhaltsbereiche. Das Ausbleiben der Interessensveränderung ist möglicherweise auf die noch fehlende Differenzierung zurück zu führen. Die Ergebnisse dieser Studie unterstreichen die Bedeutsamkeit zusätzlicher Förderung besonders begabter Kinder bereits im Grundschulbereich (vgl. Callahan, Moon, Oh, Azano, & Hailey, 2014; Vaughn, Feldhusen, & Asher, 1991).

Mittwoch
11. März

14:00 Uhr

GB 03/49

Chair(s)

Preckel, Franzis

Niempel, Christoph; Mustafic, Maida; Greiff, Samuel; Petway, Kevin T.; Roberts, Richard D.

Die ‚Dunkle Seite der Kreativität‘: Sagt Kreativität negative Veränderungen im ethischen Entscheiden von Schülerinnen und Schülern vorher?

Theoretischer Hintergrund: Kreativität wird als Fähigkeit definiert, neuartige und nützliche Ideen generieren zu können (z.B. Ward, 2004). Ethisches Entscheiden kann als die Bereitschaft und Fähigkeit verstanden werden, die Bedürfnisse anderer bei der eigenen Entscheidungsfindung zu berücksichtigen (z.B. Moran, 2014). Beide Fähigkeiten wurden wiederholt als zentrale Fähigkeiten identifiziert, die Schülerinnen und Schülern (SuS) unterstützen, erfolgreich modernen Anforderungen einer zunehmend komplexer werdenden Welt zu bewältigen (z.B. Casner-Lotto & Barrington, 2006). Einige Forscherinnen und Forscher sind allerdings skeptisch, ob Kreativität, im Gegensatz zu anderen Fähigkeiten wie dem ethischen Entscheiden, tatsächlich eine genuin förderungswürdige Fähigkeit darstellt, welche ohne Abstriche in unseren Schulen angeregt werden sollte. Experimentelle Befunde zur sogenannten ‚Dunklen Seite der Kreativität‘ zeigen, dass Kreativität unerwünschtes, unethisches Verhalten begünstigt (z.B. Gino & Ariely, 2012). Die vorliegende Studie untersucht zum ersten Mal, ob Kreativität mit ethischem Entscheiden in einer Stichprobe von SuS längsschnittlich negativ assoziiert ist.

Fragestellung: Diese Studie war zwei Fragestellungen gewidmet: (1) Inwieweit hängt Kreativität mit einer negativen Veränderung im ethischen Entscheiden zusammen? (2) Inwieweit sind die beobachteten Zusammenhänge unabhängig von den zugrundeliegenden kognitiven Fähigkeiten der SuS (schlussfolgerndes Denken)?

Methode: Es wurden sowohl Selbsteinschätzungen als auch Lehrereinschätzungen als Maße für Kreativität und ethisches Entscheiden genutzt. Insgesamt wurden 1869 US-amerikanische SuS (48.4% männlich) über zwei Messzeitpunkte hinweg untersucht. Von einer Substichprobe von 417 SuS lagen zusätzlich Werte zum schlussfolgernden Denken vor. Insgesamt wurden die SuS von 245 Lehrkräften beurteilt. Das durchschnittliche Alter der SuS betrug 12.26 (SD = 0.94) zum ersten und 12.83 Jahre (SD = 0.75) zum zweiten Messzeitpunkt.

Die Items zur selbst- und lehrereingeschätzten Kreativität und ethischem Entscheiden basieren auf dem International Personality Item Pool (IPIP; ipip.ori.org). Jede der eingesetzten Skalen bestand jeweils aus sechs Items. Als Test zum schlussfolgerndem Denken wurde der CTP 4 (ERB, 2012) eingesetzt.

Für die Datenanalyse wurde die Methode der Strukturgleichungsmodellierung gewählt. Zunächst wurden konfirmatorische Faktorenanalysen sowie eine Messinvarianztestung durchgeführt. Zur Berechnung der Fragestellung wurde ein cross-lagged model (Little, Preacher, Selig, & Card, 2007) berechnet, welches jeweils die autoregressiven und kreuzverzögerten längsschnittlichen Assoziationen zwischen selbst- und lehrereingeschätzter Kreativität und selbst- und lehrereingeschätztem ethischen Entscheiden über zwei Messzeitpunkte enthielt (Modell A). In einem weiteren Modell wurde zusätzlich die Fähigkeit zum schlussfolgernden Denken der SuS als zeitunabhängige Kovariate mitaufgenommen (Modell B, basierend auf der Substichprobe von 417 SuS; Kommentar: bzgl. der Fragestellung sind die Ergebnisse von Modell A, berechnet mit N=1869, und Modell B, berechnet mit N=417, vergleichbar).

Ergebnisse: Vorabanalysen: Konfirmatorischen Faktorenanalysen zeigten, dass selbsteingeschätzte Kreativität, lehrereingeschätzte Kreativität, selbsteingeschätztes ethisches Entscheiden und lehrereingeschätztes ethisches Entscheiden vier distinkte Faktoren pro Messzeitpunkt darstellen (Messzeitpunkt 1: CFI = .96, RMSEA = .05, 90% CI [.04, .05], SRMR = .04; Messzeitpunkt 2: CFI = .96, RMSEA = .05, 90% CI [.04, .05], SRMR = .04). Die Messinvarianztestung ergab, dass ein metrisch invariantes Modell angenommen werden kann.

Hauptanalysen: Ohne Berücksichtigung des schlussfolgernden Denkens der SuS fanden wir keinerlei Hinweise auf einen etwaig negativen Effekt von Kreativität auf längsschnittliche Veränderungen im ethischen Entscheiden (Modell A: CFI = .96, RMSEA = .03, 90% CI [.03, .03], SRMR = .04). Nach Kontrolle des schlussfolgernden Denkens der SuS zeigte die lehrereingeschätzte Kreativität einen negativen Zusammenhang mit dem lehrereingeschätzten ethischen Entscheiden zum nachfolgenden Zeitpunkt (Modell B: CFI = .94, RMSEA = .04, 90% CI [.04, .04], SRMR = .05). Zusammengenommen zeigten die Ergebnisse der vorliegenden Studie, dass Kreativität keinen generellen Prädiktor von negativen Veränderungen im ethischen Entscheiden von SuS darstellt. Allerdings weisen die Ergebnisse auch darauf hin, dass die gefundenen längsschnittlichen Zusammenhänge nicht unabhängig vom schlussfolgernden Denken sind und bedeutsame Unterschiede zwischen Subgruppen von SuS unterschiedlicher kognitiver Fähigkeiten diskutiert werden müssen.

Hinweis: Das Manuskript zur vorliegenden Studie wird im Oktober 2014 zur Veröffentlichung in einem peer-reviewed journal eingereicht

Mittwoch
11. März

14:00 Uhr

GB 03/49

Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen I

In dem vom BMBF geförderten Programm „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor (KoKoHs)“ (www.kompetenzen-im-hochschulsektor.de) forschen seit 2011 bundesweit 24 Projektverbände mit ca. 70 Einzelprojekten zur theoretischen Modellierung von akademischen fachbezogenen und fachübergreifenden Kompetenzen, die in verschiedenen Studiendisziplinen erworben werden sowie an deren Operationalisierung und validen Erfassung. In diesem Symposium werden wichtige Ergebnisse aus diesem neuen Forschungsprogramm präsentiert.

Die Beiträge umfassen Arbeitsergebnisse aus den beiden zentralen Strängen fachspezifische und generische Kompetenzen. Sie decken alle im Forschungsprogramm repräsentierten Studiendomänen ab (Bildungs- bzw. Erziehungswissenschaft, Psychologie, Lehrerbildung, Wirtschafts- und Ingenieurwissenschaften) und liefern Resultate für die Kompetenzmodellierung und -messung in den jeweiligen, bislang nur wenig erforschten Fachdisziplinen. Der inhaltliche Fokus aller Beiträge liegt zum einen auf der theoretischen Modellierung des zu erfassenden Konstrukts und den dazu entwickelten Instrumenten. Zum anderen werden dazu jeweils empirische Ergebnisse v.a. aus Validierungsstudien (z.B. zur theoretisch angenommenen Kompetenzstruktur, zur konvergenten bzw. diskriminanten Validität in Relation zu anderen Konstrukten wie Intelligenz etc.) sowie Analysen zu individuellen (z.B. Geschlecht) und institutionellen Merkmalen (z.B. Lerngelegenheiten) präsentiert, mit denen sich der erfasste Kompetenzstand vorhersagen lässt. Die präsentierten theoretischen Modellierungen, Messinstrumente und empirische Befunde werden im Rahmen des Symposiums kritisch diskutiert und Perspektiven für die weitere Forschung herausgearbeitet.

Mittwoch
11. März

14:00 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen I

Zhang, Ying; Klopp, Eric; Dietrich, Heike; Brünken, Roland; Krause, Ulrike-Marie; Spinath, Birgit; Stark, Robin; Spinath, Frank M.
Model-based Prediction of Scientific Competencies in Psychology

Theoretical background: Scientific competencies are tremendously important since science is one of the most crucial driving forces in modern society (Feist, 2012). To date, it is not yet clear which factors contribute to scientific competencies in psychology. To fill this gap, we conducted this study.

First, we developed a theoretical model of scientific competencies in psychology (Dietrich et al., 2013). Within this model, personal characteristics consisting of the cognitive, motivational, and personality-related factors are very important for efficient competency acquisition and utilization (Dietrich et al., 2013). Therefore, we focused on personal characteristics in the prediction of scientific competencies.

Regarding personal characteristics, we integrated a range of constructs. As a cognitive factor, we expected that intelligence plays an important role for acquiring scientific competencies (Dietrich et al., 2013) because intelligence is positively associated with academic performance (Deary, Strand, Smith, & Fernandes, 2007). As a motivational factor, academic self-concept was expected to advance acquiring scientific competencies, because students with high academic self-concept perform better in academic settings (Jansen, Schroeders, & Lüdtke, 2014). Regarding personality-related factors, conscientiousness, openness, and self-control capacity should be taken into account. Referring to individual differences in persistence, responsibility, and effort (Costa & McCrae, 1992), conscientiousness is considered as a prerequisite for working in a scientifically competent manner (Dietrich et al., 2013). Besides, openness is linked to creativity and to the tendency to seek educational experience (Costa & McCrae, 1992) that are relevant for acquiring scientific competencies (Dietrich et al., 2013). Moreover, self-control capacity facilitates the academic learning process (Bertrams & Dickhäuser, 2009), we therefore expected that self-control capacity favors acquiring scientific competencies. Finally, we expected that students who have absolved more semesters should have been exposed to more study contents building a central base of scientific competencies. To sum it up, we considered not only personal characteristics, but also the number of study semesters in the prediction of scientific competencies.

Research questions: The aim of this study was to investigate the predictive validity of the cognitive, motivational, and personality-related factors after controlling the number of study semesters.

Methods: Our sample consisted of 77 bachelor and master students (M = 21.3 years old, Male: 20.2%) who were studying psychology. The current study semester of the participants ranged from 1 to 5. Between 2012 and 2013, the participants completed a set of tests comprising of variables corresponding to personal characteristics. In 2014, these participants completed the test of scientific competencies. We applied three stepwise regression analyses to examine the predictive validity of the cognitive, motivational and personality-related variables respectively. In each regression model, we entered the number of study semester as a control variable in the first step. In the second step, we entered intelligence in model 1, academic self-concept in model 2, as well as conscientiousness, openness and self-control capacity in model 3.

Results: The number of study semesters explained 10.9 % of variance in scientific competencies. Intelligence explained 11.5 % variance in scientific competencies over and above the number of study semesters. Academic self-concept explained 6.7 % variance in scientific competencies, once the number of study semesters was held constant. Inconsistent with our theoretical model, personality-related predictors showed no predictive validity in our sample.

In our relatively small sample, the results confirmed that variance in scientific competencies could be explained by predictors corresponding to personal characteristics, although the effect sizes tended to be small. Even after controlling the effect of the number of study semesters, intelligence and academic self-concept contributed to scientific competencies. Among the predictors, intelligence explained the most variance. Surprisingly, none of the personality-related variables could explain variance in scientific competencies in our sample.

Mittwoch
11. März

14:00 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen I

Rott, Benjamin; Leuders, Timo; Ostermann, Andreas

Facetten mathematischer Forschungskompetenz – epistemologische Überzeugungen und kritisches Denken

Theoretischer Hintergrund: Im Laufe ihrer Ausbildung müssen Studierende des Lehramtes wissenschaftliche Kompetenzen in verschiedenen, heterogenen Bezugswissenschaften erwerben: in den Bildungswissenschaften, aber auch in den Disziplinen der von ihnen gewählten Schulfächer. Wissenschaftsbezogene Kompetenzen solcher Art können einerseits nach ihrem Domänenbezug und andererseits nach den Facetten Wissen und Können sowie Überzeugungen unterschieden werden. Innerhalb des komplexen Kompetenzbündels, das im Projekt LeScEd (Learning the Science of Education) in bildungswissenschaftlichen Studiengängen untersucht wird, widmen wir uns im hier berichteten Teilprojekt (1) dem kritischen Denken und (2) Überzeugungsstrukturen von Lehramtsstudierenden. Wir konzentrieren uns hierbei auf die Wissenschaft Mathematik, da Überzeugungen von Lernenden oft wenig reflektiert sind und in starkem Gegensatz zur reichhaltigen Diskussion in der Mathematikphilosophie stehen.

(1) Als eine Facette mathematikbezogener Forschungskompetenz, wie sie auch für ein Studium des Lehramtes bedeutsam ist, fokussieren wir auf das kritische Denken. Stanovich und Stanovich (2010, S. 210 ff.) verorten kritisches Denken in einem dreistufigen Modell des Denkens als Prozess der kritischen und reflektierten Überwachung von Problemlöseprozessen. Mit diesem Konstrukt lässt sich beispielsweise erklären, warum Personen mit den gleichen Voraussetzungen auf der Stufe des algorithmischen Denkens insbesondere bei komplexeren Aufgaben unterschiedlich gut abschneiden (vgl. ebd., S. 212 ff.).

(2) Im Bereich der Einstellungen und Überzeugungen fokussieren wir auf epistemologische Überzeugungen, d.h. Überzeugungen zur Natur des Wissens und zum Wissenserwerb, als weitere Facette bildungswissenschaftlicher Forschungskompetenz. Es handelt sich hierbei um einen seit vielen Jahren beforschten Gegenstand der Psychologie und Fachdidaktik, dessen Einfluss auf die Aufnahme und Verarbeitung von Informationen sowie auf Lern- und Lehrprozesse unbestritten sind (Hofer & Pintrich, 1997). In der Regel werden mehrere Dimensionen unterschieden (vgl. Hofer & Pintrich, 1997), wobei Lernende im Verlauf ihrer schulischen und universitären Ausbildung in jeder Dimension Entwicklungen durchlaufen können von eher starren, absolutivistischen Annahmen hin zu flexiblen, vernetzten und evaluativistischen Überzeugungen.

Insgesamt lässt sich allerdings konstatieren, dass sich sowohl die theoretische Fundierung als auch die empirische Erforschung dieses Forschungsbereichs als sehr heterogen darstellt (Bromme, Kienhues & Stahl, 2008). Als akzeptiert gilt mittlerweile, dass epistemologische Überzeugungen domänenspezifisch ausgeprägt sind und daher entsprechend erforscht werden sollten (Hofer, 2000). Neuere Studien deuten darauf hin, dass es nicht ausreicht, ausschließlich die Orientierung solcher Überzeugungen (z.B. „mathematisches Wissen ist sicher / unsicher“) zu erfassen (Greene & Yu, 2014). In Erweiterung der bisherigen psychologischen Messpraxis sollten zusätzlich zu den Überzeugungs-Orientierungen deren Begründungen erfasst werden (Rott, Leuders & Stahl, 2014). Fragestellung: Wie stehen epistemologische Überzeugungen mit anderen Aspekten forschungsbezogener Kompetenzen in Verbindung? Untersucht werden insbesondere die Beziehungen (1) der Fähigkeit zum kritischen Denken mit (2) den epistemologischen Überzeugungen von zukünftigen Mathematiklehrkräften, wobei letztere sowohl in ihren Orientierungen als auch in ihren Begründungen erfasst werden.

Methode: In unserem Beitrag präsentieren wir die Ergebnisse einer Studie an zwei Universitäten (PH Freiburg, n = 280, sowie Uni Duisburg-Essen, n = 187), in der wir (1) die Fähigkeit zum kritischen Denken mit (2) den epistemologischen Überzeugungen von zukünftigen Mathematiklehrern in Beziehung setzen. Auf der Basis qualitativer und quantitativer Vorstudien (Rott et al. 2014) wurden hierfür entsprechende Tests sowie Fragebögen mit geschlossenen und offenen Items entwickelt.

Ergebnisse: Die Ergebnisse deuten auf eine langsame, aber stetige Entwicklung der Fähigkeit zum mathematisch-kritisches Denken und eine stärker reflektierte Auseinandersetzung mit epistemologischen Fragen im Verlauf des Studiums hin. Des Weiteren zeigt sich ein Zusammenhang zwischen beiden Aspekten, wobei insbesondere die Begründung von Überzeugungen diese Zusammenhänge wesentlich stärker prägt ist als die Orientierung der Überzeugungen. In dem Vortrag werden die Ergebnisse in Bezug auf die beiden Konstrukte sowie ihre Diskriminanzvalidität diskutiert und es werden Rückschlüsse auf die Bedeutung der Konstrukte innerhalb der Wissenschaftskompetenz gezogen.

Mittwoch
11. März

14:00 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen I

Neumann, Irene; Rösken-Winter, Bettina; Lehmann, Malte; Durchhardt, Christoph; Heinze, Aiso; Nickolaus, Reinhold

Diagnose mathematischer Kompetenzen von Ingenieurstudierenden unter besonderer Berücksichtigung des Problemlösens

Theoretischer Hintergrund und Fragestellung: In ingenieurwissenschaftlichen Studiengängen sind bereits im ersten Studienjahr hohe Abbruchquoten zu verzeichnen (vgl. Heublein et al., 2012). Diese resultieren nicht zuletzt aus einer Überforderung der Studierenden in mathematischen Lehrveranstaltungen, vermutlich da es den Lernenden nicht gelingt, ihre schulisch erworbenen mathematischen Kompetenzen an der Hochschule effektiv zu nutzen bzw. diese nicht ausreichend anschlussfähig sind (vgl. Kajander & Lovric, 2005; Luk, 2005). Bisher wenig untersucht ist dabei, welche spezifischen Kompetenzfacetten im Bereich der Ingenieursmathematik für die Probleme der Studierenden zu Studienbeginn zentral sind. Ausgehend von den mathematischen Anforderungen im Ingenieursstudium (z.B. SEFI, 2013) kann der mathematischen Problemlösekompetenz eine essenzielle Rolle zugesprochen werden. Bedingungsfaktoren für erfolgreiches mathematisches Problemlösen – außerhalb des ingenieurwissenschaftlichen Kontexts – wurden in einer Vielzahl von Arbeiten untersucht und können zur Analyse herangezogen werden. Dazu gehören etwa eine adäquate mathematische Wissensgrundlage, flexibel aktivierbare Heuristiken zur Suchraumeinschränkung sowie Metakompetenzen hinsichtlich des Ablaufs eines Problemlöseprozesses (vgl. Schoenfeld, 1985; Bruder & Collet, 2011).

Methode: Für die explorative Analyse der spezifischen mathematischen Schwierigkeiten von ingenieurwissenschaftlichen Studierenden zu Studienbeginn wurde eine quantitativ-qualitative Forschungsstrategie verfolgt. Dazu wurde in zwei Studien ein Test aus dem Projekt KOM@ING (BMBF-Programm KoKoHs) eingesetzt, der schulisch erworbene und für das ingenieurwissenschaftliche Studium relevante mathematische Kompetenzen erfasst. In Studie 1 wurde dieser Test 1069 Studienanfängerinnen und Studienanfängern vorgelegt (792 Universität, 237 Fachhochschule, 40 keine Angabe; Studiengänge: Maschinenbau 754, Bauingenieurwesen 119, Fahrzeugtechnik 86, Rest sonstige oder keine Angabe). Die Daten wurden mithilfe eines eindimensionalen IRT-basierten Partial-Credit-Modells ausgewertet. In Studie 2 bearbeiteten 34 Ingenieurstudierenden (alle Universität, Studiengang Maschinenbau) zunächst den Test individuell und etwa eine Woche später eine Auswahl von 16 Testitems (darunter Problemlöseitems) erneut in Zweier- oder Dreiergruppen. Die Studierenden sollten während der Gruppenarbeitsphase laut denken, um die kognitiven Prozesse bei der Bearbeitung der Aufgaben zugänglich zu machen. Die Gruppenarbeitsphase wurde videografiert, transkribiert und gemäß der qualitativen Inhaltsanalyse (Mayring, 2000) anhand eines Kategoriensystems ausgewertet, das theoriebasiert deduktiv entwickelt und im Rahmen einer Pilotierung verfeinert wurde. Eine Doppelkodierung der Daten von zwei Gruppen zeigte eine hohe Interraterreliabilität (Cohen's kappa=.90).

Ergebnisse & Diskussion: Die Auswertung der Daten von Studie 1 ergab einen zufriedenstellenden Fit der Items zum Partial-Credit-Modell (WMNSQ zwischen 0.91 und 1.13) sowie eine ebenfalls zufriedenstellende Reliabilität (EAP/PV=.78; WLE=.76). Die Verteilungen von Itemschwierigkeiten und Personenfähigkeiten zeigten eine gute Passung, sodass der eingesetzte Test insgesamt als adäquat für die Stichprobe angesehen werden kann. Die Schwierigkeitsschätzer dienten zur Auswahl der Items für die Gruppenarbeitsphase aus Studie 2. Die Teilnehmenden aus Studie 2 unterschieden sich nicht signifikant in ihren Fähigkeiten von denen aus Studie 1 (MWLE1=-0.02, SDWLE1=1.13; MWLE2=-0.11, SDWLE2=1.38), sodass sie das Leistungsspektrum adäquat abdecken.

Die quantitative Auswertung in Studie 1 machte deutlich, dass viele Ingenieursstudierende gemessen an den Studienanforderungen eine eher geringe mathematische Kompetenz aufweisen. Dies betrifft insbesondere Fachhochschulstudierende. In der bisher vorliegenden qualitativen Auswertung in Studie 2 konnten drei Typen von nicht erfolgreichen Problemlöseprozessen identifiziert werden. Typ 1 umfasste einen adäquaten Umgang mit Heuristiken, scheiterte jedoch an fehlendem prozeduralen Wissen zu mathematischen Lösungsverfahren. Bei Typ 2 konnte bei bekannten Problemtypen Wissen über Lösungsschemata abgerufen werden, es war aber kein adaptiver Einsatz der Heuristiken bei unbekanntem Problemtypen möglich. Bei Typ 3 zeigten sich fehlende Metakompetenzen im Sinne einer inadäquaten Herangehensweise an Probleme als entscheidend.

Die bisherigen Ergebnisse machen deutlich, dass die Schwierigkeiten der Erstsemesterstudierenden ganz unterschiedliche Ursachen aufweisen. Entsprechend sollten Unterstützungsmaßnahmen im Rahmen von Vorkursen adäquate Diagnoseangebote mit zugehörigen Förderangeboten umfassen. Für Letztere kommen z.B. die Förderung von prozeduralem mathematischem Wissen, die Dekontextualisierung von Problemlöseheuristiken oder Metawissen über die Phasen eines Problemlöseprozesses in Frage.

Mittwoch
11. März

14:00 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen I

Vogel, Sylvia; Jenßen, Lars; Dunekacke, Simone; Grassmann, Marianne; Koinzer, Thomas; Wedekind, Hartmut; Schmude, Corinna; Blömeke, Sigrid

Entwicklung und Validierung eines Leistungstests zur Messung professioneller Kompetenz von angehenden frühpädagogischen Fachkräften im Bereich Mathematik

Theoretischer Hintergrund: Struktur, Niveau und Entwicklung professioneller Kompetenz von frühpädagogischen Fachkräften ist nach wie vor ein Forschungsdesiderat (Anders, 2012). Dies ist u.a. darauf zurückzuführen, dass die standardisierte Erfassung von Kompetenzen sich im Bereich der Frühpädagogik als eine äußerst anspruchsvolle Aufgabe erweist, da hier weniger das geplante und systematische gestalten von Lehr-Lernsituationen im Mittelpunkt steht, sondern Lernen in informellen und alltäglichen Settings stattfindet und damit stärker als im schulischen Kontext situationsspezifische Kompetenzen der Fachkräfte erfordert (Liegler, 2008).

Im Rahmen des vom BMBF geförderten Projekts KomMa – Struktur, Niveau und Entwicklung professioneller Kompetenz von Erzieherinnen und Erziehern im Bereich Mathematik wurden erstmals Testinstrumente zur Erfassung des professionellen Wissens angehende frühpädagogischer Fachkräfte im Bereich Mathematik entwickelt. Die strukturelle Konzeptualisierung der Tests orientiert sich dabei an den Arbeiten von (Shulman, 1986) und unterscheidet zwischen dem mathematischen Fachwissen, dem mathematikdidaktischen und dem allgemein-pädagogischen Wissen.

Fragestellung: Der Beitrag geht auf die Entwicklung von Instrumenten zur Messung professionellen Wissens vor dem Hintergrund situierter und informeller Lehr-Lern-Settings ein und beschreibt dabei insbesondere die inhaltliche Ausdifferenzierung der theoretisch angenommenen Wissensfacetten. Des Weiteren wird auf die Validierung der entwickelten Instrumente eingegangen. Dabei werden überblicksartig verschiedene Schritte der Validierung vorgestellt, die im Projekt unternommen wurden (z.B. (Dunekacke, Jenßen, & Blömeke, 2015; Jenßen, Dunekacke, & Blömeke, 2015). Ausführlicher wird auf die faktorielle Validität eingegangen und Ergebnisse zur dimensional Struktur des professionellen Wissens angehender frühpädagogischer Fachkräfte berichtet.

Methode: Im Rahmen des Beitrages werden Ergebnisse aus der sog. Hauptstudie, bei der knapp 2.000 angehende frühpädagogische Fachkräfte in den beiden in Deutschland derzeit möglichen Ausbildungswegen der fachschulischen Ausbildung bzw. des fachhochschulischen Studiums getestet wurden. Außerdem wurde der Ausbildungszeitpunkt berücksichtigt, indem Teilnehmerinnen und Teilnehmer am Beginn und am Ende der Ausbildung befragt wurden.

Die Analyse erfolgt mit der Software MPlus im Rahmen einer Zwei-Ebenen-Multiple-Group-Analyse, bei denen die vier beschriebenen Gruppen (Fachschule vs. Fachhochschule und Beginn vs. Ende der Ausbildung) sowie die Lehrveranstaltungen bzw. Klassen, im Rahmen derer die Testungen stattfanden, berücksichtigt werden. Für die vier genannten Gruppen wird außerdem auf die Frage der Messinvarianz eingegangen.

Ergebnisse: Die Ergebnisse belegen, dass auch bei frühpädagogischen Fachkräften ein dreidimensionales Modell, welches zwischen dem mathematischen Fachwissen, dem mathematikdidaktischen und dem allgemein-pädagogischen Wissen unterscheidet signifikant besser zu den Daten passt, als ein eindimensionales Modell. Dabei korrelieren das mathematische und das allgemein-pädagogische Wissen erwartungsgemäß am niedrigsten miteinander (0.61), das mathematische und das mathematikdidaktische Wissen in mittlerer Höhe (0.70) sowie das mathematikdidaktische und das allgemein-pädagogische Wissen hoch (0.92) miteinander.

Damit ähnelt die Korrelationsstruktur grundsätzlich der von angehenden Lehrkräften (Blömeke, Kaiser, & Lehmann, 2008), unterscheidet sich jedoch in der Stärke der Zusammenhänge. Im Rahmen des Beitrages wird insbesondere der starke Zusammenhang zwischen dem mathematikdidaktischen und dem allgemein-pädagogischen Wissen diskutiert. Es werden mögliche Gründe aus der Perspektive der methodischen Konzeptualisierung des Tests und der inhaltlichen Ausgestaltung der pädagogischen Arbeit im Elementarbereich diskutiert.

Mittwoch
11. März

14:00 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen I

Gerick, Julia

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 1

Mit der International Computer and Information Literacy Study (ICILS 2013) werden erstmalig computer- und informationsbezogene Kompetenzen im Sinne von ICT Literacy von Schülerinnen und Schülern der Sekundarstufe I international vergleichend, repräsentativ und computerbasiert, mittels Live-Software-Umgebung und eigens für die Studie entwickelter Aufgabenformaten untersucht und ein empirisch geprüftes Kompetenzstufenmodell auf der Datengrundlage der Testergebnisse von rund 60 000 Schülerinnen und Schülern aus weltweit 21 Bildungssystemen vorgelegt (vgl. Fraillon, Schulz & Ainley, 2013). Die Studie stellt für Deutschland höchst relevante Informationen über den Kompetenzstand von Achtklässlerinnen und Achtklässlern in dieser fächerübergreifenden Schlüsselkompetenz, medial vermittelte Informationen zu verstehen, zu nutzen und zu kommunizieren zu Verfügung und untersucht die Rahmenbedingungen des Kompetenzerwerbs zur Verfügung (vgl. Bos, Eickelmann, Gerick et al., 2014, in Druck). ICILS 2013 erweitert damit den Kanon der regelmäßig in Studien international vergleichend erfassten Kompetenzen in den Bereichen Mathematik, Lesen und Naturwissenschaften um den im 21. Jahrhundert immer relevanter werdenden Bereich der computer- und informationsbezogenen Kompetenzen.

Im Rahmen dieses zweiteiligen Symposiums werden zentrale Befunde von ICILS 2013 sowie vertiefende Analysen vorgestellt. Die Inhalte des Symposiums gehen dabei über die nationale Berichtslegung (vgl. Bos, Eickelmann, Gerick et al., in Druck, 2014) hinaus und nehmen thematisch besonderen Bezug zum Umgang mit Heterogenität in diesem Kompetenzbereich auf. Die Beiträge nutzen unterschiedliche quantitative methodische Zugänge, die der Mehrebenenstruktur und Stichprobengröße der Studie Rechnung tragen.

Der erste Teil des Symposiums geht mit vier Beiträgen zunächst auf die Anlage und Durchführung der Studie, die theoretische Verortung sowie auf Befunde zum internationalen Vergleich der Schülerkompetenzen und zu schulischen Rahmenbedingungen ihres Erwerbs ein. Eickelmann, Bos und Gerick fokussieren in ihrem Beitrag auf die methodische Anlage sowie die Beschreibung der Erhebungsinstrumente der Studie und geben somit einen Überblick über das Untersuchungsdesign von ICILS 2013. Anschließend stellen Goldhammer, Senkbeil, Eickelmann und Bos die im Rahmen von ICILS 2013 entwickelten Kompetenzstufen dar, betrachten vergleichend verschiedene Konzepte von ICT-Literacy und nehmen darin eine theoretische Verortung der computer- und informationsbezogenen Kompetenzen vor. Bos, Eickelmann und Gerick stellen nachfolgend die Befunde des internationalen Vergleichs der computer- und informationsbezogenen Kompetenzen von Achtklässlerinnen und Achtklässlern vor und fokussieren dabei vertiefend für Deutschland auch auf Schulformunterschiede. Der letzte Beitrag in diesem ersten Teil des Symposiums betrachten Gerick, Eickelmann, Schaumburg und Kahnert die schulischen Lehr- und Lernbedingungen sowie ihre Bedeutung für computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern.

Mittwoch
11. März

14:00 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 1

Chair(s)
Bos, Wilfried
Eickelmann, Birgit
Gerick, Julia

Eickelmann, Birgit; Bos, Wilfried; Gerick, Julia; Kahnert, Julia

Die International Computer and Information Literacy Study (ICILS 2013): Anlage, Ziele und Instrumentierung der Studie

Aufgrund des gesellschaftlichen Wandels und der fortschreitenden Technisierung aller Lebens- und Arbeitsbereiche besteht ein großer Bedarf hinsichtlich der Erforschung der fächerübergreifenden computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern. Im Gegensatz zu bereichsspezifischen Kompetenzen, die bereits umfangreich in den letzten Jahren untersucht wurden, stellt sich in Bezug auf den vorgenannten Kompetenzbereich zunächst die Frage nach einer geeigneten Konzeptionierung und Operationalisierung sowie einer darauf aufbauenden Entwicklung von Testinstrumenten, die mit der Studie ICILS 2013 vorgelegt werden.

Mit ICILS 2013 werden erstmalig vor dem Hintergrund eines theoretischen Rahmenmodells computer- und informationsbezogene Kompetenzen von weltweit mehr als 60.000 Schülerinnen und Schülern computerbasiert und international vergleichend in 21 Bildungssystemen, darunter drei Benchmark-Teilnehmer, gemessen (Bos, Eickelmann, Gerick et al., 2014, in Druck; Fraillon, Schulz & Ainley, 2013; Fraillon, Ainley, Schulz, Friedmann & Gebhardt, 2014, in Druck). Das Konstrukt der computer- und informationsbezogenen Kompetenzen, das der Studie zugrunde liegt, umfasst zwei zentrale Teilbereiche: (1) Informationen sammeln und organisieren sowie (2) Informationen erzeugen und austauschen. Weiterhin wird mit ICILS 2013 untersucht, unter welchen Rahmenbedingungen der Kompetenzerwerb stattfindet. Dazu wird der Schülertest, an dem pro teilnehmende Schule 20 zufällig ausgewählte Schülerinnen und Schüler der achten Jahrgangsstufe teilnehmen, um einen Schülerfragebogen, der u.a. Fragen zur schulischen und außerschulischen Nutzung digitaler Medien, zum Wissen über sowie Einstellungen zu neuen Technologien enthält, ergänzt. Zusätzlich werden schriftliche Befragungen von Lehrpersonen, die in der 8. Jahrgangsstufe unterrichten, Schulfragebögen eingesetzt.

Der vorliegende Beitrag fokussiert auf die Grundlagen der Studie ICILS 2013. Er stellt überblicksartig entlang eines theoretischen Rahmenmodells (vgl. Eickelmann, Bos, Gerick & Kahnert, 2014, in Druck) die theoretische und empirische Anlage der Studie vor und gibt Einblicke in die zentralen Aspekte der Konzeption hinsichtlich der computerbasierten Testmodule und der Aufgabenkonstruktion. Zur Erfassung der computer- und informationsbezogenen Kompetenzen der Achtklässlerinnen und Achtklässler werden in ICILS 2013 computerbasierte Tests eingesetzt. In einem Multimatrix-Design bearbeitet jede Schülerin bzw. jeder Schüler zwei der vier 30-minütigen Testmodule. Innerhalb der Module kommen drei verschiedene Itemformate zum Einsatz: (1) konventionell angelegte nicht interaktive Testitems (information-based response tasks, u. a. Multiple-Choice- und Drag-and-Drop-Aufgaben); (2) Performanzaufgaben (skills tasks), deren Lösung die Nutzung von Software- oder Computeranwendungen erfordert und (3) Autoreaufgaben (authoring tasks), die den komplexesten Aufgabentyp darstellen und bei denen die Schülerinnen und Schüler Informationsprodukte (z.B. Präsentationen) unter Verwendung von computerbasierten Live-Software-Anwendungen bearbeiten müssen. Diese Testanwendungen wurden speziell für ICILS 2013 entwickelt. Für die Bearbeitung der Autoreaufgaben ist teilweise die gleichzeitige Nutzung verschiedener Programme notwendig (z.B. E-Mail, Internetseiten, Tabellenkalkulation, Textverarbeitung oder Multimedia-Software), um den Realitätsbezug der Aufgaben zu gewährleisten. Alle Aufgaben innerhalb eines Testmoduls sind durch ein alltagsnahes Thema miteinander verknüpft (z.B. Planung eines außerschulischen Sportangebotes).

Mit diesem Beitrag werden die Anlage, Ziele und Instrumentierung der Studie ICILS 2013 vorgestellt. Neben dem theoretischen Rahmenmodell werden Beispielaufgaben aus den computerbasierten Testmodulen erläutert. Diese werden den Teilbereichen und Aspekten des Konstruktes der computer- und informationsbezogenen Kompetenzen zugeordnet, auf den Kompetenzstufen verortet und ihre Lösungshäufigkeiten auf nationaler und internationaler Ebene vorgestellt.

Mittwoch
11. März

14:00 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 1

Chair(s)
Bos, Wilfried
Eickelmann, Birgit
Gerick, Julia

Goldhammer, Frank; Senkbeil, Martin; Eickelmann, Birgit; Bos, Wilfried

Das Konzept der computer- und informationsbezogenen Kompetenzen in ICILS 2013: Definition und Einordnung

Die International Computer and Information Literacy Study (ICILS 2013) definiert informations- und computerbezogene Kompetenzen als „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, S. 17). Somit steht diese Definition wie andere aktuelle Definitionen von ICT-Literacy für die gleichzeitige Bewältigung von technischen wie auch kognitiven Anforderungen im Kontext computergestützter Informationsverarbeitung. In diesem (theoretischen) Beitrag wird zunächst in die Rahmenkonzeption von ICILS 2013 eingeführt, auf deren Grundlage die ICILS 2013-Messinstrumente entwickelt wurden. Konzeptionell ist das erfasste Konstrukt in zwei Teilbereiche organisiert (vgl. Senkbeil, Goldhammer, Bos, Eickelmann, Schwippert & Gerick, 2014, in Druck): Der Teilbereich I, Informationen sammeln und organisieren, repräsentiert rezeptive Aspekte, d.h. Wissen über und Verständnis von Computernutzung, auf Informationen zugreifen und Informationen bewerten, sowie Informationen verarbeiten und organisieren. Der Teilbereich II, Informationen erzeugen und austauschen, repräsentiert produktive Aspekte, d.h. Informationen umwandeln, erzeugen, kommunizieren und austauschen sowie Informationen sicher nutzen.

Zur Einordnung des in ICILS 2013 gemessenen Kompetenzkonstruktes wird es im vorliegenden Beitrag vor dem Hintergrund verwandter Konstruktdefinitionen und zugehöriger Testkonzeptionen diskutiert und die Ansätze miteinander verglichen. Zu den einschlägigen Konstrukten im Bereich des internationalen Large Scale-Assessments zählen die Digitale Lesekompetenz, wie sie im Programme for International Student Assessment (PISA) 2009 und 2012 (OECD, 2013) erfasst wurde, sowie Technologiebasiertes Problemlösen im Programme for the International Assessment of Adult Competencies (PIAAC; OECD, 2009). Auf nationaler Ebene ist weiterhin das Konzept der technologischen und informationsbezogenen Literacy zu nennen, das im Rahmen der National Educational Panel Study (NEPS) definiert wurde (Senkbeil, Ihme & Wittwer, 2013). Außerdem hat das Forschungsverbundprojekt CavE-ICT-PISA eine theoretische Rahmenkonzeption als Synthese der in der Literatur vorgestellten Konzeptionen von ICT-Skills entwickelt (Engelhardt et al., 2013) und versteht dabei ICT-Skills als kognitive Fertigkeiten, die sich aus traditionelleren Fertigkeiten herausgebildet haben, um auf ICT-Anforderungen reagieren und Informations- und Kommunikationstechnologien für eigene Zwecke erfolgreich einsetzen zu können.

Der Beitrag beschreibt ausgehend von den konzeptionellen Betrachtungen die in ICILS 2013 auf empirischer Basis gebildeten Kompetenzstufen zur inhaltlichen Interpretation der in einer Schülerpopulation erreichten Testwerte. Insgesamt wurden [Anzahl] Kompetenzstufen datengeleitet und auf Grundlage des Anforderungsniveaus der konstruierten Testaufgaben entwickelt. Sie reichen von rudimentärem Wissen über und Nutzen von Computeranwendungen bis hin zum sicheren Bewerten und Organisieren selbstständig ermittelter Informationen und Erzeugen von inhaltlich sowie formal anspruchsvollen Informationsprodukten. Die inhaltliche Bedeutung der Kompetenzstufen wird im Beitrag mit beispielhaften Testaufgaben aus den ICILS 2013 eingesetzten computerbasierten Schülertests für Achtklässlerinnen und Achtklässler illustriert.

Mittwoch
11. März

14:00 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 1

Chair(s)
Bos, Wilfried
Eickelmann, Birgit
Gerick, Julia

Bos, Wilfried; Eickelmann, Birgit; Gerick, Julia

IEA-ICILS 2013: Ergebnisse des internationalen Vergleichs computer- und informationsbezogener Kompetenzen von Achtklässlerinnen und Achtklässlern und Perspektiven für Deutschland

Mit der International Computer and Information Literacy Study (ICILS) 2013 werden erstmalig computerbasiert computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern der achten Jahrgangsstufe in 21 Bildungssystemen vergleichend untersucht. Im Gegensatz zu in Deutschland regelmäßig durchgeführten internationalen Vergleichsstudien mit dem Schwerpunkt auf bereichsspezifischen Kompetenzen, fokussiert ICILS 2013 auf fächerübergreifende Kompetenzen, von denen angenommen wird, dass sie im Zuge des Wandels zur Wissens- und Informationsgesellschaft von zentraler Bedeutung sind, um erfolgreich am Leben im Alltag, in der Schule sowie im Beruf und in der Gesellschaft teilhaben zu können (vgl. Bos, Eickelmann, Gerick et al., 2014, in Druck).

In dem der Studie zu Grunde liegenden theoretischen Rahmenmodell von ICILS 2013 sind die computer- und informationsbezogenen Kompetenzen der Schülerinnen und Schüler als Outcome verortet, welcher durch verschiedene Voraussetzungs- und Prozessfaktoren beeinflusst werden können. Dieser Beitrag fokussiert auf die Ebene des Outcome (vgl. Bos, Eickelmann & Gerick, 2014, in Druck) und bildet den Ausgangspunkt für die weiteren Beiträge im Rahmen dieses zweiteiligen Symposiums.

Folgenden Fragestellungen werden in diesem Beitrag betrachtet: (1) Welche Unterschiede ergeben sich in Bezug auf computer- und informationsbezogene Kompetenzen von Achtklässlerinnen und Achtklässlern im internationalen Vergleich? (2) Welche Unterschiede in den Schülerleistungen hinsichtlich der computer- und informationsbezogenen Kompetenzen ergeben sich für Deutschland zwischen Schulformen?

Die erste Fragestellung, die mit Hilfe der internationalen ICILS-2013-Stichprobe mit rund 60.000 Schülerinnen und Schüler untersucht wird, knüpft an die erste der vier internationalen Forschungsfragen von ICILS 2013 an, die den Ländervergleich fokussiert (vgl. Fraillon, Schulz & Ainley, 2013; Fraillon et al., 2014, in Druck). Die zweite Forschungsfrage greift die Möglichkeit auf, vertiefend für Deutschland Schulformunterschiede hinsichtlich der Kompetenzen von Achtklässlerinnen und Achtklässlern an Gymnasien und an anderen Schulformen der Sekundarstufe I zu untersuchen.

Die Ergebnisdarstellung fokussiert im Rahmen des internationalen Vergleichs einerseits auf das Leistungsniveau der Schülerinnen und Schüler („Länderranking“) sowie auf die Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen. Darüber hinaus werden für Deutschland vertiefend differenziert nach Schulform das Leistungsniveau sowie die Kompetenzstufenverteilung betrachtet.

Die Ergebnisse werden vor dem Hintergrund aktueller Entwicklungen in Deutschland und in den beteiligten Bildungssystemen hinsichtlich des Lehrens und Lernens mit neuen Technologien – u.a. Rahmenpläne, Konzepte zur IT-Ausstattung, Lehreraus- und -fortbildung – diskutiert und eingeordnet. Darüber hinaus erfolgt perspektivisch mit Blick auf Deutschland die Ableitung möglicher Implikationen und Herausforderungen.

Mittwoch
11. März

14:00 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 1

Chair(s)
Bos, Wilfried
Eickelmann, Birgit
Gerick, Julia

Gerick, Julia; Eickelmann, Birgit; Schaumburg, Heike; Kahnert, Julia

Die Bedeutung von schulischen Lehr- und Lernbedingungen für die computer- und informationsbezogenen Kompetenzen von Achtklässlerinnen und Achtklässlern im Kontext von ICILS 2013

Vor allem für den Bereich des Lernens mit digitalen Medien, der sehr ressourcenintensiv ist, unterliegt schulisches Lernen gegebenen Rahmenbedingungen, die sich hemmend oder förderlich auf Lehr- und Lernprozesse auswirken können. Durch technologische Entwicklungen werden immer neue Herausforderungen an Schulen und Lehrpersonen herangetragen, moderne und gleichzeitig den pädagogischen Anforderungen angemessene Lehr- und Lernbedingungen zu schaffen (vgl. u.a. Eickelmann, 2010; Gerick & Eickelmann, 2014; Schaumburg, 2003; Schulz-Zander, 2001).

Dieser Beitrag fokussiert vor diesem Hintergrund auf der Grundlage der Studie ICILS 2013 in einem ersten Schritt auf eine Beschreibung relevanter Lehr- und Lernbedingungen auf Schulebene im internationalen Vergleich. Neben der IT-Ausstattung und dem IT-Support werden auch die Fort- und Weiterbildungspraxis sowie Aspekte der Lehrerkooperation in Bezug auf digitale Medien in den Blick genommen (vgl. Gerick, Schaumburg, Eickelmann & Kahnert, 2014, in Druck). In dem theoretischen Rahmenmodell zum Erwerb computer- und informationsbezogener Kompetenzen, welches der Studie ICILS 2013 zugrunde liegt, werden die schulischen Lehr- und Lernbedingungen als Voraussetzungen auf der Schul- und Klassenebene verortet. Von ihnen wird angenommen, dass sie unmittelbar Einfluss auf die Prozesse, wie z.B. der Nutzung von digitalen Medien im Unterricht durch Lehrpersonen sowie Schülerinnen und Schüler haben (vgl. Fraillon, Schulz & Ainley, 2013; Eickelmann, Bos, Gerick & Kahnert, 2014, in Druck). Da sich die Zuordnung einzelner Aspekte zu dem Bereich der Voraussetzungen allerdings nicht immer trennscharf von dem Bereich der Prozesse vornehmen lässt – nicht alle Bedingungsfaktoren sind den Schulen nur vorgegeben, sondern können vielmehr auf Einzelschulebene gestaltet werden (vgl. u.a. Eickelmann, 2010) – wird mit diesem Beitrag der Frage nachgegangen, wie sich eine direkte Verbindung zwischen den oben genannten Merkmalen der schulischen Lehr- und Lernprozesse und den Schülerleistungen hinsichtlich der computer- und informationsbezogenen Kompetenzen abbilden lässt. Daher betrachtet dieser Beitrag in einem zweiten Schritt das Zusammenspiel der oben genannten schulischen Lehr- und Lernbedingungen und ihrer Bedeutung für die fächerübergreifende Schlüsselkompetenz der computer- und informationsbezogenen Kompetenzen.

Insgesamt bearbeitet dieser Beitrag im Kontext der schulischen Lehr-Lernbedingungen des Erwerbs computer- und informationsbezogener Kompetenzen die folgenden zwei Fragestellungen: (1) Wie stellen sich die schulischen Lehr- und Lernbedingungen in diesem Kontext in Deutschland und im internationalen Vergleich dar? (2) Welche Bedeutung haben schulische Lehr- und Lernbedingungen für die computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern der achten Jahrgangsstufe in Deutschland?

Die erste Fragestellung zum internationalen Vergleich wird auf Grundlage der internationalen ICILS-2013-Datenbasis von rund 60.000 Schülerinnen und Schülern und etwa 35.000 Lehrpersonen bearbeitet und fokussiert auf den internationalen Vergleich der oben genannten Lehr- und Lernbedingungen mittels deskriptiver Analysen. Vertiefend wird im Rahmen der zweiten Fragestellung für Deutschland mittels Mehrebenenpfadmodell auf der Grundlage von Daten auf Schul- und Schülererebene mit der Software Mplus 7 das Zusammenspiel der genannten schulischen Lehr- und Lernbedingungen – IT-Ausstattung und -Support, Fortbildungen und die Rolle von Lehrerkooperationen – betrachtet. Es wird untersucht, ob sich Effekte der schulischen Lehr- und Lernbedingungen auf die computer- und informationsbezogenen Kompetenzen der Schülerinnen und Schüler identifizieren lassen. Die Befunde werden vor dem Hintergrund des Forschungsstandes und hinsichtlich weiterführender Implikationen und Entwicklungsbedarfe diskutiert.

Mittwoch
11. März

14:00 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 1

Chair(s)
Bos, Wilfried
Eickelmann, Birgit
Gerick, Julia

Manzel, Sabine

Konflikt und Diskurs in der politischen und zivilgesellschaftlichen Bildung

Der schulische Bildungsauftrag umfasst den Auftrag, Kinder und Jugendliche zur Demokratie zu erziehen und ihnen Kompetenzen zu vermitteln, die politische und zivilgesellschaftliche Teilhabe ermöglichen. Von Schulen wird in diesem Zusammenhang der produktive Umgang mit Differenz, die konstruktive Lösung von Konflikten und eine Kultur der Anerkennung erwartet (BLK-Abschlussbericht, 2007, S.10f.). In den gesellschaftswissenschaftlichen Fächern werden Toleranz, Argumentationsfähigkeit und Verhandlungsfähigkeit zum Ausgleich von Interessenkonflikten als originärer Bestandteil der Fachkompetenz im Politikkompetenzmodell (Detjen et al. 2012) betrachtet. Auf der unterrichtspraktischen Ebene diskutieren Schüler/-innen z.B. in Rollenspielen kontrovers über politische Streitfragen. Konflikte werden somit im Politikunterricht nicht nur als Inhalte zum Gegenstand (Van de Vliert, 1997), sondern auch fachspezifisch simuliert in politischen Debatten. Kontroversen Diskussionen wird ein hohes Potential für kognitiv aktivierende Lernprozesse (Kunter & Trautwein, 2013) zugeschrieben.

Im Disziplinen-Cluster „Erziehungswissenschaft und Didaktiken der Gesellschaftswissenschaften“ geht es u.a. um die Frage, ob die Adaption eines psychologischen Messinstruments zu den vier Konfliktstilen „Vermeiden, Nachgeben, Dominieren und Integrieren“ (Rahim, 1983) geeignet ist, nicht nur allgemeine Interaktionsprozesse im (demokratischen) Klassenraum zu beschreiben und zu analysieren, sondern auch fachspezifisch zur Aufklärung von diskursiver Partizipation in Argumentationsphasen im Politik-/SoWi-Unterricht beitragen kann.

Es bedarf neben einer Operationalisierung und validen Messung von theoretischen Konfliktdimensionen auch der empirischen Klärung der Zusammenhangsstruktur zwischen Konfliktstilen, demokratischem Klassenklima, diskursiver Partizipation im Politikunterricht, Motivation und Wissensentwicklung. In der demokratiebezogenen Bildungsforschung und der Politikdidaktik liegen systematische Studien dazu vor. Hintergrundvariablen wie Gender und kulturelles Kapital zeigen signifikante Effekte auf die Ausbildung von Konfliktstilen und diskursiver Beteiligung im Fachunterricht.

Die Vielfalt des Symposiums zeigt sich ebenso in der Methode der Datenerhebung und -auswertung. Die Quellen der Daten sind vielfältig: BLK-Programm „Demokratie lernen und leben“, Schüler/-innen der Sekundarstufen allgemeinbildender Schulen in Niedersachsen, Ba-Wü und NRW. Zu den Instrumenten zählen standardisierte Fragebögen, die quantitativ statistisch ausgewertet werden (Item-Selektion in ConQuest, Strukturgleichungen und Mehrebenenanalysen in MPlus), die Adaption eines Inventars zu Konfliktstilen aus der US-Organisationspsychologie (konfirmatorische Faktorenanalyse zur Überprüfung der Dimensionalität), standardisierte Videodaten einer Fish-Bowl-Diskussion, die sowohl qualitativ als auch quantitativ analysiert werden (siehe Einzelabstracts).

Folgende Einzelbeiträge beschäftigen sich mit Konflikt- und Diskursstilen sowie Einflussvariablen wie Gender, Migrationshintergrund, kulturellem Kapital, Gewalterfahrung und Wissen: (1) Abs, Roczen, Bender & Filsecker: Entwicklung eines Instruments zur Erfassung von Konfliktstilen bei Schüler/-innen, (2) Weißeno & Landwehr: Konfliktstile und Motivation im Politikunterricht, (3) Gronostay: Prozessanalyse von Schüler/-innen-Argumentationen im politischen Fachunterricht – Ergebnisse einer Videostudie, (4) Oberle & Forstmann: Demokratisches Unterrichtsklima und Geschlecht als Faktoren der diskursiven Beteiligung von Schüler/-innen im politischen Fachunterricht, (5) Biedermann, Diskutant

Mittwoch
11. März

14:00 Uhr

HGB 40

Chair(s)

Manzel, Sabine

Abs, Hermann Josef; Filsecker, Michael

Entwicklung eines Instruments zur Erfassung von Konfliktstilen bei Schüler/-innen

Theoretischer Hintergrund: Die Bereitschaft Konflikte gewaltfrei und mit der Absicht eines Ausgleichs der Interessen aller Beteiligten zu lösen, stellt ein wesentliches Element sowohl prosozialer Erziehung als auch politischer Bildung dar (Detjen et al. 2012). Dagegen zeigen sich in Schulen zum Teil problematische Konflikthaltungen, wie solche, die den Anderen zu dominieren trachten, und auch Tendenzen Konflikte gänzlich zu vermeiden (z.B. Reinhardt, 2006). Jedoch fehlt es in Deutschland bislang an Instrumenten um die Einstellungen von Schülern und Schülerinnen valide und reliabel zu erfassen. Forschung zur Entwicklung von Einstellungen zu Konflikten ist auf solche Instrumente angewiesen. Für den Bereich der professionellen Interaktion von Berufstätigen in wirtschaftlichen Unternehmen wurden entsprechende Instrumente schon seit längerem erprobt (Rahim, 1983; Bilsky & Wülker, 2000). Dabei werden bei Erwachsenen zumindest vier Konfliktstile, nämlich „Dominieren“, „Nachgeben“, „Vermeiden“ und „Integrieren“ unterschieden.

Fragestellung: Noch nicht untersucht wurde, inwiefern das Konzept der Konfliktstile mit seinen spezifischen Ausprägungen aus der Organisationspsychologie in einer Population von Schüler/-innen die Entwicklung eines Messinstruments anleiten kann. Der Vortrag geht der Frage nach, ob 26 konfliktbezogene Items, die für Kontext Evaluation eines Schulentwicklungsprogramms sprachlich und inhaltlich adaptiert wurden, psychometrischen Kriterien einer differenziellen Messung unterschiedlicher Konfliktstile im Sinne des theoretischen Konstrukts genügen.

Methode: Die Stichprobe besteht aus 3.822 Schüler/-innen der Jahrgänge 8 bis 10 aus 65 Schulen. Nachdem, die Skalen im Zuge einer früheren Untersuchung lediglich mittels einer Hauptkomponentenanalyse und Reliabilitätsanalyse bestimmt worden waren (Abs et al. 2007), dient nun eine konfirmatorische Faktorenanalyse dazu, zunächst die Eindimensionalität der Skalen zu prüfen und diese hinsichtlich der jeweils integrierten Items zu optimieren. Sodann werden verschiedene alternative Modelle mit allen Items getestet, um zu prüfen, ob tatsächlich von vier getrennte Dimensionen (unterschiedliche Konfliktstilen) angenommen werden darf oder ob Modelle, die eine geringere Anzahl von Dimensionen bzw. übergeordnete Dimensionen annehmen, eine höhere Modellgüte aufweisen.

Ergebnisse: Die Ergebnisse verweisen zunächst auf die berechnete Annahme der Eindimensionalität der vier ursprünglich konzipierten Skalen. Eine Optimierung durch Weglassen einzelner Items ist möglich. Die gemeinsame Analyse der vier Konfliktstile führt zu einer guten Modellanpassung, bei der jedoch zwei der Dimensionen, nämlich „Nachgeben“ und „Integrieren“, eine hohe Korrelation aufweisen. Deshalb werden vergleichende konfirmatorische Faktorenanalysen mit Modellen gerechnet, die von drei statt vier Dimensionen oder einer übergeordneten Dimension ausgehen. Das führen aber sämtlich zu einer schlechteren Modellanpassung.

Mittwoch
11. März

14:00 Uhr

HGB 40

Chair(s)

Manzel, Sabine

Weißeno, Georg; Landwehr, Barbara

Konfliktstile und Motivation im Politikunterricht

Theoretischer Hintergrund: Die Klärung konkurrierender und konfligierender Ideen und Interessen sowie das Aushandeln von Konfliktregelungen wird in der Politikdidaktik als Konfliktfähigkeit bezeichnet (Behrmann et al. 2004). Das Austragen politischer Meinungsverschiedenheiten ist eine Zieldimension des Politikunterrichts. In politischen Diskussionen im Klassenzimmer können interpersonale Konflikte entstehen. Im Sinne des Ziels politischer Bildung sollen aber politische Meinungsverschiedenheiten interessengeleitet ausgetragen werden. Konfliktfähigkeit ist aus theoretischer Perspektive keine Politikkompetenz, sondern eine Persönlichkeitsdimension (Weißeno 2015).

Eine theoretische Verortung der normativen politikdidaktischen Annahmen kann durch das Dual Concern Model von Rubin et al. (1994) erfolgen. Es stellt die Strategiewahl in Konflikten als interessengeleitet dar. Das Interesse eines am Konflikt Beteiligten orientiert sich an der Durchsetzung der eigenen Ziele und an den Belangen der anderen. Dabei gibt es individuelle Unterschiede im Umgang mit Konflikten. Das Modell von Deutsch (1976) geht hingegen davon aus, dass die Strategiewahl motivationsgeleitet ist. Die Beteiligten fassen den Konflikt als kooperative oder als Konkurrenzsituation auf. Rahim (1992) stellt die Konfliktstile – vergleichbar mit den Strategien im Rubin-Modell – als Formen des Umgangs in Organisationen dar. Konfliktstile bieten Handlungsalternativen. Die jeweiligen Handlungsfolgen sind abzuwägen.

Das dispositionale Pendant zu den Handlungserwartungen bei politischen Diskussionen ist die subjektive Einschätzung der eigenen lernrelevanten Kompetenzen und Fähigkeiten (fachspezifisches Selbstkonzept) (Wild et al. 2006). Wer sich für Politik interessiert, kann in politischen Diskussionen seinen Standpunkt vertreten. Mit dem politischen Interesse (persönliches Interesse) und dem Selbstkonzept sehen sich die Schüler/-innen mehr oder weniger in der Lage, interessengeleitet Diskussionen zu führen. Das Wissen über die eigenen Fähigkeiten führt zu einer erhöhten Unterrichtsbeteiligung (Trautwein 2003). Schülerinnen geben an, Auseinandersetzungen in stärkerem Maß zu vermeiden und verglichen mit Schülern eher nachzugeben. Schüler setzen sich dagegen signifikant häufiger in Konfliktsituationen durch (Abs et al. 2007; Weisweiler et al.).

Fragestellungen: In diesem Beitrag gehen wir dem Zusammenhang der motivationalen Variablen mit den für das BLK-Programm adaptierten vier prototypischen Konfliktstilen nach dem Rahim Organizational Conflict Inventory (ROCI-II) nach und prüfen, inwieweit sich das fachspezifische Selbstkonzept und das politische Interesse auf die Konfliktstile auswirken. Zweitens prüfen wir, ob sich Geschlechtsstereotype für die Reaktionstendenz in politischen Diskussionen ausdifferenzieren lassen.

Methode: Im Beitrag gehen wir theoriegeleitet den Hypothesen nach. Erstens prüfen wir, inwieweit sich Effekte des fachspezifischen Selbstkonzepts mit einer Skala (6 Items) von Krampen (1998) und des politischen Interesses mit der Standardfrage aus der politischen Soziologie auf die jeweiligen Konfliktstile darstellen lassen. Dabei gehen wir davon aus, dass sich die wahrgenommene fachspezifische persönliche Motivation auf die Auswahl einer geeignet erscheinenden Handlungstendenz in politischen Diskussionen auswirkt. Zweitens prüfen wir, ob Mädchen andere Konfliktstile bevorzugen als Jungen. Die Hypothesen prüfen wir mit einer Strukturgleichung auf der Basis einer Befragung von 669 Schüler/-innen der 9. Klassen an Realschulen.

Ergebnisse: Die Sensitivität der Instrumente des Rahim-Modells scheint auch für Schüler/-innen gewährleistet. Zu den Geschlechterstereotypen zeigen Strukturgleichungsmodelle erwartungskonforme Befunde, d. h., es lassen sich ein dominanter Konfliktstil häufiger bei Jungen, ein integrierender und nachgebender häufiger bei Mädchen beobachten. Das fachspezifische Selbstkonzept wirkt sich positiv auf den integrativen und etwas auf den dominanten Konfliktstil aus. Das politische Interesse wirkt sich schwach auf alle Dimensionen aus, aber am stärksten und negativ auf den dominierenden Konfliktstil.

Mittwoch
11. März

14:00 Uhr

HGB 40

Gronostay, Dorothee

Prozessanalyse von Schüler/-innen-Argumentationen im politischen Fachunterricht – Ergebnisse einer Videostudie

Theoretischer Hintergrund: Unterrichtsmethoden zur Initiierung von argumentativen Lehr-Lernprozessen (z.B. Pro-Contra-Debatte, Fishbowl-Diskussion) sind für den politischen Fachunterricht von zentraler Bedeutung. Mündliches Argumentieren wird dabei als Facette kommunikativer politischer Handlungsfähigkeit verstanden, die der unterrichtlichen Förderung bedarf. Argumentationsprozesse im Unterricht dienen auch der Entwicklung anderer Kompetenzdimensionen, insbesondere der Elaboration und Anwendung von Fachwissen sowie der reflexiven Urteilsbildung (Detjen, Massing, Richter & Weißeno, 2012).

Vor dem Hintergrund (sozio)-kognitiver Lerntheorien in Anlehnung an Piaget und Vygotsky ist davon auszugehen, dass diskursive Argumentation ein hohes Potential für kognitiv aktivierende Lernprozesse birgt (Palincsar, 1998). Als stärker kognitiv aktivierend gelten insbesondere solche Unterrichtssettings, die die Anwendung und den Transfer von Fachwissen erforderlich machen und in denen Lernende aktiv verschiedene Lösungsmöglichkeiten (oder Positionen) zu einer komplexen Problemstellung diskutieren (Kunter & Trautwein, 2013). Für den politischen Fachunterricht existieren bislang jedoch kaum Studien, die Einblick in die Tiefenstruktur von Schüler/-innen-Argumentation geben (Petrik, 2013).

In argumentativen Lehr-Lernsettings können die einzelnen, von den Schüler/-innen vorgebrachten argumentativen Aussagen als beobachtbare Indikatoren der ihnen zugrunde liegenden kognitiven Prozesse aufgefasst werden (Sadler & Zeidler, 2004). Die Tiefenstruktur der Argumentation lässt sich dann anhand der Argumentationsstruktur, der eingebrachten Fachinhalte sowie der Transaktivität der Diskussion beschreiben.

Transaktivität meint die wechselseitige Bezugnahme auf Beiträge anderer Diskursteilnehmer/-innen, die eine notwendige Voraussetzung für ko-konstruktive Argumentationen darstellt (Felton & Kuhn, 2001; Stegmann, Weinberger & Fischer, 2011). Die Argumentationsstruktur lässt sich anhand der Unterscheidung funktionaler Begründungselemente (z.B. Konklusion, Fakt, Schlussregel oder Ausnahmebedingung) in Anlehnung an Toulmin (1958) analysieren. Im didaktischen Kontext wird die Qualität der Argumentation dabei in der Regel mit zunehmender Anzahl von Argumentationselementen sowie mit dem Vorkommen bestimmter Argumentationselemente (zum Beispiel Einwänden und Ausnahmebedingungen) assoziiert (Osborne, Erduran & Simon, 2004). Die Komplexität der Fachinhalte lässt sich anhand des Vernetzungsniveaus bemessen, wobei zum Beispiel das Aufzeigen von Zusammenhängen einem höheren Niveau entspricht als das Nennen von Einzelfakten (Kauertz, Fischer, Mayer, Sumfleth & Walpuski, 2010).

Ziele: Die Studie zielt auf die deskriptive Analyse der Tiefenstruktur von Argumentationsprozessen im politischen Fachunterricht. Als Beschreibungskategorien dienen die funktionale Struktur, die Transaktivität und fachinhaltliche Komplexität mündlicher Schüler/-innen-Argumentationen im Politikunterricht. Auch Zusammenhänge zwischen den fokussierten Dimensionen werden untersucht.

Forschungsdesign und Stichprobe: Es handelt sich um eine Videostudie zur Prozessqualität argumentativer Lehr-Lernprozesse. Zehn Schulklassen der Jahrgangsstufen acht und neun an Gymnasien in NRW erhalten Unterricht zur Parteiverbotsthematik (2 Doppelstunden à 90 Minuten) und führen abschließend eine kontroverse Diskussion (Methode: Fishbowl-Diskussion). In wechselnden Kombinationen diskutieren vier Schüler/-innen der Klasse, wobei eine vorgegebene Argumentationsrichtung zu vertreten ist.

Methode: Die Schüler/-innen-Argumentationen werden mit Hilfe eines Kategoriensystems zur Analyse der Struktur der Argumentation (Toulmin, 1958, Riemeier, von Aufschnaiter, Rogge & Fleischhauer, 2012), der Transaktivität (Felton & Kuhn, 2001; Stegmann, Weinberger & Fischer, 2011) und der fachinhaltlichen Komplexität (Manzel & Gronostay, 2013,) kodiert. Die Kodierung wird anhand der videografierten und transkribierten Schüler/-innen-Diskussionen mit der Software MAXQDA durch jeweils zwei Kodierer/-innen vorgenommen. Die Auswertung erfolgt quantitativ.

Ergebnisse: Ergebnisse liegen zum Zeitpunkt der GEBF-Tagung im März 2015 vor.

Mittwoch
11. März

14:00 Uhr

HGB 40

Chair(s)

Manzel, Sabine

Oberle, Monika; Forstmann, Johanna

Demokratisches Unterrichtsklima und Geschlecht als Faktoren der diskursiven Beteiligung von Schüler/-innen im politischen Fachunterricht

Theoretischer Hintergrund: Die diskursive Beteiligung von Schüler/innen im Politikunterricht dient nicht nur dem erfolgreichen Wissenserwerb, sondern auch der Entwicklung kommunikativer Handlungskompetenzen als Teilfacette der Politikkompetenz (Detjen et al. 2012). Voraussetzung für politische Diskussionen im Unterricht sind zum einen die Schaffung entsprechender Diskursgelegenheiten, zum anderen die Bereitschaft der Lernenden, diese wahrzunehmen und selbst aktiv daran teilzunehmen. Faktoren, die diese Teilhabe beeinflussen, sind für die Politische Bildung bislang kaum erforscht. Vermutungen bestehen hinsichtlich geschlechtsspezifischer Fachunterrichtsbeteiligung zu Lasten der Mädchen (vgl. Oberle 2013), systematische Studien hierzu fehlen jedoch bislang. Dass das sog. „demokratische Klassenklima“ bzw. „diskursive Unterrichtsklima“ Wissenserwerb und politische Wahlbereitschaft fördert, darauf weisen die CiVED-Studien (Torney-Purta et al. 2001) hin. Dass diese Effekte sowohl auf die individuelle Wahrnehmung des Unterrichtsklimas als auch die geteilte Klassenwahrnehmung zurückzuführen sind, konnte Watermann (2003) zeigen.

Fragestellung: Inwiefern das demokratische Unterrichtsklima die eigene aktive Beteiligung an politischen Diskussionen im Fachunterricht beeinflusst, ist bislang nicht erforscht. Auch bezüglich geschlechtsspezifischer Muster der Beteiligung im Politikunterricht bestehen weitreichende Forschungsdefizite. Die Studie geht der Frage nach, ob die Beteiligung der Schüler/-innen an Diskussionen im alltäglichen Fachunterricht durch a) das demokratische Unterrichtsklima (individuelle sowie von der Klasse geteilte Wahrnehmung) sowie b) das Geschlecht der Lernenden beeinflusst wird. Berücksichtigt werden dabei die geschlechtsspezifische Zusammensetzung der Klasse wie auch das Geschlecht der Lehrkraft. Kontrolliert werden darüber hinaus die sozio-demographischen Hintergrundvariablen Alter, kulturelles Kapital des Elternhauses und Migrationshintergrund der Schüler/-innen sowie die Schulform, da die begründete Annahme besteht, dass auch diese in Zusammenhang mit der Beteiligung an Diskussionen im Politikunterricht und/oder dem wahrgenommenen Unterrichtsklima stehen. Hierauf zurückzuführende „Scheineffekte“ sollen damit ausgeschlossen werden.

Methode: Untersucht wird ein Sample aus N = 714 Schülerinnen und Schülern der Jahrgänge 9 und 10 aus insgesamt 29 Klassen (Gymnasien und Gesamtschulen, überwiegend Niedersachsen; 67 % Gymnasium; 52,5 % weiblich; Altersdurchschnitt 15,4 Jahre). Die Datenerhebung erfolgte 2013 und 2014 mittels standardisiertem Fragebogen durch geschulte Versuchsleiter/-innen im Klassenverband. Messmodelle der interessierenden Konstrukte wurden ebenso wie latente Regressionen und Strukturgleichungen in MPlus 7.0 gerechnet, unter Berücksichtigung der hierarchischen Struktur der Daten. Dabei kamen überwiegend 4-stufig likert-skalierte Items zum Einsatz. Die latent modellierten Konstrukte weisen eine akzeptable Reliabilität und Datenpassung auf (Erhebungsinstrumente: Beteiligung im Politikunterricht: 4 Items, Cronbach's Alpha= .80; CFI= 1.00; RMSEA= .01; WRMR= .25; Chi²= 2.12(2)ns, teilweise adaptiert nach Kessels 2002; demokratisches Unterrichtsklima: 5 Items, Cronbach's Alpha= .78; CFI= 0.99; RMSEA= .05; WRMR= .55; Chi²= 14.05(5)*, nach Torney-Purta et al. 2001; kulturelles Kapital: „Bücherfrage“ mit 5 Antwortoptionen (ebd.); Migrationshintergrund: Geburtsländer Schüler/-in und Eltern, zu Hause gesprochene Sprachen).

Ergebnisse: Erste Datenanalysen bestätigen beide zentralen Annahmen: Das demokratische Unterrichtsklima hat einen positiven Effekt auf die aktive Beteiligung der Schüler/innen im Politikunterricht ($\beta = .273$), und die berichtete Aktivität fällt für Jungen signifikant höher aus als für Mädchen ($\beta = .354$). Die Varianz der Unterrichtsaktivität zwischen den Klassen (ICC = .04) spricht für die Untersuchung von Effekten auf Klassenebene. Die Ergebnisse der detaillierten Analysen unter Einbeziehung der kontrollierten Hintergrundvariablen sowie der Mehrebenenanalysen (in MPlus) werden in dem Vortrag präsentiert.

Mittwoch
11. März

14:00 Uhr

HGB 40

Chair(s)

Manzel, Sabine

Bremm, Nina

Herausforderungen und Strategien des konstruktiven Umgangs mit Heterogenität an sozial deprivierten Standorten

Schulen sind sozialräumliche Kontexte eingebettet und werden durch diese geprägt. So können Studien zeigen, dass die soziokulturelle Zusammensetzung der Schülerschaft bzw. die Sozialstruktur im Umfeld der Schule (zum Kontextbegriff Ditton, 2013) die Schulqualität und die Leistungsfähigkeit ihrer SchülerInnen mitbestimmen (Baumert et al., 2006; Faber, 2013). Gleichzeitig treffen SchülerInnen in belasteten Kontexten vermehrt auf sozialräumlich und familiär bedingte Restriktionen, wie etwa eine eingeschränkt hilfreiche elterliche Unterstützung in Bezug auf schulische Lerninhalte oder schwierige regionale Arbeitsmarktsituationen. Damit einher können Phänomene eines negativen Selbstkonzept oder einer wahrgenommenen Perspektivlosigkeit gehen, die durch familiales Armutserleben verstärkt werden (Groh-Samberg, 2009). Jedoch können belastete Lebenskontexte auch Opportunitäten eröffnen, bspw. mit Blick auf überdurchschnittliche Bildungsaspirationen, die durch positive Rollenmodelle bestärkt werden können, oder herkunftsbedingte Bildungsressourcen, deren Aktivierung Chance und Herausforderung für Schulen darstellt.

International zunehmend auch national existiert ein verstärktes Interesse an Schulen, die aufgrund ihrer Lage mit hohem Urbanisierungsgrad in ethnisch und/oder sozial segregierten Stadtteilen mit einer Kumulation von Bildungsrisiken konfrontiert sind (Baumert et al., 2006). Diese Schulen sind besonders herausgefordert, den heterogenen Bildungsvoraussetzungen ihrer SchülerInnen konstruktiv zu begegnen und diese zu entwickeln. Im Rahmen des Symposiums werden Herausforderungen solcher Schulen thematisiert und verschiedene Strategien im Umgang mit herkunftsbedingter Vielfalt bezugnehmend auf unterschiedliche Akteure des Bildungssystems vorgestellt.

Im Beitrag „Kriteriengeleitete und datengestützte Vernetzung von Schulen der Sekundarstufe I in der Metropole Ruhr“ wird das Projekt „Potentiale entwickeln – Schulen stärken“ vorgestellt. Dabei wird thematisiert, wie sich Kontextmerkmale sowie Merkmale der Schul- und Unterrichtsgestaltung an Schulen in herausfordernder sozialräumlicher Lage darstellen. Zudem werden Ergebnisse der kriteriengeleiteten Klassifikationen von Schulnetzwerken vorgestellt.

Die Beiträge „Sprachsensible Unterrichts- und Schulentwicklung – eine Frage der Einstellung?“ und „sehr sehr gutes Deutsch ist nicht genug – sprach-, kultur- und herkunftsbezogen rassialisierende Argumentationen bei PH-Lehrkräften in Österreich“ greifen Aspekte der sprachlichen Bildung von SchülerInnen bzw. Studierenden mit Migrationshintergrund auf. Dabei fokussiert der erstgenannte Beitrag Schul- und Unterrichtsentwicklung, indem neben der Einstellung der Lehrkräfte zur sprachlichen Vielfalt und dem selbstberichteten Umgang mit sprachlicher Heterogenität in der Unterrichtspraxis auch die Verankerung fächerübergreifender Sprachbildung auf organisationaler Ebene betrachtet wird. Der zweitgenannte Beitrag vertieft die Frage nach der Einstellung von Lehrkräften an PH's anhand hermeneutischer Sekundäranalysen.

Der Beitrag „Schulleitungshandeln in Schulen in schwieriger Lage. Ein systematischer Vergleich von Deutschland und den USA“ greift die Rolle von Schulleitungen effektiver Schulen in sozialräumlich benachteiligter Lage auf. Dabei werden international vergleichend in Form eines kontrastiven Fallvergleichs von Deutschland und den USA Unterschiede des Schulleitungshandelns unter Berücksichtigung bildungspolitischer und schulgesetzlicher Rahmenbedingungen herausgearbeitet.

Mittwoch
11. März

14:00 Uhr

HGB 50

Webs, Tanja; Wisberg, Eva; van Holt, Nils; Holtappels, Heinz Günter; van Ackeren, Isabell

Kriteriengeleitete und datengestützte Vernetzung von Schulen der Sekundarstufe I in der Metropole Ruhr

Befunde rezenter Schulleistungsvergleichsstudien zeigen, dass die Qualität der pädagogischen Arbeit an der Einzelschule ausschlaggebend dafür ist, inwiefern die Bildungserfolge von Schülerinnen und Schülern durch individuelle Lernvoraussetzungen und schulische Kontextbedingungen beeinflusst werden (vgl. Baumert et al., 2006; Muijs et al., 2004). Insbesondere Schulen in der Metropole Ruhr arbeiten in einem herausfordernden Umfeld, da sich die Bevölkerungsstruktur in bestimmten Regionen durch hohe Anteile an Arbeitslosen und Migranten – vor allem bei jungen Erwachsenen – auszeichnet. Dennoch leisten Schulen angesichts der regionalspezifischen Charakteristika des Schulstandortes und der heterogenen Schülerschaft bereits vielerorts vorbildliche Arbeit (vgl. Regionalverband Ruhr, 2012). Um systemisch und synergetisch im Umgang mit ähnlichen Kontextbedingungen zu wirken, wird daher zur stärkeren interschulischen Vernetzung einzelner Bildungsangebote und Initiativen geraten (vgl. ebd.). Gestützt wird diese Empfehlung durch Forschungsergebnisse, die verdeutlichen, dass Schulnetzwerke den Beteiligten Raum für intensiven Austausch über Erfahrungen, Informationen und Wissen bieten sowie die Teilhabe an anderen Praxen ermöglichen (vgl. Berkemeyer et al., 2009; Järvinen & van Holt, 2011). Da schulische Netzwerke die Lernprozesse der Akteure von- und miteinander begünstigen, gelten sie insgesamt als innovationsförderlich, um Schulentwicklungsprozesse voranzutreiben.

Vor diesem Hintergrund liegt das Ziel des Projekts Potenziale entwickeln — Schulen stärken im Allgemeinen darin, Schulen der Sekundarstufe I in schwieriger sozialräumlicher Lage in der Ruhrmetropole evidenz- und netzwerkbasierend weiterzuentwickeln. Konkret werden die am Projekt beteiligten Schulen für die Zusammenarbeit an eigenen Entwicklungsvorhaben vernetzt. Im Rahmen der Netzwerkarbeit werden kontextspezifische und bedarfsorientierte Unterstützungsangebote bereitgestellt. So sollen Schulen – ausgehend von ihren spezifischen Stärken und Schwierigkeiten – unterstützt werden, ihre Ressourcen auszuschöpfen und ihre Problemlösefähigkeit zu stärken. Durch die Vernetzung der Schulen können zudem bestehende Ansätze aufgegriffen und weiterentwickelt sowie neue Ansätze gemeinsam konzipiert, erprobt und reflektiert werden. Zur Begleitung und Unterstützung der Schulnetzwerke und Projektschulen arbeiten unterschiedliche Kooperationspartner aus Wissenschaft (Institut für Schulentwicklungsforschung an der TU Dortmund und Universität Duisburg-Essen) und Bildungsadministration (Qualitäts- und Unterstützungsagentur – Landesinstitut für Schule NRW) zusammen.

Um die skizzierten Potenziale interschulischer Vernetzung und Kooperation bestmöglich zu nutzen, werden die am Projekt beteiligten Schulen anhand der Kontextmerkmale (z.B. sozialräumlicher Besonderheiten, Charakteristika der Schülerkomposition) sowie der Merkmale der Gestaltungsqualität auf Schul- und Unterrichtsebene (z.B. Lehrerverbände, individuelle Förderung) zu überregionalen Netzwerken zusammengeschlossen. Auf diese Weise sollen Schulen systematisch, d.h. kriteriengeleitet und datengestützt über die Regionen hinweg vernetzt werden.

In Anbetracht der Projektziele wird im Rahmen des Vortrags folgenden Fragestellungen nachgegangen: (1) Wie stellen sich Kontextmerkmale sowie Merkmale der Schul- und Unterrichtsgestaltung an Schulen in besonders herausfordernder sozialräumlicher Lage dar?, (2) Inwiefern lassen sich die heterogenen schulexternen und pädagogisch-organisatorischen Merkmale der Schulen klassifizieren bzw. inwieweit lassen sich Gruppen von Schulen mit ähnlichen Ausprägungen und Profilen der Kontextmerkmale und der Merkmale der Schul- und Unterrichtsgestaltung identifizieren?

Die Datengrundlage bilden die Angaben unterschiedlicher schulischer Akteure der 40 Projektschulen, wie Schulleitungen, Lehrerkollegien, Schülerinnen und Schüler sowie Eltern, die im Kontext der fragebogengestützten Ausgangserhebung generiert wurden.

Die Ergebnisse der multiperspektivischen, empirisch-quantitativen Analysen werden dazu genutzt, Schulen mit ähnlichen Entwicklungsbedingungen, -bedarfen und -interessen zu identifizieren. Darauf basierend werden kriteriengeleitet und datengestützt (1) Netzwerke zusammengestellt, (2) Ziele und Schwerpunkte der Netzwerkarbeit abgeleitet sowie (3) im Rahmen der Netzwerkarbeit passgenaue Unterstützungsmaßnahmen angeboten.

Abschließend werden Chancen und Grenzen einer datengestützten Ermittlung schulischer Entwicklungsbedarfe und -interessen im Hinblick auf die Planung und Initiierung von Schulentwicklungsprozessen an Einzelschulen und damit hinsichtlich der Förderung der Schulqualität und -wirksamkeit diskutiert. Zudem wird erörtert, inwiefern die systematische, überregionale Vernetzungsstrategie im Projektkontext nachhaltige Kooperationsstrukturen schafft, in denen sich Schulen auch über die Projektlaufzeit hinaus entwickeln können.

Mittwoch
11. März

14:00 Uhr

HGB 50

Demski, Denise; Racherbäumer, Kathrin; van Ackeren, Isabell

Sprachsensible Unterrichts- und Schulentwicklung – eine Frage der Einstellung?

Theoretischer Hintergrund: Während einerseits Ergebnisse aus Leistungsvergleichsstudien Schülerinnen und Schülern aus eingewanderten Familien in deutschen Schulen relativ systematisch unterdurchschnittliche Leistungen bescheinigen (z.B. Pöhlmann et al. 2012), zeigt der Diskurs der interkulturellen Erziehungswissenschaft reichhaltige Ansätze auf, wie diese Situation verändert werden soll (Fürstenau 2013, S. 220). Dabei geraten insbesondere Schulen in sozialräumlich deprivierter Lage in den Fokus, da hier Sprachprobleme von Kindern mit und ohne Migrationshintergrund von den Lehrerinnen und Lehrern als zentrale Herausforderung benannt werden (Racherbäumer & van Ackeren 2014). Insgesamt wird ein konstruktiver und sozial gerechter Umgang mit der Heterogenität bzw. Diversität der Schülerinnen und Schüler gefordert, wozu zunächst implizite Normalitätserwartungen sowie Defizitzuschreibungen auf Seiten der Lehrkräfte überwunden werden müssen. Der sozialpsychologischen Theory of reasoned action (Ajzen & Fishbein 1980) bzw. ihrer Weiterentwicklung, der Theory of planned behavior (Ajzen 1991) folgend, kann eine Verhaltensintention als wichtigster Prädiktor für das tatsächliche Verhalten angesehen werden. Intentionen werden wiederum insbesondere durch verhaltensbezogene Einstellungen sowie subjektive Normen beeinflusst, so dass der intendierte wertschätzende Umgang mit der sprachlichen Heterogenität der Schülerinnen und Schüler eine entsprechende Haltung voraussetzt. Neben der sprachsensiblen Unterrichtsgestaltung aller Fächer werden ferner organisationale Rahmenbedingungen der Schule als wesentliche Gelingensbedingungen einer ganzheitlichen, durchgängigen Sprachbildung formuliert. Hierzu gehören z.B. eine Professionalisierung der Lehrkräfte durch entsprechende Fortbildungen, eine prozesssteuernde Begleitung durch die Schulleitung (Implementierung im schulinternen Curriculum, Bündelung einzelner Maßnahmen, Institutionalisierung professioneller Lerngemeinschaften etc.), die Kooperation mit externen Partnern sowie die Evaluation der durchgeführten Maßnahmen (Havighorst 2013).

Fragestellung: Im Vortrag werden die Einstellungen der befragten Lehrkräfte zu Heterogenität und Mehrsprachigkeit sowie ihre selbstberichtete sprachensible Unterrichtspraxis dargestellt. Bisher liegen vergleichsweise wenige Befunde zu Faktoren vor, die eine sprachensible Unterrichtsgestaltung befördern oder behindern. Aus diesem Grunde werden auch personale Merkmale fokussiert (z.B. Einstellungen zu Heterogenität und Mehrsprachigkeit, Kenntnisse im Bereich der Sprachbildung, demographische Merkmale), die mit einer sprachsensiblen Unterrichtspraxis korrelieren.

Zudem soll im Hinblick auf die Einstellung zu und den Umgang mit einer sprachlich heterogenen Schülerschaft der Einfluss des organisationalen Kontextes herausgearbeitet werden. Lassen sich in Bezug auf die Einstellungen und die Unterrichtspraxis der Lehrkräfte Unterschiede zwischen den betrachteten Einzelschulen feststellen? Inwieweit existieren an den betreffenden Schulen Absprachen zur fächerübergreifenden Sprachbildung und zeigen sich Zusammenhänge zwischen getroffenen Absprachen auf Ebene der Fachkonferenzen und der Unterrichtspraxis der Lehrkräfte?

Methode: Im Rahmen der wissenschaftlichen Begleitung des Projektes Sprachensible Schulentwicklung wurden von April bis Oktober 2014 in einer standardisierten Online-Befragung Lehrkräfte (N=519) aus 33 nordrhein-westfälischen Schulen der Sekundarstufe I zu ihrer Einstellung hinsichtlich Heterogenität und Mehrsprachigkeit sowie ihrer Unterrichtspraxis befragt. Deskriptive Statistiken werden berichtet und Korrelationsanalysen geben Hinweise auf die Zusammenhänge der betrachteten Konstrukte. Intra-Class-Korrelationen werden berechnet, um die Varianz sprachsensibler Unterrichtspraxis auf Organisationsebene aufzuzeigen.

Ergebnisse: Erste Ergebnisse zeigen, dass die überwiegende Mehrheit der Lehrkräfte ihre Schülerinnen und Schüler dazu anhält, im Unterricht ausschließlich Deutsch zu sprechen (M=4,51 auf einer fünfstufigen Likertskala); Herkunftssprachen werden hingegen kaum in das Unterrichtsgespräch einbezogen (M=1,74). Tendenziell sind sich die befragten Lehrkräfte eher sicher, Kinder nicht-deutscher Herkunft genügend zu fördern (M=3,36). Sie geben an, die Unterrichtssprache so einfach wie möglich zu halten (M=3,57) und Sachverhalte in der Alltagssprache zu beschreiben (M=3,51).

Im Hinblick auf eine organisationale Verankerung einer fächerübergreifenden Sprachbildung lässt sich feststellen, dass nicht in allen Schulen verbindliche Absprachen getroffen wurden. Die große Streuung dieser Items lässt zudem darauf schließen, dass auch innerhalb der Einzelschulen Unklarheit bezüglich des Vorliegens derartiger Absprachen vorzuherrschen scheint.

Mittwoch
11. März

14:00 Uhr

HGB 50

Döll, Marion

„sehr sehr gutes Deutsch“ ist nicht genug – sprach-, kultur- und herkunftsbezo-gen rassialisierende Argumentationen bei PH-Lehrkräften in Österreich

Im Jahr 2011 startete das österreichische Bildungsministerium (damals bm:ukk, heute bmbf) unter dem Namen „Diversität und Mehrsprachigkeit in pädagogischen Berufen“ eine Initiative zur Erhöhung des Anteils von Studierenden mit Migrationshintergrund an Pädagogischen Hochschulen. Ausschlaggebend hierfür waren informelle Erfahrungsberichte von PH-Lehrenden, die einerseits einen geringen Anteil von Studierenden mit Migrationshintergrund unter den StudienanfängerInnen beklagten und andererseits in Bezug auf diese Gruppe besonders viele drop-outs beobachtet haben wollten. Da man zu geringe Deutschkompetenzen, v.a. im produktiven schriftlichen Bereich, als Grund vermutete, wurde ein Projekt zur Entwicklung eines förderdiagnostischen Verfahrens zur Diagnose und Unterstützung der Entwicklung wissenschaftlicher Textkompetenz in Auftrag gegeben (Knappik 2013, Knappik/Dirim/Döll 2014). Zur Klärung des Passungsverhältnisses zwischen vorliegenden Kompetenzmodellen, die sich in aller Regel auf Textkompetenzen, die im Rahmen zehensemestriger Fachstudiengänge an Universitäten angeeignet werden (u.a. Steinhoff 2007), beziehen, und den spezifischen Merkmalen der im Laufe eines sechssemestrigten Lehramtsstudiums an einer österreichischen Pädagogischen Hochschule anzueignenden wissenschaftlichen Textkompetenz wurden mit insgesamt 35 in einem Top-Down-Verfahren ausgewählten Lehrenden von sieben Pädagogischen Hochschulen semi-strukturierte ExpertInneninterviews durchgeführt. Im Zentrum der Interviews standen Fragen zu den sprachlichen Anforderungen eines PH-Studiums und sprachlichen Aspekten, die von den Studierenden besonders herausfordernd erlebt werden. Ausgehend von der Tatsache, dass der Aufbau von wissenschaftlicher Textkompetenz unabhängig von Sprach- und Migrationsbiografie für alle Studierende eine zentrale und herausfordernde Aufgabe darstellt, wurde eine Differenzierung zwischen ein- und migrationsmehrsprachig aufgewachsenen Studierenden sowie zwischen Studierenden mit und ohne Migrationshintergrund seitens der InterviewerInnen erst in der letzten Frage vorgeschlagen. Die Ergebnisse der Inhaltsanalyse (nach Meuser/Nagel 2009) zeigen, dass die Lehrenden keine bedeutsamen Unterschiede in der Textkompetenz von ein- und mehrsprachigen Studierenden berichten – ein Befund, der durch die Ergebnisse einer linguistischen Analyse eines Korpus von 115 von Studierenden derselben sieben Hochschulen verfassten Texte auf Verwendung fach- und bildungssprachlicher Mittel gestützt wird (Knappik/Dirim/Döll 2014).

Im Rahmen der inhaltsanalytischen Bearbeitung der Daten fiel auf, dass die Mehrheit der befragten Hochschullehrkräfte den klar umrissenen und vorher von ihnen eingehaltenen Kontext der sprachlichen Herausforderungen eines Studiums an einer österreichischen PH verließen, wenn sie am Ende des Interviews um die Darstellung der Differenzen zwischen ein- und mehrsprachig aufgewachsenen Studierenden gebeten wurden. Statt auf etwaige spezifische sprachliche Hürden einzugehen, wurden verschiedene persönliche Hypothesen zur Erklärung des geringen Anteils von migrationsbedingt mehrsprachigen AbsolventInnen vorgebracht, die einerseits auf diffuse Kulturkonzepte rekurrend kulturalisierende Argumentationen umfassen und andererseits verschiedene Standards bei der Bewertung sprachlicher Kompetenzen von ein- und migrationsbedingt mehrsprachig aufgewachsenen Studierenden in Form des native-speakerism (Holliday 2006) deutlich werden lassen.

Diese Beobachtungen zum Anlass nehmend werden die Daten derzeit einer hermeneutischen Sekundäranalyse unterzogen, um Formen des kultur- und herkunftsbezogenen Rassismus (Rommelspacher 2009), des Neo-Linguizismus (Dirim 2010) und deren Einbettung in die Argumentationshaushalte der PH-Lehrkräfte herauszuarbeiten. Gegenstand des Beitrags ist die Vorstellung der Ergebnisse der skizzierten explorativen Analyse und die Diskussion ihrer Bedeutung im Kontext von bildungspolitischen Forderungen nach „mehr Lehrpersonen mit Zuwanderungsgeschichte“ (Presse- und Informationsamt der Bundesregierung 2007) und mehr „Diversität und Mehrsprachigkeit in pädagogischen Berufen“ (bmbf 2013).

Mittwoch
11. März

14:00 Uhr

HGB 50

Klein, Esther Dominique

Schulleitungshandeln in Schulen in schwieriger Lage Ein systematischer Vergleich von Deutschland und den USA

Theoretischer Hintergrund: In der Forschung zu effektiven Schulen in schwieriger Lage wird vielfach die Bedeutung der Schulleitung betont; als zentral wird in diesem Zusammenhang neben der Entwicklung von Strategien zur Organisations- und Personalentwicklung (Transformational Leadership, Leithwood et al., 2006) auch die Rolle der Schulleitung als Instructional Leader (z.B. Hallinger, 2003), welcher Ziele und Strategien für die Verbesserung des Unterrichts entwickelt und den Lehrkräften in beratender und unterstützender Funktion zur Seite steht, diskutiert.

Mit wenigen Ausnahmen (z.B. Racherbäumer et al., 2013) bezieht sich die vorliegende Forschung zu effektivem Führungshandeln in Schulen in schwieriger Lage allerdings bislang kaum auf Deutschland, sondern erfolgt v.a. in den USA und England. Führungshandeln ist aber in landesspezifische Steuerungssysteme und kulturelle Traditionen eingebettet (Hallinger & Leithwood, 1998). Es ist davon auszugehen, dass differenzielle Rahmenbedingungen und Hierarchiegefüge sowie landesspezifische Entwicklungsaufgaben von Schulen in schwieriger Lage dazu führen, dass das Rollenverständnis der Schulleitung sowie ihre Interdependenzbeziehungen mit anderen Akteuren, und damit auch ihre Einflussmöglichkeiten, in verschiedenen Ländern divergieren.

Fragestellung: Inwiefern sich Rahmenbedingungen, Akteure und Handlungsstrategien von effektiven Schulleitungen in Schulen in schwieriger Lage in Deutschland und in anglophonen Ländern unterscheiden, ist bislang allerdings nicht systematisch gegenübergestellt worden. Diese Studie greift dieses Desiderat in einer kontrastierenden Fallstudie auf, um die rechtlichen, strukturellen und organisationalen Rahmenbedingungen von schulischem Führungshandeln in Schulen in schwieriger Lage in den USA und Deutschland systematisch zu beschreiben. Ziel ist es, empirisch fundierte Hinweise zur Rolle der Schulleitung im Schulentwicklungsprozess, zu den Interdependenzverhältnissen zwischen Schulleitung und weiteren Akteuren und zur konkreten Vorbereitung von Schulleitungsakteuren auf die Arbeit an Schulen in schwieriger Lage herauszuarbeiten und damit Bedingungen für die Realisierung effektiven Schulleitungshandelns ableiten zu können. Die herausgearbeiteten Befunde können als Grundlage für sich anschließende Studien zur vergleichenden Analyse tatsächlichen Führungshandelns in deutschen und US-amerikanischen Schulen in schwieriger Lage dienen.

Methodisches Vorgehen: Da beide Staaten föderal strukturiert sind, werden jeweils hinsichtlich ihrer bildungspolitischen Ausrichtung sowie demographischer Rahmenbedingungen gegensätzliche Bundesstaaten bzw. -länder in den Blick genommen, um die Bandbreite an möglichen Ausdifferenzierungen abbilden zu können. Fokussiert werden dabei Nordrhein-Westfalen und Bayern sowie Kalifornien und Texas. Mittels kategoriengeleiteter Dokumentenanalysen von Schulgesetzen, Dienstvorschriften, Informationsbroschüren, Standards sowie weiteren relevanten Dokumenten werden die rechtlichen, organisatorischen und strukturellen Bedingungen des Führungshandelns innerhalb von Schulen analysiert. Zudem werden ebenfalls im Rahmen von Dokumentenanalysen Rahmenbedingungen von Schulen in schwieriger Lage in den Blick genommen. Diese Analysen werden durch teilstandardisierte Experteninterviews mit Vertretern der Schulbehörden in den vier Systemen ergänzt, welche inhaltsanalytisch ausgewertet werden. Die Analysekategorien umfassen z.B. die Bereiche Tätigkeitsfelder, Interdependenzbeziehungen mit anderen Akteuren, Unterstützungsstrukturen für Schulen in schwieriger Lage, etc.

Ergebnisse: Erste Auswertungen weisen auf zentrale Unterschiede zwischen Deutschland und den USA beispielsweise mit Blick auf die Rollenzuschreibung an die Schulleitung hin. Während etwa in den USA sowohl in den herangezogenen Dokumenten als auch in den Interviews die Rolle der Schulleitung als „Instructional Leader“ durchgehend hervorgehoben und anderen Aufgaben vorangestellt wird, lässt sich eine klare Zuschreibung dieser Rolle in Deutschland lediglich vereinzelt ablesen, was z.B. anhand divergierender Arbeitszeitmodelle, aber auch mit Blick auf die länderspezifische Aufteilung von Funktionen und Befugnissen zwischen Schulleitung, lokaler Schulaufsicht und internen Akteuren erklärt werden kann. Hieraus dürften sich wiederum auch Unterschiede in der Selbstzuschreibung durch die Schulleitung ergeben. Zudem deuten sich Unterschiede in der Autonomie der Schulleiter/innen bei der Entscheidung über Entwicklungsstrategien sowie in der finanziellen und programmatischen Unterstützung von Schulen in schwieriger Lage an. Weitere Ergebnisse sollen auf der Tagung vorgestellt werden.

Mittwoch
11. März

14:00 Uhr

HGB 50

Ehlert, Antje

Das Arbeitsgedächtnis im Vor- und Grundschulalter: prognostische Aussagekraft, Trainierbarkeit, Struktur

Die Grundlage aller Beiträge bildet der gemeinsame Blick auf das Arbeitsgedächtnis nach Baddeley. Dabei liegt das zentrale Interesse darin, Zusammenhänge zwischen Arbeitsgedächtnis und domainspezifischen Kompetenzen im Vor- und Grundschulalter zu analysieren.

So fokussieren Brandenburg & Hasselhorn die prognostische Aussagekraft von Arbeitsgedächtnisleistungen in Bezug auf die Überwindung einer Lesestörung. In einer Längsschnittstudie werden Grundschul Kinder mit einer diagnostizierten Lesestörung Mitte der 3. und Ende der 4. Klasse untersucht. Studienziel ist es, eine Aussage zu treffen, inwiefern die Funktionstüchtigkeit des Arbeitsgedächtnisses zur Prognose persistierender bzw. vorübergehender Lesestörungen herangezogen werden kann und wie der prognostische Nutzen anerkannter Prädiktoren, wie die phonologische Bewusstheit und Benennungsgeschwindigkeit, im Vergleich zum Arbeitsgedächtnis ist.

Berner, Lehner & Seitz-Stein untersuchen die Trainierbarkeit des visuell-räumlichen Arbeitsgedächtnisses und dessen Einfluss auf die mathematische Vorschulkompetenz. In einem Experimental-Kontrollgruppendesign werden die Experimentalgruppe mit einer statisch-visuellen Anforderung und die Kontrollgruppe mit Geschichtenvorlesen und Ausmalen trainiert. Die Autorengruppe analysiert die Trainierbarkeit von visuell-räumlichen Arbeitsgedächtnisprozessen und deren Einfluss auf die mathematische Kompetenzentwicklung. Dabei werden trainingsnahe Effekte sowie 10-wöchige Trainingseffekte berichtet.

Die Beiträge von Ehlert, Hannuschke & Fritz sowie Hannuschke, Ehlert & Fritz beruhen auf der gleichen Studie mit Grundschulkindern, deren Leistungen im Längsschnitt jeweils Ende 1.-3. Klasse erhoben wurden. Im Fokus steht die Arbeitsgedächtnisstruktur in Abhängigkeit spezifischer mathematischer Leistungsfähigkeiten. So untersucht Hannuschke et al. die mathematisch hochleistenden und Ehlert et al. die mathematisch niedrigleistenden Grundschul Kinder. Es werden die altersabhängige Struktur des Arbeitsgedächtnisses und die längsschnittliche Veränderung dieser Struktur analysiert.

Obgleich alle Studien das Baddeley-Arbeitsgedächtnis fokussieren, als Längsschnittstudien konzipiert sind und dabei das Vor- und Grundschulalter im Blick haben, zeigen die Einzelbeiträge deutliche Unterschiede. Zum einen repräsentieren sie die mögliche Breite an wissenschaftlichen Fragestellungen rund um das Arbeitsgedächtnis, hinsichtlich der Entwicklung, der Struktur, der Trainierbarkeit oder des Einflusses auf bzw. die Prognose für die Schriftsprache und die mathematische Kompetenz. Zum anderen verdeutlichen die Beiträge die methodische Vielfalt, das Arbeitsgedächtnis zu untersuchen, beispielsweise mittels Trainingsstudien oder über Leistungsdaten. Abhängig von der Fragestellung werden differente Auswertungsmethoden gewählt, wie Varianzanalysen mit Messwiederholungen, hierarchisch logistische Regressionen oder explorative und konfirmatorische Faktorenanalysen. Wird die Verwertbarkeit der Ergebnisse betrachtet, weisen die Beiträge ebenfalls heterogene Zielausrichtungen auf. So gehören Befunde zu Arbeitsgedächtnisstrukturen eher der Grundlagenforschung an, wobei der Schwerpunkt auf mathematisch niedrig- und hochleistende Kinder wiederum Ableitungen für Erklärungsansätze zur Entstehung von Leistungsdifferenzen möglich machen. Hingegen werden Befunde zur Prognose von Lernstörungen bedeutsam die Diagnostik beeinflussen, wie auch Befunde zur Trainierbarkeit Auswirkungen auf die Konzeption von Förderansätzen haben werden.

Mittwoch
11. März

14:00 Uhr

HGC 20

Chair(s)
Ehlert, Antje**Das Arbeitsgedächtnis im Vor- und Grundschulalter: prognostische Aussagekraft, Trainierbarkeit, Struktur**

Brandenburg, Janin; Hasselhorn, Marcus

Lässt sich durch das Arbeitsgedächtnis vorhersagen, ob Kinder ihre Lesestörung im Verlauf der Grundschuljahre überwinden?

Kinder mit Lesestörung zeigen oftmals spezifische Besonderheiten im Arbeitsgedächtnis: Während für die zentrale Exekutive massive Defizite berichtet wurden, scheinen Beeinträchtigungen in der phonologischen Schleife nur geringfügig und im Notizblock gar nicht zu bestehen (Brandenburg et al., 2013). Erste Arbeiten belegen zusätzlich den prognostischen Nutzen einer Arbeitsgedächtnisdiagnostik für Kinder mit Leseschwierigkeiten: Beispielsweise konnten Fischbach, Preßler und Hasselhorn (2012) zeigen, dass eine vorschulische Erfassung der phonologischen und exekutiven Arbeitsgedächtnisfunktionen dabei hilft zu prognostizieren, ob ein Kind später von Lernschwierigkeiten im Lesen betroffen ist oder nicht. Unklar ist jedoch, inwiefern das Arbeitsgedächtnis auch zur Prognose persistierender bzw. vorübergehender Leseschwierigkeiten herangezogen werden kann. Eine solche Vorhersage ist besonders wünschenswert, läge hierin doch eine Möglichkeit frühzeitig zu identifizieren, welche lesechwachen Kinder besonderer Förderung bedürfen. Ziel der vorliegenden Studie war es daher zu überprüfen, ob die Kenntnis der Funktionstüchtigkeit des Arbeitsgedächtnisses eine Vorhersage darüber erlaubt, ob ein lesebeeinträchtigtes Kind im weiteren Entwicklungsverlauf seine Lernstörung überwindet.

An der Studie nahmen 81 Kinder mit einer diagnostizierten Lesestörung (Leseleistung: $PR < 16$) teil. Zum ersten Messzeitpunkt (t1, Mitte der 3. Klasse) bearbeiteten die Kinder Aufgaben zur phonologischen Schleife, zur zentralen Exekutive sowie zur phonologischen Bewusstheit und Benennungsgeschwindigkeit. Am Ende der 4. Klasse (t2) wurde die Leseleistung erneut überprüft und die Kinder danach eingeteilt, ob das Lesedefizit persistiert (t2-Leseleistung: $PR < 16$) oder überwunden wird (t2-Leseleistung: $PR > 25$). Von den 81 Drittklässlern mit Lesestörung zeigten am Ende der 4. Klasse 41 Kinder eine persistierende Lesestörung, während 25 sie überwunden hatten. 15 Kinder konnten nicht klassifiziert werden und wurden daher von weiteren Analysen ausgeschlossen.

Mittels hierarchisch logistischer Regression wurde untersucht, inwiefern die Prädiktormasse Arbeitsgedächtnis, phonologische Bewusstheit und Benennungsgeschwindigkeit geeignet sind, zwischen Kindern zu differenzieren, deren Lesestörung am Ende der 4. Klasse noch persistiert bzw. überwunden wurde. Im ersten Analyseblock wurden der IQ und die Leseausgangsleistung zu t1 in die Regression eingeführt, im zweiten Analyseblock folgten die spezifischen Prädiktoren.

Bereits auf Basis des IQs und der Leseausgangsleistung konnten 65% der Kinder (Nagelkerkes $R^2 = .23$) in die korrekte Gruppe der Überwinder bzw. Persistierer klassifiziert werden. Durch Hinzunahme der spezifischen Prädiktoren konnte diese Klassifizierung signifikant erhöht werden, $\chi^2(4) = 13.71$, $p = .006$, $R^2 = .46$, wobei die Gruppe der Persistierer mit 78% signifikant besser klassifiziert werden konnte, als die Gruppe der Überwinder (65%). Signifikante Beiträge zur Vorhersage lieferten die zentrale Exekutive und die Benennungsgeschwindigkeit, jedoch nicht die phonologische Schleife und die phonologische Bewusstheit.

In einem weiteren Analyseschritt wurden die spezifischen Varianzanteile von zentrale Exekutive und Benennungsgeschwindigkeit untersucht: Die Benennungsgeschwindigkeit konnte zusätzlich zur zentralen Exekutiven eine Varianzaufklärung von 7% leisten. Umgekehrt zeigte sich für die zentrale Exekutive ein zusätzlicher Varianzanteil von 12%, wenn sie erst nach der Benennungsgeschwindigkeit in die Regression aufgenommen wurde.

Diese Befunde haben Implikationen für die Prognose persistierender Lesestörungen. Neben dem bereits gut belegten Einfluss der Benennungsgeschwindigkeit, wurde erstmals der prognostische Nutzen der zentralen Exekutive untersucht: Lesebeeinträchtigte Kinder, die in der 3. Klasse über vergleichsweise schwach ausgeprägte zentral-exekutive Arbeitsgedächtnisfunktionen verfügten, waren eineinhalb Jahre später signifikant häufiger von persistierenden Lernstörungen betroffen. Dieser Befund ist interessant, wenn man bedenkt, dass bei lesebeeinträchtigten Kindern bislang vorwiegend die phonologische Informationsverarbeitung im Fokus der Forschung stand. Zentral-exekutive Prozesse bei schwachen Lesern sollten daher zukünftig verstärkt untersucht werden. Einschränkend ist jedoch zu bedenken, dass aufgrund der geringen Stichprobengröße nur große Effekte aufgedeckt werden können. Dies unterstreicht einerseits die Bedeutsamkeit der gefundenen Effekte, führt andererseits aber dazu, dass zum gegenwärtigen Zeitpunkt nicht geschlussfolgert werden kann, dass phonologische Schleife und phonologische Bewusstheit zur Prognose einer persistierenden Lesestörung gänzlich ungeeignet sind.

Mittwoch
11. März

14:00 Uhr

HGC 20

Berner, Valérie-Danielle; Lehner, Susanne; Seitz-Stein, Katja

VSSP-Training und Effekte auf mathematische Basiskompetenzen im Vorschulalter

Unbestritten ist der hohe Einflussfaktor basaler Arbeitsgedächtnisprozesse auf die Lernprozesse und den späteren schulischen Erfolg von Schülerinnen und Schülern (Alloway & Alloway, 2010; Alloway, 2011). In verschiedenen Studien konnte die Rolle des Arbeitsgedächtnisses auf bereichsspezifische Vorläuferfertigkeiten nachgewiesen werden. Während eine hohe Bedeutung der phonologischen Schleife beim Erlernen von Lesen und Schreiben beigemessen wird, sind für den Erwerb mathematischer Kompetenzen der visuell-räumliche Notizblock (VSSP) und die zentrale Exekutive wesentlich (Krajewski, Schneider & Nieding, 2008). Empirische Befunde zeigen ferner, dass Lese- und Rechtschreibschwierigkeiten mit Defiziten in der phonologischen Schleife einhergehen und Rechenschwierigkeiten mit einer Beeinträchtigung im visuell-räumlichen Notizblock und in der zentralen Exekutive (Schuchardt, Maehler & Hasselhorn, 2008; Passolunghi, 2011). Während zur Vorhersage und Prävention von Lese-Rechtschreibschwierigkeiten bereits zahlreiche Studien vorliegen, besteht größerer empirischer Forschungsbedarf für den Bereich mathematischer Kompetenzen (Krajewski, Nieding & Schneider, 2008). So liegen bislang nur wenige Studien mit längsschnittlichem Design vor, die die Förderung basaler Arbeitsgedächtnisprozesse systematisch untersuchen und evaluieren (z. B. Wass, Scerif & Johnson, 2012). Vielversprechend scheint es, die Vorhersage und Prävention von mathematischen Kompetenzen bei Vorschulkindern zu fokussieren (ebd.), verfügen doch Kinder in den ersten Lebensjahren bereits über numerische Basisfertigkeiten, die die weitere Entwicklung mitbestimmen (Krajewski, Nieding et al., 2008).

Vor diesem Hintergrund wird im vorliegenden Beitrag diese Forschungslücke aufgegriffen und die Fördermöglichkeit von visuell-räumlichen Arbeitsgedächtnisprozessen und deren Einfluss auf mathematische Kompetenzen bei Vorschulkindern untersucht.

Datengrundlage bilden zwei Teilstudien. Die Stichprobe umfasste für die erste Teilstudie $n = 22$ Kinder und für die zweite Teilstudie $n = 33$ Kinder. Die insgesamt $N = 55$ Vorschulkinder trainierten zur Hälfte mit der Matrix-Aufgabe, einer basalen statisch-visuellen Arbeitsgedächtnisaufgabe (dem Testaufgabenansatz folgend). Die andere Hälfte der Kinder bildete die aktive Kontrollgruppe, die Geschichten vorgelesen bekam und Ausmalbilder bearbeitete. Das statisch-visuelle Training fand ebenso wie die Beschäftigung in der Kontrollgruppe neun Mal in ca. 10-15 minütigen Einzelsitzungen statt.

Erfasst wurden neben der statisch-visuellen auch die dynamisch-räumliche Komponente des VSSP und die beiden anderen Bereiche des Arbeitsgedächtnisses (Hasselhorn et al., 2012), die phonologische Schleife und die zentrale Exekutive. Indikatoren für mathematische Basiskompetenzen wurden in Anlehnung an Ennemoser, Krajewski und Sinner (2011) erhoben. Neben dem Vor- und Nachtest fand eine ca. 10 wöchige Follow up – Erhebung statt, ein Jahresfollow up steht noch aus.

In Varianzanalysen mit Messwiederholungen zeigte sich, dass für die Kinder beider Fördergruppen Leistungszuwächse im Bereich des visuell-räumlichen Arbeitsgedächtnisses sowie im Bereich der mathematischen Basiskompetenzen nachgewiesen werden konnten. Die gewonnenen Befunde geben Hinweise zur Förderung des visuell-räumlichen Arbeitsgedächtnisses und dessen Einfluss auf mathematische Kompetenzen, die abschließend im Hinblick auf ihre praktische Bedeutsamkeit diskutiert werden.

Mittwoch
11. März

14:00 Uhr

HGC 20

Chair(s)
Ehlert, Antje

Das Arbeitsgedächtnis im Vor- und Grundschulalter: prognostische Aussagekraft, Trainierbarkeit, Struktur

Hannuschke, Marianne; Ehlert, Antje; Fritz-Stratmann, Annemarie

Veränderung der Arbeitsgedächtnisstruktur über die Zeit bei Mathe-Hochleistenden

Das Arbeitsgedächtnis wird als ein Erklärungsfaktor für schulische Leistungen thematisiert. Forschungsbefunde belegen dessen Bedeutsamkeit für den Rechen- und Schriftspracherwerb (Alloway & Gathercole, 2006) und konnten Einflüsse verschiedener Fehlfunktionen im Arbeitsgedächtnis auf das Vorliegen von Lernstörungen nachweisen (u.a. Pickering, 2006). Das klassische Arbeitsgedächtnismodell nach Baddeley (1986) postuliert eine dreigliedrige Struktur, bestehend aus einer Hauptkomponente (zentrale Exekutive, ZE) und zwei Subkomponenten (visuell-räumlicher Notizblock, NB; phonologische Schleife, PS). Diese Struktur konnte bei verschiedenen Stichproben und ab einem Alter von 6 Jahren relativ altersinvariant nachgewiesen werden (u.a. Alloway, Gathercole & Pickering, 2006). Vergleiche verschiedener Arbeitsgedächtnisstrukturmodelle bei Vier- bis Fünfzehnjährigen ergaben bei dieser dreigliedrigen Faktorenstruktur eine bessere Passung als bei einer zweigliedrigen (Gathercole, Pickering, Ambridge & Wearing, 2004). Im Gegensatz dazu geht Baddeley (2000) jedoch davon aus, dass die Zusammenhänge zwischen den Komponenten mit steigendem Alter enger werden sollten, da die begrenzte Kapazität der Subkomponenten durch eine stärkere Zusammenarbeit mit der ZE verbessert wird. Jarvis und Gathercole (2003) fanden diesbezüglich bei 11- und 14-Jährigen eine bessere Passung einer Zweier-Struktur, bei der ein verbaler und ein nicht-verbaler Arbeitsgedächtnisfaktor extrahiert wurden. Als eine Ursache für solche konträren Befunde kann angeführt werden, dass dreigliedrige Modelle bisher v.a. bei unselegierten Stichproben nachgewiesen werden konnten, die sich sowohl hinsichtlich des Alters als auch bezüglich des Vorliegens von Lernstörungen unterschieden (Schuchardt, Roick, Mähler & Hasselhorn, 2008). Studien, die spezifische Altersbereiche fokussierten, fanden wie Jarvis und Gathercole (2003) eher zweigliedrige Strukturen.

Die Struktur-Äquivalenz-Hypothese thematisiert die Gleichheit der Arbeitsgedächtnisstruktur unabhängig vom Vorliegen von Lernstörungen und relativ unabhängig vom Alter. Deren Nachweis ist grundlegend für die Interpretation von spezifischen Beeinträchtigungen im Arbeitsgedächtnis als ursächlichen Faktor für Lernschwierigkeiten. Gelingt der Nachweis solcher Strukturäquivalenzen nicht, müssen Lernstörungen eher vor dem Hintergrund von Strukturdefiziten diskutiert werden. Zur Prüfung der Strukturäquivalenz untersucht die Studie (1) die Arbeitsgedächtnisstruktur bei in Mathematik hochleistenden SchülerInnen in Klasse 1 bis 3 und (2) die Veränderung der Strukturen über die Zeit.

Ausgehend von 444 Schulkindern, deren mathematischen Leistungen und Arbeitsgedächtnisfähigkeiten (neun Subtests: Zahlen-, Wörter-, Pseudowörternachsprechen vorwärts zur Erfassung der PS, Corsi-Blöcke und Matrizen beim NB und Zahlen- und Wörternachsprechen rückwärts, sowie Listening Span und Zählspanne bei der ZE) jeweils am Ende der 1. bis 3.Klasse untersucht worden sind, wurde eine Subgruppe von n = 98 mathe-hochleistenden SchülerInnen gebildet, deren Faktorenstruktur der Arbeitsgedächtnisaufgaben für jeden Messzeitpunkt separat analysiert worden sind. Für die Zugehörigkeit zur Mathe-Hochleistungsgruppe mussten Leistungen zum ersten Messzeitpunkt und zu mindestens einem weiteren vorliegen, die im oberen Quartil der Stichprobe lagen. Es durften zu keinem Zeitpunkt Leistungen erreicht worden sein, die im unteren Quartil lagen. Im ersten Schritt wurde die Arbeitsgedächtnisstruktur mithilfe eines explorativen Vorgehens (exploratorische Faktorenanalysen mit SPSS) betrachtet und im zweiten Schritt modelliert und strukturprüfend (konfirmatorische Faktorenanalysen mit dem R-Package lavaan) auf ihre Passung hin analysiert. Die Passung des Strukturmodells nach Baddeley, sowie die Passung alternativer zwei-, drei- und viergliedriger Faktorenmodelle, die sich aus den exploratorischen Faktorenanalysen ergeben haben, wurden zu allen drei Zeitpunkten separat bezüglich gängiger Fit-Indices eingeschätzt und gegebenenfalls aufgrund schlechter Modifikationsindices verbessert. Im dritten Schritt wurde die Passung des klassischen Modells mit dem am besten zu den Daten passenden Alternativmodell verglichen, wobei der AIC als Informationsindex dem BIC vorgezogen worden ist, da der letztere sensitiver gegenüber der Modellsparsamkeit ist.

Das Arbeitsgedächtnismodell nach Baddeley wies zu allen Messzeitpunkten eine sehr gute Passung auf, entsprach aber nicht dem bestpassenden Modell für die Daten. Bei den Mathe-Hochleistenden zeigte sich eine bessere Passung eines Zwei-Faktorenmodells. Die Faktoren entsprachen einem phonologischen und einem visuell-räumlichen Faktor, die hoch korreliert waren, wobei sich die Ladungsmuster der manifesten Variablen zwischen den Messzeitpunkten leicht veränderten. Lediglich die hohe Ladung des Kunstwörternachsprechens lud zum dritten Messzeitpunkt unerwartet auf dem visuell-räumlichen Faktor.

Mittwoch
11. März

14:00 Uhr

HGC 20

Chair(s)
Ehlert, Antje

Das Arbeitsgedächtnis im Vor- und Grundschulalter: prognostische Aussagekraft, Trainierbarkeit, Struktur

Ehlert, Antje; Hannuschke, Marianne; Fritz-Stratmann, Annemarie

Längsschnittliche Analyse der Arbeitsgedächtnisstruktur bei rechenschwachen Grundschulkindern

Ogleich unterschiedliche Modelle des Arbeitsgedächtnisses diskutiert werden, hat sich in den letzten Jahrzehnten das Modell nach Alan D. Baddeley (1986, 2006) in der Forschung, insbesondere in der Entwicklungspsychologie und der pädagogischen Psychologie, durchgesetzt. Auch für die Erklärung von Lernstörungen hat sich das Modell als aufschlussreich erwiesen, da Studien zeigen konnten, dass Lernstörungen mit verschiedenen Fehlfunktionen im Arbeitsgedächtnis einhergehen (Alloway & Gathercole, 2006; Pickering, 2006; Schuchardt, Mähler & Hasselhorn, 2008). Insgesamt wird eine dreigliedrige Arbeitsgedächtnisstruktur angenommen, die sich in zwei Subkomponenten (phonologische Schleife und visuell-räumlicher Notizblock) und eine Hauptkomponente (zentrale Exekutive) differenziert. Seit 2000 (Baddeley, 2000) wird eine vierte Komponente angenommen (episodischer Puffer), die allerdings bislang noch wenig erforscht ist. Verschiedene Studien belegen, dass die dreigliedrige Struktur bereits bei 6-jährigen Kindern und bis ins hohe Erwachsenenalter vorliegt (Alloway, Gathercole & Pickering, 2006; Gathercole et al., 2004; Kane et al., 2004). Diese altersinvariante Arbeitsgedächtnisstruktur wurde jedoch an unselegierten Stichproben nachgewiesen. Studien, die spezifische Altersbereiche fokussieren, finden beispielsweise bei 11- und 14-jährigen Kindern eine 2-Faktoren-Struktur (Jarvis & Gathercole, 2003). Wird die Arbeitsgedächtnisstruktur von Kindern mit unterschiedlichen Lernstörungen betrachtet, weisen Befunde wiederum auf eine dreigliedrige Struktur hin (Schuchardt, Roick, Mähler & Hasselhorn, 2008). Einschränkend ist jedoch anzumerken, dass es sich bei den untersuchten Stichproben um altersheterogene Kinder (1.- bis 4. Klasse) mit diversen Lernstörungen handelt.

Um Lernstörungen vor dem Hintergrund spezifischer Fehlfunktionen bzw. Beeinträchtigungen im Arbeitsgedächtnis zu erklären, muss die Struktur-Äquivalenz-Hypothese belegt sein. Dies meint, dass Kinder mit und ohne Lernstörungen äquivalente Arbeitsgedächtnisstrukturen aufweisen. Liegen solche Strukturäquivalenzen nicht vor, müssen Lernstörungen eher vor dem Hintergrund von Strukturdefiziten diskutiert werden. Entsprechend fokussiert die eigene Untersuchung auf die folgenden Fragen: (A) welche Arbeitsgedächtnisstrukturen weisen rechenschwache Kinder zu unterschiedlichen Altersbereichen auf und (B) wie verändert sich diese Struktur im Längsschnitt.

444 Erstklässler wurden jeweils Ende des Schuljahres zu insgesamt 3 Messzeitpunkten (1. - 3. Klasse) hinsichtlich ihrer mathematischen Leistung und ihrer Arbeitsgedächtnisfähigkeiten untersucht. Basierend auf dieser Stichprobe wurde eine mathematisch schwach leistende Gruppe (n = 98) gebildet. Entsprechend den Kriterien der Gruppenbildung mussten die Rechenleistungen der Kinder zum 1. Messzeitpunkt und mindestens zu einem weiteren Messzeitpunkt im untersten Quartil liegen; zu keinem Messzeitpunkt durften sich die Rechenleistungen im obersten Quartil befinden. Das Arbeitsgedächtnis wurde mit insgesamt 9 Aufgaben untersucht, wobei die phonologische Schleife mit 3 (Zahlen-, Wörter-, Pseudowörternachsprechen vorwärts), der visuell-räumliche Notizblock mit 2 (Corsi-Blöcke, Matrix) und die zentrale Exekutive mit 4 Maßen (Zahlennachsprechen rückwärts, Wortsprache rückwärts, Listening Span, Zählspanne) analysiert wurden. In einem ersten Schritt wurden die Daten pro Messzeitpunkt mittels einer explorativen Faktorenanalyse geprüft (frei, mit vorgegebener Faktorenanzahl). In einem zweiten Schritt wurde die Struktur der Daten konfirmatorisch untersucht (R-Package lavaan). Dabei wurde die Passung der Daten pro Messzeitpunkt hinsichtlich des klassischen AG-Modells nach Baddeley und der Faktorenstruktur der explorativen Faktorenanalysen geprüft. Pro Messzeitpunkt lagen somit das klassische 3-Faktoren-Modell und alternativ ein zwei-, drei- und viergliedriges Faktorenmodell vor. Nicht in jedem Fall genügten die Modelle den Gütekriterien. Für den dritten Schritt, dem Modellvergleich, wurden das klassische Modell und das am besten zu den Daten passende alternative Modell ausgewählt. Da insgesamt eine eher explorativ-datenbasierte Vorgehensweise gewählt wurde, nahm beim Modellvergleich der AIC eine größere Bedeutung als der BIC ein.

Zu jedem Messzeitpunkt konnte das klassische Arbeitsgedächtnismodell nach Baddeley modelliert werden und genügte immer den Gütekriterien einer konfirmatorischen Faktorenanalyse. Allerdings war dieses Modell nicht zu jedem Messzeitpunkt auch das Modell, welches am besten zu den Daten passte. Vielmehr zeigte sich eine eher instabile Arbeitsgedächtnisstruktur, die zwischen zwei und drei beteiligten Arbeitsgedächtniskomponenten variierte. Zusätzlich unterschieden sich über die Zeit hinweg die Zusammenhänge und die Ladungshöhen zwischen den verschiedenen manifesten und latenten Variablen.

Mittwoch
11. März

14:00 Uhr

HGC 20

Chair(s)
Ehlert, Antje

Das Arbeitsgedächtnis im Vor- und Grundschulalter: prognostische Aussagekraft, Trainierbarkeit, Struktur

Wirth, Joachim

Lehren und Lernen: Zum 60. Geburtstag von Detlev Leutner

Detlev Leutner prägt die empirische Bildungsforschung seit mehr als dreißig Jahren, sei es u.a. als langjähriges Mitglied der AEPF, der Projektgruppen „Fahranfängervorbereitung“ und „Begleitetes Fahren“ der Bundesanstalt für Straßenwesen (BASt), der PISA 2012 Problem Solving Expert Group, als Mitinitiator und Sprecher des DFG-Schwerpunktprogramms 1293 „Kompetenzmodelle“, als Mitglied im Vorstand des KMK-Instituts zur Qualitätsentwicklung im Bildungswesen (IQB), als Mitglied im DFG-Senat und nicht zuletzt als Gründungsmitglied der GEBF. Zu seinen Forschungsschwerpunkten und -interessen zählen u.a. Selbstreguliertes Lernen, Lernen mit Multimedia, Problemlösen, frühe Mathematik, Professionswissen von Lehrkräften, Computer-/Simulation-Based Training, Evaluation von Ausbildungsprogrammen sowie Forschungsmethodik und Diagnostik.

Im letzten Jahr ist Detlev Leutner sechzig Jahre alt geworden. Zu diesem Anlass stellen vier Arbeitsgruppen aktuelle empirische Untersuchungen vor. Die Arbeitsgruppen verstehen sich als „akademische Nachfahren“ Detlev Leutners. Ihre Denkweisen, ihre Forschungen und ihre Interessen wurden von Detlev Leutner geprägt. Und so zeigt sich in allen vier Beiträgen die Handschrift Detlev Leutners, auch wenn er nicht bei allen vorgestellten Forschungsarbeiten direkt beteiligt war.

Die quasi-experimentelle Untersuchung von Malone, Hiltz und Brünken dient der Förderung von Fahranfängern und liefert empirische Befunde, die für die Entwicklung eines Multimediatrainings für Fahranfänger zur Gefahrenentdeckung und -bewertung genutzt werden können. Plass, Biles und Homer präsentieren zwei experimentelle Studien aus dem Bereich Game Studies. Sie untersuchen die Wirksamkeit unterschiedlicher instruktorischer Zielstellungen und Rückmeldungen im Rahmen pädagogischer Computerspiele auf die motivationale Zielorientierung und den Lernerfolg und finden Interaktionseffekte u.a. mit situationalem Interesse. Schmeck und Opfermann stellen ebenfalls zwei experimentelle Studien vor, thematisch jedoch im Bereich Selbstreguliertes Lernen und Lernen mit Multimedia angesiedelt. Sie überprüfen das Multimedia-Prinzip des sinnstiftenden Zeichnens auf seine Lernförderlichkeit und finden dafür deutliche empirische Evidenz. Stebner, Schiffhauer, Schmeck, Marschner und Wirth evaluieren in einer experimentellen und einer quasi-experimentellen Studie ein Training zum selbstregulierten Lernen für Gymnasiasten der fünften Jahrgangsstufe. Die empirischen Befunde bestätigen, dass nicht nur unter laborartigen, sondern auch unter realen Schul- und Unterrichtsbedingungen das selbstregulierte Lernen am besten durch ein kombiniertes Training kognitiver sowie metakognitiver Strategien gefördert wird.

Die vier Beiträge geben exemplarisch einen (keinesfalls umfassenden) Einblick in die verschiedenen Forschungsfelder, die auch durch Detlev Leutners Forschungen bearbeitet und gestaltet wurden. Diskutiert werden die vier Beiträge von Detlev Leutner.

Mittwoch
11. März

14:00 Uhr

HGC 30

Chair(s)

Wirth, Joachim
Brünken, Roland

Malone, Sarah; Hiltz, Jana; Brünken, Roland

Entdeckung oder Bewertung - Ergebnisse einer Eyetrackingstudie als Grundlage für Fahranfängerassessment und -training

Fahranfänger weisen verglichen mit erfahreneren Autofahrern ein drastisch erhöhtes Unfallrisiko auf, das sich innerhalb der ersten Monate selbstständigen Fahrens deutlich reduziert (z.B. Maycock, Lockwood, & Lester, 1991; Schade, 2001). Als ursächlich dafür werden vor allem Defizite angenommen, die auf der mangelnden Erfahrung und Übung der Risikogruppe beruhen (vgl. Anfängerrisiko, z.B. Leutner, Brünken, & Willmes-Lenz, 2009). Unter anderem mithilfe neuer Medien, die in Fahrausbildung und Fahrerlaubnisprüfung zunehmend eingesetzt werden, wird angestrebt die Unfallgefahr für Fahranfänger durch valide Selektion und realitätsnahes Training zu reduzieren.

Innerhalb eines Forschungsprojektes, das von der Bundesanstalt für Straßenwesen gefördert wird, soll ein Multimediatraining für Fahranfänger konzipiert und evaluiert werden. Insbesondere der Möglichkeit dynamisches Material zu präsentieren und ökologisch valide Anforderungen an die Fahrschüler zu stellen, wird ein hohes Potenzial zur Verbesserung der Fahranfängervorbereitung beigemessen (z.B. Malone & Brünken, 2013). Für die Strukturierung eines solchen Lernprogrammes eignet sich das Four-Component Instructional Design (4C/ID)-Modell (van Merriënboer, Clark, & Croock, 2002), das sich explizit auf das Training komplexer kognitiver Fertigkeiten, wie dem Autofahren bezieht. Das Modell zielt vorrangig auf die Vermittlung von Handlungswissen ab, wodurch Expertiseaufbau – als Ergebnis von Schemakonstruktion und -automatisierung – stärker gefördert werden kann.

Dabei bietet sich vor allem die Möglichkeit, eine Fähigkeit standardisiert zu erfassen und zu trainieren, die erwiesenermaßen substantiell für sicheres Fahren ist und die bei Fahranfängern oft noch defizitär zu sein scheint: die Gefahrenwahrnehmung (Hazard Perception). Dabei handelt es sich um das schnelle Erkennen von Gefährdungen und das Reagieren auf Gefahrenreize im Verkehrsumfeld (McKenna & Crick, 1994), das in sogenannten Hazard Perception Tests gemessen werden kann. Die Aufgabe der Testteilnehmer besteht in der Regel darin, eine rasche, fahrungspezifische Reaktion (z.B. Tastendruck) auf Gefahrenhinweise zu zeigen, die in dargebotenen Verkehrsszenarien auftauchen. Fahranfänger schneiden darin oft deutlich schlechter ab als erfahrene Fahrer (z.B. Scialfa et al., 2012; Scialfa et al., 2011). Um optimale Prüf- und Trainingsmethoden für Fahranfänger zu entwickeln, ist es notwendig zu identifizieren, bei welchen Teilprozessen der Gefahrenwahrnehmung die Fahranfänger besondere Schwächen aufweisen. Nach einem Modell von Grayson et al. (2003) können beim Risikoverhalten vier Teilprozesse unterschieden werden: hazard detection (Gefahrenentdeckung), threat appraisal (Gefahrenbewertung), action selection (Handlungsauswahl) und implementation (Ausführung). Die üblichen Hazard Perception Tests beinhalten allerdings nur die Gefahrenentdeckung und die Gefahrenwahrnehmung, die in einer Messung konfundiert sind. Hystegge et al. (2010) zeigten in ihrer Studie mit einem statischen Hazard Perception Test, dass diese beiden Prozesse mit Hilfe der Blickbewegungsmessung voneinander getrennt gemessen werden können. Es zeigte sich, dass erfahrene und unerfahrene Fahrer auf Fotos gezeigte Verkehrsgefahren etwa gleich schnell fokussierten (Gefahrenentdeckung), Fahranfänger aber länger dafür brauchten auf die Gefahren zu reagieren (Gefahrenbewertung).

In der vorliegenden Studie wurde überprüft, ob sich die Ergebnisse von Hystegge et al. (2010) bei Einsatz eines Hazard Perception Tests mit dynamischen Verkehrsszenarien replizieren lassen.

An der Studie nahmen 22 erfahrene Autofahrer und 15 Fahrschüler teil. Den Teilnehmern wurden 32 animierte Verkehrsszenarien aus der Fahrerperspektive präsentiert, die Hinweise auf potenzielle Gefahren beinhalten konnten. Auf Gefahrenreize sollte per Tastsendruck schnellstmöglich reagiert werden. Die Blickbewegungen der Teilnehmer wurden dabei mithilfe eines Tobii 300 Eyetrackers aufgenommen.

Im Gegensatz zu Hystegge et al. (2010) zeigte sich in den Blickbewegungsmessungen, dass die erfahrenen Fahrer mehr Gefahrenhinweise entdeckten und diese auch früher fokussierten als die Fahrschüler. Zudem reagierten erfahrene Fahrer auf mehr Gefahren, nachdem sie sie gesehen hatten; Sie bewerteten die Hinweise also eher als gefährlich als die Fahrschüler. Allerdings erfolgte die Gefahrenbewertung bei Fahrschülern und erfahrenen Fahrern etwa gleich schnell.

Für das zu entwickelnde Fahrtraining bedeuten die Ergebnisse, dass gezielt Trainingseinheiten zu Gefahrenentdeckung und -bewertung entwickelt werden sollten. Unentbehrlich ist dabei ein intensives Training der Gefahrenentdeckung mit dynamischem Material, da unerfahrene Fahrer insbesondere in der bewegten Umwelt Gefahrenreize zu selten entdecken.

Mittwoch
11. März

14:00 Uhr

HGC 30

Plass, Jan; Biles, Melissa; Homer, Bruce

The Impact of Different Types of Digital Badges in a Geometry Game on Student Motivation and Learning

Badges, markers of achievements within video games, can motivate learners (Amory et al., 1999; Kapp, 2012), but may decrease intrinsic motivation. Achievement goal theory was used to examine effects of badges on motivation and learning in an educational video game. Achievement goal theory postulates two primary goal types: mastery, which focus on learning and development of abilities, and performance, which focus on demonstrating or validating abilities relative to others (Dweck & Leggett, 1988). Feedback provided to students can influence their goal orientation (Butler, 2000), with mastery orientation generally associated with higher achievement (Harackiewicz et al., 2002). This suggests that in-game badges could affect goal orientation and students' learning and motivation.

Study 1: The relation between badges and students mastery goal orientation was examined. Middle school students (N = 52) completed a pretest of geometry and the mastery achievement goal orientation subscale from the Patterns of Adaptive Learning Scales (PALS; Midgley et al., 2000). Students were randomly assigned to play either the badges or no-badges version of Noobs vs. Leets, a digital game to teach geometry (Plass et al., 2012). After playing for approximately 25 minutes, a posttest of geometry was given.

The No-badges group performed significantly better on the posttest than the badges group, $F(1, 21) = 6.28$, $p = .018$, $\eta_p^2 = .17$. This effect was mitigated by a condition by mastery goal interaction, $F(1, 31) = 5.49$, $p = .03$, $\eta_p^2 = .15$: Mastery orientation was positively correlated with knowledge posttest score in the no-badges condition, with no relation in the badges condition. This suggests that badges can help students with low mastery goal orientation, but may hinder students with high mastery goal orientation, see Figure 1; <http://tinyurl.com/GEBF-Fig-1>.

Study 2: Study 2 examined effects of interest and badge design on learning outcomes. Middle school students (N = 85) were randomly assigned to play a performance, mastery or no badges version of the geometry game. Pretest and posttest of geometry knowledge, and the Situational Interest Survey were given (Linnenbrink-Garcia et al., 2010).

Students in the performance badges condition performed significantly better on the knowledge posttest than students in the Mastery Badges condition ($p = .007$); no-badges condition was not significantly different from either of the other conditions, $F(2, 84) = 3.60$, $p = .032$, $\eta_p^2 = .089$. There was also a significant interaction between Badges and Situational Interest, $F(1, 84) = 4.50$, $p = .01$, $\eta_p^2 = .11$. Learners with higher situational interest performed better with mastery badges; learners with low situational interest did worse with mastery badges. There was no relation between situational interest and learning outcomes in the other conditions (Figure 2; <http://tinyurl.com/GEBF-Fig-2>).

Discussion: Taken together, these results suggest badges do not always help in educational video games: Types of badges interact with students' interest and motivation to affect learning. This is consistent with more general findings on external rewards in learning. Future research will need to be conducted to generalize our findings to other badge ecosystems, other learner populations, and other subject matters.

Mittwoch
11. März

14:00 Uhr

HGC 30

Chair(s)

Wirth, Joachim
Brünken, Roland

Schmeck, Annett; Opfermann, Maria

The Generative Drawing Principle in Multimedia Learning

Theoretischer Hintergrund: Die Forschung zum multimedialen Lernen hat gezeigt, dass Bilder als effektive Ergänzung zu einem Text nicht immer nur vorgegeben sein müssen, sondern dass es auch lernförderlich sein kann, die Lernenden aufzufordern, zu den gelesenen Inhalten Bilder selbst zu erstellen (vgl. Leutner, Opfermann & Schmeck, 2014; van Meter & Fioretto, 2013). Im Gegensatz zum Lernen mit vorgegebenen Bildern fördert das selbständige Erstellen von Bildern das elaborierte Verständnis auch über eine selbstregulatorische Komponente: Lernende, die Bilder selbst erstellen, anstatt sich ein vorgegebenes Bild nur anzuschauen, können selbst überprüfen, ob sie die wesentlichen Inhalte eines Lernstoffes soweit verstanden haben, dass sie diese in einer Zeichnung wiedergeben können. Dies führt dann im Optimalfall zur Konstruktion eines integrierten mentalen Modells des zu lernenden Sachverhaltes. Empirische Befunde zeigen allerdings, dass das selbstständige Erstellen von Bildern zu einem Text effektiver ist, wenn Lernende in der Lage sind, qualitativ gute Bilder zu erstellen, ohne ihr Arbeitsgedächtnis bei diesem Visualisierungsprozess kognitiv zu überlasten (für einen Überblick vgl. Leutner & Schmeck, 2014; van Meter & Garner, 2005). Entsprechend postuliert das Prinzip des sinnstiftenden Zeichnens (sog. generative drawing principle): "People learn better from a science text when they are asked to draw illustrations representing the main ideas of the text" (Schwamborn, Mayer, Thillmann, Leopold & Leutner, 2010; S. 878). Die bisherige Forschungslage liefert erste Evidenz für die Gültigkeit des Prinzips des sinnstiftenden Zeichnens; allerdings sind weitere Studien notwendig, um seine Generalisierbarkeit zu überprüfen.

Fragestellung und Methode: Aus den oben genannten Gründen wurde in zwei experimentellen Studien der Effekt des selbständigen Erstellens von Bildern auf das Verständnis eines biologischen Sachtextes (bestehend aus sieben Textabschnitten) an Schülerinnen und Schülern der 8. Klasse (Exp. 1: N = 48; Exp. 2: N = 164) überprüft. Als abhängige Variable wurde das aus dem Text gewonnene Verständnis mit einem Multiple-Choice Test und einem Zeichentest erfasst. Als unabhängige Variable diente die Art der Textbearbeitungsinstruktion. In Experiment 1 wurden die Lernenden entweder instruiert, zu jedem Textabschnitt ein Bild zu zeichnen, das die wesentlichen Inhalte dieses Textabschnittes wiedergibt (Zeichengruppe), oder sie lasen lediglich den Text (Kontrollgruppe). In Experiment 2 wurde ein 2*2-Design mit den Faktoren „Zeicheninstruktion“ (ja/nein) und „vorgegebenes Bild“ (ja/nein) realisiert. Lernende mussten hierbei analog zu Experiment 1 nur einen Text lesen (Kontrollgruppe), lasen den Text und erstellten dazu eigene Bilder (Zeichengruppe), lasen den Text und erhielten dazu vorgegebene Bilder (Bildpräsentationsgruppe) oder lasen den Text, zeichneten eigene Bilder und sollten diese dann mit vorgegebenen Bildern vergleichen (Bildpräsentations-Zeichengruppe).

Ergebnisse und Diskussion: Für Experiment 1 zeigte sich, dass Lernende, die instruiert wurden, selbst Bilder zu erstellen (Zeichengruppe), im Vergleich zu Lernenden, die nur den Text lasen (Kontrollgruppe), signifikant besser im Multiple-Choice Verstehentest ($d = 0.85$) sowie im Zeichentest ($d = 1.15$) abschnitten. In Experiment 2 zeigte sich ein ähnliches Ergebnismuster. Die Lernenden der Zeichengruppe schnitten im Multiple-Choice Verstehentest signifikant besser ab als die Lernenden der Kontrollgruppe ($d = .0.52$), der Bildpräsentationsgruppe ($d = 0.49$) und der Bildpräsentations-Zeichengruppe ($d = .57$). Im Zeichentest schnitten die Lernenden der Zeichengruppen signifikant besser ab als die Lernenden der Kontrollgruppe ($d = 1.87$) sowie die Lernenden der Bildpräsentationsgruppe ($d = 0.68$).

Insgesamt liefern die Ergebnisse weitere Evidenz für die Gültigkeit des Prinzips des sinnstiftenden Zeichnens. Lernende aufzufordern, zu den gelesenen Inhalten Bilder selbst zu erstellen, erwies sich nicht nur lernförderlicher als das alleinige Lesen des Textes; sondern es war auch effektiver als das Lernen mit vorgegebenen Bildern.

Mittwoch
11. März

14:00 Uhr

HGC 30

Chair(s)

Wirth, Joachim
Brünken, Roland

Stebner, Ferdinand; Schiffhauer, Silke; Schmeck, Annett; Marschner, Jessica; Wirth, Joachim

Von interner zu externer Validität: Entwicklung und Evaluation eines Trainings zum selbstregulierten Lernen aus Sachtexten und durch Experimentieren

Selbstreguliertes Lernen ist eine anspruchsvolle Lernform, bei der Lernende eigenständig Aufgaben analysieren, sich Ziele setzen, ihren Lernfortschritt reflektieren und gegebenenfalls korrigierende Maßnahmen zur Optimierung des eigenen Lernens durchführen müssen (Schreiber, 1998; Spörer & Brunstein, 2006). Im Zeitalter neuer Technologien und Kommunikationsmittel verliert Wissen schnell an Aktualität und steigert damit die Bedeutung der Bereitschaft und der Fähigkeit der Lernenden, sich im Sinne des selbstregulierten Lernens über den gesamten Lebenslauf neues Wissen und neue Fähigkeiten eigenständig anzueignen (Bjork, Dunlosky & Kornell, 2013). Diesbezüglich muss es das Ziel der Schulen sein, ihre Schülerinnen und Schüler auf diese Herausforderungen vorzubereiten und die notwendigen Kompetenzen zu vermitteln. Während die Bedeutung des selbstregulierten Lernens für alle Fächer groß ist, wurde sich in diesem Projekt aufgrund der empirischen Befundlage zunächst mit den Naturwissenschaften beschäftigt (vgl. Prenzel & Parchmann, 2003).

Ziel des Projekts war es, auf Basis von Vorarbeiten durch u.a. Leutner, Barthel und Schreiber (2001) sowie Leopold, den Elzen-Rump und Leutner (2006) ein lernförderliches und praxisnahes Training zum selbstregulierten Lernen aus Sachtexten und durch Experimentieren für Fünft- und Sechstklässler zu entwickeln. Als theoretische Grundlage diente das Selbstregulationsmodell von Schreiber (1998), nach dem im Sinne des Hucklepacktheorems (Klauer, 2000) untergeordnete (z.B. Lesestrategien) mit übergeordneten Strategien (z.B. metakognitive Kontrollstrategien) kombiniert werden, um die Ausführungsqualität der untergeordneten Lernstrategien zu erhöhen.

In diesem Beitrag werden zwei Studien präsentiert, die im Rahmen des Projektes „Ganz In - mit Ganzttag mehr Zukunft. Das neue Ganzttagsgymnasium NRW“ durchgeführt wurden.

In der ersten Studie wurden zunächst je ein Training zum selbstregulierten Lernen aus Sachtexten sowie zum selbstregulierten Lernen durch Experimentieren empirisch auf Lernförderlichkeit geprüft. In beiden Trainings wurden die unter- und übergeordneten Lernstrategien kombiniert vermittelt. Die Trainings erstreckten sich über ein Schulhalbjahr hinweg mit je 90 Minuten Trainingszeit pro Woche. Dazu wurden 243 Gymnasiasten (Alter $M=10.7$ Jahren; $SD=0.49$; 51% weiblich) von zwei Projektschulen im Sinne eines experimentellen $2 \times 2 + 1$ -Designs mit vier Trainings- und einer Kontrollgruppe randomisiert auf fünf Lerngruppen pro Schule aufgeteilt. Neben demographischen Daten und Kontrollvariablen wie Intelligenz, Motivation und räumliches Vorstellungsvermögen wurden für das Lesen und Experimentieren jeweils das Strategiewissen, die Strategieanwendung sowie der Wissenserwerb mit verschiedenen z.T. standardisierten Instrumenten erhoben. Das Training wurde von Mitarbeiterinnen und Mitarbeitern der Universitäten durchgeführt, die regelmäßig durch die Lerngruppen rotierten, um Trainiereffekte zu vermeiden.

Die Ergebnisse der Evaluation zeigen sowohl für das Lesen als auch für das Experimentieren deutliche Vorteile der Gruppen, die ein kombiniertes Training erhalten, im Vergleich zu den Kontrollbedingungen, die entweder nur ein Training untergeordneter Strategien oder gar keine Strategien vermittelt bekommen. Die Lernförderlichkeit des Trainings konnte durch diese experimentelle Studie empirisch bestätigt werden.

In der zweiten Studie wurden die beiden Trainings zunächst in einem gemeinsamen Training integriert, so dass Lernende in diesem einen Training beides lernen: selbstreguliertes Lernen aus Sachtexten und durch Experimentieren. Neben der Lernförderlichkeit sollte dieses Mal zudem die Praktikabilität überprüft werden, weshalb dieses Training über ein Schulhalbjahr hinweg von vier Lehrerinnen und Lehrern, die währenddessen mit Prozessbegleitung von den Mitarbeiterinnen und Mitarbeitern der Universität unterstützt wurden, durchgeführt wurde. Es handelte sich um ein quasi-experimentelles Kontrollgruppendesign, bei dem in der EG 98 und in der KG 106 Gymnasiasten aus je vier Parallelklassen (Alter $M=10.34$ Jahren; $SD=0.56$; 45% weiblich) unterrichtet wurden. Es wurden dieselben Instrumente eingesetzt wie in der ersten Studie.

Analog zu Studie 1 verfügen am Ende des Halbjahres die Schülerinnen und Schüler der Trainingsgruppe über signifikant mehr Strategiewissen, wenden die Strategie erfolgreicher an und erlangen dadurch mehr inhaltliches Wissen. Darüber hinaus konnte die Praktikabilität des Trainings im Schulalltag bestätigt werden, wodurch das gesamte Projekt als großer Erfolg gewertet werden kann.

Die Ergebnisse beider Studien bestätigen durch den erfolgreichen Transfer des Trainings von laborartig kontrollierten Bedingungen in die wenig kontrollierte, reale Schulpraxis die interne sowie externe Validität des Trainings.

Mittwoch
11. März

14:00 Uhr

HGC 30

Chair(s)

Wirth, Joachim
Brünken, Roland

Wagner, Daniela

Modellierung und Erfassung des Zusammenhangs zwischen selbstreguliertem Lernen und leistungsrelevanten Konstrukten im Kontext pädagogischer Forschung

Das Symposium umfasst vier Beiträge zum Thema selbstreguliertes Lernen, wobei sowohl die theoretische Überprüfung als auch die Erfassung dieses Konstrukts fokussiert werden.

Es werden Untersuchungen aus der Perspektive der Fachdisziplinen Psychologie und Bildungswissenschaften präsentiert. Die Forschung zur theoriegeleiteten Modellierung und Erfassung selbstregulierten Lernens hat in den letzten Jahren an Bedeutung gewonnen. Gerade die Frage, welche Komponenten die Kompetenz zum selbstregulierten Lernen abbilden, hat es notwendig gemacht, die Zusammenhgangsstruktur selbstregulierten Lernens mit anderen relevanten Konstrukten zu überprüfen.

Ein anderes sehr kontrovers diskutiertes Thema im Bereich der Selbstregulationsforschung bezieht sich auf die Erfassung und Messung dieses Konstrukts. Gerade die Kritik an Selbstberichtsverfahren hat dazu geführt, dass verstärkt eine Verzahnung von Online und Offline-Verfahren angestrebt wird und bei der Messung selbstregulierten Lernens sowohl qualitative als auch quantitative Standards berücksichtigt werden.

Die erfolgreiche Ausführung von Strategien selbstregulierten Lernens steht in engem Zusammenhang mit motivationalen Voraussetzungen, die in Anlehnung an das sozial-kognitive Selbstregulationsmodell von Zimmerman (2000) in der Phase der Handlungsplanung eine entscheidende Rolle spielen. Im Rahmen der ersten beiden Beiträge wird daher die Zusammenhgangsstruktur zwischen selbstreguliertem Lernverhalten und zentralen motivationalen Variablen auf latenter Ebene überprüft. Da sich die Motivation im Rahmen von Modellen des selbstregulierten Lernens hauptsächlich auf Einstellungen und Erwartungen bezieht, fokussiert der erste Beitrag von Dörrenbächer und Perels den Zusammenhang von selbstreguliertem Lernen und Komponenten der Motivationsregulation und erweitert das zugrundeliegende Modell somit um volitionstheoretische Komponenten. Im zweiten Beitrag von Leidinger und Perels steht hingegen der reziproke Zusammenhang von selbstreguliertem Lernen und Selbstwirksamkeitsüberzeugungen im Vordergrund. Ausgehend von bisherigen Befunden der pädagogischen Interventionsforschung werden in beiden Beiträgen zusätzlich die Effekte der jeweiligen Konstrukte auf unterschiedliche Leistungsmaße differenziert nach Altersgruppen (Studierende, Grundschüler der vierten Klassenstufe) betrachtet.

Bellhäuser, Roth und Schmitz stellen solch ein neues Messinstrument vor (SRL@U), das auf seine Testgütekriterien überprüft wurde und den Prozess des selbstregulierten Lernens ganzheitlich abdeckt. Der Beitrag fokussiert die Kriteriumsvalidität, zu deren Bestimmung verwandte Konstrukte herangezogen wurden.

Der Beitrag von Ogrin und Schmitz erweitert die Erfassung selbstregulierten Lernens im Sinne des Multi-Method-Ansatzes durch den Einsatz von Lerntagebüchern, um Lernhandlungen zeitnah als Folge von states abbilden zu können. Die Untersuchung befasst sich mit der Übereinstimmung von Angaben zum selbstregulierten Lernen, die sowohl aus Offline- (Fragebogen) als auch aus Online-Verfahren (Lerntagebücher) gewonnen wurden. Zusätzlich wurden bei der Messung einerseits quantitative und andererseits qualitative Standards angelegt, um die Passung der eingesetzten Strategien zur Lernsituation sowie die Umsetzungsgüte zu überprüfen.

Mittwoch
11. März

14:00 Uhr

HGC 40

Chair(s)

Perels, Franziska

Modellierung und Erfassung des Zusammenhangs zwischen selbstreguliertem Lernen und leistungsrelevanten Konstrukten im Kontext pädagogischer Forschung

Dörrenbächer, Laura; Perels, Franziska

Struktureller Zusammenhang von selbstreguliertem Lernen, Motivationsregulation und Leistung

Die Fähigkeit zum selbstregulierten Lernen stellt eine zentrale Kompetenz im Rahmen des lebenslangen Lernens dar (Wirth & Leutner, 2008) und wird meist in sozial-kognitiven Theorien verankert, die beim Lernprozess das Zusammenspiel von Verhalten, Kognitionen und Umweltbedingungen in den Vordergrund stellen (Bandura, 1986; Zimmerman, 2000). Selbstreguliertes Lernen wird dabei relativ einheitlich als Kompetenz definiert, die kognitive, metakognitive sowie motivationale Komponenten umfasst (Landmann, Perels, Otto & Schmitz, 2009), wobei letztere sich hauptsächlich auf Einstellungen wie z. B. die intrinsische Motivation oder Ergebniserwartungen beziehen (Zimmerman, 2000). Betrachtet man den Lernprozess hingegen aus volitionstheoretischer Sicht, so sind vor allem die Regulation der Motivation (Corno, 1993) bzw. die Selbstkontrolle durch Unterdrückung vorherrschender Verhaltenstendenzen zur Zielerreichung (Duckworth & Seligman, 2006) für den Umgang mit Lernschwierigkeiten notwendig. Daher besteht die Frage, ob die Motivationsregulation als Bestandteil des selbstregulierten Lernens aufzufassen ist oder ein eigenes Konstrukt darstellt, das aber eng mit selbstregulativen Kompetenzen zusammenhängt (Wolters, 2003). Ziel der vorliegenden Untersuchung war die Analyse der Zusammenhangsstruktur des selbstregulierten Lernens und der Motivationsregulation. Dabei sollte mittels Strukturgleichungsmodellierung untersucht werden, ob sich die Motivationsregulation in das Messmodell des selbstregulierten Lernens integrieren lässt oder ob eine Zweifaktoren-Lösung bessere Fit-Indices liefert. Verschiedene Studien belegen den Zusammenhang der beiden Konstrukte mit Leistung (z. B. Sitzmann & Ely, 2011; Schwinger, Steinmayr & Spinath, 2012), weswegen die Leistungsvorhersage (unter Kontrolle kognitiver Fähigkeiten) in der vorliegenden Untersuchung anhand des Modells mit den besseren Fit-Indices analysiert werden sollte.

Grundlage der Studie bildet eine Fragebogenerhebung bei 263 Studierenden (M Alter = 23,5 Jahre; w = 186). Das selbstregulierte Lernen wurde in Anlehnung an das Traitmodell von Hong und O'Neil (2001) sowie Zimmerman (2000) durch die Abfrage der beiden Komponenten Motivation (intrinsische Motivation, Zielorientierung) sowie Metakognition (Strategie- und Zeitplanung, Selbstevaluation) erfasst, während die Motivationsregulation durch die Subskalen Zukunftsorientierung (Zimbardo & Boyd, 1999), akademischer Belohnungsaufschub (Bembenutty & Karabenick, 1998) und Prokrastination (invertiert; Tuckman, 1991) erfasst wurde. Diese drei Traits haben sich in verschiedenen Studien als wichtige Motivationsregulatoren herausgestellt (Bembenutty & Karabenick, 2004; de Bilde, Vasteenkiste & Lens, 2011; Park & Sperling, 2012). Die Reliabilitäten der Subskalen wiesen zufriedenstellende bis gute Werte auf ($.70 < \alpha < .89$). Die akademische Leistung wurde über die Abiturnote erfasst, während die kognitiven Fähigkeiten über einen Intelligenztest (CFT 20-R; Weiß, 2008) sowie einen Arbeitseffizienztest erfasst wurden (AET; Conzelmann & Kersting, 2012).

Wurden die Motivationsregulation und das selbstregulierte Lernen (bestehend aus Motivation und Metakognition) als zwei getrennte Faktoren modelliert, zeigten sich schlechte Modellfitindices ($\chi^2(14) = 97.2$, $p < .001$, $\chi^2/df = 6.9$, RMSEA = .15, CFI = .80, SRMR = .14). Die Unterordnung des Faktors Motivationsregulation unter den Faktor 2. Ordnung selbstreguliertes Lernen verbesserte den Modellfit signifikant und wies zufriedenstellende Güteindices auf ($\chi^2(12) = 27.0$, $p < .001$, $\chi^2/df = 2.25$, RMSEA = .07, CFI = .96, SRMR = .04). Die Abiturnote wurde in diesem Modell sowohl durch den Faktor selbstreguliertes Lernen als auch durch die Kovariate kognitive Fähigkeiten signifikant vorhergesagt und es gab einen indirekten Effekt der kognitiven Fähigkeiten über selbstreguliertes Lernen auf die Abiturnote.

Die vorliegende Untersuchung deutet darauf hin, dass die Motivationsregulation als eine Teilkomponente des selbstregulierten Lernens neben motivationalen und metakognitiven Komponenten aufzufassen ist (Wolters, 2003), weshalb eine Überarbeitung der motivationalen Komponente bei Modellen des selbstregulierten Lernens anzustreben und empirisch zu überprüfen ist. In zukünftigen Studien kann in Anlehnung an die State-Trait-Unterscheidung untersucht werden, ob sich dieselbe Modellstruktur bei State-Daten zeigt, die selbstreguliertes Lernen und die Motivationsregulation z. B. täglich und situationsbezogen über Lerntagebücher erfassen. Außerdem können Interventionsstudien zur Förderung des selbstregulierten Lernens dahingehend evaluiert werden, ob die Komponenten der Motivationsregulation ebenfalls eine Veränderung zeigen, selbst wenn sie nicht trainiert wurden.

Mittwoch
11. März

14:00 Uhr

HGC 40

Leidinger, Manuela; Perels, Franziska

Was selbstregulierte Grundschüler glauben zu können. Selbstreguliertes Lernen und Selbstwirksamkeitsüberzeugungen als Prädiktoren schulischer Leistung im Primarbereich

Empirische Studien belegen, dass für den schulischen Kontext die Förderung von Selbstregulationsstrategien unmittelbar einhergeht mit schulischem Lern- und Leistungserfolg (z.B. DeCorte, Mason, Depaepe & Verschaffel, 2011). Selbstreguliertes Lernen ist dabei definiert als ein zyklisches Zusammenwirken kognitiver, metakognitiver und motivationaler Komponenten (Perry & Winne, 2006), die sich in Anlehnung an das sozial-kognitive Modell der Selbstregulation von Zimmerman (2000a) innerhalb der verschiedenen Phasen des Lernprozesses wiederfinden (Schmitz & Wiese, 2006).

Für den akademischen Kontext erschließen sich entsprechend diesem Modell verschiedene metakognitive und motivationale Strategien, die Schüler dazu befähigen, ihr Lernverhalten zu regulieren (Puustinen & Pulkkinen, 2001). Dignath, Büttner und Langfeldt (2008) geben in ihrer Metaanalyse einen Überblick zur Effektivität entsprechender Strategietrainings auf die akademische Leistung. Für Interventionen im Bereich Mathematik ermittelten sie dabei Effektstärken von $d = .96$, für den Bereich des Leseverstehens Effektstärken von $d = .44$.

Die erfolgreiche Ausführung metakognitiver und motivationaler Selbstregulationsstrategien steht dabei in enger Interaktion mit motivationalen Voraussetzungen, von denen vor allem den Selbstwirksamkeitsüberzeugungen (Bandura, 1997) eine bedeutende Rolle zugesprochen wird (Schunk & Ertmer, 2000; Zimmerman, 2000b). In verschiedenen Studien konnte aufgezeigt werden, dass Selbstwirksamkeitsüberzeugungen für das Lernen und die Persönlichkeitsentwicklung von zentraler Bedeutung sind (Richardson, Abraham & Bond, 2012). Für den akademischen Bereich berichten Marsh, Hau, Artelt, Baumert und Peschar (2006) Korrelationen von $r = .27$ zwischen akademischer Selbstwirksamkeit und mathematischer Leistung sowie Korrelationen von $r = .29$ zwischen akademischer Selbstwirksamkeit und Leseverstehen.

Für die Entwicklung von Selbstwirksamkeitsüberzeugungen belegen Studien einen positiven Effekt selbstregulativer Fähigkeiten (Perels, Gürtler & Schmitz, 2005; Schunk & Pajares, 2004). Schunk und Zimmerman (2007) ebenso wie Zimmerman und Cleary (2006) postulieren hierbei einen reziproken Zusammenhang zwischen den beiden Konstrukten, da umgekehrt auch die Entwicklung selbstregulativer Fähigkeiten entschieden durch die Selbstwirksamkeitsüberzeugungen einer Person beeinflusst wird. Zimmerman und Kitsantas (2007) sprechen hierbei von "self-efficacy for self-regulated learning".

Vor diesem Hintergrund bestand die Zielsetzung der vorliegenden Studie darin, das Zusammenwirken der drei Variablen selbstreguliertes Lernen, Selbstwirksamkeitsüberzeugungen und Leistung näher zu untersuchen. Dazu wurde einerseits die Vorhersagekraft der beiden Faktoren Selbstwirksamkeitsüberzeugungen und selbstreguliertes Lernen auf die akademische Leistung von Schülern in den Domänen Mathematik und Leseverstehen modelliert und andererseits der postulierte reziproke Zusammenhang der beiden Prädiktorvariablen empirisch überprüft.

Die Untersuchung wurde mit 309 Schülern der vierten Klassenstufe durchgeführt, die im Durchschnitt 9 Jahre alt waren ($SD = 0.63$; range = 8-11 Jahre). Das selbstregulierte Lernverhalten sowie die Selbstwirksamkeitsüberzeugungen wurden zu zwei Messzeitpunkten mittels Fragebogenerhebung erfasst. Die internen Konsistenzen der Subskalen fielen zu beiden Messzeitpunkten zufriedenstellend aus (Cronbach's α T1: $.69 - .88$; Cronbach's α T2: $.71 - .92$). Zur Leistungserfassung wurden standardisierte Testverfahren für Mathematik („DEMAT 3+“; Roick, Gölitz & Hasselhorn, 2004) und Leseverstehen („HAMLET 3-4“; Lehmann, Peek & Poerschke, 1997) eingesetzt.

Der Faktor selbstregulierten Lernens wurde in Anlehnung an Zimmerman (2000a) sowie Hong und O'Neil (2001) als Faktor 2. Ordnung mittels der Variablen Metakognition und Motivation abgebildet, während der Faktor Selbstwirksamkeitsüberzeugungen über die drei Indikatoren allgemeine, schulspezifische und domänenspezifische Selbstwirksamkeitsüberzeugungen aufgenommen wurde. Die hieraus resultierenden Faktorenmodelle wurden konfirmatorisch überprüft und lieferten jeweils zufriedenstellende Fitstatistiken.

Für beide Leistungsbereiche konnte nachgewiesen werden, dass die Selbstwirksamkeitsüberzeugungen einen höheren prädiktiven Wert ($\beta = .59$) für die Leistungsvorhersage besitzen und zudem positiv durch die selbstregulativen Fähigkeiten einer Person vorhergesagt werden können ($\beta = .27$). Die Ergebnisse liefern somit wichtige Implikationen für die Konzeption zukünftiger Interventionen im pädagogischen Kontext.

Mittwoch
11. März

14:00 Uhr

HGC 40

Bellhäuser, Henrik; Roth, Anne; Schmitz, Bernhard

SRL@U: Validierung des Fragebogens zur Erhebung des selbstregulierten Lernens an der Universität

Selbstreguliertes Lernen (SRL) ist eine Kompetenz, die insbesondere im Studium von hoher Relevanz ist, da ein robuster Zusammenhang zu Studienerfolg gezeigt werden konnte (Richardson, Abraham, & Bond, 2012). Aufgrund des großen Forschungsinteresses an diesem Thema findet sich in der Literatur eine Vielzahl von entsprechenden Messverfahren (Roth, Ogrin & Schmitz, 2014). Weiteste Verbreitung im deutschen Sprachraum fand hierbei das Inventar zur Erfassung von Lernstrategien im Studium (LIST; Wild, Schiefele & Winteler, 1992). Aus der Perspektive des Prozessmodells des selbstregulierten Lernens (Schmitz & Wiese, 2006) wird im LIST allerdings vornehmlich die aktionale Phase des Lernens (z.B. Monitoring, kognitive Lernstrategien, ressourcenorientierte Lernstrategien) betrachtet, während die präaktionale Phase (Zielsetzung, Planung, Selbstmotivation, Emotionsregulation) und die postaktionale Phase (Reflexion, Attribution, Bezugsnormorientierung, Vorsatzbildung) weitgehend unberücksichtigt bleiben.

Unser Ziel war daher die Entwicklung eines neuen Messinstruments, das alle Phasen des Prozessmodells gleichermaßen abdeckt. Der Fragebogen zur Erhebung des selbstregulierten Lernens an der Universität (SRL@U; Bellhäuser & Schmitz, in Vorbereitung) wurde in mehreren Überarbeitungsschleifen entwickelt. Dabei wurden in einem ersten Schritt mögliche Items aus bestehenden Fragebögen gesammelt sowie neue Items formuliert. Diese erste Version des SRL@U mit 105 Items wurde an einer Stichprobe von N1 = 222 eingesetzt. Auf Basis theoretischer Überlegungen sowie einer Reliabilitätsanalyse wurde die Auswahl von Items überarbeitet (103 Items) und in einer Studie mit N2 = 312 Studierenden eingesetzt. In einem weiteren Überarbeitungsschritt wurde erneut eine Reliabilitätsanalyse durchgeführt sowie die Anfälligkeit der Items für soziale Erwünschtheit berücksichtigt, wodurch ein Fragebogen mit 62 Items und 11 Subskalen entstand. Diese Version des Fragebogens wurde in drei weiteren Studien (N3 = 195, N4 = 249, N5 = 182) eingesetzt, sodass insgesamt schon Daten von mehr als 1.000 Studierenden vorliegen, die zur Bestimmung der Testgütekriterien verwendet werden können.

Im vorliegenden Vortrag soll die Kriteriumsvalidität des SRL@U im Vordergrund stehen. Dabei wurden in allen fünf Studien diverse Konstrukte erhoben, unter anderem Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1999), Big Five Persönlichkeitseigenschaften (Rammstedt & John, 2005), Intelligenz (Wortschatztest WST, Schmidt & Metzler, 1992; Raven Advanced Progressive Matrices Test, Arthur & Day, 1994), Lern- und Leistungsmotivation (SELLMO, Spinath, Stiensmeier-Pelster, Schöne, & Dickhäuser, 2002), akute oder frühere ADHS-Symptomatik (ASRS-V1.1, WHO, 2003; Wender Utah Ratings Scale, Retz-Junginger et al., 2002), Abiturnote, Leistung in einem Mathematiktest sowie geplante und reale Lernzeit in einem Online-Vorkurs und Zufriedenheit mit der eigenen Leistung in besagtem Kurs.

Erste Analysen weisen auf erwartungskonforme Zusammenhänge zwischen dem Gesamtscore im SRL@U und den Validierungskriterien hin: So konnte eine Korrelation mit Selbstwirksamkeitserwartung von $r = .447$ ($p < .001$) und mit Gewissenhaftigkeit von $r = .598$ ($p < .001$) gezeigt werden.

Weitere Analysen werden auch auf Ebene der Subskalen durchgeführt und schließen auch nicht-lineare Zusammenhangshypothesen ein. So wird beispielsweise vermutet, dass der Zusammenhang zwischen Intelligenz und Selbstregulationskompetenz kurvilinear verläuft: Im niedrigen Intelligenzbereich erwarten wir eher niedrige Ausprägungen von SRL; allerdings erwarten wir im hohen Intelligenzbereich ebenfalls eher niedrige SRL-Werte, da bei hoher Intelligenz der Einsatz von Lernstrategien oftmals nicht nötig ist, um gute Schulleistungen zu erreichen und daher wenig SRL praktiziert wird (Dresel & Haugwitz, 2005).

Mittwoch
11. März

14:00 Uhr

HGC 40

Chair(s)

Perels, Franziska

Modellierung und Erfassung des Zusammenhangs zwischen selbstreguliertem Lernen und leistungsrelevanten Konstrukten im Kontext pädagogischer Forschung

Ogrin, Sabine; Schmitz, Bernhard

Im Großen wie im Kleinen? Übereinstimmung von Schülerangaben zum selbstreguliertes Lernen im Querschnitt und als Prozess

Während und nach der Schulzeit bestimmen Fähigkeit und Bereitschaft, sich neues Wissen anzueignen, über selbstständiges Arbeiten und den beruflichen Erfolg (Brunstein & Spörer, 2001). Metaanalysen konnten zeigen, dass selbstreguliertes Lernen – insbesondere Zielsetzung, Reflexion und Selbstwirksamkeitserwartung – großen Einfluss auf gute Leistungen haben (Richardson, Abraham & Bond, 2012). Kern des selbstregulierten Lernens (SRL) ist dabei das systematische Herangehen an die Lernhandlung unter Anwendung metakognitiver, motivationaler und kognitiver Lernstrategien mit dem Ziel, die Effektivität des Handelns zu erhöhen (Zimmerman, Bonner & Kovach, 1996).

Wirth und Leutner (2008) fordern zur Messung des SRL den Einsatz von qualitativen Standards und Online-Standards. Qualitative Standards umfassen nicht die Erhebung der Quantität der eingesetzten Strategien, sondern die Passung der Strategie zur Lernsituation sowie die Umsetzungsgüte. Online-Standards bilden die Nähe der Messung zum eigentlichen Verhalten als ein sehr wichtiges Kriterium für die Validität der erhobenen Daten ab. Beide Standards werden jedoch nur sehr selten bei der Messung des SRL eingesetzt (Roth, Ogrin & Schmitz, 2014).

Strukturierte Tagebücher bieten die Möglichkeit, beiden Standards zu entsprechen (Klug, Ogrin, Keller et al., 2011). Mehrfach über viele Lernhandlungen eingesetzt bilden sie das Verhalten zeitnah als Abfolge von states ab und werden der Definition des SRL als Prozess (z.B. Schmitz & Wiese, 2006) gerecht. Sie ermöglichen weiterhin, die Aussagen der Lernenden über die eingesetzten Strategien auf ihre Passung zu Aufgabe und Zielsetzung zu untersuchen.

Ziel der vorliegenden Studie war es daher, zum einen die Übereinstimmung von Angaben zum SRL aus Fragebögen (quantitativ, offline) und aus Lerntagebüchern (qualitativ, online) zu untersuchen und zum anderen die Passung der eingesetzten Lernstrategien und Rückschlüsse der Lernenden zu überprüfen.

Es nahmen 75 Lehrer der Haupt- und Werkrealschulen in Baden-Württemberg mit ihren insgesamt 1.408 Schülern der 5. Klasse an der Interventionsstudie „Lernen mit Plan“ teil. 47 dieser Lehrer (801 Schüler) wurden geschult, über einen Zeitraum von sechs Wochen das SRL im in Kombination mit Mathematischem Modellieren (Maaß, 2007) zu fördern. 28 Lehrer (527 Schüler) fungierten mit ihren Klassen als Wartekontrollgruppe. Die Hälfte der Schüler war männlich. Das mittlere Alter der Schüler betrug elf Jahre. Die Schüler der Experimentalgruppe wurden angehalten, das Lerntagebuch über den Interventionszeitraum von 42 Tagen auszufüllen. Für die Tagebuchbearbeitung an mind. 28 Tagen erhielten sie für jeden ausgefüllten Tag 0,50€ (max. 20 €). Insgesamt wurden die Daten von 415 Schülern aus 34 Klassen in die Analysen eingeschlossen.

Das SRL wurde mittels Fragebogen und Lerntagebuch erfasst. Beides war angelehnt an Otto (2007). Fragebogen und Lerntagebuch umfassten u.a. die Planung und Motivation, das Aufschiebeverhalten und Self-Monitoring sowie die Anstrengung, Konzentration und Selbstwirksamkeit. Darüber hinaus wurden offene Fragen zur Zielsetzung, Reflexion des Lernvorgangs und Vorsatzbildung beantwortet.

Die Übereinstimmung von Fragebogen- und Tagebuchdaten im Sinne einer monotrait-multimethod Validität erwies sich als mittel bis niedrig. So betrug beispielsweise die Korrelation der Angaben zur Planung aus Fragebogen und Tagebuch $r = .30$ ($p < .01$) und zur Motivation $r = .23$ ($p < .01$). Die Angaben zur Konzentration und zum Monitoring korrelieren nicht signifikant miteinander. Es zeigte sich jedoch in beiden Messinstrumenten ein höherer Zusammenhang zwischen den Strategien einer Lernphase als über die verschiedenen Komponenten wie z.B. die Metakognition. Diese Ergebnisse werden in Zusammenhang mit den qualitativen Daten diskutiert.

Mittwoch
11. März

14:00 Uhr

HGC 40

Chair(s)

Perels, Franziska

Modellierung und Erfassung des Zusammenhangs zwischen selbstreguliertem Lernen und leistungsrelevanten Konstrukten im Kontext pädagogischer Forschung

Jäkel, Nils

Die Wirksamkeit des bilingualen Sachfachunterrichts auf dem Prüfstand

In der Fremdsprachendidaktik wird bilingualer Sachfachunterricht als eines der vielversprechendsten Fremdsprachenlernprinzipien bzw. -konzepte des 21. Jahrhunderts angesehen (García, 2009) und die Zahl der Schulen, die ihn im Wahlbereich anbieten, ist in den letzten zehn Jahren stark gestiegen; in NRW haben in 2013 mehr als 40% aller Gymnasien, knapp 30% aller Gesamtschulen und 20% aller Realschulen eine Form bilingualen Unterrichts angeboten. In quantitativ angelegten Studien (z.B. Köller et al., 2012; Nold et al., 2008; Zydariß, 2007) erwiesen sich die am bilingualen Sachfachunterricht teilnehmenden Schüler gegen Ende der Sekundarstufe I als fremdsprachlich deutlich kompetenter als ihre gleichaltrigen Mitschüler in Regelklassen. Neben den zuvor erwähnten Untersuchungen überwiegen (explorative) querschnittliche Studien mit geringen Probandenzahlen und qualitativen Designs. Zudem deuten die fallstudienartigen Daten einiger Forscher auf nicht zu unterschätzende Verzerrungseffekte hin (z.B. Fehling, 2008; Bredenbröcker, 2000; Burmeister, 1994), die durch die Eingangsselektion der Schüler und ihrer Vorbereitung auf den bilingualen Sachfachunterricht in den Klassen 5 und 6 verursacht werden. Um CLIL- und Regelschüler vergleichen zu können, brauchen wir longitudinale Studien, die für vielfältige Störvariablen und Bedingungsfaktoren kontrollieren.

Der bilinguale Sachfachunterricht mit seinem dualen fremd- und sachfachlichen Fokus und seinen multiperspektivischen Zielen, Methoden und Inhalten ist interdisziplinär angelegt und die daraus resultierende Lehr-Lernsituation erfordert komplexe Studiendesigns. Das Symposium umfasst vier auf quantitativen Studien basierende Beiträge zur sachfachlichen Kompetenz (Geschichtswissen), fremdsprachliche (Sprachkompetenzen) und lernpsychologische (akademisch-fremdsprachliches Selbstkonzept, Lernstrategien) Aspekte des Lernens in der Primar- und Sekundarstufe thematisieren. Das beitragsübergreifende Erkenntnisinteresse liegt in der kritischen Überprüfung der Wirksamkeit des bilingualen Sachfachunterrichts, wenn verzerrende Einflussfaktoren kontrolliert bzw. die jeweiligen Kompetenzbereiche longitudinal untersucht werden. Dallinger et al. stellen in ihrem Beitrag Ergebnisse zur sachfachlichen Kompetenz von Schülern der 8. Klasse im bilingualen Geschichtsunterricht vor, einem bisher stark vernachlässigt Forschungsfeld. Rumlich und Dicke werden in ihrem Beitrag Ergebnisse aus einer longitudinalen Studie zum Themenkomplex des Selbstkonzepts im Regelunterricht Englisch und bilingualer Sachfachunterricht vorstellen. Schwanke et al. stellen in ihrem Beitrag Ergebnisse zum Einfluss von Migrationshintergrund und Geschlecht im bilingualen Sachfachunterricht der 1.-6. Klasse in Tübingen vor. In seinem Beitrag zum Einfluss von Lernstrategien und Selbstwirksamkeit auf die Sprachkompetenz im Regelunterricht und bilingualer Sachfachunterricht stellt Nils Jäkel seine Forschungsergebnisse aus einer Querschnittsstudie vor.

Die beiden Diskutanten aus den Bereichen der Sprachlehr/lernforschung (Prof. Dr. Lena Heine) und der pädagogischen Psychologie (Prof. Dr. Jens Möller) sind ausgewiesene Experten auf Ihren Gebieten und bringen durch eigene Forschungstätigkeiten im Bereich des Lernens mit Fremdsprachen als Arbeitssprachen eine optimale kritische Außenperspektive in dieses Symposium.

Mittwoch
11. März

14:00 Uhr

HGC 50

Dallinger, Sara; Jonkmann, Kathrin; Hollm, Jan

(Wirklich) „Zwei Fliegen mit einer Klappe“? – Die Effekte bilingualen Geschichtsunterrichts auf die Englisch- und Sachfachkompetenzen von Schülerinnen und Schülern

Seit der Etablierung bilingualer Unterrichtsprogramme in Deutschland besteht eine kontinuierlich geführte Debatte zum Mehrwert dieser Unterrichtsform für die Schülerinnen und Schüler (Breidbach & Viebrock, 2012). Grundlage für diese Diskussion stellen Theorien aus dem Zweitspracherwerb (z. B. input hypothesis von Krashen, 1985; threshold hypothesis von Cummins, 1979) sowie kognitionspsychologische Erkenntnisse zu kognitiven Korrelaten von Bilingualismus dar (Adesope, Lavin, Thompson & Ungerleider, 2010). Deren Übertragung auf den Kontext des bilingualen Unterrichts gibt Grund zur Annahme, dass sich sowohl Fremdsprachen- als auch Sachfachkompetenzen der bilingual unterrichteten Lerner günstiger entwickeln könnten als bei einsprachig unterrichteten Schülerinnen und Schülern. Bisherige Untersuchungen kamen zu entsprechenden Ergebnissen, wonach sich bilingualer Unterricht positiv auf die Fremdsprachenkompetenzen auswirkt (Köller, Leucht & Pant, 2012; Nold, Hartig, Hinz & Rossa, 2008; Zydariß, 2007). Im Gegensatz dazu sind jedoch kaum Studien bekannt, die sich mit den Effekten dieser Unterrichtsform auf die Sachfachkompetenzen beschäftigt haben (Bonnet, 2012; Bernhardt, 2011; Coyle, Hood & Marsh, 2010) und hierbei mithilfe adäquater Forschungsmethoden generalisierbare und belastbare Aussagen treffen konnten (Pérez-Cañado, 2012; Bruton, 2011). Dies liegt mitunter darin begründet, dass bisherige Untersuchungen zur Wirksamkeit bilingualen Unterrichts überwiegend querschnittlich durchgeführt wurden und weder Unterschiede in den individuellen Lernvoraussetzungen der Schülerinnen und Schüler noch bezüglich klassenbezogener Merkmale (z. B. Kompositionseffekte) berücksichtigten. Dies schien zu einer tendenziellen Überschätzung der Effekte dieser Unterrichtsform auf die Schülerleistungen geführt zu haben (Rumlich, 2014).

Der vorliegende Beitrag beantwortet daher die Frage, ob sich die Kompetenzen der Lerner im bilingualen Geschichtsunterricht nicht nur in Englisch, sondern auch in Geschichte günstiger entwickeln als bei einsprachig unterrichteten Schülerinnen und Schülern, unter Kontrolle zentraler Faktoren auf der Schülerebene (Geschlecht, Migrationshintergrund, sozioökonomischer Status, figurale und verbale Grundfähigkeiten, Motivation, akademisches Selbstkonzept) und der Klassenebene (mittlere Leistungen, Unterrichtsqualität in Englisch und Geschichte).

Hierfür wurde in der Studie „Competences and Motivation in Bilingual Instruction in History“ (COMBIH) eine längsschnittliche Untersuchung mit einem ersten Messzeitpunkt zu Beginn des bilingualen Geschichtsunterrichts in der 8. Klasse sowie einem zweiten am Ende des Schuljahres durchgeführt. Es wurden Daten von 1806 Achtklässlern aus 37 Gymnasien erhoben, wovon 703 Schülerinnen und Schüler aus 30 Klassen bilingual unterrichtet wurden und 1103 Lerner aus 45 Klassen einsprachigen Sachfachunterricht erhielten. Die Lerner nahmen an einem figuralen und verbalen Grundfähigkeitstest teil (KFT; Heller & Perleth, 2010) sowie an Leistungstests zur allgemeinen englischen Sprachfähigkeit (C-Test; vgl. Grotjahn, 2011), zum englischen Hörverstehen (Köller, Knigge & Tesch, 2010) und zum Sachwissen in Geschichte, welcher auf Basis des entsprechenden Bildungsplans entwickelt wurde (KM, 2004). Weiterhin füllten die Lerner einen Fragebogen zu ihrem familiären Hintergrund, Motivation und akademischem Selbstkonzept (Wigfield & Eccles, 2000) sowie zur Unterrichtsqualität in Englisch und Geschichte aus (Kunter, Baumert, Blum, Klusmann, Krauss & Neubrand, 2011). Die Daten wurden mehrebenenregressionsanalytisch ausgewertet, um Unterschiede in den Leistungen zwischen den Schülergruppen sowohl auf Individual- als auch auf Klassenebene kontrollieren zu können (Raudenbush & Bryk, 2002).

Die Analysen ergaben, dass Schülerinnen und Schüler aus bilingual unterrichteten Klassen unter Kontrolle der Ausgangsleistungen, der individuellen Lernvoraussetzungen und der Merkmale auf Klassenebene eine deutlich bessere Leistungsentwicklung in ihren Englischkompetenzen zeigten als die einsprachig unterrichteten Lerner. Hierbei ergaben sich Vorsprünge in Höhe von mindestens einviertel Schuljahre. Bezüglich der Verbesserung des Sachwissens in Geschichte zeigten sich parallele Entwicklungsverläufe zwischen Schülerinnen und Schülern aus bilingual und einsprachig unterrichteten Klassen. Dies deutet darauf hin, dass der bilinguale Unterricht mit einer sehr positiven Entwicklung der Fremdsprachenleistungen einhergeht, aber keinen Mehrwert für das Sachfach für die Lerner bereithält, zumal dies zumeist mit einem um eine Stunde höheren Unterrichtsvolumen unterrichtet wird. Die vorliegenden Befunde können als Basis für zukünftige Entscheidungen hinsichtlich der Ausweitung bilingualer Unterrichtsprogramme dienen.

Mittwoch
11. März

14:00 Uhr

HGC 50

Schwanke, Katrin; Steinlen, Anja; Piske, Thorsten

Werden Kinder mit Migrationshintergrund oder Mädchen in bilingualen Programmen in Grundschule und Gymnasium benachteiligt?

In der Presse werden immer wieder Stimmen laut, die von einer Benachteiligung bestimmter Lernergruppen während der Schulzeit berichten. Im besonderen Fokus stehen hierbei bereits seit mehreren Jahren Kinder mit Migrationshintergrund und aktuell auch Mädchen, nachdem bis vor wenigen Jahren noch den Jungen besondere Aufmerksamkeit geschenkt wurde.¹

Da die solchen Einschätzungen zugrundeliegenden Daten zu Schülerleistungen bisher vor allem aus Regelschulen stammen, möchten wir der Frage nachgehen, ob sich Beobachtungen zu den Leistungen bestimmter Schülergruppen auch für den bilingualen Unterricht (BU) bestätigen lassen. In Deutschland ist seit Beginn der 1990er Jahre ein stetiger Zuwachs an Grund- und weiterführenden Schulen mit bilingualen Angeboten zu verzeichnen. Bereits im Jahr 2006 orientierten sich dabei über 600 Schulen am vor allem aus Kanada bekannten Immersionsansatz (KMK 2006), nach dem die Fremdsprache als Unterrichtssprache in verschiedenen Fächern eingesetzt wird. Gleichzeitig steigt in Deutschland der Anteil an Schülerinnen und Schülern mit Migrationshintergrund. Im Jahr 2011 lag er bei Kindern und Jugendlichen im Alter von bis zu 15 Jahren bereits bei rund 32% (Statistisches Bundesamt, 2012: 56). Dies legt nahe, dass auch immer mehr Kinder mit Migrationshintergrund eine bilinguale Schule besuchen werden. Bisher haben sich nur wenige Studien mit dem Einfluss des Migrationshintergrundes auf die sprachlichen und schulischen Leistungen im bilingualen Grundschulkontext auseinandergesetzt (siehe aber z.B. Steinlen & Piske 2013). Die Frage, ob insbesondere Mädchen oder Jungen in einem solchen Programm Vorteile erfahren, ist unseres Wissens bisher nicht aufgegriffen worden.

Für die weiterführende Schule existieren dagegen – neben der bekannten DESI-Studie (DESI-Konsortium 2008) – weitere Studien, in denen die Englischleistungen verschiedener Merkmalsgruppen miteinander verglichen worden sind (z.B. bilingual unterrichtete Klassen vs. Klassen mit Regelunterricht, Jungen vs. Mädchen, +/- Migrationshintergrund) (z.B. Köller 2010, Zydatiś 2007). Diese Arbeiten beziehen sich aber vor allem auf Neuntklässler. Eine systematische Untersuchung der Auswirkungen bilingualen Unterrichts von der Grundschule in die Sekundarstufe I am Gymnasium unter Berücksichtigung einer möglichen Wechselwirkung zwischen Geschlecht und Migrationshintergrund steht noch aus.

In der vorgestellten Studie wird die Entwicklung der rezeptiven englischen Wortkenntnisse bei Kindern der 1.-6. Klasse in einer bilingualen Grundschule bzw. einem bilingualen Gymnasium in Tübingen untersucht. Als Testinstrument wird der BPVS II (British Picture Vocabulary Scale, Dunn et al. 1997) eingesetzt. In diesem Test müssen die Schülerinnen und Schüler aus vier Bildern das Bild auswählen, das zu einem mündlich vorgegebenen englischen Wort passt. Um kognitive Leistungen zu erfassen, wird mit den Kindern außerdem ein Aufmerksamkeits- und ein Test zur nonverbalen Intelligenztest durchgeführt.

In der vorgestellten Studie wird vornehmlich auf die folgenden drei Fragestellungen eingegangen: (1) Wie entwickeln sich die sprachlichen Kompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund im BU und zwar sowohl in der Grundschule als auch auf dem Gymnasium? (2) Wie entwickeln sich bestimmte sprachliche Kompetenzen von Mädchen und Jungen im BU? (3) Zeigen sich Effekte einer Interaktion zwischen den Variablen „Migrationshintergrund“ und „Geschlecht“?

¹ So fanden sich in den letzten sieben Jahren zum Beispiel folgende Überschriften: „Mädchen und Migranten verschlechtern deutsches Ergebnis“ (Focus 2.4.14 zu den neuesten PISA-Ergebnissen), „Werden Kinder mit Migrationshintergrund oder Mädchen in der Schule benachteiligt?“ (Elitenetzwerk Bayern 16.08.2012, zur Notengebung im deutschen Schulsystem), „Gleichstellung von Jungen und Mädchen unterschiedlicher Herkunft in der Schule“ (Genderkompetenzzentrum 2010), „Ausgebremste Jungs“ (Hinz&Kunzt 172/Juni 2007), „Risse in der Meritokratie: Was wird in deutschen Schulen eigentlich bewertet?“ (Sciencefiles 22.7.2012), „Mädchen Top – Jungen Flop?“ (Heise Online 12.04.2008), „Nach Pisa: Was Migranten in der Schule fehlt“ (Tagesspiegel 4.12.2012), „Deutsches Schulsystem hält Migranten klein“ (Zeit Online 4.12.2013), „Migranten-Kinder starten in den Schulen durch“ (Die Welt 23.06.12).

Mittwoch
11. März

14:00 Uhr

HGC 50

Rumlich, Dominik; Dicke, Theresa

Die Entwicklung eines höheren fachbezogenen Selbstkonzepts als Erklärung für die fremdsprachlichen Leistungen der SchülerInnen im bilingualen Unterricht? Eine empirische-quantitative Evaluation eines theoretischen Postulats

Der regelrechte Hype um bilingualen Sachfachunterricht (bili SFU; engl. CLIL) in den letzten 20 Jahren scheint auch von wissenschaftlicher Seite gerechtfertigt, da in vielen Fallstudien (z.B. Bredenbröker, 2000) sowie einigen large-scale Studien (z.B. Köller et al., 2012; Nold et al., 2008) die Teilnehmenden deutlich höhere (fremdsprachliche) Kompetenzen aufweisen als ihre Peers in Regelklassen. Dennoch fehlen „eindeutige, empirisch abgesicherte Ergebnisse hinsichtlich der zentralen Annahmen und Ansprüche des bilingualen Unterrichts und seiner Überlegenheit“ (Vollmer, 2010, S.51). Mögliche Ursachen hierfür sind der komplexe Forschungskontext im Spannungsfeld der Sachfachdidaktiken, der Englischdidaktik und der empirischen Bildungsforschung, die Eingangsselektion der bili SFU SchülerInnen (z.B. Bonnet, 2012; Fehling, 2008) sowie deren Vorbereitung in den Klassen 5 und 6 mit zwei zusätzlichen Englischstunden pro Woche: Diese und weitere Faktoren üben einen verzerrenden Einfluss auf querschnittliche Messungen aus und sind bislang nicht ausreichend kontrolliert worden.

Diese Selektions- und Präparationseffekte sind auch ein Problem in Studien des fachbezogenen Englisch-Selbstkonzepts (Daley, Onwuegbuzie & Bailey, 1997; Mills, Pajares & Herron, 2007), dessen implizierte Steigerung (z.B. Zydariß, 2007, S.148) aufgrund der Unterrichtsprinzipien des integrierten Sach- und Sprachlernens als Erklärung für die hohen sprachlichen Kompetenzen der Teilnehmenden herangezogen wird; auf theoretischer Ebene werden die Zuwächse der Sprachkompetenzen häufig affektiv-motivational über die Erwartungskomponente im Erwartungs-Wert-Modell (z.B. Eccles [Parsons] et al., 1983; Eccles, 2005) erklärt (Zaubauer et al., 2009).

Das Erkenntnisinteresse der vorliegenden Studie besteht einerseits in der Untersuchung der a priori existierenden Unterschiede in Bezug auf sprachliche Kompetenzen, fachbezogenes Selbstkonzept etc. und die sich daraus potentiell ergebende Verzerrungen bisheriger Studienergebnisse. Zudem ermöglicht die Kontrolle der Ausgangslagen, die Leistungsfähigkeit des bilingualen Sachfachunterrichts im Hinblick auf die fremdsprachliche und affektiv-motivationale Entwicklung der teilnehmenden SchülerInnen deutlich fehlerbereinigter zu evaluieren und wichtige Erkenntnisse im Bereich einer essentiellen interdisziplinären Forschungslücke zu erlangen.

Um sowohl die Verzerrungseffekte durch Präparation und Selektion als auch den tatsächlichen Einfluss des CLIL-Unterrichts näher zu untersuchen, wurde in den Jahren 2011-2013 eine quasi-experimentelle Längsschnittstudie (N=1.398) in NRW durchgeführt. Dazu wurde am Ende der 6.Klasse, d.h. unmittelbar vor dem Beginn des bili SFU, sowie zwei Jahre danach u.a. allgemeine englische Sprachkompetenz (Amt für Schule, Hamburg, 1998; 2000), verbale kognitive Fähigkeiten (KFT; Heller & Perleth, 2000), fachbezogenes Selbstkonzept (DESI, Wagner et al., 2009; Ländervergleich Sprachen, Sachse et al., 2012; KESS 7, Bos et al, 2009) erhoben. Die Stichprobe umfasste SchülerInnen bilingualer Zweige (N=503), nicht-bilingualer Zweige (der gleichen Schulen; N=473) sowie SchülerInnen von Regelschulen ohne bilingualen Unterricht als unverzerrte Kontrollgruppe (N=182).

Die Sprachkompetenzdaten wurden mit einem eindimensionalen Rasch-Modell skaliert (Conquest, Wu et al., 2007) und zusammen mit den weiteren Daten mit Mplus 7 (Muthén & Muthén, 1998-2012) ausgewertet. Die Kennwerte der Rasch-Items sind gut ($0.8 < WMSQ < 1.25$; EAP/PV Reliabilitäten $> .90$; Item-Separationsreliabilitäten $> .99$). Das manifeste Gesamtmodell geschätzt mit MLR (type=complex) weist ebenfalls gute Fit-Werte auf ($R^2=.61$; $X^2(10)=17.76$, $p=.06$; CFI=.99; TLI=.97; RMSEA=.03, .00<90% CI<.05; SRMR=.01).

Der Effekt des bilingualen Treatments auf das fachspezifische Selbstkonzept nach zwei Jahren bilingualen Unterricht ist klein ($\beta=.11$, $SE=.06$, $p<.05$), es wird maßgeblich vom Selbstkonzept der Lernenden am Ende der Klasse 6 determiniert ($\beta=.47$, $SE=.04$, $p<.001$). Die Daten zeigen zudem, dass sich zwar das Selbstkonzept signifikant auf die allgemeine Sprachkompetenz auswirkt ($\beta=.19$, $SE=.04$, $p<.001$), für das Treatment „bilingualer Unterricht“ gibt es jedoch keinen signifikanten Haupteffekt ($\beta=.05$, $SE=.17$, $p=.78$) auf die Sprachkompetenzen entgegen der Ergebnisse anderer Studien.

Bei der Berechnung manifester multipler Regressionsmodelle vergleichbar mit denen vorangegangener Studien ohne Kontrolle der Ausgangslage bzw. der Positiv- und Negativselektion wird der Haupteffekt des bilingualen Unterrichts – fälschlicherweise – signifikant. Diese und weitere Ergebnisse lassen die Vermutung zu, dass in bisherigen, v.a. querschnittlichen Studien mit nur einer Messung, die Effekte des bili SFU auf das Selbstkonzept und die Fremdsprachenkompetenz überschätzt wurden.

Mittwoch
11. März

14:00 Uhr

HGC 50

Jäkel, Nils

Selbstwirksamkeit, Lernstrategiegebrauch und Sprachstand von Schülern des bilingualen Sachfachunterrichts und Englisch Regelunterrichts

Der Bilinguale Sachfachunterricht (bili SFU, engl. CLIL) stellt durch seinen dualen Fokus auf Fremdspracherwerb und das Erlernen von Sachfachwissen höhere Anforderungen an Schüler als dies im Regelunterricht Englisch der Fall ist. Um die Schüler stärker im Lernprozess des bili SFU zu unterstützen und ihnen Werkzeuge an die Hand zu geben, die ihnen bei der Bewältigung dieser erhöhten Anforderungen helfen können, wurden Sprachlernstrategien als ein vielversprechendes Instrument herausgestellt (vgl. Coyle, 2007; Mehisto, Marsh, Frigolos, 2008; Järvinen, 2007). Studien zu Sprachlernstrategien weisen auf einen positiven Zusammenhang von erfolgreichem Erlernen von Sprachen und Strategiegebrauch hin (vgl. Griffiths, 2003; Green & Oxfröd 1995). Dies lässt sich auch in Bezug auf den Bilingualen Sachfachunterricht vermuten. Obwohl im Bereich des bili SFU demnach ein besonderes Interesse an Sprachlernstrategien bestehen müsste (vgl. Coyle, 2007), wurde der Zusammenhang zwischen Strategiegebrauch und Sprachlernerfolg im bili SFU in Deutschland bisher nicht untersucht. Ein weiterer essentieller Faktor für den schulischen Erfolg sind individuelle Selbstwirksamkeitsüberzeugungen. Selbstwirksamkeit hat sich sowohl in pädagogisch-psychologischen Studien aber auch in Fremdspracherwerbsstudien als ein wichtiger Prädiktor schulischen bzw. akademischen Erfolgs erwiesen (Bandura, 1997; Mills, 2007, Yang, 1999). Die vorliegende Studie untersucht neben dem Lernstrategiegebrauch daher auch die Selbstwirksamkeit der Schüler in Bezug auf den Fremdspracherwerb und ihre Auswirkung auf den Sprachstand in Englisch von Gymnasiasten im Regelunterricht Englisch sowie im bili SFU der 9. Klasse.

Ziel der Studie ist es, herauszufinden (1) welche Faktoren Lernstrategiegebrauch vorhersagen; (2) wie der Strategiegebrauch den Sprachstand beeinflusst; und (3) wie die Selbstwirksamkeit zum Sprachlernen sich auf den Sprachstand auswirkt, während für Alter, Geschlecht, sozioökonomischen Status der Eltern, kognitive Fähigkeiten (CFT-20; Weiß, 2008), L1 (Muttersprache deutsch / nicht-deutsch) sowie Zugehörigkeit zu regulären Englisch- bzw. bilingualen Klassen kontrolliert wird. Der Lernstrategiegebrauch wurde mit Rebecca Oxfords 50-Item Strategy Inventory for Language Learning erhoben (SILL; Oxford, 1990). Der Sprachstand wurde durch eine vier C-Tests umfassende Testbatterie (Cronbach's $\alpha=0.845$; Zydaniß, 2005) sowie Schulnoten in Englisch ermittelt. Die C-Test Batterie wurde mittels einer konfirmatorischen Faktorenanalyse (CFA) validiert. Die Selbstwirksamkeit der Schüler in Bezug auf das Englischlernen wurde anhand einer fünf Item umfassenden Skala erhoben (Cronbach's $\alpha=0.725$) und durch eine CFA validiert.

Die Hypothesen wurden anhand eines Strukturgleichungsmodells (SEM) überprüft ($\chi^2(73) = 138,519$; $p=0.001$, CFI=0.954, RMSEA=0.049, PCLOSE=0.547). Die Ergebnisse zeigen, (1) dass Mädchen signifikant mehr und CLIL Schüler weniger Lernstrategien verwenden und dass eine höhere Selbstwirksamkeit höheren Strategiegebrauch vorhersagt. (2) Wider Erwarten übte Lernstrategiegebrauch einen negativen Einfluss auf den Sprachstand aus, nachdem für alle Störvariablen statistisch kontrolliert wurde. Dieses Ergebnis bestätigt Gardner et al. (1997), die in ihrer Studie ebenfalls einen negativen Einfluss von Lernstrategiegebrauch auf den Sprachstand fanden. Die Ergebnisse zeigen weiter, (3) dass eine höhere Selbstwirksamkeit in Bezug auf das Englischlernen den Sprachstand signifikant positiv beeinflusst. Die latente Variable Selbstwirksamkeit wurde von Geschlecht, SES, L1 und Zugehörigkeit zum bili SFU bzw. englischen Regelklassen positiv vorhergesagt. Mädchen, Schüler mit höherem SES, Schüler mit einer anderen L1 als Deutsch, und Schüler des bili SFU berichteten eine höhere Selbstwirksamkeit. Dies bedeutet, dass die Quantität des Lernstrategiegebrauchs in der vorliegenden Studie nicht wie in anderen Studien (z.B. Green & Oxford, 1995; Nguyen & Godwyll, 2010) und entgegen unserer Hypothese keinen positiven Einfluss auf den Sprachstand. Stattdessen verschlechterte erhöhter Lernstrategiegebrauch die Note und reduzierte die Leistungen im C-Test. Selbstwirksamkeit hingegen führte zu signifikant besseren C-Test Ergebnissen und besseren Noten. Fraglich ist, inwieweit der SILL das Strategieverhalten von jugendlichen Lernern mit gehobenem Sprachstand adäquat erfasst. Insbesondere CLIL Schüler könnten Strategien gebrauchen, die nicht im SILL aufgeführt sind oder sie könnten den Strategiegebrauch ggf. so stark automatisiert haben, dass sie den Gebrauch nicht erinnern.

Mittwoch
11. März

14:00 Uhr

HGC 50

Bernhardt, Raphael; Frey, Andreas

Linking eines papierbasierten Tests mit einem computerisierten adaptiven Test unter Berücksichtigung von Itempositionseffekten

Die Kompetenzmessung innerhalb einer Studie kann durch unterschiedliche Testformen (z.B. computerbasiertes oder papierbasiertes Testen) erfolgen. Um auf einer gemeinsamen Metrik zu berichten, werden häufig die Itempools der Testformen mittels Item-Linking-Strategien basierend auf der Item-Response-Theorie (IRT) zusammengebracht. Beim Linking gibt es die Annahme zur Invarianz von Itemparametern über verschiedene Testformen hinweg. In der Regel finden Aspekte wie Itempositionseffekte oder die Art des jeweils verwendeten Testheftdesigns keine Beachtung, obwohl anzunehmen ist, dass diese die Invarianzannahme verletzen könnten. Mit der vorliegenden Studie wird ein Ansatz zur Berücksichtigung von Itempositionseffekten beim Linking vorgestellt und an einem empirischen Datensatz illustriert.

Im Rahmen der Entwicklung drei computerisierter adaptiver Tests im Projekt „Messung allgemeiner Kompetenzen-adaptiv“ wurden in einer Kalibrierungsstudie die Itemparameter mit Hilfe eines computerbasierten Tests (CBT) mit fester Itemreihenfolge und einem komplexen Testheftdesign geschätzt. In einer zusätzlichen Studie wurden Itempositionseffekte für die Itemparameter des CBT ermittelt. In der Pilotierungsstudie für den adaptiven Test wurde für jede Domäne zusätzlich ein papierbasierter Test (PPT) mit fester Itemreihenfolge ohne Testheftdesign kalibriert. Es gab somit für jede Domäne lediglich ein Testheft. Der Itempool für den PPT hat je nach Domäne zwischen 88% und 100% gemeinsame Items mit dem Itempool des CBT. Es liegt damit ein nicht äquivalentes Gruppendesign mit der Verwendung gemeinsamer Items (vgl. Kolen & Brennan, 2014) bei der Nutzung unterschiedlicher Testformen mit unterschiedlichen Testheftdesigns vor. Die Kalibrierung der unterschiedlichen Testformen erfolgte zu unterschiedlichen Zeitpunkten mit unterschiedlichen Stichproben. Daraus ergeben sich folgende Fragen: (1) Wie lassen sich die Itempools beider Testformen angemessen miteinander verbinden? (2) Welche Auswirkungen hat die Verwendung von Itempositionseffekten auf die Auswahl der Linkitems? (3) Welche Auswirkungen hat die Linkingprozedur auf die Personenparameterverteilung und die Reliabilität des PPT?

N = 1632 (46% weiblich) Schülerinnen und Schülern (SuS) beruflicher Schulen bekamen einen CBT und N = 528 (38% weiblich) einen PPT vorgelegt. Mit dem Ziel beide Testformen zu verlinken, wurden zuerst alle gemeinsamen Items beider Testformen als Linkitems angesehen. Die Linkingprozedur basiert auf vier Schritten. Schritt 1: Die Antworten auf die Items beider Testformen wurden separat mit einem Rasch-Modell skaliert. Schritt 2: Die Itemschwierigkeiten der Linkitems des PPT wurden mittels Mean/Mean-Methode auf die Metrik der computerbasierten Testung transformiert (vgl. Kolen & Brennan, 2014). Schritt 3a: Itempositionseffekte für die Zielmetrik des CBT wurden ermittelt. Schritt 3b: Items mit signifikanten Unterschieden in der Schwierigkeit zwischen beiden Testformen wurden identifiziert. Lag der Schwierigkeitsparameter des CBT außerhalb des 95% Konfidenzintervalls des Schwierigkeitsparameters des PPT, wurde die Hypothese der Itemparameterinvarianz zwischen den Testungen verworfen und das entsprechende Item nicht mehr als Linkitem angesehen. Aufgrund der unterschiedlichen Testheftdesigns wurden vor dem Vergleich der Itemschwierigkeiten beider Testformen Itempositionseffekte auf die Itemschwierigkeiten des CBT addiert. Schritt 4: Die Antworten auf die Items des PPT wurden erneut mit einem Rasch-Modell skaliert. Die ausgewählten Linkitems aus Schritt 3 wurden als Anker verwendet. D.h. die Itemschwierigkeiten der Linkitems aus dem CBT inklusive Positionseffekte wurden als Itemschwierigkeit im PPT verwendet. Die restlichen Itemschwierigkeiten wurden frei geschätzt.

Durch die Anwendung der vorgestellten Methode konnten beispielsweise für den PPT der Domäne Naturwissenschaft (N = 173; 34% weiblich) 83% der ursprünglichen Linkitems genutzt werden. In der Domäne Lesen (N = 179; 42% weiblich) konnten durch die Berücksichtigung der Positionseffekte 10% mehr Linkitems als ohne Berücksichtigung der Positionseffekte identifiziert werden. Die vorgestellte Linkingprozedur, insbesondere die Verwendung der Positionseffekte, hat keine signifikanten Auswirkungen auf die Personenparameterverteilung und die Reliabilität des PPT der verwendeten Studie.

Mit Hilfe des vorgestellten Ansatzes können Itempositionseffekte beim Linking zweier Testformen und somit eine mögliche Invarianz der Itemparameter bei unterschiedlichen Testheftdesigns berücksichtigt werden.

Mittwoch
11. März

14:00 Uhr

GC 03/142

Chair(s)

Bernhardt, Raphael

Engelhardt, Lena; Goldhammer, Frank; Naumann, Johannes; Frey, Andreas; Hartig, Katja; Horz, Holger; Kuchta, Kathrin; Wenzel, S. Franziska C.

Konstruktvalidierung durch die Manipulation der Aufgabenschwierigkeit in simulationsbasierten ICT-Aufgaben

Ein kompetenter Umgang mit Informations- und Kommunikationstechnologien (ICT) stellt eine wichtige Voraussetzung für die erfolgreiche Teilhabe in unserer heutigen Gesellschaft dar. ICT Fertigkeiten (Skills) werden als „survival skill“ (Eshet-Alkalai, 2004) und als eine der Schlüsselkompetenzen für lebenslanges Lernen (European Commission, 2007) bezeichnet und stellen demnach ein wichtiges Bildungsziel dar. Um ICT-Fertigkeiten verhaltensbasiert zu erheben, wurden auf Grundlage einer theoretischen Rahmenkonzeption simulationsbasierte Testaufgaben entwickelt. Aufbauend auf der Konzeption des International ICT Literacy Panel (2002) ist auch hier die kognitive Operation zentral. Es wird unterschieden, ob eine Aufgabe den Zugriff (access), das Verwalten (manage), Integrieren (integrate), Bewerten (evaluate) oder Erstellen (create) von Information erfordert. Mit Informationen umzugehen war auch schon vor der Existenz von ICT notwendig und es wird angenommen, dass sich benötigte ICT-spezifische Fertigkeiten aus traditionellen Fertigkeiten wie Problemlösen und Medienrezeption, gemeinsam mit technischem Wissen, entwickelt haben. Aus diesen Bereichen vorhandene Theorien und Modelle wurden herangezogen um die bei der Nutzung von ICT ablaufenden kognitiven Prozesse zu beschreiben und weiterführend Merkmale herauszuarbeiten, die zur Schwierigkeit einer Aufgabe beitragen. Solche Merkmale sind zum Beispiel, abgeleitet aus dem Bereich der Medienrezeption, die Menge an Informationen auf einer Internetseite oder farbliche Hervorhebung oder vergrößerte Schrift zur Unterstützung der Informationsverarbeitung. Aus dem Bereich des Problemlösens lässt sich die Anzahl notwendiger Handlungsschritte als schwierigkeitsbestimmend ableiten und aus dem Bereich des technischen Wissens das Ausmaß in dem Wissen bezüglich Gefahren im Internet oder Gestaltungsmöglichkeiten beim Erstellen von Dokumenten zur Lösung der Aufgabe benötigt wird.

Zur Untersuchung, ob die theoretisch definierten ICT-spezifischen Merkmale die Aufgabenschwierigkeit beeinflussen, wurden entwickelte Items hinsichtlich dieser potentiell schwierigkeitsbestimmenden Merkmale systematisch verändert. Hat die Änderung dieser Aufgabenmerkmale einen Effekt auf die Lösungswahrscheinlichkeit, so kann dies als Beleg der Validität der Interpretation der Testwerte gewertet werden (Embretson, 1983), sofern diese nicht zu einer Änderung des Konstrukts führen. Es wurden 40 Testaufgaben, acht pro kognitive Operation, systematisch verändert mit dem Ziel sowohl schwerere als auch leichtere Items zu erhalten. In jedem veränderten Item tritt das jeweilige schwierigkeitsbestimmende Merkmal in stärkerem Maße, gleichbleibend oder in geringerem Maße auf als im Originalitem. Das zugehörige Originalitem bildet die Referenz für die zu prüfende Schwierigkeitsveränderung durch die Merkmale, ebenso wie alle veränderten Items in denen ein spezifisches Merkmal nicht verändert wurde. Die Testaufgaben wurden ca. N=1000 15-jährigen Schülerinnen und Schülern verschiedener Schulformen vorgelegt. Die Teilnahme war freiwillig und die Untersuchung fand während der Unterrichtszeit im Klassenkontext statt. Den Schülerinnen und Schülern wurden zufällig entweder nur Original- oder veränderte Items zugeteilt. Dazu wurde ein Rotationsdesign verwendet, sodass jeder Person bis zu 40 Items vorgelegt wurden. Für die Originalitems liegen ca. 200 Antworten pro Item und für die veränderten Items ca. 120 Antworten pro Item vor. Zusätzlich haben alle Schülerinnen und Schüler Fragen zum Nutzungsverhalten von ICT beantwortet.

Analysen basierend auf einem Teildatensatz weisen darauf hin, dass für originale und veränderte Items das Rasch-Modell gilt. Ein Vergleich der geschätzten Itemschwierigkeiten zeigt Änderungen in erwarteter Richtung sowie eine breitere Testinformationskurve für den veränderten Itemsatz. Da sich durch die Veränderung zwar die Schwierigkeit nicht aber das gemessene Konstrukt ändern soll, wird der Zusammenhang der Leistungsdaten mit Nutzungsgewohnheiten für beide Itemsätze ermittelt und verglichen. Weitergehend sollen erklärende IRT-Modelle angewandt werden um Schwierigkeitsänderungen durch die veränderten Merkmale zu erklären. Während Aufgabenmerkmale wie technisches Wissen spezifisch für den ICT Bereich sind, sollten weniger ICT spezifische Merkmale auch auf weitere Inhaltsbereiche übertragbar sein. Diese vorläufigen Ergebnisse sollen mit vollständigem Datensatz überprüft werden. Die Ergebnisse liefern Erkenntnisse hinsichtlich der Konstruktvalidität und können durch gezielte Veränderung der Aufgabenmerkmale zur weiteren Aufgabenkonstruktion und –Anpassung genutzt werden. Ein geeigneter Itempool mit Items bestimmter Schwierigkeiten wird zum Beispiel für Adaptives Testen benötigt, wodurch bei simulationsbasierten Aufgaben das Problem langer Bearbeitungszeiten gelöst werden kann.

Mittwoch
11. März

14:00 Uhr

GC 03/142

Lüdtke, Oliver; Robitzsch, Alexander; Trautwein, Ulrich

Ein interpersoneller Ansatz zur Modellierung der Akkuratheit von Fähigkeitsselbsteinschätzungen von Schülerinnen und Schülern

Inwiefern die Selbsteinschätzung schulischer Fähigkeiten verzerrt ist und ob eine positive Verzerrung eventuell mit einer verbesserten Anpassung einhergeht, ist ein kontrovers diskutiertes Thema in der pädagogisch-psychologischen Forschung. In dem vorliegenden Beitrag wird ein statistisches Modell vorgestellt, das auf einem interpersonellen Ansatz (Kwan, John, Kenny, Bond & Robins, 2004) basiert, der es ermöglicht zwei bisher in der Forschung miteinander konfundierte Komponenten der Selbstwahrnehmung voneinander zu trennen: die mittlere Tendenz, mit der andere Personen wahrgenommen werden (Perceiver-Effekt), und die mittlere Tendenz, mit der man von anderen Personen wahrgenommen wird (Target-Effekt).

Der interpersonelle Ansatz setzt voraus, dass ein sogenanntes round-robin Design (Kenny, Kashy & Cook, 2006) vorliegt, in dem jeder Schüler bzw. jede Schülerin von allen anderen Schülern z.B. hinsichtlich ihrer Fähigkeit in einem Fach beurteilt wird (z.B. „Wie gut ist der Schüler X in Mathe?) und zusätzlich auch noch sich selbst beurteilt („Wie gut bist Du selbst in Mathe?“; Selbsteinschätzung). In der TRAIN-Studie (Trautwein et al., 2010) wurde dieses Design im Rahmen eines sozialen Netzwerkes in der 5. Jahrgangsstufe für die Fächer Mathematik und Englisch in 50 Klassen umgesetzt. Zusätzlich liegen die Fachleistungen in den beiden Fächern vor (Testleistung und Noten).

Im Mittelpunkt des Vortrags steht die statistische Modellierung der verschiedenen Komponenten der Selbstwahrnehmung mit Hilfe des Social Relations Modells (SRM; Kenny, 1994). Das SRM kann als ein Mehrebenenmodell mit gekreuzten Zufallseffekten interpretiert werden (siehe Snijders & Kenny, 1999). Es wurde gezeigt, dass sich das SRM besonders flexibel in einem Bayes-Ansatz als Mehrebenenmodell spezifizieren lässt (Lüdtke, Robitzsch, Kenny & Trautwein, 2013; Lunn, Jackson, Best, Thomas & Spiegelhalter, 2012). Im SRM wird das individuelle Urteil eines Schülers A über einen Mitschüler B (z.B. die Einschätzung der Fähigkeit eines Mitschülers in Mathe) in drei Komponenten zerlegt: (i) mittlere Tendenz, mit der Schüler A die Fähigkeit anderer Schüler beurteilt (Perceiver-Effekt), (ii) mittlere Tendenz, mit der Mitschüler B von anderen Schülern als fähig beurteilt wird (Target-Effekt) und (iii) ein Effekt, der spezifisch für die Beurteilung des Mitschülers B durch Schüler A ist (Relationship-Effekt). Insgesamt werden im SRM fünf Parameter geschätzt (drei Varianzen für die Perceiver-, Target- und Relationship-Effekte sowie zwei Korrelationen).

In dem Vortrag wird eine Erweiterung des SRM für bivariate Outcomes präsentiert (Fächer Mathematik und Englisch). Das bivariate SRM umfasst insgesamt 16 Parameter (6 Varianzen und 10 Korrelationen). Es zeigt sich für die Train-Studie, dass ein Konsensus zwischen den Schülerinnen und Schülern besteht, wer von den Mitschülern gut in Mathematik und Englisch ist (15 bzw. 13% der Varianz in den individuellen Urteilen liegen zwischen den Targets; Kenny, 1994). Des Weiteren werden Schüler, die von ihren Mitschülern als gut in Mathe wahrgenommen werden auch als gut in Englisch wahrgenommen. Interessanterweise wird ein Schüler, der als besonders gut in Mathe von einem Mitschüler wahrgenommen wird, auch als besonders gut in Englisch wahrgenommen ($r = .56$).

Ausgehend von dem bivariaten SRM wird gezeigt, wie Kovariablen in das Mehrebenenmodell aufgenommen werden können. Dies ermöglicht erstens die Berechnung des von Kwan et al. (2004) vorgeschlagenen Index zur Selbstüberschätzung (siehe auch Kenny, 1994), der die Selbsteinschätzung von dem Perceiver- und Target-Effekt bereinigt (d.h., wie ein Schüler im Mittel andere Schüler wahrnimmt und wie er von den anderen im Mittel wahrgenommen wird). Zweitens können Zusammenhänge zu den Fachleistungen und anderen psycho-sozialen Outcomes untersucht werden. Es zeigt sich, dass die Target-Effekte, aber nicht die Perceiver-Effekte, hoch mit den Leistungen (Testleistung und Note) korrelieren. Des Weiteren sind die Perceiver- und Target-Effekte jeweils mittelhoch mit den Selbsteinschätzungen korreliert. Dies weist darauf hin, dass in den Selbsteinschätzungen mehrere Komponenten konfundiert sind. Implikationen für die Erfassung selbstbezogener Fähigkeitseinschätzungen im schulischen Kontext und die Forschung zur Genese akademischer Selbstkonzepte werden diskutiert.

Mittwoch
11. März

14:00 Uhr

GC 03/142

Chair(s)
Bernhardt, Raphael

Enders, Natalie

Modellierung der dimensionalen Struktur sozialer Kompetenz: Selbstberichtete Verhaltenstendenzen in Peerkonflikten im frühen Jugendalter

Soziale Kompetenz wird in der Erziehungswissenschaft und Psychologie als bedeutsame Prozess- und Ergebnisvariable für Bildung und psychosoziales Wohlbefinden angenommen. Ihre Förderung ist in allen deutschen Schulgesetzen verankert und die Europäische Union bestimmte sie als eine Schlüsselkompetenz für lebenslanges Lernen.

Wenngleich keine allgemein akzeptierte Definition des Konstrukts existiert, lässt sich effektives Verhalten in sozialen Situationen als zentrales Element gängiger Sozialkompetenzdefinitionen identifizieren. Allerdings wird die Operationalisierung dieser Grundannahme durch die Multidimensionalität sowie die Alters- und Situationsspezifität des Konstrukts erschwert. Obwohl einige theoretisch hergeleitete Modelle existieren, liegen bedauerlicherweise nur wenige empirisch hergeleitete Kompetenzstrukturmodelle vor. Das Dissertationsprojekt widmet sich der Bearbeitung dieses Forschungsdesiderats.

In Anlehnung an bildungswissenschaftliche Kompetenzmodelle wurde ein Strukturmodell sozialer Kompetenz hergeleitet. Hierzu wurde der Untersuchungsbereich exemplarisch auf Peerkonflikte zwischen gleichgeschlechtlichen Fünftklässler_innen eingegrenzt. Die Nutzung eines verhaltensanalytischen Ansatzes zur Operationalisierung des Konstruktbereichs machte drei konsekutive Untersuchungen erforderlich:

Die erste Untersuchung diente der inhaltlichen Beschreibung und Systematisierung von Konfliktsituationen. Unter Verwendung eines Fragebogens mit offenem Antwortformat wurden $n = 147$ Fünftklässler_innen um die Beschreibung einer für sie schwierig zu lösenden Konfliktsituation mit einem gleichgeschlechtlichen Peer gebeten. Die inhaltliche Analyse der Fragebögen ließ deutliche Unterschiede in den von Mädchen und Jungen berichteten Situationen erkennen, weswegen diese getrennt voneinander nach den ihnen zugrunde liegenden Kernkonflikten klassifiziert wurden. Für die Mädchen konnten sieben und für die Jungen sechs Konflikttypen bestimmt werden, die aufgrund ihrer Vorkommenshäufigkeit als am relevantesten betrachtet werden.

Die zweite Untersuchung hatte die Bestimmung und Klassifikation von Verhaltensstrategien zum Ziel. Hierzu wurden aus den Situationsbeschreibungen der Fünftklässler_innen aus der ersten Untersuchung Vignetten konstruiert, welche die relevantesten Mädchen- und Jungenkonflikte abbildeten. Diese Vignetten wurden einer weiteren Stichprobe von $n = 124$ Schüler_innen schriftlich vorgelegt. Die Kinder wurden gebeten, mögliche Verhaltensreaktionen für die Situationen aufzuschreiben. Diese Verhaltensstrategien wurden inhaltlich analysiert und zehn voneinander verschiedene Strategietypen identifiziert.

Die dritte Untersuchung diente der Modellierung der dimensionalen Struktur sozialer Kompetenz. Zur Operationalisierung des Konstrukts wurden auf der Basis der ersten beiden Untersuchungen zwei geschlechtsspezifische Fragebögen erstellt. Die zehn Verhaltensstrategien wurden in Itemform dargeboten und um ein geschlossenes Antwortformat ergänzt. Dabei waren die Mädchenitems unter sieben und die Jungenitems unter sechs Vignetten gruppiert, in denen die verschiedenen Konfliktsituationen beschrieben wurden. Die auf diese Art und Weise konstruierten Instrumente umfassten 62 Items in der Mädchen- und 53 Items in der Jungenversion.

Mittels explorativer Faktorenanalysen wurden die den Verhaltensweisen zugrunde liegenden latenten Dimensionen bestimmt. An einer weiteren Stichprobe erfolgte ein konfirmatorischer Replikationsversuch dieser dimensional Struktur. Da die Situationsspezifität sozialer Kompetenz in den Fragebögen über Vignetten repräsentiert ist, wurde anhand von konfirmatorischen Multitrait-Multioccasion-Analysen überprüft, ob ein Einfluss der Vignetten auf die unter ihnen dargebotenen Verhaltensantworten (Items) innerhalb einer Verhaltensstrategie besteht.

Die dimensionale Struktur der Fragebögen wurde an $n = 269$ Mädchen und $n = 282$ Jungen hergeleitet. Es resultierte ein dreifaktorielles Modell (prosozial, aggressiv und vermeidend) für die Mädchen und ein zweifaktorielles Modell (prosozial und aggressiv) für die Jungen. Die konfirmatorische Replikation der Modelle erfolgte an einer zusätzlichen Stichprobe von Fünftklässler_innen ($n = 283$ Mädchen und $n = 299$ Jungen). Bei beiden Geschlechtern zeigen sich sowohl für das Modell ohne als auch das Modell mit Methodenfaktor akzeptable Kennwerte von RMSEA und SRMR ($.05 < RMSEA < .06$; $.07 < SRMR < .08$), während CFI und TLI nicht ausreichend hoch ausfallen ($CFI < .97$; $NNFI < .97$). Was den Modellvergleich betrifft, fallen bei beiden Geschlechtern Chi-Quadrat-Differenzen-Test und AIC zugunsten des situationsspezifischen CFA-Modells aus, während der Vergleich der BIC gegen diese Hypothese spricht.

Die Bedeutung der Ergebnisse für empirische Zugänge zur Modellierung sozialer Kompetenz wird diskutiert.

Mittwoch
11. März

14:00 Uhr

GC 03/142

Chair(s)

Bernhardt, Raphael

Naumann, Alexander; Hartig, Johannes

Zusammenhang der Instruktionssensitivität von Tests und Items

Ergebnisse von Schülerinnen und Schülern in curriculums- oder standardbasierten Leistungsmessungen dienen regelmäßig als Basis für Rückschlüsse über Unterricht (Hascher & Schmitz, 2010). Jedoch kann der Zusammenhang von Testleistung und Unterricht aufgrund der Eigenschaften eines Tests stark variieren (Grossman, Cohen, Ronfeldt, & Brown, 2014), das heißt, die eingesetzten Instrumente können unterschiedlich stark dazu in der Lage sein, Effekte des Unterrichts aufzufangen, also unterschiedlich instruktionssensitiv sein. Jedoch wurde bislang wenig untersucht, inwiefern die Items die Instruktionssensitivität eines Tests beeinflussen.

Instruktionssensitivität ist definiert als die psychometrische Eigenschaft eines Tests oder eines einzelnen Items, Effekte des Unterrichts auffangen zu können (Polikoff, 2010). Jedoch ist der Zusammenhang der Instruktionssensitivität von Tests und Items nach wie vor unklar. Unsere Studie zielt darauf ab, den Einfluss der Instruktionssensitivität der Items eines Tests auf die Instruktionssensitivität des Tests als Ganzen zu untersuchen. Wir vermuten, dass ein Test umso instruktionssensitiver erscheint, je höher die mittlere Instruktionssensitivität der Items ausgeprägt ist.

Die Instruktionssensitivität eines Tests wird üblicherweise durch die Regression von Testscores auf Unterrichtsmerkmale untersucht (z.B. Ing, 2008). Je höher der Zusammenhang von Testscore und Unterrichtsmerkmal, umso höher die Instruktionssensitivität des Tests. Dagegen fokussieren Untersuchungen der Instruktionssensitivität einzelner Items die Itemschwierigkeit und Trennschärfe (Haladyna, 2004). Naumann, Hochweber und Hartig (2013) schlugen ein längsschnittliches Mehrebenen-IRT-Modell vor, das bestehende Ansätze integriert, um klassenspezifische Veränderungen von Itemschwierigkeiten über Messzeitpunkte zu schätzen. Die klassenspezifischen Itemschwierigkeitsveränderungen werden als normalverteilt angenommen, wobei Mittelwert und Varianz als statistische Indikatoren für die Sensitivität eines Items dienen. Das heißt, es werden zwei Arten von Sensitivität unterschieden, (a) globale Sensitivität, die angibt, inwiefern sich die Schwierigkeit eines Items über Klassen innerhalb einer Stichprobe im Mittel über die Zeit verändert, und (b) differenzielle Sensitivität, die angibt, inwiefern die Veränderung der Itemschwierigkeit über Klassen hinweg variiert. Zudem kann die Verteilung der Itemschwierigkeitsveränderung um Regressionsterme ergänzt werden, so dass Unterrichtsmerkmale als Prädiktoren für die klassenspezifischen Itemschwierigkeitsveränderungen herangezogen werden können, um die differenzielle Sensitivität der Items zu erklären.

Auf Basis dieses IRT-Modells wurden dichotome Schülerantworten auf jeweils 25 Items für zwei Messzeitpunkte mittels R simuliert. Dazu werden Fähigkeitsparameter für $N=2000$ Individuen in 100 Klassen aus einer Standardnormalverteilung gezogen. Für den ersten Messzeitpunkt werden Itemschwierigkeiten so gewählt, dass sie eine Range von 3 Logits um die mittlere Fähigkeit der Stichprobe abdecken. Die Items sind in allen Klassen gleichermaßen schwierig. Für den zweiten Messzeitpunkt wird angenommen, dass alle Items im Mittel gleichermaßen leichter werden (identische globale Sensitivität), die Veränderung jedoch über Klassen hinweg variiert (varierende differenzielle Sensitivität). Zusätzlich wird die Itemschwierigkeitsveränderung durch eine standardnormalverteilte Klassenvariable Z beeinflusst. Der Einfluss von Z auf die klassenspezifische Schwierigkeitsveränderung variiert dabei über die 15 Items hinweg zufällig im Intervall $[-1, 1]$. Anschließend werden die Schülerantworten jedes Messzeitpunkts mittels eines einfachen Rasch-Modells skaliert. Die Fähigkeitsparameter für T_2 werden mehrebenenregressionsanalytisch auf das Klassenmerkmal Z sowie die Fähigkeit zum ersten Messzeitpunkt regediert.

Über 1000 Replikationen hinweg wurde der Zusammenhang des Einflusses von Z auf die Testleistung zu T_2 (Instruktionssensitivität auf Testebene) mit dem mittleren Einfluss von Z auf die klassenspezifischen Itemschwierigkeitsveränderungen (differenzielle Instruktionssensitivität auf Itemebene) untersucht. Die Ergebnisse unterstützten unsere Hypothese. Ein Test scheint umso instruktionssensitiver hinsichtlich eines Klassenmerkmals zu sein, je höher die mittlere Instruktionssensitivität der Items hinsichtlich desselben Merkmals ausgeprägt ist ($r=.94$, $p<0.001$).

Zusammenfassend unterstreichen die Ergebnisse die Wichtigkeit, die Instruktionssensitivität der Items eines Tests zu beachten. Nicht nur können Tests unterschiedlich instruktionssensitiv hinsichtlich desselben Merkmals sein, sondern das Ausmaß ihrer Instruktionssensitivität hängt maßgeblich mit den Testitems zusammen. In der Testkonstruktion scheinen daher Informationen über die differenzielle Instruktionssensitivität von Items nützlich, wenn ein Test als Basis für Rückschlüsse über Unterricht dienen soll. Nach wie vor ist jedoch weitere Forschung zum Einfluss der globalen Sensitivität von Items auf die Instruktionssensitivität von Test erforderlich.

Mittwoch
11. März

14:00 Uhr

GC 03/142

Chair(s)

Bernhardt, Raphael

Niedermeyer, Inga; Jordan, Anne-Katrin; Heinze, Aiso; Grüßing, Meike; von Seeler, Torben; Rogalski, Karin
Mathe macht stark – Erste Evaluationsergebnisse eines Programms zur Förderung leistungsschwacher Schülerinnen und Schüler im Anfangsunterricht

In der TIMSS 2011-Studie (Bos, Wendt, Köller & Selter 2012) zeigte sich, dass ein Fünftel der Schülerinnen und Schüler aus Deutschland am Ende der Grundschulzeit nur die unteren beiden Kompetenzstufen erreicht, d.h. nur über rudimentäre mathematische Kompetenzen verfügt. Bekannt ist, dass ein substanzieller Anteil der Varianz der Mathematikleistungen zu diesem Zeitpunkt durch mathematische Fähigkeiten im Vorschulalter erklärt werden kann (Krajewski & Schneider, 2006). Für das Bildungssystem stellt sich entsprechend die Frage der Implementierung von effektiven Fördermaßnahmen, die bereits ab Beginn der Grundschulzeit zur Diagnose und Förderung eingesetzt werden können.

Das Programm „Mathe Macht Stark - Grundschule“ des Instituts für Qualitätsentwicklung an Schulen Schleswig-Holstein (IQSH) reagiert auf diese Problematik und umfasst ein kohärentes Unterstützungsprogramm für Grundschulen bestehend aus Materialien, zusätzlichen Lehrerstunden, Fortbildungen und Beratungsangeboten. Ziel des Programms ist die systematische unterrichtsbegleitende Diagnose und Förderung grundlegender arithmetischer Kompetenzen von leistungsschwachen Schülerinnen und Schülern in den ersten beiden Grundschuljahren. Lernschwierigkeiten in Mathematik sollen so erkannt und der Anteil von Kindern mit geringen mathematischen Kompetenzen reduziert werden.

Um die Wirksamkeit des Programms zu überprüfen, werden die Leistungen der Schülerinnen und Schüler von 40 Grundschulen in Schleswig-Holstein bis zum Ende der Grundschulzeit verfolgt, so dass einerseits die Nachhaltigkeit der Förderung untersucht und andererseits eine kriteriale Überprüfung hinsichtlich der Bildungsstandards Mathematik vorgenommen werden kann. Es wurde ein quasiexperimentelles Design mit drei Gruppen gewählt: 20 Schulen erhalten neben dem Material und dazugehörigen Fortbildungen zwei zusätzliche Lehrerwochenstunden, um die Förderung umzusetzen (Gruppe 1). Zehn Schulen bekommen die Materialien und die Fortbildungen, aber keine zusätzlichen Lehrerwochenstunden (Gruppe 2) und zehn weitere Schulen außerhalb des Programms setzen die in Schleswig-Holstein üblicherweise vorgesehene Förderung von leistungsschwachen Schülerinnen und Schülern um und dienen als Kontrollgruppe (Gruppe 3). Die Schulen wurden so ausgewählt, dass sie im Hinblick auf die soziale, kulturelle und leistungsbezogene Zusammensetzung der Schülerschaft (z.B. SES, Migrationsanteil, Anzahl Inklusionskinder) vergleichbar sind. Insgesamt sind ca. 2500 Schülerinnen und Schüler aus 120 Klassen beteiligt.

Im Verlauf des ersten Schuljahres wurden fünf schriftliche Kurztests zu spezifischen Inhaltsbereichen des Arithmetikunterrichts eingesetzt. Darüber hinaus wurden Intelligenz (CFT) sowie Lerneingangsvoraussetzungen in den Bereichen Mathematik (Hamburger Rechentest) und Sprache (Münsteraner Screening) erfasst, um auf dieser Grundlage über Matching-Verfahren eine faire Vergleichbarkeit zu gewährleisten. Jeweils am Ende des Schuljahres wird eine adaptierte Version des Mathematiktests des Projekts „Persönlichkeits- und Lernentwicklung von Grundschulkindern“ (Lipowsky, Faust & Karst, 2011) eingesetzt. Diese Tests erfassen die Leistungsentwicklung über das gesamte Schuljahr und sind über die Klassenstufen verankert, wodurch eine längsschnittliche Skalierung möglich ist. Die Intensität der individuellen Förderung wird jährlich über Lehrerfragebögen erfasst.

Auf Basis der Daten aus dem ersten Schuljahr, soll es im Vortrag zum einen um die Frage gehen, ob es Unterschiede zwischen den drei Gruppen in der Kompetenzentwicklung im Bereich Arithmetik gibt bzw. welche Faktoren diese Kompetenzentwicklung beeinflussen. Zum anderen soll die Wirkung speziell für diejenigen Kinder, die im Rahmen des Programms intensiv gefördert wurden, untersucht werden.

In Strukturgleichungsmodellen zeigt sich unter Berücksichtigung der Mehrebenenstruktur der Daten und der Hintergrundvariablen, dass nach dem ersten Schuljahr bei Betrachtung aller Schülerinnen und Schüler keine signifikanten Unterschiede zwischen den drei Gruppen deutlich werden. Um nur die Schülerinnen und Schüler zu untersuchen, die auch tatsächlich im Rahmen des Programms gefördert wurden, wird mit Hilfe des propensity score matchings (Ho, Imai, King & Stuart, 2011) eine geeignete Vergleichsgruppe ermittelt. Die Ergebnisse werden bis zum Vortrag vorliegen und dort vorgestellt.

Mittwoch
11. März

14:00 Uhr

GC 03/146

Lindmeier, Anke; Grüßing, Meike; Heinze, Aiso

„Warum ist das so?“ – Vorläuferfähigkeiten für mathematisches Argumentieren im Kindergarten erfassen

Mathematisches Argumentieren ist eine der fünf allgemeinen mathematischen Kompetenzen in den Bildungsstandards Mathematik und damit ein wesentliches Ziel der Grundschule (KMK, 2004). Neben der selbstständigen Prüfung mathematischer Aussagen (im engeren Sinne mathematisches Argumentieren) umfasst dies auch Vorläuferfähigkeiten (z.B. mathematische Zusammenhänge erkennen, Begründungen nachvollziehen). In den amerikanischen Bildungsstandards werden diese Vorläuferfähigkeiten bereits im Vorschulbereich thematisiert (NCTM, 2000). Hintergrund ist, dass dem Argumentieren im Unterricht als instruktionalem Merkmal auch eine spezielle Rolle beim Erwerb konzeptuellen Verständnisses zugesprochen wird (Kuhn, 1993). Für die Situation in Deutschland ist unklar, ob und in welchem Umfang Kinder zu Beginn der Grundschulzeit die Kompetenz des mathematischen Argumentierens bereits ausgebildet haben. Aufgrund der vergleichsweise gering ausgeprägten mathematischen Lerngelegenheiten im Kindergarten in Deutschland dürften hier eher die o.g. Vorläuferfähigkeiten, denn mathematisches Argumentieren im engeren Sinne erwartet werden.

Ziel unserer Studie ist die Modellierung und Messung der Vorläuferfähigkeiten zum mathematischen Argumentieren von Kindern im letzten Kindergartenjahr. Neben der Beschreibung geeigneter Teilaspekte und deren Operationalisierung durch materialbasierte Interviews stellt sich die Frage nach der Eigenständigkeit und Ausprägung dieser Fähigkeiten. Dies führt zu den Forschungsfragen (1) über welche (Vorläufer-)Fähigkeiten zum mathematischen Argumentieren Kinder im letzten Kindergartenjahr verfügen und (2) ob diese Fähigkeiten von mathematischem Wissen sowie der Arbeitsgedächtnisleistung als Indikator für kognitive Grundfähigkeiten differenziert werden können.

In der vorliegenden Studie wurden ein Test zu (Vorläufer-)Fähigkeiten zum mathematischen Argumentieren entwickelt und bei $N = 120$ Kindern im Alter von $M = 70,6$ Monaten ($SD = 5,5$ Monate) eingesetzt. Der Test zu (Vorläufer-)Fähigkeiten zum mathematischen Argumentieren umfasste 14 Items mit einer internen Konsistenz von $.69$ (Cronbachs α), was noch als zufriedenstellend bewertet wurde. Mit einer mittleren Lösungsrate von $M = .28$ ($SD = .19$) ist der Test für die Gesamtstichprobe als eher schwer einzustufen.

Zur differenzierten Analyse der Testcharakteristika wurde bei einer Teilstichprobe von $N = 82$ Kindern ($M = 68,9$ Monate, $SD = 4,3$ Monate) an einem weiteren Testtag die Arbeitsgedächtnisleistung (HAWIK IV, Subtest „Zahlen nachsprechen“, Petermann & Petermann, 2008) sowie mathematisches Wissen (Eigenentwicklung, 17 Items, Cronbachs $\alpha = .71$) erhoben. Der Test zum mathematischen Wissen war ebenfalls als materialbasiertes Interview gestaltet und umfasste genau die mathematischen Inhalte, die im Argumentationstest genutzt werden sollten. Zur Kontrolle von sprachlichen Kompetenzen wurde ein bildbasierter Sprachindikator verwendet, um sprachauffällige Kinder identifizieren zu können. Insgesamt wurden 7 Kinder aus der Auswertung ausgeschlossen, da diese bereits in nicht-mathematischen Alltagssituationen keine einfache Implikation wiedergeben konnten. Die verbliebenen 75 Kinder unterscheiden sich im Argumentationstest nicht von der Gesamtstichprobe ($M = .26$; $SD = .18$) und erzielten im Test zum mathematischen Wissen im Mittel eine Lösungsrate von $M = .58$ ($SD = .16$).

Die (Vorläufer-)Fähigkeiten zum mathematischen Argumentieren korrelieren mit dem mathematischen Wissen ($r = .58$, $p < .001$) bzw. der Arbeitsgedächtnisleistung ($r = .29$, $p = .01$), jedoch nicht mit dem Alter der Kinder ($r = .04$, $p = .75$). Im Regressionsmodell kann nur das mathematische Wissen als signifikante Einflussgröße bestätigt werden, die ca. 34% der Varianz der (Vorläufer-)Fähigkeiten zum mathematischen Argumentieren erklären kann, sodass diese Konstrukte entsprechend als unabhängig bewertet werden können. Auch wenn in dem Argumentationstest im Wesentlichen Vorläuferfähigkeiten im Sinne des Herstellens von Zusammenhängen bzw. des Nachvollziehens von Begründungen abgebildet wurde, so wird deutlich, dass es sich bei diesem Konstrukt um mehr als mathematisches Begriffswissen handelt. In einer Folgestudie soll die Entwicklung dieser Fähigkeiten im Verlauf des letzten Kindergartenjahres untersucht werden, um ihre Stabilität bzw. Variabilität im Übergang zur Grundschulzeit einschätzen zu können. Darüber hinaus stellt sich die Frage nach der spezifischen Prädiktionskraft für die mathematische Argumentationskompetenz im Grundschulalter.

Mittwoch
11. März

14:00 Uhr

GC 03/146

Dumont, Hanna; Leonie, Mach; Juliane, Rothe

Entwicklung und Validierung eines Fragebogens zur Erfassung des Grads an individueller Förderung in der Grundschule

In der derzeitigen Debatte um die zunehmende Heterogenität der Schülerschaft spielt die Frage, wie Lehrkräfte den heterogenen Lernausgangsvoraussetzungen von Schülerinnen und Schülern im Unterricht gerecht werden können, eine große Rolle. In diesem Zusammenhang wird häufig von der Notwendigkeit einer individuellen Förderung der einzelnen Schülerinnen und Schüler gesprochen. Das Konzept der individuellen Förderung erscheint dabei „in der öffentlichen Diskussion nahezu unhinterfragt als Schlüssel zur Lösung bildungspolitischer und pädagogischer Probleme“ (Klieme & Warwas, 2011, S. 805). Obwohl die Idee der Anpassung des Unterrichts an die individuellen Bedürfnisse der Schülerinnen und Schüler keineswegs neu ist – in der Pädagogischen Psychologie im Rahmen der Aptitude-Treatment-Interaction-Forschung unter dem Stichwort Adaptive Teaching diskutiert (Cronbach, 1967), in der Erziehungswissenschaft seit langem in reformpädagogischen Ansätzen und dem Konzept der Inneren Differenzierung enthalten (Trautmann & Wischer, 2008) – liegen bislang kaum empirische Studien zur Verbreitung von individueller Förderung im deutschen Schulsystem sowie zur Wirksamkeit vor.

Die empirische Erfassung von individueller Förderung im Unterricht wird möglicherweise dadurch erschwert, dass es sich dabei um kein klar umrissenes Konzept der Unterrichtsgestaltung handelt, sondern vielmehr einen Sammelbegriff für eine Vielzahl unterschiedlicher Maßnahmen und Strategien im Unterricht darstellt (Fischer, 2014). Vor diesem Hintergrund wird im vorliegenden Beitrag der Frage nachgegangen, ob es möglich ist, individuelle Förderung als ein Tiefenmerkmal von Unterricht, d.h. unabhängig von den konkret eingesetzten Methoden, durch Schülerangaben zu erfassen – wie es bereits für andere Dimensionen von Unterrichtsqualität üblich ist (vgl. Lüdtke, Trautwein, Kunter & Baumert, 2006; Helmke, 2010). Dazu wird die Entwicklung und Validierung eines Fragebogens zur Erfassung des Grads der individuellen Förderung im Unterricht vorgestellt. Fokussiert wurde dabei die Grundschule, da Maßnahmen der individuellen Förderung aufgrund der heterogeneren Schülerschaft hier häufiger auftreten sollten als in der Sekundarstufe und es aus jüngerer Zeit Hinweise darauf gibt, dass die Erfassung von Unterrichtsqualität durch Angaben von Grundschülerinnen und -schülern möglich ist (Fauth, Decristan, Rieser, Klieme & Büttner, 2013).

Für die Entwicklung des Fragebogens wurde zunächst ein Itempool auf der Basis von Items aus bisherigen Studien sowie selbst entwickelten Items erstellt (Beispielitems: „Verschiedene Schüler arbeiten mit unterschiedlichen Büchern und Materialien“, Beantwortung auf einer Likert-Skala mit 1 = stimmt gar nicht, 2 = stimmt eher nicht, 3 = stimmt eher, 4 = stimmt genau). Im Anschluss daran wurden im Rahmen von Experteninterviews mit neun Lehrkräften aus Schulen, in deren Schulprogramm individuelle Förderung verankert war, 20 Items ausgewählt. Diese Items wurden gemeinsam mit bereits etablierten Skalen zur Erfassung der Unterrichtsqualität in einer Stichprobe von 479 Schülerinnen und Schülern der 3. und 4. Jahrgangsstufe in 41 Lerngruppen an neun Schulen in Nordrhein-Westfalen und Baden-Württemberg eingesetzt. Bei fünf der ausgewählten Schulen handelte es sich um Schulen, die sich durch ein Konzept der individuellen Förderung im Unterricht auszeichneten.

Nach einer umfassenden Item- und Reliabilitätsanalyse konnten neun der eingesetzten 20 Items zu einer reliablen Skala zusammengefasst werden (Trennschärfen $> .50$, Cronbach's $\alpha = .86$, ICC (1) = .71, ICC (2) = .97). Für die Validität der Skala ließen sich auch deutliche Hinweise finden. So lag der Mittelwert der Skala in den Schulen mit Schwerpunkt individueller Förderung mit $M = 2.93$ ($SD = .64$) wesentlich höher als in den Regelschulen mit $M = 1.91$ ($SD = .45$), was einer standardisierten Mittelwertsdifferenz von $d = 1.83$ entspricht. Im Gegensatz dazu unterschieden sich die beiden Schultypen auf den anderen Unterrichtsskalen deutlich weniger ($d = .19$ bis $d = .59$). Darüber hinaus zeigten sich theoretisch plausible Korrelationsmuster mit den bisherigen Unterrichtsskalen, so korrelierte beispielsweise die neu entwickelte Skala negativ mit Rigidität und Strukturierung und stark positiv mit der wahrgenommenen Autonomieunterstützung. Der entwickelte Fragebogen erscheint somit für bislang fehlende empirische Studien zur Wirksamkeit von individueller Förderung geeignet.

Mittwoch
11. März

14:00 Uhr

GC 03/146

Chair(s)

Niedermeyer, Inga

Mathematik und individuelle Förderung im Kindergarten und in der Grundschule

Blaurock, Sabine

Außerschulische Bildung von Erstklässlern

Organisierte Freizeitaktivitäten sind in der aktuellen Forschungsdiskussion als Bereich der außerschulischen Bildung in den Fokus des Interesses gerückt. Fähigkeiten, wie z.B. soziale Kompetenzen, gehen auf Aneignungsprozesse im Rahmen der Freizeitaktivitäten zurück (Metsäpelto & Pulkkinen, 2014). Ihnen wird eine hohe Bedeutung für eine gelingende Schulkarriere zugeschrieben. Dass überwiegend privat finanzierte Aktivitäten von der sozialen Herkunft der Kinder abhängen, konnte in einigen Studien gezeigt werden (z.B. Mühler & Spieß, 2008). Welche Rolle dabei die elterlichen Förderorientierungen und die häusliche Anregungsqualität bereits in der frühen Kindheit spielen, ist bisher unbekannt.

Der vorliegende Beitrag fragt danach, welche Faktoren der frühen häuslichen Lernumwelt die Intensität der außerschulischen Aktivitäten von Erstklässlern bestimmen.

Das Qualitätsmodell der häuslichen Lernumwelt (HLE), das seine theoretische Fundierung aus einer sozialökologischen Perspektive bezieht (Bronfenbrenner & Morris, 2006), stellt den konzeptuellen Rahmen dar. Die HLE besteht demnach aus einem mehrdimensionalen Konstrukt aus Struktur-, Orientierungs- und Prozessmerkmalen. Strukturmerkmale umfassen relativ beständige und überdauernde Rahmenbedingungen in einer Familie, wie den Bildungsstand der Eltern und den sozioökonomischen Status. Unter Orientierungen werden generelle Werte und Einstellungen bezüglich der Bildung und kindlichen Entwicklung verstanden. Die Prozessmerkmale beinhalten alle unmittelbaren Merkmale der Interaktionen des Kindes mit seiner sozialen und materiellen Umwelt. Für Struktur- Orientierungs- und Prozessmerkmale wird ein Beziehungsmuster angenommen, wobei die Prozesse als insbesondere bedeutsam für die Vorhersage der kindlichen Kompetenzen angesehen werden (Tietze, Roßbach & Grenner, 2005).

Datengrundlage ist die Längsschnittstudie „Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter“ (BiKS) 3-10, in der 547 Kinder über die gesamte Kindergarten- und Grundschulzeit hinweg wissenschaftlich begleitet wurden.

Informationen über familiäre Hintergrundmerkmale und die elterlichen Förderorientierungen konnten durch Befragungen der Eltern gewonnen werden. Die Prozessmerkmale wurden während der gesamten Kindergartenzeit jährlich erfasst. Die eher globalen Aspekte wurden in der Kindergartenzeit mit der deutschen Version der Home Observation for Measurement of the Environment - Early Childhood (HOME-EC; Caldwell & Bradley, 1984) erhoben. Außerdem kam ein selbst entwickeltes Live-Rating, die Familieneinschätzungsskala (FES-Kiga; Kuger, Pflieger & Roßbach, 2005), zur Erfassung des elterlichen Unterstützungsverhaltens in einer Eltern-Kind-Interaktion im häuslichen Umfeld zum Einsatz. Für die Kindergartenzeit wurden drei instrumentenübergreifende Skalen aus der FES-Kiga und der HOME-EC in Anlehnung an das Rahmenmodell von Klieme et al. (2006) gebildet. Globale Anregungsaspekte wurden unter der Skala „Allgemeines Klima“ und domänenspezifische unter den Skalen „Anregung in Literacy“ und „Anregung in Numeracy“ zusammengefasst.

Die außerfamilialen Aktivitäten wurden im ersten Grundschuljahr in einem Elternfragebogen erfasst. Die Eltern konnten angeben, wie häufig Ihr Kind an Vereinsaktivitäten bzw. außerschulischem Unterricht, z.B. Sportunterricht, Theatergruppe und Fremdsprachenunterricht, teilnimmt. Für die Analysen wurde aus den sieben vorgegebenen Aktivitäten ein additiver Index gebildet.

In die hierarchischen linearen Regressionsmodelle wurden die Struktur-, Orientierungs- und Prozessmerkmale der Familie schrittweise zur Vorhersage der Inanspruchnahme außerfamilialer Aktivitäten aufgenommen.

Im Ergebnis zeigt sich einerseits, dass familiäre Hintergrundmerkmale, wie Schulbildung und Einkommen der Eltern an Bedeutung für das Ausmaß bildungsrelevanter Aktivitäten verlieren, wenn die schulvorbereitende Förderorientierung und die familiäre Anregungsqualität in der frühen Kindheit berücksichtigt werden. Andererseits besitzt die familiäre Anregung, vor allem im Bereich „Anregung in Literacy“, zwar die stärkste Erklärungskraft. Der Migrationshintergrund einer Familie, die Anzahl der Kinder in einem Haushalt sowie eine schulvorbereitende Förderorientierung der Eltern behalten jedoch ihre eigenständigen Effekte. Eine chancengerechte Inanspruchnahme einer außerschulischen Bildung von Erstklässlern kann schon vor dem Schulbeginn, z.B. durch Familienbildungsangebote zur Förderung der Eltern-Kind-Interaktion, beeinflusst werden.

Mittwoch
11. März

14:00 Uhr

GC 03/146

Resa, Elisabeth; Turani, Daniel; Wieduwilt, Nadine; Tuffentsammer, Mathias

Welche Rolle spielen Fortbildungen für die realisierte pädagogische Prozessqualität? Ergebnisse der wissenschaftlichen Evaluation des Bundesprogramms „Schwerpunkt-Kitas Sprache & Integration“

Theoretischer Hintergrund: Der erfolgreiche Erwerb von Sprache ist Voraussetzung für das Wahrnehmen von Bildungsangeboten und gesellschaftliche Teilhabe. In Deutschland ergeben sich vor allem für Kinder aus bildungsfernen Familien und Kinder mit Migrationshintergrund bereits beim Eintritt in die Grundschule Nachteile (Dubowy et al., 2008; Weinert et al., 2010), die vorrangig durch mangelnde Sprachkenntnisse begründet sind. Vor diesem Hintergrund fördert das BMFSFJ im Rahmen des Bundesprogramms „Schwerpunkt-Kitas Sprache & Integration“ Kindertageseinrichtungen, um sie bei der Aufgabe der alltagsintegrierten sprachlichen Bildung zu unterstützen. Im Rahmen des Bundesprogramms nimmt die Weiterqualifizierung der Kita-Teams durch interne und externe Fortbildungen eine zentrale Rolle ein.

In der internationalen Forschung zur frühpädagogischen Qualität wird davon ausgegangen, dass Aspekte der Strukturqualität und Orientierungsqualität die pädagogische Prozessqualität beeinflussen, welche wiederum Einfluss auf die kindliche Entwicklung hat (Kuger & Kluczniok, 2008; Pianta et al., 2005). Als ein Aspekt der Orientierungsqualität auf Einrichtungsebene wird die Fortbildungsaktivität des Teams verstanden, für die bisher vergleichsweise wenige Erkenntnisse in Bezug auf die Zusammenhänge zur pädagogischen Prozessqualität vorliegen. Auf der Grundlage eines Modells professioneller Kompetenzen (Anders, 2012), welches verschiedene Kompetenzfacetten pädagogischer Fachkräfte differenziert und professionelle Kompetenzen als grundsätzlich ansprechbar und veränderbar beschreibt, muss davon ausgegangen werden, dass (teamorientierte) Fortbildungsaktivitäten eine zentrale Bedeutung in der Weiterqualifizierung pädagogischer Fachkräfte einnehmen. Erste Ergebnisse der Evaluation des Bundesprogramms lieferten empirische Evidenz für die Bedeutsamkeit der Fortbildungsaktivität für die pädagogische Prozessqualität.

Fragestellung: Vor diesem Hintergrund wird in der aktuellen Studie untersucht (1) ob sich verglichen mit dem ersten Messzeitpunkt Veränderungen in der Fortbildungsaktivität der Einrichtungen zum Messzeitpunkt zwei zeigen, (2) ob sich eine zunehmende Fortbildungsaktivität der Einrichtungen in einer verbesserten pädagogischen Prozessqualität niederschlägt und (3) wie Fortbildungen gestaltet sein müssen, damit sie zur Qualitätsentwicklung der Einrichtungen beitragen. Im Rahmen der Evaluationsstudie werden folgende Stichprobengruppen miteinander verglichen:

- zwei Gruppen von Einrichtungen, die am Programm teilnehmen und jeweils eine unterschiedliche Form zusätzlicher fachlicher Unterstützung erhalten,
- eine Gruppe von Einrichtungen, die ohne zusätzliche fachliche Unterstützung am Programm beteiligt ist sowie
- eine Vergleichsgruppe von Einrichtungen, die nicht am Bundesprogramm teilnimmt.

An der Evaluation beteiligten sich zum ersten Messzeitpunkt insgesamt 335 Einrichtungen; zum Messzeitpunkt zwei konnten 322 Einrichtungen erneut erhoben werden.

Methode: Die Evaluation ist als quasi-experimentelle Längsschnittstudie mit drei Messzeitpunkten angelegt und verfolgt einen multimethodischen Ansatz. Die vorliegenden Ergebnisse basieren auf Daten des ersten und zweiten Messzeitpunktes, die durch Online-Befragungen pädagogischer Fachkräfte sowie Qualitätsbeobachtungen und Interviews in den Einrichtungen erhoben wurden. Zur Feststellung der sprachförderlichen Prozessqualität wurden Teile der KES-E (Roßbach & Tietze, in Vorbereitung), KES-R und KRIPS-R (Tietze et al., 2007a; 2007b), der DO-RESI (Fried & Briedigkeit, 2008) sowie eine deutsche Version der TILRS (Girolametto et al., 2000) eingesetzt. Für die Beantwortung der Fragestellungen werden deskriptive Auswertungsformen und regressionsanalytische Verfahren genutzt.

Ergebnisse: Hinsichtlich der internen Fortbildungsaktivität der Programmgruppen haben sich zum ersten Messzeitpunkt signifikante Unterschiede zwischen den Gruppen ($p < 0.05$) gezeigt. Anhand der Daten des ersten Messzeitpunktes konnte zudem ein signifikant positiver Zusammenhang ($r = .19^{**}$, $p < 0.01$) zwischen der internen Fortbildungsaktivität und der pädagogischen Prozessqualität (KES-E) nachgewiesen werden. Darüber hinaus hat sich gezeigt, dass der Einfluss der Untersuchungsgruppe auf die pädagogische Prozessqualität teilweise durch die interne Fortbildungsaktivität vermittelt wird. Für den zweiten Messzeitpunkt ist ein signifikanter Anstieg ($p < 0.05$) der Anzahl interner Fortbildungen bei einer der drei Programmgruppen zu verzeichnen, wohingegen bei den anderen beiden Programmgruppen nur geringe Veränderungen zu verzeichnen sind. Aufbauend auf diesen Ergebnissen werden die Daten aus Messzeitpunkt zwei vertiefend analysiert.

Mittwoch
11. März

14:00 Uhr

GC 03/146

Sari, Duygu; Zinn, Bernd

Die Heterogenität der epistemologischen Überzeugungen von Auszubildenden der Metalltechnik und ihre Effekte auf die Fachkompetenz

Theoretischer Hintergrund: Dass die epistemologischen Überzeugungen von Lernenden einen Einfluss auf die Lernleistung nehmen, wurde für den allgemein bildenden Bereich schon mehrfach belegt. In mehreren Studien konnte gezeigt werden, dass diese handlungswirksam sind und entwickelte Überzeugungen zu Wissen und Wissenserwerb positiv mit den schulischen Leistungen einhergehen (vgl. z.B. Hofer & Pintrich 1997; Trautwein & Lüdtke 2008). In der beruflichen Kompetenzforschung ist hingegen wenig bekannt, insbesondere fehlen systematische Untersuchungen zum Ausgangskonstrukt im Zusammenhang mit der Fachkompetenz. Dass die epistemologischen Überzeugungen im beruflichen Handeln gegenüber anderen Handlungszusammenhängen im menschlichen Dasein eine besondere Charakteristik aufweisen und daher auch domänenspezifisch betrachtet werden müssen, kann grundlegend angenommen werden (vgl. Buehl & Alexander 2006; Zinn 2013).

Fragestellung: Ziel der empirischen Studie ist es, die epistemologischen Überzeugungen von gewerblich-technischen Auszubildenden der Metalltechnik empirisch aufzuklären. Im Fokus der Studie stehen (1) die Begründung eines systematischen Beschreibungswissens zum Entwicklungsstand der epistemologischen Überzeugungen von Auszubildenden im Berufsfeld Metalltechnik, (2) die Optimierung des Instruments WÜGTA zur Erfassung der domänenspezifischen epistemologischen Überzeugungen von Auszubildenden und (3) die Untersuchung, inwiefern die Überzeugungen zu Wissen und Wissenserwerb von Auszubildenden eine prognostische Validität für die metalltechnische Fachkompetenz am Ende des ersten Ausbildungsjahres besitzen.

Methode: Zur Untersuchung der Forschungsfragen wurden Auszubildende in der einjährigen Berufsfachschule Metalltechnik (n=356) zu den epistemologischen Überzeugungen, dem fachspezifischen Vorwissen, der mathematischen Fähigkeiten, der kognitiven Leistungsfähigkeit, dem Fähigkeitsselbstkonzept und dem Berufsinteresse befragt. Die Studie wurde im Schuljahr 2013/14 durchgeführt und umfasst eine Eingangserhebung zum Ausbildungsbeginn und eine Abschlusstestung zum Ende des ersten Ausbildungsjahres. Es wurden standardisierte quantitative Messinstrumente eingesetzt und die Analyse erfolgte mittels deskriptiver Analysen und multivariater statistischer Modellierungen.

Ergebnisse: Die Studienergebnisse zeigen mehrere intra- und interindividuelle Unterschiede in Abhängigkeit der betrachteten Lernermerkmale auf und belegen einen bedeutungsvollen Zusammenhang zwischen dem Entwicklungsstand der epistemologischen Überzeugungen und dem fachspezifischen Wissen von Auszubildenden. Zudem liefert die Studie ein valides Instrument (WÜGTA II) zur Erfassung der domänenspezifischen epistemologischen Überzeugungen von Auszubildenden der Metalltechnik. Die Forschungsergebnisse unterstützen den Aufbau eines systematischen Erklärungs- und Beschreibungswissens zu den epistemologischen Überzeugungen von Auszubildenden und liefern Ansatzpunkte für eine evidenzbasierte Förderung der domänenspezifischen Überzeugungen zum Wissen und Wissenserwerb im Kontext der Genese beruflicher Handlungskompetenz.

Mittwoch
11. März

14:00 Uhr

GC 03/149

Chair(s)
Sari, Duygu

Föste, Dennis; Gabriel, Nagy; Jan, Retelsdorf

Drop-out-Determinanten in (dualen) mathematisch-naturwissenschaftlichen Berufsausbildungen

Formale Ausbildungsabschlüsse sind in Deutschland von langfristiger Bedeutung für berufliche und soziale Integration von Jugendlichen (Konietzka, 2010; Böhnke, 2005). Die Lösung des Ausbildungs-vertrags stellt somit bezogen auf die Lebens- und Aufstiegschancen von Jugendlichen ein kritisches Ereignis dar. Ist eine bestimmte soziale Gruppe hiervon verstärkt betroffen, trägt dies zur (Re-) Produktion sozialer Ungleichheit bei.

Aus der Perspektive betroffener Betriebe kommen Vertragslösungen zudem zumeist Ressourcenverlusten gleich. Umso mehr gilt dies für Betriebe, die für den eigenen Fachkräftebedarf ausbilden. Vor dem Hintergrund des sich abzeichnenden Mangels an beruflich qualifizierten Fachkräften im MINT-Segment (MINT=Mathematik, Informatik, Naturwissenschaften & Technik) (Anger et al., 2013) erscheinen vorzeitige Lösungen von Ausbildungsverträgen in diesem Segment besonders folgenreich. Unser Fokus richtet sich daher auf die Bedingungsfaktoren von Vertragslösungen in (dualen) mathematisch-naturwissenschaftlichen Berufsausbildungen.

In der Literatur zu den Ursachen von vorzeitigen Vertragslösungen finden sich nur wenige breit angelegten quantitativen Studien (Beicht & Walden, 2013: 2) und keine Panelerhebungen (BiBB, 2014: 162). Eine aktuelle multivariate Analyse weist Migrationshintergrund, Geschlecht (jeweils gerichtete Effekte) sowie eine ‚Diskrepanz zwischen Wunsch- und Ausbildungsberuf‘ als Risikofaktoren für einen Ausbildungsabbruch aus (Beicht & Walden, 2013).

Übertagen auf das MINT-Segment ergeben sich hieraus folgende Fragestellungen: Erweist sich auch im Bereich der mit MINT-Ausbildungsberufe die Zugehörigkeit zu einer bestimmten sozialen Gruppe als Risikofaktor für vorzeitige Vertragslösungen? Variiert das Vertragslösungsrisiko für Jugendliche mit Migrationshintergrund etwa in Abhängigkeit von der Zugehörigkeit zu einer bestimmten Herkunftsgruppe?

Lassen sich entsprechender Effekte auch unter Kontrolle anderer in der Literatur genannter Einflussfaktoren feststellen? Welche Rolle spielt diesbezüglich die Realisierung einer Wunschausbildung?

Die Analyse wird auf Grundlage der Daten der ersten beiden Messzeitpunkte des Projekts „Mathematisch-naturwissenschaftliche Kompetenzen in der beruflichen Erstausbildung“ (ManKobE) durchgeführt (Retelsdorf et al., 2013). Unsere Analyse konzentriert sich dabei auf die Auszubildenden der gewerblich-technischen Berufe (n=1287) sowie die angehenden Industriekaufleute (n=654). Die zu jedem Messzeitpunkt von den Berufsschulen aktualisierten Trackingdaten (zuletzt Sommer 2014) ermöglichen es Ausbildungsabbrüche im Verlauf und unabhängig von Stichprobenausfällen auf Probandenebene zu untersuchen. Mit den Daten des ersten Messzeitpunktes sowie den Trackingdaten (Vorabversion) für die ersten beiden Ausbildungsjahre wird eine Untersuchung von Vertragslösungen im Längsschnitt möglich. Der Beobachtungszeitraum ist somit auf die ersten 24 Monate der Ausbildung begrenzt. In dieser Zeit finden typischerweise die meisten Vertragslösungen statt (BiBB, 2014).

Die multivariaten Analysen bezüglich möglicher Risikofaktoren für eine vorzeitige Ausbildungsvertragslösung werden als Proportional-Hazards-Model mit robusten Standardfehlern gerechnet. So wird es möglich rechtzensierte Daten etwa in Folge eines Testgruppenausfalls zu berücksichtigen.

Die drei größten Herkunftsgruppen im Sample stammen aus der ehemaligen Sowjetunion, Türkei und dem ehemaligen Jugoslawien und werden entsprechend in den Analysen berücksichtigt. Die Probanden wurden zum ersten Messzeitpunkt offen nach ihren Wunschberufen gefragt. Im Falle einer Übereinstimmung der Angaben mit dem derzeitigen Ausbildungsberuf wurde dies nachverkodet. Als Kontrollvariablen dienen unter anderen Abbruchintention zum ersten Messzeitpunkt, der Schulabschluss und die Ausbildungsberufe.

In den Analysen zeigt sich zunächst ein bivariater Zusammenhang zwischen Ausbildungsstatus und dem nach Herkunftsregionen unterschiedenen Migrationshintergrund. Die Unterschiede zwischen Jugendlichen ohne Migrationshintergrund und den Jugendlichen verschiedener anderer Herkunftsregionen bestehen dabei sowohl bezogen auf die Häufigkeit von Vertragslösungen, als auch hinsichtlich deren Qualität. So sind die Jugendlichen verschiedener anderer Herkunftsregionen stärker von Vertragslösungen betroffen und verlassen zugleich häufiger das berufliche Schulsystem, ohne eine weitere Qualifizierung. Zudem erwiesen sich die zum ersten Messzeitpunkt erhobene Abbruchsintention und das Fehlen einer Hochschulzugangsberechtigung als prädiktiv für die späteren Vertragslösungen. Die Realisierung eines Wunschberufs steht zwar in keinem direkten Zusammenhang mit den Vertragslösungen, hängt jedoch mit der Abbruchsintention zusammen.

In der multivariaten Analyse, zeigt sich im umfassendsten Modell ebenfalls ein Effekte von Abbruchintention. Bezogen auf den Migrationshintergrund bleibt ein Effekt für die Gruppe der Jugendlichen aus dem ehemaligen Jugoslawien bestehen.

Mittwoch
11. März

14:00 Uhr

GC 03/149

Chair(s)
Sari, Duygu

Förster-Kuschel, Jana

Wie nehmen betriebliche Ausbilder Heterogenität von Auszubildenden wahr und wie gehen sie damit um? – Ergebnisse einer Interviewstudie in ausgewählten kaufmännischen Berufen

Auszubildende in der dualen Ausbildung unterscheiden sich oft deutlich hinsichtlich ihres Alters, ihres Schulabschlusses, ihrer beruflichen Vorbildung sowie ihrer sozial-kulturellen Merkmale (Seeber, 2010). Mit Blick auf die demografische Entwicklung und damit verbunden einem Rückgang an Schulabgängern sowie veränderten Bildungsentscheidungen von Jugendlichen (Tonhäuser & Büker, 2014), kann davon ausgegangen werden, dass sich die Heterogenität von Auszubildenden weiter erhöhen wird (Ebbinghaus, 2011; Westhoff & Ernst, 2011). Insbesondere ausbildende Unternehmen, die sich aufgrund zunehmender Schwierigkeiten bei der Besetzung offener Ausbildungsplätze auch für leistungsschwächere Jugendliche öffnen, sehen sich mit wachsender Heterogenität ihrer Auszubildenden konfrontiert (Brünner, 2014).

Im Kontext des Lernens bezieht sich der Begriff Heterogenität auf Unterschiede zwischen Lernern. Theoretisch gibt es unendlich viele Merkmale, nach denen eine Gruppe als heterogen klassifiziert werden kann (Trautmann & Wischer, 2011). Nach Strasser (2011) lässt sich Heterogenität von Lernern anhand primärer Strukturmerkmale (z. B. Alter, soziale Herkunft) und durch sie beeinflusste sekundäre Merkmale, d. h. kognitive, motivationale und affektive Merkmale, die das Lernen maßgeblich beeinflussen (z. B. Helmke, 2004; Baumert u. a., 2002), beschreiben.

Heterogenität wurde und wird insbesondere für die Institution Schule diskutiert. Der Lernort Betrieb findet demgegenüber diesbezüglich bisher kaum Beachtung, so dass bisher noch nicht geklärt ist, in welchen (für den betrieblichen Lernprozess relevanten) Merkmalen sich Auszubildende dabei genau unterscheiden. Dem Lernort Betrieb kommt aber im Rahmen der dualen Ausbildung ein zentraler Stellenwert zu, da Auszubildenden in realen Situationen aktiv handelnd Wissen erwerben und wertvolle Erfahrungen sammeln können. Dies kann sich wiederum positiv auf die Motivation, den Transfer des Erlernten und die Persönlichkeitsentwicklung der Auszubildenden auswirken (Noß, 2000). Das Lernen im Betrieb wird maßgeblich durch die Arbeit des betrieblichen Ausbildungspersonals geprägt. Der Umgang mit heterogenen Auszubildenden verlangt dieser Personengruppe großes pädagogisches Können ab. Haupt- und insbesondere nebenberufliche Ausbilder verfügen aber i. d. R. über keine systematische pädagogische Ausbildung und greifen daher bei ihrer Ausbildertätigkeit hauptsächlich auf eigene Erfahrungen und ihre subjektiven Theorien zurück (z. B. Bahl, 2012). Vor diesem Hintergrund und unter Qualitätsaspekten gilt es zu klären, welche Auswirkungen sich aus der Heterogenität von Auszubildenden für das betriebliche Ausbildungspersonal ergeben und wie der Umgang mit Heterogenität adäquat erfolgen kann. Dafür ist es zunächst erforderlich zu klären, wie das betriebliche Ausbildungspersonal Heterogenität von Auszubildenden wahrnimmt und wie es mit der wahrgenommenen Heterogenität von Auszubildenden umgeht. In einer Interviewstudie, die von Mai bis August 2014 durchgeführt worden ist, wird versucht, erste Antworten auf diese beiden Fragen aus Sicht haupt- und nebenamtlicher Ausbilder zu beantworten.

Den theoretischen Rahmen für die Untersuchung bildet ein modifiziertes Modell der Ausbilder-Auszubildenden-Interaktion nach Verstege (2007; in Anlehnung an Nickel (1976)) sowie Becker, Oldenbürger und Piehl (1987). Das Modell betont die Bedeutung der wechselseitigen Wahrnehmung zwischen Ausbilder und Auszubildendem für das Verhalten des jeweils anderen. Neben der Wahrnehmung des jeweils anderen wirken der soziokulturelle Bezugsrahmen sowie intrapersonale Faktoren (Kognition, Motivation und Emotion) auf das Verhalten des Ausbilders bzw. des Auszubildenden. Soll nun die komplexe Struktur des Ausbilderverhaltens hinsichtlich des Umgangs mit Heterogenität von Auszubildenden analysiert werden, gilt es, diese drei Aspekte zu untersuchen. Da der Forschungsstand zu Heterogenität von Auszubildenden aus der Sicht des betrieblichen Ausbildungspersonals als defizitär betrachtet werden kann, wurde in der Interviewstudie folgender Frage nachgegangen: Wie nehmen haupt- und nebenberuflichen Ausbilder Heterogenität von Auszubildenden wahr und wie gehen sie damit um?

In der Untersuchung wurden 14 Ausbilder/innen in ausgewählten kaufmännischen Berufen (Bankkaufmann, Bürokaufmann, Kaufmann im Einzel- und Großhandel sowie Industriekaufmann) mittels halbstrukturierter Interviews befragt. Die Interviews werden zum jetzigen Zeitpunkt inhaltsanalytisch ausgewertet. Auf der Tagung sollen erste Ergebnisse der Interviewstudie präsentiert werden.

Mittwoch
11. März

14:00 Uhr

GC 03/149

Chair(s)
Sari, Duygu

Borzikowsky, Christoph; Nagy, Gabriel; Bernhardt, Fabian

Die Bedeutung der Personen-Umwelt-Kongruenz und der sozialen Identifikation für den akademischen Erfolg in der beruflichen Ausbildung

Trotz der kontinuierlichen Verbesserung von Lernmethoden empfindet immer noch etwa ein Viertel der Jugendlichen eine starke Belastung durch die Schule (Currie et al., 2012). Infolgedessen brechen immer mehr Schülerinnen und Schüler ihre komplette Schullaufbahn ab (Hammond, Linton, Smink & Drew, 2007). Dieser Befund ist alarmierend, da Jugendliche ohne Schulabschluss nachweislich über geringere Fachkompetenzen, geringere Lernmotivationen und geringere Jobchancen als ihre Altersgenossen mit einem Schulabschluss verfügen (Eckstein & Wolpin, 1999; Montmarquette, Viennot-Briot & Dagenais, 2007). Nach Mansel und Hurrelmann (1992) sind hohe Abbruchsquoten insbesondere in beruflichen Ausbildungsgängen zu verzeichnen, in denen Berufsschülerinnen und -schüler mit einer Doppelbelastung aus Berufsschule und Betrieb konfrontiert werden.

Diese Studie verfolgte deshalb das Ziel, Abbrüche in der beruflichen Ausbildung näher zu erforschen. Dabei stand zum einen die Frage im Zentrum, welche psychologischen Variablen auf Schülerseite die Entscheidung zum Ausbildungsabbruch beeinflussen und zum anderen, wie der akademische Erfolg der Berufsschülerinnen und -schüler in positive Bahnen gelenkt werden könnte.

Die Erforschung dieser bildungswissenschaftlichen Fragestellungen erfolgte durch die Kombination aus organisationspsychologischen und sozialpsychologischen Ansätzen. Auf der einen Seite haben aktuelle Studien aus dem Arbeits- und Organisationsbereich die Kongruenz zwischen Personen und ihrer Umwelt für den akademischen Erfolg betont (Li, Yao, Chen & Wang, 2013; Schmitt, Oswald, Friede, Imus & Merritt, 2008). Für eine hohe Arbeitszufriedenheit müssen sich demnach die Anforderungen des Arbeitsplatzes möglichst gut mit der Eignung des Bewerbers decken. Hingegen resultiert eine Nichtübereinstimmung in Stress- und Belastungsreaktionen, was die vielen Abbrüche in der beruflichen Ausbildung erklären könnte.

Auf der anderen Seite haben verschiedene sozialpsychologische Studien zur Arbeits- und Lernmotivation die Bedeutung der sozialen Identifikation (wahrgenommene Mitgliedschaft in einer sozialen Gruppe, Tajfel & Turner, 1986; Turner, Hogg, Oakes, Reicher & Wetherell, 1987) für leistungsbezogene Variablen herausgestellt. So konnte beispielsweise van Knippenberg (2000) einen positiven Einfluss einer hohen Identifikation mit dem eigenen Betrieb auf die Arbeitsmotivation und -leistung nachweisen (vgl. auch van Knippenberg & Ellemers, 2003). Übertragen auf das Problem der steigenden Abbruchsquoten in der beruflichen Ausbildung könnte somit auch eine geringe Identifikation mit dem eigenen Ausbildungsberuf ursächlich für diese Entwicklung sein.

Für die Überprüfung dieser Hypothesen wurden 235 Berufsschülerinnen und -schüler (46 männlich, 189 weiblich, $M(\text{Alter}) = 21.47$ Jahre, $SD(\text{Alter}) = 3.63$ Jahre) an fünf deutschen Berufsschulen mit Hilfe eines Papier- und Bleistift-Fragebogens zu ihrer wahrgenommenen Personen-Umwelt-Kongruenz (vier Subskalen) und zu ihrer Identifikation mit dem eigenen Ausbildungsberuf befragt. Als abhängige Maße enthielt der Fragebogen Items zur Abbruchsintention, zur Zufriedenheit mit dem eigenen Ausbildungsberuf und zur Lernmotivation.

In einem ersten Schritt wurden die vier Subskalen zur Personen-Umwelt-Kongruenz und die Identifikationsskala hinsichtlich ihrer Dimensionalität überprüft. Eine konfirmatorische Faktorenanalyse ergab, dass die 5-Faktorenlösung im Vergleich zu anderen Modellen die beste Anpassung an die beobachteten Daten lieferte ($CFI = .977$, $TLI = .968$, $RMSEA = .046$, $SRMR = .041$). In einem zweiten Schritt diente dieses 5-Faktorenmodell zur Vorhersage der Abbruchsintention, der Zufriedenheit mit dem eigenen Ausbildungsberuf und der Lernmotivation. Mit Hilfe von latenten Regressionsanalysen konnte nachgewiesen werden, dass zwei von vier Subskalen der Personen-Umwelt-Kongruenz, insbesondere jedoch die Identifikation mit dem eigenen Ausbildungsberuf die drei abhängigen Maße signifikant vorhersagen konnten (latentes R^2 zwischen .26 und .53). So zeigte sich die Identifikation mit dem eigenen Ausbildungsberuf als signifikanter negativer Prädiktor für die Abbruchsintention ($\beta = -.41$), als signifikanter positiver Prädiktor für die Zufriedenheit mit dem eigenen Ausbildungsberuf ($\beta = .54$) und als signifikanter positiver Prädiktor für die Lernmotivation ($\beta = .26$).

Zusammenfassend konnten in dieser Studie somit zwei psychologische Variablen identifiziert werden, die auf Schülerseite die Entscheidung zum Ausbildungsabbruch beeinflussen. Hierbei stellte sich insbesondere die Identifikation mit dem eigenen Ausbildungsberuf als bedeutsamer Prädiktor heraus, der genutzt werden könnte, um Ausbildungsabbrüche zu reduzieren und leistungsrelevante Variablen auf Schülerseite (z.B. die Lernmotivation) zu verbessern.

Mittwoch
11. März

14:00 Uhr

GC 03/149

Chair(s)
Sari, Duygu

Zinn, Bernd; Sari, Duygu; Wyrwal, Matthias

Förderung lernschwacher Auszubildender in der Metalltechnik

Theoretischer Hintergrund: Viele Jugendliche haben bedingt durch Leistungsdefizite, Motivationsprobleme und unzulänglich entwickelten strategischen Fähigkeiten große Schwierigkeiten die Ausbildungsanforderungen zu erfüllen. Beiträge der empirischen Lehr-Lernforschung zu Benachteiligten und insbesondere zur Erforschung von adaptiven Fördermöglichkeiten sind ein generelles Desiderat berufs- und wirtschaftspädagogischer Forschung (vgl. z.B. Bojanowski 2005). Im bautechnischen Bereich wurde das berufsbezogene Strategietraining BEST entwickelt und im Rahmen von zwei Interventionsstudien positiv, im Hinblick auf die bautechnische Fachkompetenz, evaluiert (Norwig et al. 2010, Petsch et al. 2012). Das berufsbezogene Strategietraining BEST umfasst eine kombinierte Strategieförderung von allgemeinen und bautechnischen Problemlösestrategien (ebd.). Im Rahmen einer von der Baden-Württemberg Stiftung finanzierten und vom Netzwerk Bildungsforschung unterstützten Studie wurde das berufsbezogene Strategietraining auf den metalltechnischen Bereich adaptiert. Im Theorieteil des Vortrags wird das Fördertraining im Bezugsfeld des metalltechnischen Fachwissens erläutert. Im zweiten Schritt werden die Effekte des Trainings auf die Entwicklung des Fachwissens und der Motivationsentwicklung dargestellt.

Fragestellung: Im Beitrag wird von der Entwicklung des berufsbezogenen Strategietrainings im Berufsfeld Metalltechnik sowie der Überprüfung seiner Wirksamkeitsmechanismen auf die metalltechnische Fachkompetenz berichtet. Hierbei wird geklärt inwieweit das Fördertraining einen Effekt auf die Fachkompetenzentwicklung hat und welche Schülergruppen hierdurch gefördert werden. Zudem wird erörtert, wie sich das integrative berufsbezogene Strategietraining auf die Motivations- und Interessenentwicklung sowie das Fähigkeitsselbstkonzept der Auszubildenden auswirkt.

Untersuchungsdesign/ Methode: Die Studie beinhaltet ein klassisches Experimental- und Kontrollgruppendesign mit Auszubildenden (n = 391) an zwölf Berufsfachschulen in Baden-Württemberg. Dafür wurden die Ausbildungsberufe des Anlagenmechanikers für Sanitär, Heizung und Klima und des Metallbauers gewählt, da gerade in diesen Berufen kognitiv schwächere Schüler angesiedelt sind. Die Intervention erfolgte über das gesamte Schuljahr 2013/14 mit insgesamt 30 Unterrichtsstunden à 90 Minuten. Zur Kompetenzfeststellung wurde eine Eingangs-, Zwischen- und Abschlusstestung durchgeführt. Mit quantitativen Instrumenten wurde das berufsspezifische Vorwissen, die metalltechnischen Kompetenzen, die kognitive Leistungsfähigkeit, die Basiskompetenzen (Mathematik, Lesen), das Fähigkeitsselbstkonzept und Fachinteresse erhoben. Die Datenauswertung erfolgte mit deskriptiven und multivariaten statistischen Methoden.

Ergebnisse: Die Befunde liefern eine Aussage zur Wirkung des Trainings auf die Fachkompetenz- und Motivationsentwicklung sowie auf das Fähigkeitsselbstkonzept der Schülerinnen und Schüler. Die Eingangstestung zeigt, dass die mehrheitlich kognitiv schwachen Auszubildenden beider Berufsgruppen (Anlagenmechaniker und Metallbauer) mit ungünstigen Eingangsvoraussetzungen in die Berufsausbildung starten. Demnach bringen die Jugendlichen ein geringes fachspezifisches Vorwissen mit und zeigen in Bezug auf ihre metalltechnischen Kompetenzen bei einfachen mathematischen Operationen Schwierigkeiten. Über den Trainingsverlauf hinweg, konnten in beiden Berufsgruppen signifikante Unterschiede (mit mittlerer bis starker Effektstärke) auf die Fachkompetenzentwicklung zugunsten der Experimentalgruppe konstatiert werden. Es kann zudem gezeigt werden, dass gerade kognitiv schwächere Schüler vom Strategietraining profitieren. Auf den Motivationsverlauf und das Fähigkeitsselbstkonzept hat das Fördertraining keinen erkennbaren Einfluss. Bezogen auf die Entwicklung des fachspezifischen Interesses kann, bei der am Fördertraining teilnehmenden Kohorte der Metallbauer, ein üblicherweise stattfindender Interessensrückgang verhindert werden.

Mittwoch
11. März

14:00 Uhr

GC 03/149

Chair(s)
Sari, Duygu

Bosse, Stefanie; Henke, Thorsten; Jäntsich, Christian; Lambrecht, Jennifer; Spörer, Nadine

Alles bleibt anders! Zur Veränderung inklusiver Einstellung und Selbstwirksamkeit von Grundschullehrkräften

Die pädagogische Weiterentwicklung vieler Schulen zielt seit dem Inkrafttreten der UN-Behindertenrechtskonvention 2009 in Deutschland auf die Umsetzung inklusiven Unterrichts ab. Dies fordert für alle im Schulprozess beteiligten Akteure eine Veränderungsbereitschaft in Bezug auf ihr Denken und Handeln. Das Ziel vieler bildungspolitischer Maßnahmen ist seitdem darauf gerichtet, Lehrkräfte und ganze Schulen bei der Ausgestaltung inklusiven Unterrichts zu unterstützen, so dass diese Innovation sowohl in Bezug auf die Kompetenzen als auch bezüglich des Wohlbefindens der Schülerinnen und Schüler positive Wirkungen erzielt.

Die Erforschung der Einstellung zum inklusiven Unterricht ist insofern entscheidend, weil Studien einen Zusammenhang zwischen den Einstellungen und dem professionellen Handeln von Lehrkräften sehen (Ajzen & Fishbein, 2000; Hartinger, Kleickmann & Hawelka, 2006). Dabei wird angenommen, dass die Selbstwirksamkeit zum einen bedeutsam für die Veränderungsfähigkeit von Lehrerinnen und Lehrern in Innovationsprozessen ist (Edelstein, 2002) und zum anderen auch Einflussfaktor für die inklusive Einstellung ist (Jordan, Kircaali-Iftar & Patrick Diamond, 1993; Sharma, Loreman & Forlin, 2012; Soodak, Podell & Lehmann, 1998; Weisel & Dror, 2006). Vielfältige Studien benennen darüber hinaus Trainings- und Fortbildungsmaßnahmen als Prädiktoren der inklusiven Einstellung (Avramidis & Kalyva, 2007; de Boer, Pijl & Minnaert, 2011).

Wenngleich mittlerweile vielfältige Instrumente zur Erfassung der Einstellung und der Selbstwirksamkeit entwickelt wurden (Antonak & Larrivee, 1995; Forlin, Cedillo, Romero-Contreras, Fletcher & Hernández, 2010; Kopp, 2009; Kunz, Luder & Moretti, 2010) und die Zusammenhänge zwischen den beiden Merkmalen querschnittlich betrachtet werden, so fehlen derzeit längsschnittliche Untersuchungen, die Aussagen über die Stabilität der Einstellung und Selbstwirksamkeit vornehmen. Dieses Desiderat wird mit der vorliegenden Studie aufgegriffen. Das Ziel unserer Untersuchung ist die Prüfung der Variabilität bzw. der Stabilität der inklusiven Einstellung und Selbstwirksamkeit über die Zeit. Darüber hinaus sollen prädiktive Effekte der Selbstwirksamkeit auf die Einstellung untersucht werden. Schließlich werden weitere personenbezogene Lehrermerkmale im Hinblick auf mögliche Zusammenhänge zur inklusiven Einstellung und Selbstwirksamkeit geprüft. Mittels einer Längsschnittsbefragung mit drei Messzeitpunkten wurden die folgenden Hypothesen geprüft: (1) Inklusive Einstellung und Selbstwirksamkeit sind relativ stabile Merkmalsdispositionen in dem Sinn, dass frühere Ausprägungen der Merkmale spätere Ausprägungen vorhersagen (Stabilitätshypothese). (2) Inklusive Einstellung und Selbstwirksamkeit hängen positiv zusammen (Querschnittshypothese). (3) Unter Kontrolle der querschnittlichen Zusammenhänge hat die Selbstwirksamkeit einen prädiktiven Effekt auf die später gemessene Einstellung (Kreuzkorrelationshypothese). (4) Die Nutzungshäufigkeit von Fortbildungen und eine positiv eingeschätzte Nützlichkeit dieser Fortbildungen haben einen positiven Einfluss auf die inklusive Einstellung (Beeinflussungshypothese).

Die vorliegende Studie bezieht sich auf Daten aus dem Forschungsprojekt „Lernen und Lehren in der inklusiven Grundschule“, bei dem N = 250 Lehrkräfte aus 84 Grundschulen zu drei Messzeitpunkten an Online-Befragungen teilgenommen haben. Die Messung der Einstellung und der Selbstwirksamkeit erfolgte mittels Skalen des Instruments KIESEL (Bosse & Spörer, 2014; Cronbachs α zwischen .73 und .90). Darüber hinaus zeigten konfirmatorische Faktorenanalysen, dass für beide Konstrukte die Items hoch auf den jeweiligen Faktoren laden und dass von einem jeweiligen Generalfaktor ausgegangen werden kann. Die Beurteilung der besuchten Fortbildungen erfolgte mittels adaptierten Skalen von Klein-Heßling und Röder (2009).

Zur Prüfung der Hypothesen wurden Längsschnittstrukturgleichungsmodelle betrachtet. Die Vergleichbarkeit der Messungen und damit der Invarianz der latenten Variablen wurde anhand verschiedener Messinvarianzbedingungen auf die Einstellungs- und Selbstwirksamkeitsskalen getestet. Anschließend wurde ein latentes autoregressives Modell geprüft, um mögliche kreuzverzögerte Effekte der Selbstwirksamkeit auf die Einstellung nachzuweisen. Erste Analysen zeigen keine statistisch bedeutsamen Zusammenhänge sowohl der Einstellung als auch der Selbstwirksamkeit zu sozio-demografischen Variablen. Derzeit werden die Fortbildungshäufigkeit und die Einschätzung der Nützlichkeit der Fortbildungen als Prädiktoren der inklusiven Einstellung und Selbstwirksamkeit regressionsanalytisch geprüft.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Urton, Karolina; Wilbert, Jürgen; Hennemann, Thomas

Auswirkungen des Erlebens von Respekt, Selbstbestimmung und Selbstwirksamkeit auf das Belastungserleben von Lehrkräften an inklusiven Grundschulen

Der Weg hin zu einem inklusiven Schulsystem kann als schulischer Innovationsprozess gesehen werden, der zum Ziel hat ein Klima der Wertschätzung und Anerkennung von Vielfalt zu schaffen. Diese grundlegenden Veränderungen der schulischen Organisationsstrukturen stellen eine Herausforderung für die Lehrkräfte, Kollegien und Schulleitungen dar. Zentrales Anliegen der Studie ist es in diesem Kontext zu betrachten wie ein respektvoller Umgang, das Erleben von Selbstbestimmung und Selbstwirksamkeit das Belastungserleben der Lehrkräfte beeinflusst. Diese psychologischen Faktoren nehmen wir als bedeutsam für den schulischen Inklusionsprozess an, da sich zeigen ließ, dass der respektvolle Umgang der Führungskraft mit den Mitarbeitern als einer der wichtigsten Arbeitswerte angesehen wird und als solcher einen bedeutsamen Faktor für die Bewältigung von Veränderungsprozessen in Organisationen darstellt (van Quaquebeke, Zenker & Eckloff, 2009). Darüber hinaus führt das Erleben von Respekt auch zu einer höheren Arbeitszufriedenheit der Mitarbeiter (Zenker & van Quaquebeke, 2006). Im Zusammenhang mit dem Respekterleben steht das Erleben von Selbstbestimmtheit (Eckloff, van Quaquebeke & Witte, 2008). Selbstbestimmtheit zeichnet sich durch die Befriedigung der psychologischen Bedürfnisse nach Autonomie, Kompetenz und Verbundenheit aus (Deci & Rayen, 2000) und ist auf Individuumsebene bedeutsam für die Bewältigung komplexer Veränderungsprozesse (Eckloff et al., 2008) sowie das Erleben von Integrität und Wohlbefinden (Deci & Rayen, 2000). Weiterhin stellt für das Gelingen der Implementierung und Fortführung der Inklusion die erlebte Selbstwirksamkeit der Lehrerinnen und Lehrer eine maßgebliche Rolle dar, da sie die Akzeptanz der Kinder mit sonderpädagogischem Förderbedarf als Teil der Schulgemeinschaft wesentlich beeinflusst (Soodak, Podell & Lehman, 1998). Darüber hinaus führen die individuelle und die kollektive Selbstwirksamkeit dazu, dass Personen und Gruppen schwierige Aufgaben eher als Herausforderungen begreifen, sich herausforderndere (Reform-)Ziele stecken, größere Anstrengungen aufwenden und mehr Ausdauer bei der Bewältigung von Aufgaben zeigen. Die individuelle Selbstwirksamkeit wirkt sich auch positiv auf den Umgang mit Misserfolgen und die Erholung nach Rückschlägen aus (Bandura, 1997) und ist als eine bedeutsame personale Ressource für den Umgang mit Belastungen zu sehen (Schwarzer & Hallum, 2008; Schwarzer & Warner, 2011).

Auf dieser Grundlage gehen wir davon aus, dass sich die Kollegien bedeutsam hinsichtlich des Erlebens von Respekt, Selbstbestimmung, Selbstwirksamkeit und Belastung unterscheiden. Weiterhin nehmen wir an, dass Respekt, Selbstbestimmung und Selbstwirksamkeit sich positiv auf das Belastungserleben der Lehrkräfte auswirken und das der Respekt der Schulleitung für das Kollegium einen förderlichen Einfluss auf diese Variablen hat.

Zur Beantwortung dieser Fragestellungen wurden Daten an einer Stichprobe von 35 Schulleitungen und deren Kollegien (N = 310) in inklusiv tätigen Grundschulen erhoben. Das Erleben von Respekt, Selbstbestimmtheit, individueller und kollektiver Selbstwirksamkeit sowie Belastung wurde anhand von standardisierten Fragebögen erhoben. Um der Datenstruktur zu entsprechen, erfolgte die Analyse mittels Mehrebenenmodellierung.

Zunächst ließ sich zeigen, dass die Kollegiumszugehörigkeit einer Lehrkraft einen bedeutsamen Anteil der Varianz des Respekterlebens, der Selbstbestimmtheit, der individuellen und kollektiven Selbstwirksamkeit und des Belastungserlebens aufklärt.

Mit Blick auf das Belastungserleben der Lehrkräfte zeigt sich in den vorliegenden Daten, dass sowohl erlebter Respekt seitens des Kollegiums und der Schulleitung (Zusammenarbeit), das Erleben von Selbstbestimmtheit (Kompetenz, Verbundenheit) sowie die individuelle Selbstwirksamkeit zu einer Verringerung des Belastungserlebens führen. Auch der Respekt des Kollegiums für die Schulleitung zeigt einen positiven Zusammenhang mit dem Belastungserleben der Lehrkräfte. Hinsichtlich der weiteren Variablen ließ sich kein Zusammenhang mit dem Belastungserleben nachweisen.

Mit Blick auf die Schulleitung zeigt sich ein Zusammenhang zwischen dem Respekt der Schulleitung für das Kollegium und dem Respekterleben des Kollegiums von und für die Schulleitung, der erlebten Selbstbestimmtheit (Kompetenz) und dem individuellen Selbstwirksamkeitserleben. Hingegen zeigt sich kein bedeutsamer Zusammenhang zwischen dem Respekt der Schulleitung für das Kollegium und dem Belastungserleben der Lehrkräfte, der Selbstbestimmtheit in Bezug auf Autonomie und Verbundenheit sowie der kollektiven Selbstwirksamkeit.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Schaffus, Tina; Piezunka, Anne; Grosche, Michael

"Was ist Inklusion?": Empirische Analyse von Expertenwissen über theoretische Inklusionskonzepte

Theoretischer Hintergrund: Die Ratifizierung der UN-Behindertenrechtskonvention hat der Inklusionsforschung in Deutschland einen großen Schub gegeben. Aktuell werden eine Vielzahl an Forschungsprojekten zu Inklusion realisiert, beispielsweise die Messung von Einstellungen, Erfassung von Lernzuwächsen bei Schüler_innen, Formen von sonderpädagogischer Förderung und der Umgang mit heterogenen Lerngruppen.

Jedoch sind die einzelnen Ergebnisse der empirischen Studien bei der Vielzahl der verschiedenen untersuchten Inklusionsmodelle unter demselben begrifflichen Rahmen kaum miteinander zu vergleichen oder in Einklang zu bringen. Dies liegt vor allem an theoretischen Schwächen des Begriffs Inklusion (Grosche, in Druck). Manche Forscher_innen beschreiben Inklusion als den Zugang von Kindern mit sonderpädagogischem Förderbedarf zur Regelschule, wogegen andere unter Inklusion umfassende Veränderungen auf schulstruktureller Ebene fassen, die weit über den Zugang hinausgehen und den Fokus auf die gesamte Schülerschaft legen. Noch komplizierter ist die Verwendung anderer Begriffe, die in manchen Kontexten synonym oder aber in Abgrenzung zu Inklusion verwendet werden (z.B. Integration oder Umgang mit Heterogenität).

Die ungenaue und inkonsistente Definition von Inklusion macht die Ergebnisse der empirischen Bildungsforschung angreifbar und führt zu vielen Vorbehalten und Kontroversen auf beiden Seiten von Inklusionsgegnern und Inklusionsbefürwortern. Daher ist insgesamt eine große Unsicherheit unter den Vertreter_innen der empirischen Bildungsforschung im Bezug zum Konzept Inklusion zu konstatieren. Die empirische Untersuchung von Inklusion kann jedoch nur verbessert werden, wenn die untersuchten Inklusionskonzepte eindeutiger, trennschärfer und vor allem expliziter als bisher beschrieben werden könnten.

Fragestellung: Eine Möglichkeit, sich über verschiedene Inklusionskonzepte zu verständigen, ist die synthetisierende Befragung von Expert_innen, die schon viele Jahre an Inklusionskonzepten arbeiten, aber bislang noch nicht zu einem gemeinsamen Verständnis von Inklusion gelangt sind. Daher soll in der vorliegenden qualitativen Studie untersucht werden, welche Ideen und Vorstellungen Forscher_innen im deutschsprachigen Raum zu Inklusion haben und in welchen Dimensionen sich die einzelnen Ideen und Vorstellungen unterscheiden. Hieraus werden klarere und trennschärfere Definitionen von Inklusion entwickelt, die sich insbesondere als theoretische und operationalisierbare Grundlage für Studien der empirischen Bildungsforschung eignen.

Methode: Basierend auf verschiedenen Definitionen aus der Literatur (z.B. Luhmann, 1995; Hinz, 2004; Avenarius, 2012; Degener & Mogge-Grotjahn, 2012) wurde ein halbstandardisierter Leitfaden entwickelt. Dieser diente bei der Durchführung von elf Experteninterviews mit Forscher_innen aus ganz Deutschland als Grundlage. Bei der Auswahl der Interviewpartner_innen wurde darauf geachtet, dass sie aus unterschiedlichen Disziplinen (u.a. Erziehungswissenschaften, Sonderpädagogik, Psychologie, Soziologie, Jura) kommen, bereits zu unterschiedlichen Themen von Inklusion publiziert haben und mit unterschiedlichen empirischen Methoden arbeiten. Die halbstandardisierten Interviews wurden anhand der qualitativen Inhaltsanalyse nach Gläser und Laudel (2009) ausgewertet.

Ergebnisse: Insgesamt berichten die Expert_innen über zahlreiche verschiedene Dimensionen von Inklusion. Relevante Dimensionen sind unter anderem Realisierbarkeit von Inklusion, berücksichtigte Heterogenitätsmerkmale sowie Ziele und Legitimation von Inklusion. Während manche Forscher_innen mit Inklusion die normative Forderung nach gesellschaftlicher Teilhabe und gleichwertiger Anerkennung von Differenzen verbinden, wünschen sich andere eine unaufgeregte Diskussion über das Konzept an sich und die Frage, in welchen Fällen Inklusion oder Exklusion sinnvoll sei. In Bezug auf die Realisierbarkeit von Inklusion sind manche Forscher_innen der Meinung, dass Inklusion an manchen Schulen schon umgesetzt wird, wogegen andere an dem Gedanken festhalten, dass Inklusion ein utopisches Ideal sei und somit in der Forschung nur der Prozess abgebildet werden könnte. Unterschiedliche Vorstellungen von Inklusion spiegeln sich auch in der Frage wider, ob ein inklusives Bildungssystem voraussetze, dass die Förderschulen abgeschafft werden. Des Weiteren zeigen sich auch Unterschiede in Bezug auf das individuelle Verständnis von Inklusion und dem verwendeten Inklusionsbegriff in der Forschungspraxis.

Aus methodologischer Sicht bietet der Beitrag eine Grundlage für die Operationalisierung von Inklusion in Rahmen von quantitativen Untersuchungen. Ebenfalls ermöglicht der Beitrag auf theoretischer Ebene eine differenzierte Auseinandersetzung über die Frage, in welchen Dimensionen sich einzelne Inklusionsverständnisse trennscharf voneinander unterscheiden.

Ulandowski, Katharina; Czempel, Stefanie; Dietrich, Julia

Herausforderung schulische Inklusion: Zum Zusammenhang zwischen multiprofessioneller Kooperation und Unterrichtsqualität

Die Umsetzung der UN-Behindertenrechtskonvention, die seit 2009 in Deutschland rechtswirksam ist, verlangt tiefgreifende Veränderungen in der Konzeption von Schule, Unterricht und dem professionellen Handeln von Lehrerinnen und Lehrern (Moser, 2013). Eine der relativ neuen Herausforderungen für Fachlehrerinnen und Fachlehrer ist dabei die Kooperation mit Sonderpädagoginnen und Sonderpädagogen in der Planung, Umsetzung und Reflexion von Unterricht. Nach Kreie (2009) beeinflusst die Qualität der Kooperation der Pädagoginnen und Pädagogen das Gelingen der Inklusion erheblich. Im Kontext von inklusivem Unterricht kann multiprofessionelle Kooperation unterschiedliche Funktionen und Ausprägungen haben. Gräsel et al. (2006) differenzieren Kooperationsformen zwischen Lehrerinnen und Lehrern in (1) Austausch, (2) Arbeitsteilung und (3) Ko-Konstruktion. Dass sich Kooperation zwischen den Pädagoginnen und Pädagogen positiv auf die Unterrichtsqualität auswirkt, wird vielfach angenommen. Empirische Hinweise, die diese Annahme bestätigen, liefern u.a. die Schuleffektivitäts- und die Lehrerbefragungsforschung. Unterrichtsqualität hat viele Dimensionen; eine zentrale für schulische Inklusion ist die Umsetzung von kooperativen Lernformen im Unterricht. Gemeinsamer Unterricht gelingt, wenn Schülerinnen und Schüler mit unterschiedlichen Kompetenzen voneinander in wechselnden Sozialformen lernen können (Heimlich, 2007). Kooperative Lernformen als eine Form schülerzentrierter Unterrichtsmethoden aktivieren den einzelnen, ermöglichen soziale Kontakte zwischen den Schülerinnen und Schülern und verschaffen gleichzeitig der Lehrerin oder dem Lehrer den zeitlichen Freiraum, sich einzelnen zuzuwenden, um individuell zu fördern.

Im Kontext schulischer Inklusion ist bisher wenig erforscht, wie die Kooperationsbeziehungen zwischen den Pädagoginnen und Pädagogen unterschiedlicher Professionen ausgeprägt sind und, ob die Intensität der Kooperation die Unterrichtsqualität in Form von abwechslungsreichen Sozialformen beeinflusst.

Im Rahmen der Expertise „Gemeinsamer Unterricht im Kontext von Schul- und Unterrichtskultur in der Stadt Jena“ wurde im Sommer 2013 ein Lehrerfragebogen an Grundschulen, Gemeinschaftsschulen und Gymnasien eingesetzt. Jena kann im bundesdeutschen Kontext als Modellregion betrachtet werden, die bereits im Schuljahr 2013/14 einen relativ gesehen sehr hohen Inklusionsanteil von 80% erreicht hat. In die vorliegende Analyse wurde nur die Daten von 131 Lehrerinnen und Lehrern, davon 24 Sonderpädagoginnen und Sonderpädagogen, in die Stichproben einbezogen, die tatsächlich Unterrichtserfahrungen mit integrierten Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf haben. Im Fragebogen gaben die Pädagoginnen und Pädagogen u.a. darüber Auskunft, in welcher Form sie Unterricht planen, umsetzen und reflektieren (allein, parallel zu Sonderpädagogen/Fachlehrern, gemeinsam mit Sonderpädagogen/Fachlehrern) und wie häufig sie verschiedene Sozialformen im Unterricht einsetzen.

Korrelationsanalysen dieser Skalen zeigen Zusammenhänge zwischen den Ausprägungen der Kooperationsbeziehungen der Pädagoginnen und Pädagogen und den im Unterricht häufig eingesetzten Sozialformen. Die Hypothese, dass eine intensivere Kooperation der Pädagoginnen und Pädagogen sich positiv auf das Unterrichtsgeschehen auswirkt, scheint bestätigt zu werden.

Die Ergebnisse dieser Untersuchung geben Hinweise zur Gestaltung von Lehreraus- und -fortbildung und darüber, wie Schulen in der Umsetzung und Weiterentwicklung von Inklusion unterstützt werden können.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Lambrecht, Jennifer; Bosse, Stefanie; Henke, Thorsten; Jäntsich, Christian; Spörer, Nadine

Ohne Ressourcen keine Inklusion - Eine Analyse zur schulinternen Nutzung extern festgelegter Ressourcen an inklusiven Grundschulen

Eine Aufgabe der Bildungsforschung ist es, der Bildungspolitik wissenschaftliches Wissen als Basis für rationale Entscheidungen über Ressourcenverteilungen an Schulen bereitzustellen. Zentral hierfür ist ein Verständnis davon, wie Schulen als Organisationen funktionieren, d.h. wie Ressourcen vor Ort eingesetzt werden.

Im Mehrebenensystem Schule, als welches sich dieses aus der Governance-Perspektive betrachten lässt, agieren Akteure (Politik, Schule) auf unterschiedlichen Ebenen; zwischen ihnen bestehen starke Abhängigkeiten, aber keine klaren Befehlshierarchien (Mayntz 2008). Sowohl die Politik als auch die Schulleitungen vor Ort versuchen, Bildungsprozesse zu steuern. Dabei kann jedoch nicht davon ausgegangen werden, dass diese Akteure derselben Handlungslogik folgen (Altrichter & Maag Merki 2010, Bonsel 2010), denn die Schulen passen Maßnahmen an Bedingungen an, die in ihrem pädagogischen Umfeld vorherrschen. Das ist im Sinne der erweiterten Autonomie auch gewollt (BbgSchulG, §7), um der in den letzten Jahren zum Teil systematisch erhöhten Heterogenität der Merkmale von Schulkindern begegnen zu können. Dabei ist anzunehmen, dass die Art der Organisation der Schulen variiert.

Mit der Organisationstheorie Mintzbergs (1989) lässt sich die Struktur von Organisationen analysieren. Er unterscheidet sechs Teile (operativer Kern, mittleres Management, Unterstützungsstab, Technostruktur, strategische Spitze und Ideologie/Kultur), deren unterschiedliche Konstellationen sieben Typen von Organisationen ergeben, wobei die Schule dem Typus der Professionellen Organisation zugeordnet wird. Diese zeichne sich durch einen breiten operativen Kern und einen großen Unterstützungsstab aus, während mittleres Management und Technostruktur wenig ausgeprägt seien (Mintzberg 1989).

Im Pilotprojekt „Inklusive Grundschule“ des Landes Brandenburg, an dem 84 Grundschulen in öffentlicher oder privater Trägerschaft teilnehmen, kann in den Bereichen Lernen, Sprache und Verhalten auf die Feststellung eines sonderpädagogischen Förderbedarfs verzichtet werden. Pauschal erhält jede Schule für 5% der Gesamtschülerzahl 3,5 Lehrerwochenstunden zusätzlich. Die Verwendung obliegt den Einzelschulen.

Inwiefern dies die Struktur der Schulen beeinflusst, ist ein Forschungsdesiderat. Zur Schließung dieser Lücke wird in Bezug auf den operativen Kern und den Unterstützungsstab folgenden Fragestellungen nachgegangen: (1) In welchen Teilen der Organisation werden die zusätzlichen Ressourcen eingesetzt? (2) Gibt es verschiedene Typen inklusiver Grundschulen? (3) Steht die Stundenverwendung in systematischem Zusammenhang mit Rahmenbedingungen der Schule?

In der vorliegenden Untersuchung wurden (a) die Anzahl der wöchentlichen Arbeitsstunden der verschiedenen Professionen als Ressource betrachtet, (b) der operative Kern in Unterricht (O^U) und Nachmittagsangebot (O^N) aufgegliedert und der Unterstützungsstab nach den Dimensionen Lehrkräfte (U^L), Familie (U^F) und Organisation der Kooperation mit externen Partnern (U^E) unterschieden. Es wurden folgende Hypothesen getestet: (1) Fachkräfte verschiedener Professionen werden im O^U, O^N, U^L, U^F und U^E eingesetzt, wobei es zu Überschneidungen zwischen den Teilen kommt, (2) Der Einsatz verschiedener Professionen im O^U, O^N, U^L, U^F und U^E unterscheidet sich zwischen den Schulen, (3) Dies steht in systematischem Zusammenhang mit Rahmenbedingungen der Schule: (3i) An Schulen mit flexibler Schulanfangsphase ist der O^U besonders ausgeprägt (multiprofessionelle Teams); (3ii) An Ganztagschulen gibt es einen doppelten operativen Kern (O^U, O^N) und der U^K ist besonders ausgeprägt; (3iii) An Schulen mit einem hohen Anteil an Kindern mit festgestelltem/vermutetem Förderbedarf sind O^U und U^F besonders ausgeprägt.

Die Datengrundlage für diese Untersuchung bilden die Grundschulen des Pilotprojekts „Inklusive Grundschule“. Das Instrument zur Erhebung schulkontextueller Merkmale war ein Fragebogen, den die Schulleitungen am Ende des Schuljahres 2013/14 online bearbeiteten. An der Querschnittsbefragung nahmen N = 76 Schulleitungen teil. Erfasst wurden die wöchentlichen Arbeitsstunden der verschiedenen Professionen (5 Professionen) und die Art der Stundenverwendung (8 Items). Außerdem liegen strukturelle Daten zum Anteil der Schulkinder mit festgestelltem/vermutetem Förderbedarf sowie zur Organisation des Ganztags und der Schulanfangsphase vor.

Die Daten werden in Hinblick auf die Fragestellungen regressionsanalytisch und mittels Clusteranalysen ausgewertet. Erste Analysen zeigen eine Variation der Verteilung der Sonderpädagogerstunden zwischen den Schulen im operativen Kern (Variationskoeffizient $V_{ok} = .33$) und im Unterstützungsstab (Variationskoeffizient $V_u = .62$).

Hessels, Marco

Validität von Messungen zur Einstellung der schulischen Integration

In den meisten europäischen Ländern hat die Anzahl an Kindern mit sonderpädagogischem Förderbedarf, welche in den Regelunterricht integriert werden, in den letzten Jahren deutlich zugenommen. Der Trend geht in Richtung inklusive Schulen (European Agency, 2012). Neben Ressourcen und einer entsprechenden LehrerInnenausbildung gelten die Haltungen und Einstellungen gegenüber der schulischen Inklusion von direkt (DirektorInnen, LehrerInnen, SchülerInnen), aber ebenso indirekt betroffenen Personen (z. B. Eltern, Bevölkerung) als wesentliche Faktoren für einen erfolgreichen gemeinsamen Unterricht (z.B. de Boer et al., 2011). Demzufolge ist es auch nicht verwunderlich, dass die Forschung über die Einstellung zur schulischen Integration auch stark zugenommen hat. Beispielsweise erschienen in zwei deutschsprachigen Zeitschriften (Erziehung und Unterricht; Zeitschrift für Bildungsforschung) in diesem Jahr Themenhefte, welche sich intensiv mit dieser Thematik beschäftigen. Nun stellt sich die Frage, inwieweit die bisherigen Studienergebnisse valide sind. Beispielsweise zeigte sich in der Literatur bislang, dass die Einstellung zur schulischen Inklusion stark davon abhängt, welche Behinderung das zu integrierende Kind aufweist. Zum Beispiel ist die Einstellung gegenüber der schulischen Integration von Kindern mit körperlicher Behinderung und Kindern mit Lernbehinderung deutlich positiver als jene gegenüber Kindern mit geistiger Behinderung oder Verhaltensauffälligkeit (Gebhardt et al., 2011; Schwab et al., 2012). Zudem zeigte eine Untersuchung von Schwab et al. (2014), dass die Art und Weise, wie man nach der Einstellung fragt, die Ergebnisse beeinflusst. Wurde beispielsweise zuerst nach Segregation gefragt und im Anschluss nach Integration, so zeigte sich eine negativere Einstellung gegenüber der Integration als im umgekehrten Fall. Ziel der vorliegenden Studie ist es nun, diesen Reihenfolgeeffekt näher zu analysieren und die Validität von Erhebungen, welche sich lediglich auf die Einstellung zur schulischen Integration von SchülerInnen mit SPF beziehen (und keine Fallbeschreibungen zur Art der Behinderung liefern) zu hinterfragen. Dazu wurden insgesamt 257 Studierende an der Universität Genf befragt, welche sich im ersten Jahr des Bachelorstudiums Erziehungs- und Bildungswissenschaft befanden. Zuerst wurde jedem Teilnehmer/jeder Teilnehmerin eine von acht Versionen der ATIS Skala (Schwab et al., 2014) zugewiesen. Die Skalen wurden in einem experimentellen Design mit unterschiedlicher Reihenfolge vorgelegt, wobei bei der Hälfte der Stichprobe zuerst die Einstellung zur Integration und danach jene zur Segregation erfasst wurde. Dies variierte für vier verschiedene Arten von Behinderung (körperliche Behinderung, Lernbehinderung, geistige Behinderung, Verhaltensauffälligkeiten). Im Anschluss wurde die generelle Einstellung zur schulischen Integration (ohne Fallbeschreibung) mit dem ORI (Antonak & Larrivee, 1995) erfasst. Da die Daten in zwei Kohorten erfasst wurden (2013 sowie 2014), wurden bislang nur vorläufige Ergebnisse ausgewertet. Diese zeigen jedoch, dass die Fallbeschreibung, welche sich auf die Art der Behinderung des zu integrierenden Kindes bezog, einen signifikanten Effekt hat ($F_{3,71}=4.285$, $p<.01$). Ein Einfluss der vorher durchgeführten Erhebung mit Fallbeschreibungen auf die anschließende generelle Einstellung ließ sich jedoch nur tendenziell ($p=.07$) feststellen.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Mayer, Tanja; Koinzer, Thomas; Gruehn, Sabine; Schwarz, Judith

Choice. Schulwahl und sozio-kulturelle Passung. Einzelschulwahl, Schulqualität und Schulträgerspezifika unter besonderer Berücksichtigung von Schulen in freier Trägerschaft.

Obwohl in den zurückliegenden Jahren im deutschen Schulsystem ein kontinuierlicher Anstieg privater Schulen zu verzeichnen ist, kann der Forschungsstand zu Privatschulen nur als „unterentwickelt und wenig systematisch“ (1) bezeichnet werden. Im Rahmen des DFG-Projekts „Choice - Schulwahl und sozio-kulturelle Passung“ soll daher der Frage nachgegangen werden, warum sich Eltern für eine öffentliche resp. private Grundschule für ihr Kind entscheiden. Zudem werden diese Wahlen in einer Typologie der Passung zu spezifischen Privatschulträgern verortet. Unter Zugrundelegung von Handlungstheorien rationaler Wahl (2) und Bourdieus Habitus-Konzept (3) wird danach gefragt, ob, wie und warum Eltern bestimmte Schulen wählen. Dabei ist es zudem von Bedeutung, wie sich die privaten Schulen als akzeptierte, typisierbare und gesellschaftspolitisch relevante Alternativen zu öffentlichen Schulen bzw. als 'Ausweichschulen' (4) konfigurieren.

Ebenso wurde sowohl dem Prozess als auch den Segregationseffekten von Schulwahl bisher kaum Aufmerksamkeit entgegengebracht. Kristen (5) versteht Schulwahl als Prozess, in welchem nicht alle Eltern eine aktive Schulwahl treffen. Vor allem bildungsferne Eltern suchen oftmals keine Informationen zu Schulalternativen sondern schicken ihre Kinder auf die nächstgelegene und/oder zugewiesene Schule. Die erste Forschungsfrage des Projekts lautet daher: Welche Eltern bzw. welche Milieus wählen private bzw. öffentliche Schulen für ihr Kind und welche nicht?

Ein weiterer Fokus des Projekts liegt auf den Determinanten der Schulwahl. Untersuchungen für Österreich (6) und die Schweiz (2) zeigen, dass neben vielfältigen Faktoren der Schulqualität vor allem pragmatische Gründe (z.B. Schulweg) die Entscheidung für oder gegen eine Privatschule beeinflussen – Befunde, wie sie auch aus dem öffentlichen Schulwesen bekannt sind (7). → Wie begründen Eltern ihre Wahl vor dem Hintergrund der wahrgenommenen Schulqualität bzw. anderer schulwahlrelevanter Motive?

Zudem werden die Schulwahlmotive mit den sozio-kulturellen/habituellen Hintergründen der Eltern in eine Passung mit einem spezifischen Schultyp gebracht. Wir gehen dabei davon aus, dass sich sowohl im Feld der privaten als auch der öffentlichen Schulen unterschiedliche Typen von Schulen ausmachen lassen, die sich hinsichtlich ihrer Profilierungen unterscheiden und damit unterschiedliche Milieus ansprechen. → Welche sozio-kulturellen Charakteristika weisen Eltern auf, die eine private/öffentliche Schule für ihr Kind wählen und wie unterscheiden sich diese nach Schultyp/-profil?

Neben der Schulwahl der Eltern ist Ziel des Projektes herauszuarbeiten, ob und wie von Seiten der Schulen eine Passung hergestellt wird. Während im öffentlichen Schulwesen die Schülersauswahl durch gesetzliche Vorgaben eingeschränkt ist, ist die Schülersauswahl im Privatschulbereich ein konstituierendes Element. → Wird und wenn ja, wie von Seiten der Schulen eine Passung zu den Eltern/Kindern mit hergestellt („doppelte Passung“)?

Methode: Zur Beantwortung der Forschungsfragen werden quantitative und qualitative Verfahren in der Befragung von Eltern und Schulleitungen zur Anwendung kommen. Hierfür wurden je 95 öffentliche und private Grundschulen in Berlin (N=190) ausgewählt.

Quantitative Datenerhebung: Um die Schulwahlmotive und Hintergründe zu erfassen werden 190 Schulleiter/-innen und sowie die Eltern der zukünftigen Erstklässler des Schuljahrs 2015/16 (N= 7600) mithilfe eines Onlinefragebogens befragt. Die Erhebung wird im Frühjahr 2015 stattfinden. Die Daten werden anschließend mittels multivariater Datenanalysen (z.B. Strukturgleichungsmodelle/Mehrebenenanalysen) ausgewertet.

Qualitative Datenerhebung: Zusätzlich wird eine Substichprobe von 18 Interviewpaaren bestehend aus Schulleitung und Eltern gezogen, die der gleichen Schule zuzuordnen sind. Im Mittelpunkt der Leitfadeninterviews steht der Versuch, die tieferliegenden entscheidungsrelevanten Kriterien und Wahlmotive der Eltern sowie die entsprechenden Strategien der Schulleitungen zu rekonstruieren, um die ‚doppelte Passung‘ nachzuzeichnen. Zusätzlich werden die Schulhomepages analysiert. Die Daten werden mithilfe der qualitativen Inhaltsanalyse ausgewertet.

Die qualitativen und quantitativen Daten werden zudem anhand eines Mixed Methods Ansatzes zusammengeführt.

Ergebnisse: Da die Onlinebefragungen/Leitfadeninterviews erst im Frühjahr 2015 durchgeführt werden, liegen bislang nur Ergebnisse aus der inhaltsanalytischen Auswertung der Schulhomepages vor. Hier zeichnet sich bereits ab, dass sich die (privaten) Grundschulen entlang pädagogischer und weltanschaulicher Profile typisieren lassen und dadurch ggf. unterschiedliche Zielgruppen ansprechen.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Böse, Susanne; Neumann, Marko; Becker, Michael; Maaz, Kai

Bereit für Veränderungen? – Zur Auseinandersetzung von Schulleiterinnen und Schulleitern mit Schulreformen im Kontext der Berliner Schulstrukturreform

Die Implementation umfassender Schulreformen ist ein komplexer Prozess, dessen Gelingen von verschiedenen Faktoren abhängig ist (Fullan, 2007). Eine wesentliche Rolle für den Implementationserfolg nehmen dabei die beteiligten Akteure ein. So weist Coburn (2003) darauf hin, dass die Identifikation mit der Reform, also die Akzeptanz und Verinnerlichung der Reformziele durch die Akteure, eine Schlüsselrolle im Implementationsprozess spielt und dazu beiträgt, die Ziele umzusetzen. In diesem Zusammenhang ist insbesondere die Schulleitung zu nennen, die im Implementationsprozess eine entscheidende Rolle einnehmen. Durch ihre besondere Funktion nimmt die Schulleitung eine Vermittlungs- und Kontrollposition zwischen Administration und Einzelschule ein (Bonsen, 2010) und besitzt eine gewisse „Ausstrahlungskraft“ auf die anderen am Reformprozess beteiligten Akteure wie Lehrkräfte und Eltern (Gräsel, 2010). Empirische Forschungsarbeiten haben gezeigt, dass Schulleiterinnen und Schulleiter in großem Maße die persönlichen Zielsetzungen sowie die professionellen Überzeugungen der Lehrkräfte beeinflussen (Leithwood & Jantzi, 1999). Sie können Lehrkräfte zu Veränderungen anregen und ihnen eine klare Vision für die Veränderung geben (Geijsel, Slegers, Leithwood, & Jantzi, 2003).

Zentrales Anliegen dieses Beitrags ist es daher, die Auseinandersetzung der Schulleiterinnen und Schulleiter mit der umfassenden Schulstrukturreform in Berlin, deren hauptsächliche Neuerung die Schaffung einer neuen Schulform, der Integrierten Sekundarschule (ISS), neben dem Gymnasium ist, zu untersuchen. Neben dieser strukturellen Veränderung fand eine Modifikation des Übergangsverfahrens von der Grundschule in die weiterführenden Schulen statt. Ein Instrument, das die Beteiligten in einem Reformprozess in den Mittelpunkt stellt und deren phasenweise Auseinandersetzung mit Innovationen im Zuge eines Veränderungsprozesses beschreibt, ist der Stages of Concern-Fragebogen (SoCQ, Hall & Hord, 2011), der insgesamt sieben Stufen der Auseinandersetzung unterscheidet. Der SoCQ wurde im Rahmen der BERLIN-Studie in leicht angepasster Form eingesetzt, wobei insgesamt 196 Schulleiterinnen und Schulleiter der Gymnasien sowie Integrierten Sekundarschulen im ersten Jahr nach der Schulstrukturreform befragt worden sind. Es soll der Frage nachgegangen werden, ob sich der SoCQ eignet, das Ausmaß der Auseinandersetzung von Schulleiterinnen und Schulleitern mit einer umfassenden Schulstrukturreform abzubilden. Weiterhin soll danach gefragt werden, welche Profile der Auseinandersetzung mit der Schulstrukturreform sich für die Schulleiterinnen und Schulleiter der weiterführenden Schulen zeigen und wie sich diese vor dem Hintergrund individueller Merkmale (Geschlecht, Schulform und Berufserfahrung) dieser beschreiben lassen. Zur Beantwortung dieser Fragestellungen wurden zunächst verschiedene Modelle in konfirmatorischen Faktorenanalysen (CFA) hinsichtlich ihrer Passung auf die Daten mit dem Programmpaket Mplus 7 geprüft. Zur Identifizierung von unterschiedlichen Schulleiterprofilen der Auseinandersetzung mit der Schulstrukturreform wurden anschließend latente Profilanalysen (LPA), ebenfalls mit dem Programmpaket Mplus 7, gerechnet (Analyseoption Type = Mixture). In den CFA-Analysen zeigte sich, dass sich das theoretisch intendierte 7-phasige Konzept der Stages of Concern nicht eindeutig in den Daten finden lässt. Es konnte jedoch eine 6-faktorielle Lösung gefunden werden, die für die im vorliegenden Beitrag untersuchte Schulreform plausibel erscheint. Zudem konnten insgesamt fünf Profile der Auseinandersetzung identifiziert werden, wobei die größte Gruppe die „besorgt-motivierten Umsetzer“ bilden. Bezogen auf das Merkmal Geschlecht ließ sich feststellen, dass sich in der Profilsgruppe der „Motivierten & Hochinteressierten“ überproportional häufig weibliche Schulleiter finden ließen. Für das Merkmal Schulform konnte festgestellt werden, dass sich ISS-Schulleitungen verstärkt in den Profilsgruppen wiederfinden ließen, die sich stärker mit der Reform auseinandersetzen. Für das Merkmal Berufserfahrung zeigte sich, dass sich Schulleitungen mit einer vergleichsweise geringen Berufserfahrung ebenfalls eher in den umsetzungsorientierten Profilsgruppen finden ließen.

Es kann festgehalten werden, dass der SoCQ ein multidimensionales Instrument ist, das gerade die Breite der Auseinandersetzung mit Innovationsprozessen umfänglich darstellt und damit durchaus geeignet ist, komplexe Implementationsprozesse akteurspezifisch zu untersuchen. Auch kann es dazu beitragen, den Stand eines Implementationsprozesses anhand der Einschätzungen der Akteure zu beurteilen. Gleichwohl stellt sich die Frage, inwieweit das theoretische Konzept der Stages of Concern nur für bestimmte Arten von Innovationen geeignet ist bzw. generell anwendungsspezifisch zu modifizieren ist.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Schneider, Achim

Kompetenzerwerb durch institutionsübergreifende Zusammenarbeit

Im Sommer 2007 startete das Modellprojekt „Bildungshaus 3 – 10“ des Ministeriums für Kultus, Jugend und Sport Baden-Württemberg. Die wissenschaftliche Begleitung des Modellprojekts „Bildungshaus 3 – 10“, gefördert mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) und des Europäischen Sozialfonds (ESF) der Europäischen Union (EU), wird vom ZNL TransferZentrum für Neurowissenschaften und Lernen, Ulm durchgeführt.

Im „Bildungshaus 3 – 10“ führen pädagogische Fachkräfte aus Kindergärten und Lehrkräfte aus Grundschulen an 33 Modellstandorten gemeinsam regelmäßig stattfindende, altersgemischte und institutionsübergreifende Spiel- und Lerneinheiten durch. Außerdem sind die Modellstandorte aufgefordert, ausgehend von der Beobachtung und Dokumentation kindlicher Entwicklungsprozesse individuell auf die Kinder einzugehen und die gemeinsamen Aktivitäten weiterzuentwickeln (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2007). Durch diese intensive Kooperation entstehen auf unterschiedlichen Ebenen Gesprächsnotwendigkeiten und sind Aushandlungsprozesse erforderlich.

Ähnliche, allerdings schulinterne Austausch- und Reflexionsprozesse mit Bezug zu einer gemeinsamen Praxis sind bereits Untersuchungsgegenstand und es liegen positive Befunde im Hinblick auf Kompetenzerweiterung der Lehrkräfte und auf Schülerleistungen vor (z. B. Avalos, 2011; Dunne, Nave & Lewis, 2000; Horn & Little, 2010; Louis & Marks, 1998; Vescio, Ross & Adams, 2008). Über das institutionsübergreifende Äquivalent ist jedoch relativ wenig bekannt. Eine Studie hierzu kommt von Vogt, Zumwald, Urech und Abt (2010), die Hinweise auf einen ergänzten Blick auf das Kind sowie eine erweiterte Praxis von Lehrkräften durch Teamteaching berichten.

Im Zusammenhang mit Austausch- und Reflexionsprozessen im „Bildungshaus 3 – 10“ wird erwartet, dass es zur Kompetenzerweiterung beider Professionen kommen könnte. Die Gewinne könnten daraus entstehen, dass die Perspektive einer Profession zusätzlich durch die der jeweiligen anderen kritisch hinterfragt werden könnte und somit die Reflexionsprozesse erweitert werden. Durch die gemeinsame Praxis, die auch als Reflexionsgrundlage dient, könnten die im Reflexionsprozess erarbeiteten Ideen überprüft bzw. geübt werden.

Um mögliche Gewinne und bedingende Faktoren der Austauschprozesse zu untersuchen, wurde im Rahmen der wissenschaftlichen Begleitung ein Fragebogen auf Basis zweier Fragebögen (Der Paritätische, 2007; Honig & Schreiber, 2006) entwickelt und auf Grundlage der Erfahrung der wissenschaftlichen Begleitung an den Kontext „Bildungshaus 3 – 10“ angepasst. Der Fragebogen besteht aus 24 Fragen und erfasst neben anderen Bereichen die Planung und Durchführung des professionellen Austausches. Hierbei stehen besonders drei Fragen des Fragebogens im Fokus: erstens wurde erfragt, ob ein regelmäßiger Austausch stattfindet. Zweitens wurde die Bedeutsamkeit des Austausches für die Praktiker mit Hilfe einer vierstufigen Skala ermittelt (sehr wichtig/eher wichtig/weniger wichtig/nicht wichtig). Drittens wurde nach den Auswirkungen des Austausches gefragt. Mögliche Auswirkungen zur Teamebene (12 Items) und zur Kindebene (13 Items) wurden per Multiple Choice abgefragt. Zur Kontrolle bedingender Faktoren wurden weitere Daten (z. B. die Berufserfahrung), genutzt.

Es wurden 574 Fragebögen versandt. Der Rücklauf betrug 55,2%. Für die Auswertung konnten 311 Fragebögen (174 Fragebögen aus 45 Kindergärten und 137 Fragebögen aus 32 Grundschulen) herangezogen werden. Zur Darstellung der Ergebnisse wurden deskriptive Verfahren verwendet, für Gruppenvergleiche wurden chi² – Verfahren und Verfahren für Rangdaten (Mann-Whitney-U-Test) eingesetzt. Zur Kontrolle bedingender Faktoren wurden Korrelationen genutzt. Signifikanzprüfungen erfolgten zweiseitig.

Die Ergebnisse zeigen, dass der institutionsübergreifende Austausch stattfindet und als bedeutsam angesehen wird. Positive Auswirkungen des Austausches werden von den Beteiligten sowohl auf Teamebene als auch auf Kindebene genannt. Auf Teamebene werden Auswirkungen genannt, die auf ein erweitertes gegenseitiges Verständnis hinweisen und somit positiv auf die Zusammenarbeit wirken könnten. Auf Kindebene werden Auswirkungen genannt, die u. a. auf einen erweiterten Blick auf das Kind, sowie auf eine individualisierendere Praxis schließen lassen. Dies deutet auf einen Kompetenzerwerb durch Austausch- und Reflexionsprozesse mit Bezug zu einer gemeinsamen Praxis hin. Des Weiteren wurden Randfaktoren gefunden, wie z. B. die Berufserfahrung, die die Auswirkungen des Austausches moderieren könnten.

Mittwoch
11. März

15:45 Uhr

Foyer GB - Ost

Mayer, Jürgen; Pupkowski, Vanessa; Roesler, Mariella; Sumfleth, Elke; Walpuski, Maik; Wellnitz, Nicole

Einfluss von Interesse und Motivation in Biologie und Chemie auf Leistungsunterschiede in Kompetenztests

Naturwissenschaftliche Inhalte und deren kontextuelle Einbindung sind für Schülerinnen und Schüler nicht gleichermaßen interessant und motivierend. In Testsituationen wirken sich das Interesse bei der Bearbeitung von Aufgaben sowie die Motivation diese bestmöglich zu lösen auf die Leistungen aus (Prenzel et al., 2007; Liu et al., 2012). Um kognitive Kompetenzen durch schriftliche Tests valide messen zu können, müssen daher motivationale Konstrukte (z. B. Situationales Interesse und Testmotivation) separat theoretisch konzeptionalisiert, gemessen und bei der Interpretation von Kompetenzunterschieden berücksichtigt werden. Ziel dieses Kooperationsprojektes* ist es, den durch motivationale Einflüsse bedingten Anteil von Testleistungsvarianz zwischen den Fächern Biologie und Chemie sowie zwischen den Kompetenzbereichen Fachwissen und Bewertung, unter Berücksichtigung vier identischer Kontexte (Gesundheit, Umwelt, Technik, Natürliche Ressourcen), aufzuklären.

Bei der Untersuchung der Bedeutung von Interesse und Motivation für die Messung naturwissenschaftlicher Kompetenz sind drei differenzierte Fragestellungen leitend: (1) Welcher Anteil an Varianz wird durch die Faktoren situationales Interesse, wahrgenommene Relevanz, Erwartungs- und Wertkomponente, fachbezogenes Fähigkeitsselbstkonzept und Anstrengungsbereitschaft aufgeklärt? (2) Differieren die Zusammenhänge von kognitiven und motivationalen Faktoren zwischen den Fächern Biologie und Chemie sowie zwischen den Kompetenzbereichen Fachwissen und Bewertung? (3) Welche Bedeutung haben Kontexte in Aufgaben für die Ausprägung der motivationalen Faktoren Interesse und Anstrengungsbereitschaft sowie für die Leistung?

Zur Beantwortung der Fragen werden querschnittliche Daten von 1600 Schülerinnen und Schülern der 9. und 10. Jahrgangsstufe verschiedener Schulformen erhoben. Die kognitiven Kompetenzen werden in den Domänen Biologie und Chemie für die Kompetenzbereiche Fachwissen und Bewertung mit 196 offenen und geschlossenen Items im Multi-Matrix-Design erfasst. Als motivationale Konstrukte werden das Interesse (individuelles und situationales Interesse, persönliche Relevanz), die Motivation (Erwartungs- und Wertkomponente), die Anstrengungsbereitschaft und das Fähigkeitsselbstkonzept durch die Anpassung standardisierter Skalen (z. B. Boekaerts, 2002; Sundre, 2007) differenziert nach Fach und Kompetenzbereich, teils kontextbezogen (embedded design) (Drechsel et al., 2011), erfasst. Als Kontrollvariablen werden die sprachlichen Fähigkeiten, die kognitiven Grundfähigkeiten und der Cognitive Load erhoben. Die Datenauswertung erfolgt auf Basis der klassischen Testtheorie und der Item-Response-Theorie.

Daten der Vorstudie von 1235 Schülerinnen und Schülern zeigen in der Biologie für 152 (86 %) von 176 Items zufriedenstellende Item-Fit-Werte und Trennschärfen ($0.80 < \text{MNSQ} < 1.20$, $T \leq 1.96$, $r \geq .28$). Jedes Biologieitem wurde im Mittel von 77.19 (SD = 4.05) Schülerinnen und Schülern beantwortet. Ein Proband hat durchschnittlich 11 (SD = 1.85) Items bearbeitet. Die hohen Zusammenhänge ($r = .74$) zwischen den beiden Kompetenzbereichen verdeutlichen, dass hohe Kompetenzen der Schülerinnen und Schüler beim Umgang mit Fachwissen mit hohen Bewertungskompetenzen einhergehen. Items zum Fachwissen sind schwerer zu lösen ($M = 0.01$, $SD = 1.32$) als Items zur Bewertung ($M = -0.47$, $SD = 1.26$). Die mit einer einfaktoriellen Varianzanalyse geprüfte Nullhypothese, dass die Items aller vier Kontexte gleich schwer sind, wird mit einem mittleren signifikanten Effekt von $F(3, 148) = 2.36$ ($p = .074$, $\eta^2 = .05$) bestätigt. Tendenziell sind Items zum Kontext Technik besonders schwer zu lösen und werden von den Schülerinnen und Schülern ($N = 69$) als weniger interessant eingestuft.

Auf dem Poster werden Studiendesign, Beispielaufgaben und fachübergreifende Ergebnisse aus der Vorstudie präsentiert sowie Implikationen für die Hauptstudie abgeleitet.

* DFG-Sachbeihilfe (MA 1792/6-1; SU 187/12-1; WA 2829/5-1)

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

3

van den Bogaert, Vanessa; Marschner, Jessica; Wilde, Matthias; Wirth, Joachim

Entwicklung eines Messinstruments zum situationalen Interesse und dessen Subkomponenten mit 4 Subskalen

Einleitung: In Metaanalysen konnte der Einfluss von Interesse auf den Lehrerfolg bereits bestätigt werden (Hattie, 2013; Schiefele, Krapp, & Winterler, 1992). In zahlreichen Publikationen wird aufgezeigt, dass dem Interesse lernförderlichere Effekte zugewiesen werden können. Insgesamt verbindet die Studien das Ergebnis, dass interessenbasiert effizienter, lernstrategisch elaborierter und nachhaltiger gelernt werden kann (Schiefele & Schreyer, 1994). Die Person-Gegenstands-Theorie fungiert seit den siebziger Jahren als eine zu den psychologischen Motivationstheorien sich bewusst abgrenzende, alternative (pädagogisch-psychologische) Theorie (u.a. Krapp, 2007). Sie ist die theoretische Basis für die hier beschriebene Studie. Interesse wird als eine Relation zwischen einer Person und einem Gegenstand definiert, zusätzlich ist diese Relation durch festgelegte Merkmale charakterisiert. Ziel der vorgestellten Untersuchung ist die Entwicklung eines reliablen und validen Messinstruments zur Erfassung des situationalen Interesses und dessen Subkomponenten.

Fragestellung: In der Theorie wird zwischen individuellen und situationalen Interessen differenziert (Krapp, 1992, 2002; Krapp, Hidi, & Renninger, 1992). Situationales Interesse ist gekennzeichnet durch die Wechselwirkung zwischen Personen und spezifischen Situationsfaktoren, die erstmals oder nur kurzfristig die Aufmerksamkeit von Personen binden (Schiefele, Krapp, & Schreyer, 1993). Die theoretische Begründung, dass äußere Anreize in Lehr-Lernsettings eine kurzfristig anhaltende Interessenhandlung auslösen können nämlich situationales Interesse (Krapp & Prenzel, 1992), soll mit dem entwickelten Fragebogen empirisch überprüft werden. Angenommen wird, dass eine interessen geleitete Person-Gegenstands-Beziehung dadurch gekennzeichnet ist, dass sich eine Person ohne äußeren Druck mit denen im Rahmen von Lehr-Lernsettings bestehenden Handlungsmöglichkeiten identifiziert (intrinsischer Charakter) (Berg, 1977) und diese darüber hinaus als persönlich relevant eingestuft. Die mit der Interessenhandlung einhergehenden (Lern-) Ziele müssen demnach mit den enthaltenen Einstellungen und Werten der Entwicklung des Selbstkonzeptes kompatibel sein (wertbezogener Aspekt). Die Beweggründe für die Beschäftigung mit einem Gegenstand sind in der Regel gekennzeichnet durch das Verlangen nach Erweiterung von Wissen bzw. individueller Kompetenzen (kognitiv/epistemische Ebene) und einer positiven Gefühlslage vor, während und nach der Interessenhandlung (emotionale Tönung) (Schiefele, 1992).

Vorgehen: Auf der Basis der Person-Gegenstands-Theorie des Interesses kann folglich eine Unterscheidung des situationalen Interesses in vier Komponenten vorgenommen werden, emotionale Tönung, wertbezogener Aspekte, intrinsischer Charakter und kognitiv/epistemische Ebene. In der deutschsprachigen Literatur sind kaum Messinstrumente publiziert, die das situationale Interesse in allen vier Merkmalkategorien erfassen. Das zu pilotierende Messinstrument zum situationalen Interesse umfasst in seiner Ursprungsform 30 Items, die sich zu den vier Skalen mit ähnlicher Itemanzahl zusammenfassen lassen.

Ergebnisse: Vorläufige Ergebnisse einer Konstruktionsstichprobe im Rahmen einer Interventionsstudie im Biologieunterricht (Hofferber & Wilde, 2014) mit N=137 Schülern der sechsten Jahrgangsstufe weisen darauf hin, dass vor allem inverse Items die Analyseergebnisse verzerren. Um die Messqualität zu erhöhen wurden insgesamt neun Items der Analyse entzogen. Die übrigen 21 Items weisen eine interne Konsistenz von $\alpha = .96$ auf. Die faktorielle Validität wurde mittels einer Hauptkomponentenanalyse überprüft. Es zeigt sich eine stabile Drei-Faktoren-Lösung, die 69.61 Prozent der Varianz aufklären kann. Insgesamt liegen – von den Items der Skala emotionale Tönung abgesehen – die Faktorladungen zwischen .54 und .80. Der erste Faktor (wertbezogener Aspekt) klärt 27.5 Prozent der Varianz auf, der zweite Faktor (kognitiv/epistemische Ebene) 24.7 Prozent und der dritte Faktor (intrinsische Qualität) 17.3 Prozent.

Diskussion: In der nächsten Erhebungsphase soll mit einer größeren Analyseschichtprobe das Messinstrument in überarbeiteter Form auf Reliabilität sowie Validität geprüft werden. Die vorläufigen Ergebnisse sprechen zunächst dafür, dass die Vier-Faktoren-Lösung empirisch nicht abbildbar ist, um dies überprüfen zu können, ist die erneute Erhebung notwendig. Nur, wenn die Subkomponente: emotionale Tönung einen eigenen Faktor darstellt, kann damit die theoretische Annahme der vier Subskalen bestätigt werden. Das zu verfolgende Ziel ist die Vier-Faktoren-Lösung mit einer überarbeitenden Form des Fragebogens empirisch zu prüfen.

Basten, Melanie; Wilde, Matthias

Schriftliches Argumentieren im Biologieunterricht Können SuS der 11. Klasse ihre Meinung zu einem bioethischen Thema darlegen?

Das Thema Organspende ist ein bewertungsrelevantes Thema mit Alltagsbezug für Jugendliche ab 16 Jahren, da im Besonderen seit Einführung der Entscheidungsregelung eine individuelle Spendeentscheidung von jedem deutschen Bürger ab 16 Jahren erwünscht ist (BMG, 2012). Die Bildungsstandards im Fach Biologie sehen vor, dass deutsche Schüler mit Abschluss der 10. Klasse dazu in der Lage sind, ein bioethisch relevantes Thema zu bewerten und ein Urteil zu fällen (KMK, 2004). Dabei sollen sie in der Lage sein, ihren eigenen Standpunkt als einen Standpunkt unter mehreren zu verstehen und zu akzeptieren sowie den eigenen Standpunkt argumentativ zu belegen. Um die Bewertungskompetenz im Biologieunterricht lehr- und diagnostizierbar zu machen, wurden Kompetenzmodelle für die ethische (Reitschert & Hößle, 2007) und ökologische (Eggert & Bögeholz, 2006) Bewertungskompetenz aufgestellt (zusammenfassend Hostenbach et al., 2011).

Eigene Untersuchungen haben jedoch gezeigt, dass die schriftliche Erhebung der Bewertungskompetenz von Jugendlichen problematisch ist (zusammenfassend Basten, 2013). Schülerinnen und Schüler der 11. Klasse wurden gebeten, schriftlich zu einer Dilemmageschichte zum Thema Organspende Stellung zu beziehen. Dabei wurde der Schreibprozess durch Leitfragen, die sich am Kompetenzmodell zur ethischen Bewertung (Reitschert & Hößle, 2007) orientieren, strukturiert. Mithilfe qualitativer Inhaltsanalyse (Mayring, 2007) wurden die schriftlichen Antworten der Schüler deduktiv und induktiv in Beurteilergruppen ausgewertet. Die deduktiven Kategorien waren durch die Niveaustufenausdifferenzierungen der Teilkompetenzen der ethischen Bewertung (u.a. Reitschert, 2009) vorgegeben. Für den Kodierleitfaden wurden jedoch induktiv Kodierregeln für das spezifische Thema Organspende in einer ersten Stichprobe entwickelt und teilweise in einer zweiten Stichprobe angewendet. Hier zeigten sich Probleme hinsichtlich der Anwendbarkeit der Niveaustufen der Teilkompetenzen auf das neue Thema, heterogene Beurteilerübereinstimmungen aufgrund des schwierigen Materials und insgesamt sehr niedrige erreichte Niveaustufen der Schülerinnen und Schüler der 11. Klasse. Diskussions- und Verbesserungsansätze finden sich bei der Auswahl der Thematik, dem Aufforderungscharakter der Aufgaben, der tatsächlich erwartbaren Kompetenz von Jugendlichen im Alter von 16 Jahren und der Ausdifferenzierung des Auswertungsschlüssels.

In einem nächsten Schritt wurde nun die Strukturierung des Schreibprozesses verändert, um auszuschließen, dass die geringen von den Schülerinnen und Schülern erreichten Niveaustufen dem zu geringen Aufforderungscharakter der Aufgabe und Leitfragen geschuldet waren. In Interviews zur eigenen Position zum Thema Organspende hatten sich durchaus höhere Niveaustufen bei Schülerinnen und Schülern der 12. Klasse gezeigt (Vieregge, 2011). Demzufolge ist es möglich, dass den Schülerinnen und Schülern durch die bisherige Aufgabenstellung und Leitfragen die von ihnen erwartete Anforderung nicht klar wurde. Da sie zudem vor dem Schreiben auch nicht mit dem Kompetenzmodell vertraut gemacht wurden, konnten sie auch nicht gezielt die Anforderungen der Niveaustufungen erfüllen. Die Dilemmageschichte und die dazugehörige Aufgabenstellung wurde für die neue Untersuchung in eine Coverstory eingebunden, in der die Schüler als Freunde eines Verstorbenen dessen uneinigen Verwandten darin beraten sollen, ob die Organe des Hirntoten gespendet werden sollen oder nicht. Die Strukturierung des Schreibprozesses durch die Leitfragen zeichnet sich nun durch eine klare Verknüpfung von Pro- und Conraseiten bzw. Perspektiven mit den sich uneinigen Verwandten aus. Dadurch soll die Produktion von Argumenten beider Seiten durch die Schülerinnen und Schüler angeregt werden. Im Anschluss an die Leitfragen werden die Schülerinnen und Schüler in der neuen Version nun gebeten, einen Brief an die Verwandten des Verstorbenen zu schreiben, in dem sie ihre Gedanken und ihren Ratschlag darlegen sollen. So sollen die Schülerinnen und Schüler dazu angeregt werden, auch ihre Abwägungsprozesse bei der Urteilsfindung darzulegen. Beide erbrachten Leistungen würden zu einer Einordnung in höhere Niveaustufen der Bewertungskompetenz führen. Im Vergleich dazu wird eine zweite Version der in die Coverstory eingebetteten Dilemmageschichte untersucht, in der die Schülerinnen und Schüler sich frei Notizen machen können, bevor sie den Brief an die Ratsuchenden schreiben. Am Poster sollen erste Ergebnisse der Untersuchung vorgestellt und diskutiert werden.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Heitmann, Patricia; Ziemke, Anne; Hecht, Martin

Argumentative Schreibaufgaben für die naturwissenschaftlichen Fächer und das Fach Deutsch – Untersuchung der Fachspezifität aus der Sicht von Schülerinnen und Schülern

Theoretischer Hintergrund: Die Fähigkeit zu Argumentieren gilt als zentrales Element der Wissenskonstruktion und Meinungsbildung. In allen Unterrichtsfächern wird Argumentieren für Lernprozesse vorausgesetzt und geübt (Ludwig & Spinner, 2000), allerdings ist unklar, inwiefern das Argumentieren über unterschiedliche Fächer hinweg auf vergleichbare Weise verstanden und praktiziert wird. Zum einen müssen Teilfähigkeiten und Wissensaspekte der Schreibkompetenz – unter die das argumentierende Schreiben fällt – für die jeweiligen Fächer in den Blick genommen werden. Stephany, Linnemann und Becker-Mrotzek (2013) betrachten diese als wenig fachübergreifend, zumal beispielweise ein fachspezifischer Wortschatz beim Schreiben angewandt werden muss. Zum anderen spielen schulische Fachkulturen in Form von Vorstellungen und Erwartungen hinsichtlich eines Fachs im schulischen Alltag eine wesentliche Rolle und wirken sich handlungsleitend auf Lernprozesse aus (Hericks & Körber, 2007; Willems, 2011): Während Deutschunterricht von Lernenden als kommunikativ, kollektiv ausgerichtet und gleichzeitig individuell gestaltbar konstruiert wird (Willems, 2011), assoziieren Lernende insbesondere die naturwissenschaftlichen Fächer Chemie und Physik häufig mit Merkmalen, die stark auf inhaltliche Aspekte abzielen wie „richtige Lösung“ und „Stoffmenge“ (Haag & Götz, 2012). Bezüglich der Fachkultur lässt sich vermuten, dass der Prozess des (argumentativen) Schreibens für den Deutschunterricht aus Schülersicht eine wichtige Rolle spielt; diesem wird aber in den naturwissenschaftlichen Fächern eine geringere Bedeutung beigemessen (Steiner, 2014). Hier herrscht offenbar die Vorstellung, dass Inhalte beziehungsweise deren fachliche Korrektheit zentraler sind.

Fragestellung: Im Rahmen dieser Studie wird der Frage nachgegangen, inwiefern Gemeinsamkeiten und Unterschiede in der Argumentierfähigkeit für verschiedene Fächer bestehen, wobei zunächst geprüft wird, ob Schülerinnen und Schüler den Fachbezug einer argumentativen Schreibaufgabe wahrnehmen: Erkennen Schülerinnen und Schüler bei argumentativen Schreibaufgaben unterschiedlicher Fächer mit gleicher Thematik, dass es sich um Aufgaben eines spezifischen Fachs handelt?

Diese Forschungsfrage ist insofern relevant, als dass angenommen wird, dass nach der Identifikation des Fachs ein entsprechendes Fachschema aktiviert wird. Beispielsweise könnte das Vorhandensein naturwissenschaftlicher Informationen (unbewusst) spezielle Assoziationen aufgrund von Vorerfahrungen im Unterricht hervorrufen und zu einer stark am Inhalt orientierten Argumentation führen. Aufbauend auf den Ergebnissen dieser Studie soll nachfolgend untersucht werden, inwiefern sich fachspezifische Assoziationen auf die Qualität des Argumentierens auswirken.

Methode: Insgesamt wurden vier argumentative Schreibaufgaben entwickelt, wobei jeweils zwei Themen (Ernährung und Tattoo) je Fach (Deutsch und Naturwissenschaften) verwendet wurden. Die Aufgaben der Naturwissenschaften unterscheiden sich von denen des Fachs Deutsch dahingehend, dass diese naturwissenschaftliche Informationen beinhalten.

Die Stichprobe setzt sich aus 138 Schülerinnen und Schüler der 9. bis 12. Jahrgangsstufe zusammen. Diesen wurden jeweils zwei der vier Aufgaben unter Verwendung eines vollständig gekreuzten, balancierten Testdesigns vorgelegt. Die Anforderung bestand darin, die Aufgaben sowohl mit Hilfe von dichotomen Items als auch Items mit 5-stufigen Likert-Skalen einem bestimmten Unterrichtsfach (Deutsch oder Naturwissenschaften) zuzuordnen. Die Auswertung der Daten erfolgte mit deskriptiven und inferenzstatistischen Analysen der relativen Häufigkeiten der dichotomen Zuordnungssitems sowie der zentralen Tendenz und Mittelwertunterschieden auf den Likert-Skalen.

Ergebnisse und Ausblick: Die Befunde deuten darauf hin, dass Schülerinnen und Schüler die Fachspezifität von argumentativen Schreibaufgaben wahrnehmen, das heißt Aufgaben aus dem jeweiligen Fach werden überzufällig häufig als Aufgaben des jeweiligen Fachs erkannt. Dieses Ergebnis wird zusätzlich durch hypothesenkonforme Mittelwertdifferenzen auf den Likert-Skalen gestützt.

Die Studie zeigt auf, dass Schülerinnen und Schüler Aufgaben fachadäquat klassifizieren können. Aufbauend auf diesen Ergebnissen wird in einer Folgestudie untersucht, inwieweit die Argumentationsqualität von fachspezifisch assoziierten Annahmen der Lernenden in Form schulischer Fachkulturen abhängt. Das Konzept der geplanten Folgestudie soll ebenfalls auf dem Poster präsentiert werden.

Schuttkowski, Caroline; Schmitz, Anke; Gräsel, Cornelia; Rothstein, Björn

Die Wirksamkeit der Textkohäsion auf das Textverständnis bei Sachtexten – erste Befunde einer Studie zum Lauten Denken

Eine effiziente und auf Lesevoraussetzungen ausgerichtete Gestaltung von Schulbuchtexten ist eine der wesentlichen Einflussgrößen auf die Unterrichtsqualität (vgl. Helmke, 2012). Textkohäsion gilt als ein Merkmal verständlicher Texte. Um einen kohärenten Text zu erhalten, ist ein formaler Zusammenhalt innerhalb des Textes durch sprachliche Markierungen bedeutsam. Diese sprachlichen Realisierungen von Textverbindungen werden als Kohäsion bezeichnet (Schnitz 2006). In bisherigen Studien wurde die Wirkung der Textkohäsion auf das Textverständnis mit Verständnistests nach dem Lesen der Texte überprüft (vgl. Ozuru, Dempsey & McNamara, 2009; Rothstein et al., 2014). Wie Leser/innen mit unterschiedlich kohäsiven Texten umgehen und wie sie die einzelnen Kohäsionsmarker erfassen und nutzen, ist bislang unaufgeklärt. Daher untersuchen wir explizit Leistung und Zusammenspiel isolierter Kohäsionsmarker und fokussieren Tempora und Temporaladverbialen (Rothstein, 2014). Diese werden sowohl sprachwissenschaftlich als auch sprachdidaktisch differenziell analysiert und im Hinblick auf ihre Wechselwirkungen mit den individuellen Lesevoraussetzungen untersucht. Wir gehen davon aus, dass ein ausgeprägtes Vorwissen über temporale Kohäsionsmarker das Textverständnis erleichtert, und nehmen an, dass sich deren gezielter Einsatz in Schulbuchtexten insbesondere auf das Textverständnis von Schüler/innen mit guten Lesevoraussetzungen förderlich auswirkt.

Der vorliegende Beitrag vergleicht die Verarbeitung von Texten bei Schüler/-innen mit schlechteren und besseren Lesevoraussetzungen (Jahrgangsstufe 9 in Gesamtschulen). Im Rahmen der Analyse wird ermittelt, wie die Schüler/-innen mit unterschiedlichen Lesevoraussetzungen auf die Kohäsionsmarker in einem exemplarischen Schulbuchtext reagieren, ob sie deren Funktion kennen und ob sie ihr grammatikalisches/lexikalisches Vorwissen zielführend im Leseprozess einsetzen können. Dazu wird das introspektive Verfahren des Lauten Denkens (Ericsson & Simon 1984) genutzt, das bereits erfolgreich im Hinblick auf Denk- und Problemlöseprozesse, insbesondere im naturwissenschaftlichen Bereich oder in der Fremdsprachendidaktik praktiziert wird, jedoch nur wenig Anwendung in lese- bzw. deutschdidaktischen Fragestellungen findet. In Einzelgesprächen (N = 40) werden die Schüler/-innen mit zwei inhaltlich identischen, aber unterschiedlich kohäsiven Versionen eines Schulbuchtextes (mit und ohne Tempora und Temporaladverbialen) konfrontiert (within-Design). Sie werden dazu aufgefordert, ihre Gedanken während des Bearbeitungsprozesses laut zu äußern. Auf diese Weise werden sowohl die Wahrnehmung kohäsiver Textelemente und individuelle Probleme beim Textverstehen aufgedeckt als auch der Leseprozess direkt begleitet. Die individuellen Lesevoraussetzungen werden dabei durch die Erhebung der Lesefähigkeit (LGVT 6-12; Schneider, Schlagmüller & Ennemoser, 2007), des thematischen Vorwissens und des Wissens über Kohäsionsmarker gemessen. Ein Verständnistest gibt Aufschluss über das Textverständnis und die Wirkung der temporalen Kohäsionsmarker. Erste Befunde zur Verarbeitung der unterschiedlich kohäsiven Texte in Abhängigkeit individueller Lesevoraussetzungen werden im Vortrag präsentiert.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Kotzerke, Marei; Brinkhaus, Moti; Grabowski, Joachim

Psychologische Determinanten der Schreibkompetenz

"Seit jeher befindet sich die empirische Schreibforschung im Spannungsfeld der interdisziplinären Zuständigkeit, da es sich beim Prozess des Schreibens um einen komplexen Vorgang handelt, an dem psychologische und linguistische Faktoren beteiligt sind. Ein wichtiges Bildungsziel ist der Aufbau von Schreibkompetenz, um die Teilhabe an der Gesellschaft sichern zu können. Darunter verstehen wir die Fähigkeit, funktional angemessene Texte zu produzieren (Becker-Mrotzek, 2014). Diese Fähigkeit setzt sich aus zahlreichen kognitiven, sprachlichen, motivationalen und affektiven Komponenten zusammen. Welche spezifischen Beiträge kann die Psychologie leisten, um individuelle Ausprägungen von Schreibkompetenz vorherzusagen und didaktische Maßnahmen zu unterstützen, mit denen sich Schreibfähigkeiten geeignet vermitteln und aufbauen lassen? Ein besonderes Interesse gilt hier Schülerinnen und Schülern mit sprachlichem Migrationshintergrund, deren sprachliche, nicht aber notwendigerweise kognitiv-intellektuelle Fähigkeiten vergleichsweise schwächer ausgeprägt sind.

Einerseits erweisen sich intelligenznahe Fähigkeitsaspekte, wie bei vielen anderen Kompetenzbereichen auch, als generelle Prädiktoren für Schreibkompetenz (Grabowski, 2014a). Insofern sich solche Fähigkeitsfacetten aber nur schwer beeinflussen lassen, tragen sie für sprachdidaktische Maßnahmen wenig bei. Am anderen Ende der Prozessvoraussetzungen stehen die spezifisch einzelsprachlichen Fähigkeiten, deren Vermittlung dem Sprachunterricht zuzuordnen sind. Im vorliegenden Beitrag konzentrieren wir uns auf die Rolle psychologisch beschreibbarer Fähigkeitsaspekte, die zwischen diesen Polen allgemeiner Intellektualität und einzelsprachlicher Spezifität liegen. Damit soll aufgezeigt werden, dass und in welcher Weise die Untersuchung der Schreibkompetenz auch eine Frage der (kognitiven) Psychologie darstellt (vgl. Grabowski, 2014b). Wir beziehen uns dazu auf zwei aktuelle Forschungsprojekte im schulischen Interventionskontext.

Im interdisziplinären Forschungsprojekt „Teilkomponenten von Schreibkompetenz“ (im Rahmen der Forschungsinitiative Sprachdiagnostik und Sprachförderung „FISS“ des BMBF) wurde anhand von knapp 300 Datensätzen von Schülerinnen und Schülern aus den Jahrgangsstufen 5 und 9 an den Schultypen Haupt-, Realschule und Gymnasium nachgewiesen, dass insbesondere die Fähigkeit zur Herstellung von Kohärenz und die Fähigkeit zur Perspektivenübernahme starke und sprachunabhängige Prädiktoren von Textqualität (als Indikator für Schreibkompetenz) sind (Knopp et al., 2013). Wir berichten über die Entwicklung der zugehörigen Aufgaben (von allgemeinen kognitiven bis hin zu sprachspezifischen Anforderungen) und zeigen die textsortenübergreifende Rolle dieser Fähigkeiten anhand von Regressionsanalysen.

Im zweiten empirischen Teil des Vortrags wird ein interdisziplinäres empirisches Forschungsprojekt zur individuellen Schreibberatung (aus der Förderlinie des Mercator-Instituts für Sprachförderung und Deutsch als Zweitsprache) vorgestellt, dessen Ziel die Evaluation einer Beratungskartei ist, die zur standardisierten Rückmeldung zu Sachtexten von Schülerinnen und Schülern der sechsten Klassenstufe entwickelt wurde. Wir konzentrieren uns insbesondere auf die Darstellung der erhobenen Variablen, die im Zusammenhang mit der Entwicklung der Qualität der Schülertexte über ein halbes Jahr hinweg betrachtet wurden. Hierbei wurden neben der sprachabhängigen Leseflüchtigkeit und dem Wortschatz sprachfernere Aspekte wie die motorische Schreibflüchtigkeit und das Arbeitsgedächtnis erhoben. Anhand der Daten von etwa 150 Schülerinnen und Schülern der 6. Jahrgangsstufe an einer internationalen Gesamtschule wird untersucht, welche spezifischen Zusammenhänge diese Variablen mit der Wirksamkeit der Förderung von Schreibkompetenz durch Schreibberatung aufweisen.

Der Vortrag schließt mit einer Diskussion über die Bedeutung von psychologischen Prädiktoren der Schreibkompetenz und deren Stellenwert für die Förderung von Schreibfähigkeiten im schulischen Bildungssystem und die Entwicklung und Evaluation zugehöriger Interventionsmaßnahmen.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Stiller, Cornelia; Hahn, Stefan; Stockey, Andreas; Wilde, Matthias

Kann durch das BaNa-Kurskonzept eine autonomie- und kompetenzunterstützende Lernumgebung geschaffen werden?

Ziel naturwissenschaftlicher Grundbildung (Scientific Literacy) ist es, SchülerInnen Kompetenzen zu vermitteln, durch die diese in der Lage sind, naturwissenschaftliches Wissen nicht nur zu besitzen, sondern auch anwenden zu können (NRC, 1996; OECD 2010). Das Nachvollziehen der wissenschaftlichen Erkenntnisgewinnung durch selbstgesteuertes Experimentieren kann eine geeignete Methode darstellen, um diese Ziele zu erreichen und bei den SchülerInnen den Grundstein für eine Scientific Literacy über ein Verständnis der Scientific Inquiry und Nature of Science zu legen (Mayer & Ziemek, 2006; Pfangert-Becker, 2010; Wirth, Thillmann, & Künsting, 2008). Nach der Selbstbestimmungstheorie ist eine Voraussetzung für motiviertes Handeln, dass die drei psychologischen Grundbedürfnisse nach Autonomie, Kompetenz und sozialer Eingebundenheit erfüllt sind (Deci & Ryan, 2000). Dadurch wird im schulischen Kontext bewirkt, dass sich SchülerInnen eher mit den Lerngegenständen auseinandersetzen und gute Lernleistungen zeigen (Niemiec & Ryan, 2009).

Experimentieren ermöglicht es Lernenden, selbstständig naturwissenschaftliche Fragen zu untersuchen und zu beantworten. Diese Befriedung des psychologischen Bedürfnisses nach Autonomie ist nach Deci und Ryan (2000) ein wichtiger Faktor für einen optimalen Lernprozess. Der Grad der Selbstständigkeit beim Experimentieren an die Ausgangslagen der SchülerInnen angepasst werden, um diese nicht zu überfordern (Mayer & Ziemek, 2006), einem hohen „cognitive load“ (Kirschner, Sweller & Clark, 2006) entgegen zu wirken und die eigene Kompetenzwahrnehmung zu steigern. In einem an einer Versuchsschule entwickelten, fächerübergreifenden, einjährigen Naturwissenschaftskurs (Basiskurs Naturwissenschaften, BaNa) für die Eingangsphase der Oberstufe liegt der Fokus auf dem im Kursverlauf sukzessive gesteigertem selbstständigen Experimentieren. Ziel dieser Studie ist es zu untersuchen, ob es durch das Kurskonzept gelingt, für die SchülerInnen eine autonomie- und kompetenzunterstützende Lernumgebung zu schaffen.

Die Studie ist in einem Längsschnittdesign angelegt. Zu Beginn und am Ende des Schuljahres 2012/13 bearbeiteten die SchülerInnen der Treatmentgruppe (Teilnahme am BaNa, n=81) und der Vergleichsgruppe (Regelschule, n=126) einen Fragebogen mit Skalen zu den wahrgenommen Lernaktivitäten („SchülerInnen forschen“, „Experimentieren“, „interaktives Lehren und Lernen“, „naturwissenschaftliche Anwendungen“) und Lehr- und Lernbedingungen („Kompetenzunterstützung“, „Autonomieunterstützung“, „soziale Eingebundenheit“) (Frey, 2009). Beide Gruppen wiesen eine ähnliche Geschlechterverteilung (Treatmentgruppe: 52,8% weiblich, Vergleichsgruppe: 53,91 % weiblich) und einen vergleichbaren sozioökonomischen Status (Werte auf der ISEI-Skala: Treatmentgruppen: 53,91; Vergleichsgruppe: 53,37) auf.

Bei den Skalen zu den Lernaktivitäten zeigten sich signifikante Interaktionseffekte zwischen den Messzeitpunkten und der Gruppenzugehörigkeit in der Skala „Experimentieren“ ($F(1;204) = 26.14$, $p < .001$, partielles $\eta^2 = .15$) und „Forschen“ ($F(1;204) = 17.80$, $p < .001$, partielles $\eta^2 = .09$). Die SchülerInnen der Treatmentgruppe erlebten demnach in der Eingangsphase der Oberstufe mehr Aktivitäten, die sich auf „Experimentieren“ ($M_2=2,82$) und „Forschen“ ($M_2=2,22$) bezogen, als in ihrem vorherigen Schuljahr (Experimentieren: $M_1=2,48$, Forschen: $M_1=1,78$). Bei den SuS der Vergleichsgruppe dahingegen blieb der Anteil an experimentierenden und forschenden Aktivitäten relativ konstant bzw. sank teilweise sogar in der Eingangsphase der Oberstufe im Vergleich zum vorherigen Schuljahr („Forschen“: $M_1=1,83$, $M_2=1,81$; „Experimentieren“: $M_1=2,69$, $M_2=2,53$). In den Skalen „Anwendung“ und „Interaktiv“ gab es keine signifikanten Interaktionseffekte und Haupteffekte. Die signifikanten Interaktionseffekte in den Skalen „soziale Eingebundenheit“ ($F(1;181) = 4.37$, $p < .05$, partielles $\eta^2 = .02$) und „Autonomieunterstützung“ ($F(1;189) = 5.13$, $p < .05$, partielles $\eta^2 = .03$) deuten darauf hin, dass die SchülerInnen der Treatmentgruppe („soziale Eingebundenheit“: $M_1=2,61$, $M_2=2,94$; „Autonomieunterstützung“: $M_1=2,67$, $M_2=3,05$) im Gegensatz zu der Vergleichsgruppe („soziale Eingebundenheit“: $M_1=2,80$, $M_2=2,86$; „Autonomieunterstützung“: $M_1=2,94$, $M_2=3,05$) diese Lehr- und Lernbedingungen in der Eingangsphase positiver wahrnahmen als in ihrem naturwissenschaftlichen Unterricht bisher. In der Skala „Kompetenzunterstützung“ konnten keine signifikanten Effekte gefunden werden. Die Ergebnisse geben Hinweise darauf, dass durch das Kurskonzept günstige Lehr- und Lernbedingungen geschaffen wurden, möglicherweise auch gerade weil die Erfahrungen der SchülerInnen im naturwissenschaftlichen Unterricht an der vorherigen Schule negativer geprägt waren als an den Vergleichsschulen. Warum es durch das Kurskonzept nicht gelungen ist, auch die Wahrnehmung der Kompetenzunterstützung bei den SchülerInnen zu begünstigen, soll unter anderem am Poster diskutiert werden.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Crossley, Antony

Zum Einfluss von Hausaufgaben auf die Lernleistung im Nebenfach Physik

Einleitung und Forschungsfrage: In den Hauptfächern werden Hausaufgaben (HA) oft als obligatorisch angesehen. Die Lernwirksamkeit dieser Maßnahmen wird von LehrerInnen und auch von Eltern i.d.R. nicht in Frage gestellt (vgl. Kohler, 2013). Eine Befragung unter PhysiklehrerInnen ergab, dass Lehrkräfte im Fach Physik hingegen keineswegs diese eindeutige Positionierung vertreten. Dennoch erteilen mehr als die Hälfte der befragten Lehrkräfte regelmäßig HA im Physikunterricht (vgl. Crossley & Starauschek, 2014). Ein vergleichbarer Befund zur Häufigkeit des Einsatzes von HA findet sich auch für den Chemieunterricht (vgl. Sumfleth, Kieren & van Ackeren, 2011). Daher spielen HA auch in den Naturwissenschaften eine entscheidende Rolle für den Lernerfolg. Dies wirft für die Domäne Physik die Frage auf, ob sich zentrale Ergebnisse der Hausaufgabenforschung auf das Nebenfach Physik übertragen lassen.

Metastudien berichten über einen positiven Zusammenhang zwischen Lernerfolg und HA (vgl. Hattie, 2009). Im Detail wirkt sich insbesondere die regelmäßige Bearbeitung von HA positiv auf die Schulleistungen aus, sowie die „Qualität“ der HA, z. B. über die Passung zum Unterricht (vgl. Trautwein, 2008; vgl. Kohler, 2013). Allerdings basieren die Befunde der Hausaufgabenforschung hauptsächlich auf Daten von Studien in den Hauptfächern, z. B. dem Mathematikunterricht. Für die Naturwissenschaften im Allgemeinen und den Physikunterricht im Speziellen liegen nur sehr wenig empirische Daten vor. Dies wird insbesondere dadurch mehr als deutlich, dass in Standardwerken der deutschsprachigen Physikdidaktik der Themakomplex HA im Physikunterricht nicht diskutiert wird (vgl. Hopf et al., 2011; vgl. Kircher et al., 2009). Um diese Forschungslücke zu schließen, sollen in einem ersten Schritt die folgenden beiden Forschungsfragen untersucht werden: (1) Führt die Bearbeitung von HA im Fach Physik zu besseren Lernleistungen? (2) Führt die regelmäßige Bearbeitung von HA zu höheren Lernleistungen, als unregelmäßige bzw. die gelegentliche HA-Bearbeitung?

Design: Es liegt eine quasiexperimentelle Interventionsstudie mit PRE- und POST-Test und einer natürlichen hierarchischen Datenstruktur vor. Die Intervention erfolgte in der Wärmelehre in Klassestufe 9 und dauerte durchschnittlich sieben bis acht Unterrichtswochen. Dabei bearbeiteten SchülerInnen HA in einem Onlineportal. Über die Verwendung des Onlineportals konnte u. a. bestimmt werden, ob das HA-Angebot von den SchülerInnen überhaupt angenommen wurde, wie oft bzw. regelmäßig die gestellten HA bearbeitet wurden und die HA-Bearbeitungszeit. Die Messung der abhängigen Variablen erfolgte zu zwei Messzeitpunkten mit einem erprobten Fragebogen zu Basiskonzepten der Wärmelehre (vgl. Crossley, 2012). Verwendet wurde eine inhaltsvalide und reliable Skala mit 21 Items ($\alpha_{\text{PRE}}=.71$, $\alpha_{\text{POST}}=.77$). Zum Ausschluss von Alternativerklärungen wurden neben dem physikspezifischem Vorwissen u. a. folgende Variablen kontrolliert: Alter, Geschlecht, Intelligenz, Selbstkonzept in Physik, Interesse und Motivation an Physik (zsf. Crossley, 2012), sowie Variablen zum Umgang mit HA.

Stichprobe: An der Studie nahmen $N = 910$ SchülerInnen aus 35 Gymnasialklassen der 9. Klassenstufe teil, die aus unterschiedlichen Regionen in Baden-Württemberg rekrutiert wurden.

Ergebnisse: Etwa ein Drittel der SchülerInnen bearbeiteten keine HA. Die Klassen weisen angesichts der Homogenität der Stichprobe eine große Varianz in der abhängigen Variable auf (Intraklassenkorrelation $\text{ICC}_{\text{POST}} = .11$), sodass die hierarchischen Datenstruktur bei der Datenanalyse berücksichtigt werden muss. Die Mehrebenenanalyse zeigt, dass 1) SchülerInnen die keine HA bearbeiten, einen signifikant schlechteren Score im POST-Test erzielen als Lernende, die das HA-Angebot annehmen. 2) SchülerInnen, die das HA-Angebot regelmäßig annehmen, erzielten durchschnittlich einen signifikant höheren Score im Post-Test als Lernende, die HA selten bzw. unregelmäßig bearbeiten. Diese Ergebnisse bleiben stabil, auch wenn weitere Kontrollvariablen als Prädiktoren in das Mehrebenenmodell aufgenommen werden.

Zusammenfassung: Die Bearbeitung von HA kann auch die Leistungsentwicklung im Nebenfach Physik unterstützen. Zu klären bleibt jedoch, in wie weit sich weitere Befunde der Hausaufgabenforschung im Fach Physik reproduzieren lassen und ob sich diese Ergebnisse auch auf andere Themengebiete der Physik übertragen lassen.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Mahajan, Nimisha; Thillmann, Hubertina; Wirth, Joachim

Examining the Stability of Achievement Goals: The Role of Trait versus State Goal Orientations

Research indicates that student's approach achievement-related tasks with goals oriented toward either mastery (i.e., development of competencies) or performance (i.e., demonstration of competencies) (Ames, 1992; Dweck & Leggett, 1988). Although the literature on goal orientation has shown that pursuing each type of goal is associated with different characteristics and behaviors but, there has been a controversy surrounding the stability of goal orientation constructs (Button et al., 1996; DeShon & Gillespie, 2005; Payne, 2007). Therefore, additional research is needed to better understand the stability of goal orientation constructs in terms of whether these goals reflect dispositional tendencies or are a function of the task itself. The present study contributes to this understanding by evaluating student's trait and state goal orientation in order to determine whether there is a difference between these constructs, and the nature of this difference. The study followed a 2 (mastery goal orientation: trait vs. state) x 2 (performance goal orientation: trait vs. state) between-subjects design. 220 students (from 8th and 9th grade of German schools) did self-regulated learning by conducting experiments within a computer-based learning environment. Student's completed trait and state questionnaires designed to measure their achievement goals. The trait items assessed student's dispositional tendency while state items assessed student's goals relevant to the specific learning task (for task we took a topic from the Physics subject). Three complementary data-analytic techniques were used to examine goal stability: differential continuity, mean-level change, and individual-level change (Fryer & Elliot, 2007; Muis & Edwards, 2009). We were also interested in the extent to which students' trait vs. state mastery goal orientations and trait vs. state performance goal orientations predicted changes in their learning outcome, knowledge application after learning, and strategy use during learning and for this a series of hierarchical regression analyses were conducted.

The results from our study support the assertion that state and trait goal orientation are distinct constructs. More precisely, student's levels of mastery and performance goal orientations examined at two points in time (trait vs. state) supports our hypothesis that a difference between the two exists. When changes did occur, magnitudes of those changes ranged from moderate to large. Further, we found that situational mastery goal orientation matters. Specifically, student's who focused on learning and understating of the given task had a strongest predictive relationship with learning outcome, knowledge application after learning, and strategy use during learning.

We find these results to be noteworthy for a number of reasons. Firstly, issues regarding goal orientation stability are critically important in terms of educational research and practice. Secondly, the present research specifies that dispositional tendencies do appear to be distinguishable from situational goals. Thirdly, results from our study shed light on how often these constructs should be measured. Since we found goal orientations susceptible to change as a result of situational influences, therefore we would suggest researchers to select appropriate times for measurement of these constructs. Fourthly, results from regression analyses indicate that mastery goal orientation (both trait and state) is associated with positive outcomes and therefore, emphasis should be on structuring classrooms that foster and facilitate mastery goal orientation among learners.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Hochdörffer, Katrin; Wahle, Caroline Verena; Schrader, Friedrich-Wilhelm

Gleich und Gleich gesellt sich gern? Effekte von homo- bzw. heterogener Zusammensetzung auf Aufgabenwahrnehmung und Lernerfolg in dyadischen Schülerinteraktionen

Verschiedene empirische Studien konnten belegen, dass kooperatives Lernen positive Auswirkungen auf Leistung, Persistenz und Einstellungen von Schülerinnen und Schülern hat (Hattie, 2009). In diesem Zusammenhang stellt sich die Frage, unter welchen Bedingungen kooperatives Lernen besonders erfolgreich ist. Ein Einflussfaktor auf Lernprozesse und -ergebnisse in kooperativen Lernsettings sind die individuellen Merkmale der beteiligten Gruppenmitglieder (Nijhof & Kommers, 1985). Die Zusammenhänge zwischen homogener bzw. heterogener Gruppenzusammensetzung und Lernerfolg wurden in der empirischen Lehr-Lern-Forschung vorwiegend in den 80er und 90er Jahren untersucht (Springer, Stanne & Donovan, 1999). In diesen Studien fanden vor allem Variablen wie Geschlecht, ethnische Zugehörigkeit und kognitive Fähigkeiten Beachtung. Andere potentiell relevante Schülermerkmale, wie beispielsweise metakognitive, motivationale und Persönlichkeitsmerkmale, wurden bisher wenig berücksichtigt. Anders stellt sich die Lage jedoch im Bereich der Wirtschaftspsychologie dar. Dort existieren eine Reihe von Einzelstudien und Metaanalysen, die den Zusammenhang zwischen Teamzusammensetzung und Teamleistung untersucht haben. In diesen Studien wird zwischen offensichtlichen (z. B. Geschlecht und Alter der Teammitglieder) und weniger offensichtlichen Variablen (z. B. kognitive Fähigkeiten, Persönlichkeitsmerkmale und Überzeugungen) unterschieden. Es konnte gezeigt werden, dass sog. „deep-level“ Variablen (also weniger offensichtliche Variablen) die Leistung eines Teams vorhersagen können, und dass eine heterogene Zusammensetzung des Teams im Hinblick auf diese Variablen positiv mit der Leistung des Teams korreliert (Bell, 2007; Horwitz & Horwitz, 2007). Auf der Basis dieser Ergebnisse ist es das Ziel der vorliegenden Studie, zu untersuchen, ob sich homo- bzw. heterogen zusammengesetzte Schülerdyaden in Bezug auf Aufgabenwahrnehmung und Lernzuwachs unterscheiden. Als unabhängige Variablen sollen dabei vor allem die Unterschiede bezüglich Intelligenz, Persönlichkeitsmerkmalen und Fähigkeitsselbstkonzept der Schülerinnen und Schüler in den Blick genommen werden.

Die Stichprobe umfasst 244 Zehntklässler aus insgesamt 10 Schulklassen. Das Durchschnittsalter beträgt 15,6 Jahre (SD = 0,66). 56,1 Prozent der Studienteilnehmer sind männlich, 39,3 Prozent weiblich. Im Rahmen der Studie arbeiteten die Schülerinnen und Schüler an einer Reihe von komplexen Experimenten zum Thema „Solarenergie“. Der Lernerfolg wurde mithilfe eines eigens auf den Inhalt der Experimente abgestimmten Wissenstest gemessen. Dabei wurde der Test sowohl in einer Prä-Messung vor den Experimenten als auch nach deren Durchführung eingesetzt. Nach dem Ausschluss unvollständiger Datensätze können 86 Schülerdyaden in die Analysen einbezogen werden.

Aufgrund der dyadischen Datenstruktur und der daraus resultierenden Abhängigkeiten werden die Daten mithilfe von Multi-Level-Modellen analysiert (Kenny, Kashy & Cook, 2006). In der Literatur finden sich viele verschiedene Indices, um die Homo- bzw. Heterogenität einer Dyade zu operationalisieren. In der vorliegenden Arbeit wird in diesem Zusammenhang sowohl auf den Dyadenmittelwert, als auch auf die absolute Differenz zwischen beiden Dyadenmitgliedern zurückgegriffen. Erste vorläufige Analysen zeigen, dass eine heterogene Dyadenzusammensetzung positiv mit dem Lernerfolg korreliert. Weiterführende Ergebnisse werden auf der Konferenz präsentiert.

Mittwoch
11. März

15:45 Uhr

Foyer GB - West

Zeuch, Nina; Förster, Natalie; Souvignier, Elmar

Regelgeleitete Aufgabenkonstruktion am Beispiel eines Lesekompetenztests für zweite Klassen

Theoretischer Hintergrund: Lernverlaufsdagnostik (LVD) als eine Variante formativen Assessments stellt diagnostische Informationen über den Lernverlauf von Schülerinnen zur Verfügung, anhand derer Unterricht für die ganze Schulklasse angepasst oder auch individuelle Förderentscheidungen getroffen werden können. Weiterhin bietet LVD eine Möglichkeit zur Untersuchung von Lernprozessen und Einflussfaktoren auf die Kompetenzentwicklung. Für den Einsatz als LVD müssen Testverfahren besonders ökonomisch sein (kurz sowie automatisiert in Durchführung, Auswertung und Dokumentation; Espin, McMaster, Rose & Wayman, 2012; Souvignier, Förster & Salaschek, 2014), gleichzeitig aber zentrale Kompetenzen abbilden und inhaltsgleiche Messungen mit konstanten Anforderungen über die Zeit anbieten, damit Veränderungen in den Testwerten als Leistungsentwicklung interpretiert werden können.

Die Konstruktion solcher Tests ist besonders für den Bereich der Lesekompetenz nicht trivial. Bisher angewandte Verfahren zeigen diverse Nachteile. Im Hinblick auf die Auswahl vergleichbarer Texte verhalten Lesbarkeitsindizes oder die Verwendung unterschiedlicher Textabschnitte aus Kinderbüchern nicht zu äquivalenten Anforderungen (Christ & Ardoin, 2009). Field testing methods und stratifizierte Ziehung von Item-Stichproben aus einem kalibrierten Item-Pool erwiesen sich als ausgesprochen aufwändig (Christ & Ardoin, 2008; Strathman, Klauer & Greisbach, 2010).

Eine Alternative besteht in der regelgeleiteten Itemkonstruktion (Irvine & Kyllonen, 2002). Hier wird der Einfluss von Itemmerkmalen auf die Schwierigkeit eines Items gezielt untersucht und für die Itemkonstruktion verwendet (vgl. Freund, Hofer & Holling, 2008). Daraus ergeben sich mehrere Vorteile: (1) Hinweise auf die Konstruktvalidität (Hartig & Frey, 2012), (2) Bei bekanntem Merkmalseinfluss: Möglichkeit der gezielten Konstruktion und des gezielten Einsatzes neuer (auch äquivalenter) Aufgaben ohne erneute Kalibrierung (Holling, Bertling & Zeuch, 2009), (3) Mögliche Vorstufen für Item-Cloning (Bejar, 1993) und adaptives Testen (Glas & van der Linden, 2003), (4) Ggf. Identifikation besonders änderungssensitiver Merkmale (für LVD sehr attraktiv).

Fragestellung: Wie kann die Konstruktion von äquivalenten Tests zur Erfassung von Lesekompetenz von Zweitklässlern im Rahmen der LVD durch eine regelgeleitete Vorgehensweise unterstützt werden? Wie lässt sich der Einfluss der berücksichtigten Merkmale auf die Itemschwierigkeit sowie eventuell auf die Bearbeitungszeit bestimmen? Das nachfolgend vorgestellte Vorgehen soll Gelegenheit zur Diskussion und zum Austausch über die Itemkonstruktion bieten.

Methode: Zunächst wurden drei Parallelformen von Lesetests (jeweils 60 Items) für die 2. Klasse mit Aufgaben zum Wort-, Satz- und Textverständnis konstruiert. Das Testformat umfasste Wort-Pseudowort- sowie Satzverifikationsaufgaben und die Beurteilung kurzer Satzfolgen auf ihre Plausibilität. Auf diese Weise werden Informationen über kognitive Teilprozesse gewonnen (Richter & Christmann, 2009), die differenzierte diagnostische Informationen für die individuelle Anpassung von Fördermaßnahmen darstellen können (Müller, Križan, Hecht, Richter & Ennemoser, 2013; Müller & Richter, 2014).

Aufgrund von Literaturrecherchen (z.B. Mayringer & Wimmer, 2003; Richter et al., 2012) wurden folgende Merkmale zur Konstruktion und für die spätere Bestimmung des Einflusses auf Itemschwierigkeit und Bearbeitungszeit gewählt: (1) Wort-Pseudowort-Aufgabe: Worthäufigkeit, Wortkomplexität, Silbenanzahl, (2) Satzverifikation: Worthäufigkeit, propositionale Dichte, Schlussfolgern, (3) Plausibilitätsbeurteilungen: Worthäufigkeit, propositionale Dichte, kausale Beziehungen

Die Merkmale wurden in einer Designmatrix für jedes Item spezifiziert. Über ein spezielles Kalibrierungsdesign (Aufteilung der Gesamtstichprobe, jede Gruppe bearbeitet unterschiedliche Itemkombinationen für alle drei Skalen) wird die Konfundierung von Item und Messzeitpunkt vermieden (Klein Entink, Kuhn, Hornke & Fox, 2009).

In einer Pilotstudie mit 11 Schulklassen wurden erste Befunde zur Güte der neu konstruierten Paralleltests ermittelt, die aktuell für eine Überarbeitung der Tests und den Einsatz in einer größeren Hauptstudie überarbeitet werden.

Ergebnisse: Die genannten Aufgabenformate zeigen eine gute Eignung für die LVD (gut verständlich, mittlere Bearbeitungszeit 13 Minuten). Für die Wort- und Satzskaala war Cronbachs $\alpha > .70$, für die Textskala war Cronbachs $\alpha > .60$. Für die Hauptstudie wird die Itemanzahl zugunsten einer kürzeren Bearbeitungszeit und zur Verringerung negativer Einflüsse einzelner Items auf die Konsistenz reduziert. Der regelgeleitete Konstruktionsansatz erwies sich als hilfreich für die Erstellung von Parallelaufgaben. Der Einfluss der einzelnen Merkmale auf Itemschwierigkeit und Bearbeitungszeit wird in der Hauptstudie geprüft und auf der Tagung vorgestellt.

Bürger, Sarah; Kröhne, Ulf; Goldhammer, Frank

Eine Analyse von Mode-Effekten in Leseverständnistests - Der Einfluss von Item- und Personenmerkmalen

Der in verschiedenen Kontexten und Disziplinen immer häufiger geplante Umstieg von papierbasiertem (PBA) auf computerbasiertes Assessment (CBA) führt zu der Frage der Äquivalenz beider Testformen. Dabei ist es gerade in längsschnittlich angelegten Studien, wie dem Nationalen Bildungspanel (NEPS), für eine valide Interpretation der Daten zur Veränderungsmessung sowie zum Vergleich von Kohorten wesentlich, dass die Vergleichbarkeit der Ergebnisse bei Wechsel des Administrationsmodus gewährleistet bleibt. Dabei geht es nicht nur um die numerische Vergleichbarkeit der Rohwerte hinsichtlich des erfassten Konstruktes sondern vielmehr um die psychometrische Vergleichbarkeit der beiden Testformen. Existieren Mode-Effekte im Hinblick auf psychometrische Eigenschaften der Tests, so sind auch die mit verschiedenen Testformen gewonnenen Ergebnisse nicht mehr einfach miteinander vergleichbar (z.B. Beretvas, Cawthon, Lockhart, & Kaye, 2012; Van den Noortgate & De Boeck, 2005).

Die bisherige Befundlage zeigt heterogene Ergebnisse zum Auftreten von Mode-Effekten. Relativ homogen jedoch sind die Befunde, dass das Risiko für Mode-Effekte zwischen Domänen unterschiedlich stark ausfällt und mit der Komplexität der umzusetzenden Aufgaben steigt (Pommerich, 2004). So gelten Tests zum Leseverständnis anfälliger für Mode-Effekte, als z.B. Aufgaben aus der Mathematik, da erstere oftmals die Besonderheit einer Unitstruktur aufweisen und längere zu lesende Texte darbieten, zwischen denen zu navigieren ist. Auch unterschiedliche Formen von Antwortformaten können die Äquivalenz beeinträchtigen.

Zur Analyse von Mode-Effekten werden psychometrische Äquivalenzkriterien, wie invariante Itemschwierigkeiten und gleiche latente Konstrukte herangezogen. Dabei müssen zur Vorhersage sowie auch Erklärung von Mode-Effekten auch stets die Ausprägung von Messeigenschaften (Itemmerkmalen im jeweiligen Administrationsmodus) mit berücksichtigt werden (z.B. Kroehne & Martens, 2011). Außerdem sind interindividuelle Unterschiede zwischen den Testpersonen sowie Interaktionen mit den jeweiligen Merkmalen des Tests oder einzelner Items nicht auszuschließen bzw. entsprechende Annahmen gezielt zu prüfen.

Im NEPS erfolgen die Kompetenztestungen in künftigen Erhebungen wenn möglich computerbasiert, weshalb in experimentell angelegten Studien die Testäquivalenz als wesentliche Voraussetzung für den Moduswechsel überprüft wird. Zum Vergleich der Administrationsmodi „Papier“ vs. „Computer“ für die Domäne „Lesen“ wurden in einer Mode-Effekt Studie drei Tests für zwei verschiedene Klassenstufen (7. und 9. Klasse) computerisiert (Gehrer, Zimmermann, Artelt & Weinert, 2013). In einem „Zwischensubjektdesign“ bearbeitete die Hälfte der Schüler die Aufgaben im Testheft und die andere Hälfte am Computer. Dabei wurde allen Schülern ein weiterer papierbasierter Leseverständnistest einer niedrigeren Klassenstufe vorgelegt, welcher als Außenkriterium in der Analyse der Konstruktäquivalenz verwendet werden konnte.

Zur Untersuchung von Mode-Effekten wurden Mehrgruppen-IRT-Modelle in Mplus geschätzt und Unterschiede in der Schwierigkeit der Aufgaben gefunden. Zunächst konnte jedoch für alle drei Tests anhand des Außenkriteriums die Hypothese der Konstruktäquivalenz bestätigt werden. Im nächsten Schritt wurden Hypothesen zu schwierigkeitsverändernden Merkmalen des jeweiligen Administrationsmodus auf Item- und Unit-Ebene getestet. So konnte die Hypothese, dass Items mit komplexerem Antwortformat am Computer schwieriger zu lösen sind bestätigt werden. Auch für Items an der ersten oder zweiten Position einer Unit konnten Unterschiede in der Schwierigkeit festgestellt werden. Da zu diesen am Computer vom Text aus bereits navigiert werden musste, während sie im Heft auf derselben Seite wie der zu lesende Text dargeboten wurden, erwiesen sie sich am Computer als schwieriger. Schließlich ließ sich auf Unit-Ebene auch ein Mode-Effekt in Abhängigkeit der Anzahl der Seiten pro Lesetext finden. Dabei wiesen Units, die äquivalent zum Testheft auch am Computer nur eine Text-Seite hatten keinen Mode-Effekt auf.

Diese, in Beziehung zu Itemmerkmalen stehenden Unterschiede variieren überdies zwischen Personen. Der nächste Schritt der Analyse ist nun diese interindividuelle Variabilität über Personenmerkmale, wie z.B. ICT-Literacy zu erklären.

Die Ursachen von Mode-Effekten sowie der Umgang mit diesen Unterschieden, vor allem im Kontext des Nationalen Bildungspanels, werden abschließend diskutiert.

Vaccaro, Didier; Lankes, Eva-Maria

Der Einfluss von Entscheidungsregeln für die Reliabilität und Validität der Dateninterpretation im Evaluationsprozess

Zur Steuerung des Schulsystems wurden mittlerweile in vielen Bundesländern staatliche Evaluationssysteme implementiert, die auf Basis erfahrungswissenschaftlicher Daten rationale Grundlagen für die Verbesserung von schulischen und unterrichtlichen Prozessen schaffen (Böttcher, 2009). Aufgrund ihres engen Bezugs zu den Methoden der empirischen Bildungsforschung sind die Evaluationsverfahren den dort geltenden wissenschaftlichen Gütekriterien der Objektivität, Reliabilität und Validität verpflichtet (Yarbrough, D. B. et al. 2011). Entsprechend wird Schulqualität auf der Basis von empirischen Daten aus Fragebogenerhebungen, Unterrichtsbeobachtungen, Dokumentenanalysen etc. eingeschätzt (Döbert, Rürup & Dederig, 2008). Während anhand eines solchen Instrumentariums eine objektive und reliable Datengrundlage geschaffen werden kann, fehlt es bisher an begründbaren Maßstäben zur Bewertung solcher Ergebnisse (Gärtner & Pant, 2011). Je nach Verfahren muss der Evaluator daher an mehr oder weniger Stellen „entscheiden“, welche Wertigkeit einer vorliegenden Antwortverteilung oder Textpassage zukommt. Nach Dhar (1996) werden Entscheidungen nach unterschiedlichen Regeln getroffen, die sich u.a. danach systematisieren lassen, wie die entscheidungsrelevanten Informationen miteinander Verknüpft werden. Während nicht-kompensatorische Herangehensweisen die Informationen einzeln betrachten, werden bei kompensatorischer Strategie positive und negative Aspekte z.B. über Mittelwertsbildung gegeneinander aufgewogen. Welches Entscheidungsverfahren präferiert wird, ist u.a. von den bisherigen Erfahrungen des Entscheiders abhängig (z. B. Rohrbaugh & Shanteau, 1999). Entsprechend sollten Evaluatoren mit ähnlicher beruflicher Sozialisation bei der Bewertung von Ergebnissen dieselbe Herangehensweisen bevorzugen. Nachdem die Evaluation von Schulqualität häufig von Lehrkräften durchgeführt wird (Döbert et al., 2008) und im Schulkontext kompensatorische Verfahren dominieren (z.B. Notengebungsverfahren), dürften evaluative Entscheidungen im Kontext von Schulevaluation in erster Linie nach kompensatorischen Regeln getroffen werden. Gerade dieser Entscheidungsweg sollte jedoch nach Dhar (1996) zu hoher Entscheidungsunsicherheit und in der Folge zu vermehrten Entscheidungsaufschub führen, wenn sich aufgrund des geringen Übergewichts von positiven bzw. negativen Aspekten (z. B. Anteil gewählter Antwortoptionen „trifft nicht zu“ vs. „trifft zu“) keine klare Tendenz herauskristallisiert. Während die geteilten Sozialisationserfahrungen über ihren Einfluss auf die jeweils präferierte Entscheidungsstrategie die Objektivität und Reliabilität der Einschätzung erhöhen, bleiben bei Entscheidungsaufschub Informationen unberücksichtigt, was die Validität des Evaluationsergebnisses beeinträchtigt. Um die Plausibilität einer solchen Konstellation für den Evaluationskontext zu überprüfen, wurde eine Simulationsstudie mit anonymisierten Evaluationsdaten durchgeführt. Gemäß der Vorüberlegungen wurde davon ausgegangen, dass Evaluatoren mit Lehrerfahrung die entscheidungsrelevanten Informationen aus den Daten (a) nach einheitlichen Regeln verknüpfen (ähnliche Sozialisation), wobei sie (b) die Entscheidung umso eher aufschieben, je geringer das relative Übergewicht einer Antworttendenz ausfällt (kompensatorisch Herangehensweise). Hierfür wurden verschiedene Auswertungsergebnisse aus Fragebogenerhebung, Unterrichtsbeobachtung und Dokumentenanalysen 76 Evaluationsteams zur Einschätzung vorgelegt. Insgesamt war die Interraterreliabilität zwischen den Teams zufriedenstellend bis gut (Cohens-Kappa-Werte zwischen .59 und .72) ohne Hinweise auf subgruppenspezifische Antwortprofile (1-Klassen-Lösung bei Latent-Class-Analyse). Passend zu Dhars Annahme bzw. passend zu den Vorhersagen bzgl. kompensatorischer Herangehensweisen erhöhte sich die Tendenz zum Entscheidungsaufschub bei Antwortmustern, die kein eindeutiges relatives Übergewicht von Antworttendenzen aufwiesen. Implikationen für die Schulung von Evaluatoren sowie mögliche Ansätze nicht-kompensatorischer Verknüpfungsstrategien und deren Auswirkung auf die Validität der Ergebnisse werden diskutiert.

Mittwoch
11. März

15:45 Uhr

Foyer GC - Ost

Zitzmann, Steffen; Lüdtke, Oliver; Robitzsch, Alexander

Analyse von Kontexteffekten mit Mehrebenenstrukturgleichungsmodellen: Führt ein Bayes-Ansatz zu stabileren Schätzungen bei geringer Anzahl von Gruppen?

In der empirischen Bildungsforschung stellt die Schule bzw. Klasse einen besonders relevanten sozialen Kontext dar, den es bei der Identifikation der Determinanten wirkungsvoller Lernaktivitäten zu berücksichtigen gilt. Eine Strategie zur Erfassung von Kontextmerkmalen besteht darin, Individualvariablen – bzw. Maße einer unteren Analyseebene – auf eine höhere Analyseebene zu aggregieren. Beispielsweise können die individuellen Einschätzungen des Schulklimas oder des Unterrichts durch die Schüler auf die Schul- bzw. Klassenebene aggregiert werden. Der durch die Aggregation entstandene Mittelwert würde dann das kollektive Schulklima einer Schule erfassen. Im nächsten Schritt können dann mit Hilfe von Mehrebenenmodellen, die Effekte des Kontextes auf Schüler-Outcomes untersucht werden (z.B. Raudenbush & Bryk, 2002).

Allerdings kann der aggregierte Gruppenmittelwert (z.B. Klassenmittelwert der Schülerwahrnehmungen) in bestimmten Datenkonstellationen nur eine unreliable Messung des unbeobachteten „wahren“ Gruppenmittelwertes darstellen (Lüdtke, et al., 2008). Dies wird z.B. der Fall sein, wenn nur von wenigen Schülern eine Einschätzung vorliegt, die Varianz zwischen den Gruppen relativ gering ausfällt oder die einzelnen Items nur eine geringe interne Konsistenz aufweisen. Zur Kontrolle dieser Unreliabilität bei der Erfassung von Kontextmerkmalen wurden „doppelt-latente“ Mehrebenenstrukturgleichungsmodelle vorgeschlagen (Lüdtke et al., 2011; Marsh et al., 2012), die den Gruppenmittelwert als latentes Konstrukt modellieren und die Schätzung des Kontexteffektes sowohl für den Messfehler als auch den durch die Aggregation entstehenden Stichprobenfehler korrigieren. Dieser Ansatz ist in der Software Mplus implementiert (Asparouhov & Muthén, 2007). Simulationsstudien haben allerdings gezeigt, dass dieses „doppelt-latente“ Mehrebenenmodell selbst bei größerer Anzahl (z.B. 100) von Gruppen zu instabilen Schätzungen führen kann (Lüdtke et al., 2011).

Fragestellung: Im vorliegenden Beitrag wird ein Bayes-Ansatz zur Schätzung des „doppelt-latenten“ Mehrebenenmodells vorgestellt und evaluiert. Der Bayes Ansatz wird in der Software WinBUGS spezifiziert. Insbesondere in problematischen Datenkonstellationen (z.B. geringe Anzahl von Gruppen, gering ausgeprägte Varianzen zwischen Gruppen) werden Bayes-Verfahren zur Schätzung von Mehrebenenmodellen in der methodischen Literatur empfohlen (Gelman & Hill, 2007; Hamaker & Klugkist, 2011; Kaplan, 2014; Swaminathan & Rogers, 2008). Es wurde untersucht, ob der Bayes-Ansatz in diesen problematischen Datenkonstellationen hinsichtlich der Schätzung des Gruppeneffektes besser abschneidet als der in Mplus implementierte Maximum-Likelihood (ML)-Ansatz. Des Weiteren ist von Interesse, ob der Bayes-Ansatz vor allem bei einer kleinen Anzahl von Gruppen ein besseres Konvergenzverhalten aufweist als der ML-Ansatz.

Methode: In einer Simulationsstudie untersuchten wir das statistische Verhalten (Bias, RMSE, Konvergenz) des Bayes-Ansatzes und verglichen die Ergebnisse mit denen des ML-Ansatzes in Mplus. Das datengenerierende Modell ist ein „doppel-latentes“ random-intercept Mehrebenenmodell mit einer unabhängigen Variablen, die durch fünf Indikatoren erfasst wird und sowohl einen Effekt auf Individual- als auch Gruppenebene besitzt. Es wurden die Klassenanzahl ($k = 50, 100, 200$), die Klassengröße ($n = 5, 10, 20$) und die Intraklassenkorrelation des Prädiktors ($ICC = 0.05, 0.1, 0.2$) variiert. Insgesamt umfasste das Simulationsdesign 27 Zellen. Für jede Zelle wurden 1000 Datensätze generiert und mit beiden Ansätzen analysiert. Für die Parameter des „doppelt-latenten“ Modells wurden im Bayes-Ansatz nichtinformative Prior-Verteilungen gewählt (z.B. invers-gamma mit .001)

Ergebnisse: Der Bayes-Ansatz des doppelt-latenten Mehrebenenstrukturgleichungsmodells lieferte vor allem bei geringer ICC und kleinen Gruppen stabilere Schätzungen (deutlich kleinerer RMSE) des Gruppeneffektes als der ML-Ansatz in Mplus. Des Weiteren zeigte der Bayes-Ansatz weniger Konvergenzprobleme als der ML-Ansatz, der vor allem bei geringer ICC und kleiner Anzahl von Gruppen erhebliche Probleme aufwies. Mit zunehmender Stichprobengröße und ICC näherten sich die Ergebnisse des Bayes- und ML-Ansatzes aneinander an.

Diskussion: Der Bayes-Ansatz stellt ein vielversprechendes Verfahren zur Schätzung des doppelt-latenten Mehrebenenmodellstrukturgleichungsmodells dar, dass besonders für Anwendungen geeignet scheint, in denen nur wenig Informationen bezüglich des zu erfassen Kontextmerkmals vorliegen (z.B. geringe ICC, kleine Anzahl von Gruppen). In einem nächsten Schritt soll der Effekt der Wahl verschiedener Prior-Verteilungen auf die Schätzung von Kontexteffekten untersucht werden.

Grund, Simon; Lüdtkke, Oliver; Robitzsch, Alexander

Wie kann die globale Nullhypothese der ANOVA mit multipel imputierten Daten geprüft werden: Befunde einer Simulationsstudie

Theoretischer Hintergrund: Die Varianzanalyse (ANOVA) ist eine in der Bildungsforschung weit verbreitete Methode zur Prüfung von Mittelwertsunterschieden zwischen natürlich auftretenden Gruppen (z.B. Schulformen, unterschiedlicher Familienstatus) oder Gruppen in einem Versuchsplan (z.B. Experimental- und Kontrollgruppe). Eine besondere Herausforderung stellt die Tatsache dar, dass häufig nicht für alle Untersuchungseinheiten Beobachtungen auf der abhängigen Variablen vorliegen (missing data). In der methodischen Literatur wird die multiple Imputation (Rubin, 1987) als leistungsstarkes Verfahren zur Analyse von Datensätzen mit fehlenden Werten empfohlen. Erstaunlicherweise gibt es keine eindeutigen Empfehlungen, wie die globale Hypothese der ANOVA, dass sich die Gruppen in ihren Mittelwerten gleichen, mit multipel imputierten Daten getestet werden soll. Die viel verwendete Software SPSS bietet z.B. keinen Test der globalen Hypothese der ANOVA mit imputierten Daten an. Im vorliegenden Beitrag werden drei Verfahren verglichen, die für die Prüfung der globalen Hypothese der ANOVA mit imputierten Daten vorgeschlagen wurden. Diese Verfahren betrachten die ANOVA im Rahmen des allgemeinen linearen Modells jeweils als Regressionsmodell.

Li, Raghunathan und Rubin (1991) schlugen erstens ein Verfahren (häufig "D1") vor, welches die geschätzten Koeffizienten und ihre Standardfehler über imputierte Datensätze hinweg aggregiert. Die finale Teststatistik D1 ergibt sich aus den gemittelten Punktschätzern und einer zusätzlichen Varianzschätzung für die Koeffizienten. Zweitens diskutierten Li, Meng, Raghunathan und Rubin (1991) ein Verfahren (häufig "D2"), das schon auf Ebene der Datensätze die Teststatistik berechnet und daher keine Informationen über die Standardfehler benötigt. D2 errechnet sich daher direkt aus den Teststatistiken (hier den F-Werten), gemittelt über die einzelnen Datensätze. Ein drittes Verfahren (häufig "D3") nach Meng und Rubin (1992) soll die Vorteile beider Verfahren vereinen, indem die Regressionskoeffizienten und ihre Likelihood über Datensätze aggregiert werden. Das Verfahren ist sehr aufwendig, da ein Likelihood-Ratio-Test für jeden Datensatz und zudem für die gemittelten Regressionskoeffizienten erneut berechnet werden muss.

Fragestellung: Empfehlungen zur Verwendung dieser Verfahren sind in der Literatur nicht einheitlich. Bisherige Arbeiten heben die positiven Eigenschaften von D1 und D3 hervor (z.B. Enders, 2010; Schafer, 1997; van Buuren, 2012). D2 hingegen erwies sich in einer Simulationsstudien als anfällig gegenüber der Anzahl zu testender Parameter und gegen unterschiedlich bedeutsamen Datenausfall (Li, Meng et al., 1991). D2 besitzt allerdings den Vorteil, dass es vom Anwender leicht durchgeführt werden kann, da lediglich die Teststatistiken über die einzelnen Datensätzen aggregiert werden müssen. Die vorliegende Studie vergleicht die drei Verfahren sowie den listenweisen Fallausschluss hinsichtlich der einfaktoriellen ANOVA.

Methode: In einer umfassenden Simulationsstudie untersuchten wir Typ-I- und Typ-II-Fehlerraten der Verfahren. Es wurden die Gruppenanzahl ($k = 3, 6, 12$), Gruppengrößen ($n = 25, 50, 100$), Effektstärken ($f = 0, .10, .25, .40$), Effektmuster (Differenz vs. Trend), Design (balanciert vs. unbalanciert) sowie verschiedene Mechanismen für den Datenausfall in Abhängigkeit einer Drittvariable Z (MCAR, MAR mit variierendem Effekt von Z) in der Studie berücksichtigt. Insgesamt umfasste das Simulationsdesign 1296 Zellen. Für jede Zelle wurden 5000 Replikationen durchgeführt.

Ergebnisse: D1 weist unter Einhaltung der Typ-I-Fehlerrate die größte Teststärke auf. Das einfach zu implementierende D2 verhält sich nur geringfügig liberaler als D1 mit leicht erhöhter Typ-I-Fehlerrate bei gleichzeitig höherer Teststärke, wobei es sich mit größerer Stichprobe schnell D1 annähert. D3 liefert konservative Ergebnisse und ist vor allem in kleinen Stichproben wenig teststark. Insgesamt kann gegenüber dem listenweisen Fallausschluss mit D1 und D2 nur wenig verloren werden, gegebenenfalls aber gewonnen werden. Während Typ-I-Fehlerraten sich nur geringfügig unterscheiden, sind D1 und D2 teststärker, sobald eine Kovariable vorliegt, die für die fehlenden Werte informativ ist.

Diskussion: Wir schließen aus den Ergebnissen, dass der globale Hypothesentest der ANOVA mit jedem der drei Verfahren durchgeführt werden kann. Insbesondere die einfach zu implementierenden Verfahren D1 und D2 sind für den Anwender sehr attraktiv und können leicht mit Standardsoftware umgesetzt werden.

Mittwoch
11. März

15:45 Uhr

Foyer GC - Ost

3

Hecht, Martin; Schroeders, Ulrich; Siegle, Thilo

Ein Vorhersagemodell für Itembearbeitungszeiten zur Optimierung von Testdesigns im Paper-and-Pencil Large-scale Assessment

Theoretischer Hintergrund: Zur Konstruktion von Testdesigns im Large-scale Assessment werden akkurate Itembearbeitungszeiten benötigt, um Testhefte zusammenzustellen, die eine bestimmte Sollbearbeitungszeit aufweisen. Durch inakkurate Bearbeitungszeiten wird entweder wertvolle Testzeit nicht vollständig genutzt, wenn die Bearbeitungszeiten überschätzt wurden. Andererseits schaffen die Schülerinnen und Schüler bei Unterschätzung der Itembearbeitungszeiten die Items am Ende der Testhefte nicht mehr, was die Validität des Tests und die Parameterschätzungen negativ beeinflusst. Die beste, jedoch auch kostenintensivste Methode, um akkurate Itembearbeitungszeiten für das Testdesign zu erlangen, ist deren empirische Bestimmung in Vorstudien. Ist dieses Vorgehen (z.B. aus Kostengründen) nicht möglich, müssen die für das Testdesign nicht verzichtbaren Itembearbeitungszeiten anderweitig geschätzt werden. Hierzu eignet sich ein empirisch fundiertes Vorhersagemodell, in dem aus leicht verfügbaren Itemeigenschaften – wie beispielsweise der Anzahl an Wörtern und dem Antwortformat – die Bearbeitungszeit geschätzt werden kann.

Fragestellung: Es wird untersucht, von welchen Itemeigenschaften die Itembearbeitungszeiten abhängen, um ein Vorhersagemodell von Bearbeitungszeiten für Items zu generieren, für die keine empirische Bearbeitungszeit vorliegt. Die hierfür untersuchten Itemeigenschaften sind: Anzahl Wörter, Itemschwierigkeit, Vorhandensein von Darstellungselementen (Tabellen, Diagramme, Schemata, Fotos) und das Antwortformat (Multiple Choice, Kurzantwort, Erweiterte Antwort).

Methode: Die Stichprobe besteht aus 334 Schülerinnen und Schüler der 9. Klasse von vier Gymnasien und drei Realschulen mit einem mittleren Alter von $M = 15.5$ Jahren. Die Teilnahme an dieser Studie war freiwillig und wurde nicht incentiviert. Die Datenerhebung fand im Frühjahr 2010 statt.

Der Test besteht aus 184 Items, die die in den Bildungsstandards definierte naturwissenschaftliche Kompetenz von Schülerinnen und Schülern messen, die den mittleren Schulabschluss anstreben. Diese Items wurden in einem Incomplete Block Design zu 24 Testheften zusammengestellt, die zufällig an die Testteilnehmer verteilt wurden. Da die Bearbeitungszeiten reliabel gemessen werden sollten, stand den Schülerinnen und Schüler insgesamt ausreichend Zeit zur Verfügung. Die Messung der Bearbeitungszeiten erfolgte durch Eintragung der aktuellen Uhrzeit in das Testheft vor und nach den Aufgaben.

Die Datenaufbereitung und -auswertung erfolgte in R 2.14.0. Zwei Linear Mixed Models (LMM) wurden unter Verwendung des R-Pakets lme4 spezifiziert: Das erste Modell (Baseline Model) beinhaltet neben dem Intercept die beiden Varianzquellen Personen und Items als random effects. Dadurch wird der Einfluss von Personenvariation und Itemvariation in der Bearbeitungszeit entflochten. Im zweiten Modell (Predictor Model) werden dann die zu untersuchenden Itemeigenschaften als fixed effects hinzugefügt.

Ergebnisse: Das Baseline Model zeigt, dass ein Großteil der Variation in den Bearbeitungszeiten auf die Items zurückzuführen ist. Im Predictor Model wird 71.7% dieser Varianz durch die untersuchten Itemeigenschaften aufgeklärt. Bezüglich der jeweiligen Effekte der Eigenschaften zeigt sich ein plausibles Bild: Pro Wort verlängert sich die Bearbeitungszeit eines Items um 0.47 Sekunden. Die Bearbeitung eines Multiple Choice Items dauert 20.87 Sekunden länger als wenn nur der Itemstamm gelesen wird. Eine Kurzantwort verlängert die Bearbeitungszeit um 46.99 Sekunden; eine erweiterte Antwort um 76.56 Sekunden. Die Itemschwierigkeit und das Vorhandensein von Darstellungselementen haben keinen Effekt auf die Bearbeitungszeit. Mit Hilfe dieses empirisch abgesicherten Vorhersagemodells können zukünftig Itembearbeitungszeiten von Items, für die keine Bearbeitungszeiten vorliegen (z.B. neu entwickelte Items), geschätzt werden.

Mittwoch
11. März

15:45 Uhr

Foyer GC - Ost

Thoma, Gun-Brit; Dalehefte, Inger Marie; Köller, Olaf

Die Bedeutung bestimmter Itemcharakteristika für die Bearbeitungsstrategien bei Multiple-Choice Aufgaben - eine Eyetracking-Studie

Obwohl Multiple-Choice Aufgaben (MCA) das am weitesten verbreitete Aufgabenformat ist (Chang, Lin & Lin, 2007), ist wenig darüber bekannt, wie Personen bei der Beantwortung von MCA vorgehen und ob bestimmte Strategien zu besseren Ergebnissen führen. Befunde zeigen, dass sich die Attraktorposition sowie die Itemschwierigkeit insbesondere dann auf das Testergebnis auswirken, wenn die Antworten nicht gewusst werden (McNamara & Weitzman, 1945; Schroeder, Murphy & Holme, 2012). Um der Frage nachzugehen, welche Bearbeitungsstrategien von Personen dabei eine Rolle spielen, werden in dieser Arbeit Personen mit unterschiedlicher Expertise untersucht. Eine vielversprechende Möglichkeit, Bearbeitungsstrategien sichtbar zu machen, ist, die Blickbewegungen von Person anhand eines Eyetrackers aufzuzeichnen (Pellegrino, Chudowsky & Glaser, 2001), da sie darauf hinweisen, welche Informationen gerade verarbeitet werden (vgl. Just & Carpenter, 1976). Ein Eyetracker registriert die Augenposition und stellt so Informationen über das Blickverhalten bereit (Holmqvist et al., 2011a). Vor diesem Hintergrund ergibt sich folgende Fragestellung: Unterscheiden sich Personen mit unterschiedlicher Expertise in ihren Blickbewegungen bei der Beantwortung von MCA in Abhängigkeit der Attraktorposition und der Itemschwierigkeit?

Die Stichprobe umfasst 15 Psychologiestudierende mit thematischen Wissen hinsichtlich des Themas der MCA (Experten; Alter: $M=24.20$, $SD=1.84$) und 15 Personen, die ein Fach studieren, das keinen Bezug zum Thema der MCA hat, wie Jura und BWL (Laien; Alter: $M=23.73$, $SD=2.09$). Beide Gruppen setzen sich jeweils aus acht Frauen und sieben Männern zusammen. Im Anschluss an die Erhebung persönlicher und demografischer Daten bearbeiten die Studierenden in einem individuellen Setting einen computerbasierten Test zum Thema „Gehirn und Nervensystem“. Der Test umfasst 21 MCA und wurde anhand einer größeren Stichprobe ($N=377$) validiert (Thoma, Dalehefte & Köller, 2014). Jede MCA besteht aus einer Frage, einer richtigen Antwort und drei Distraktoren und wurde auf einem SMI Remote-Eyetracker (Computermonitor mit Infrarotkamera, 120 Hz) einzeln präsentiert. Itemschwierigkeit und Position des Attraktors variieren systematisch.

Um den Einfluss von Itemschwierigkeit und Attraktorposition auf die Bearbeitungsstrategien von MCA zu untersuchen, wurden zuerst Areas of Interest (AOI) definiert. Die Frage und jede Antwortmöglichkeit werden jeweils einer AOI zugeordnet. Anschließend wurde in Anlehnung an Holmqvist et al. (2011b) anhand der zeitlichen Reihenfolge der Fixationen der einzelnen Antwortoptionen pro MCA Sequenzen identifiziert. Sequenzen bestehend aus zwei oder drei AOI weisen auf ein konzentriertes Blickverhalten hin (focused behavior), während Sequenzen mit vier oder fünf AOI auf ein überblicksartiges Blickverhalten hindeuten (overview behavior). Es wurden deskriptive Analysen durchgeführt und Unterschiede zwischen Experten und Laien (U-Test) überprüft. Zusätzlich wurden Korrelationen zwischen focused bzw. overview behavior und Itemschwierigkeit bzw. Attraktorposition berechnet.

Die Ergebnisse lassen erkennen, dass über die 21 MCA hinweg Laien mehr focused behavior zeigen, während Experten mehr overview behavior aufweisen. Dieses Blickverhalten scheint unabhängig von der Attraktorposition zu sein, da die entsprechende Korrelation mit der Anzahl an focused bzw. overview Sequenzen weder für die Experten (focused behavior: $r=.03$, $p=.57$; overview behavior: $r=.01$, $p=.88$) noch für die Laien (focused behavior: $r=.01$, $p=.87$; overview behavior: $r=.02$, $p=.79$) signifikant ist. Hinsichtlich der Itemschwierigkeit zeigen die Befunde, dass sich das Blickverhalten zwischen Experten und Laien für focused behavior mit zunehmender Schwierigkeit ähnelt. Es gibt keine signifikanten Unterschiede bei schweren Items (focused behavior: $M_{Laien}=35.07$; $SD_{Laien}=16.55$; $M_{Experten}=26.07$; $SD_{Experten}=10.53$; $U=79.50$; $z=-1.37$, $p=0.17$), dagegen aber bei leichten Items ($M_{Laien}=63.47$; $SD_{Laien}=44.93$; $M_{Experten}=31.87$; $SD_{Experten}=19.68$; $U=55.50$; $z=-2.37$, $p=0.02$). Für overview behavior gibt es keine signifikanten Unterschiede weder für leichte noch für schwere Items.

Insgesamt deuten die Befunde darauf hin, dass Experten aufgrund ihres thematischen Wissens den Attraktor identifizieren können, indem sie die Frage und die Antwortoptionen der Reihe nach durchlesen und sich für eine entscheiden. Laien dagegen scheinen mehr zwischen der Frage und/oder den einzelnen Antwortoptionen hin und her zu springen. Es scheint somit, dass Experten und Laien unterschiedliche Strategien bei der Bearbeitung von MCA anwenden."

Koć-Januchta, Marta; Höffler, Tim; Thoma, Gun-Brit; Precht, Helmut; Leutner, Detlev

Visualisierer und Verbalisierer lernen unterschiedlich: Eine Blickbewegungsanalyse beim Lernen von computerbasierten Text-Bild-Kombinationen

Theoretischer Hintergrund: Eine Vielzahl von Studien hat prinzipiell nachgewiesen (Mayer, 2005), dass eine Kombination aus Text und Bildern lernwirksam ist und tieferes Verständnis fördert. Trotzdem bestehen große interindividuelle Unterschiede beim Lernen mit multimedialen Lernumgebungen, etwa bedingt durch unterschiedliches Vorwissen (Kalyuga, 2007), räumliches Vorstellungsvermögen (Höffler, 2010) oder kognitiven Stil (Visualisierer/Verbalisierer). Was letztere betrifft, so wurde bereits gezeigt, dass Visualisierer bessere Lernergebnisse aus Text/Bild-Kombinationen erreichen als Verbalisierer (Höffler et al., 2010) und eher auf Bilder fokussieren (Mehigan et al., 2011). Andererseits wird die visuelle Aufmerksamkeit in erster Linie vom Text gelenkt (Hannus & Hyönä, 1999; Schmidt-Weigand et al., 2010).

Fragestellung: Die vorliegende Untersuchung zielt darauf ab, die Unterschiede zwischen Verbalisierern und Visualisierern beim Lernen mit Text/Bild-Kombinationen mittels Eye-Tracking genauer zu untersuchen. Wir gingen davon aus, dass Visualisierer nicht nur mehr Zeit zum Betrachten der Bilder aufwenden würden als zum Lesen der Texte, sondern dass sie auch erfolgreicher darin sind, relevante Bereiche der Bilder zu identifizieren und ihren Blick häufiger zwischen verschiedenen relevanten Bereichen der Bilder sowie zwischen verschiedenen Bildern wechseln zu lassen. Verbalisierer hingegen sollten mehr auf den Text als auf die Bilder fokussieren.

Methode: 40 Studierende wurden individuell getestet. Zunächst füllten sie neun verschiedene Fragebögen zum kognitiven Stil (Cronbachs Alpha zwischen .77 und .94) und zwei Tests zum räumlichen Vorstellungsvermögen aus, gefolgt von Fragen zum Vorwissen bezüglich der die Untersuchung beinhaltenden Themen: die Funktionsweise einer Toilettenspülung, erlernte Hilfslosigkeit, die ATP-Synthase sowie das Knüpfen eines Knotens. Diese sehr verschiedenen Themen sollten eine größtmögliche Verallgemeinerbarkeit der Ergebnisse sicherstellen bzw. ggf. bereichs- oder wissensspezifische Unterschiede (etwa beim motorischen Lernen) aufdecken. Wir erwarteten in allen Bereichen nur niedriges Vorwissen der Probanden.

Anschließend folgte die Testphase, in der die Probanden auf einem Computerbildschirm mit Eye-Tracker nacheinander jedes der Themen anhand einer Kombination von drei Bildern und drei Textelementen erlernten. Bilder und Text beinhalteten jeweils prinzipiell dieselbe Information.

Nach Abschluss der Lernphase wurde der Lernerfolg anhand offener, bildlicher und ja/nein-Fragen gemessen, zudem war der zu erlernende Knoten zu knüpfen.

Ergebnisse: Eine Faktorenanalyse zeigte, dass alle Skalen zum kognitiven Stil auf einem Faktor luden (77% Varianzaufklärung). Dieser Faktor wurde dazu genutzt, die Probanden als visuelle oder verbale Lerner zu klassifizieren (oberes Drittel versus unteres Drittel). Wie erwartet, zeigten t-Tests, dass für alle Themen Visualisierer signifikant mehr Zeit auf die Bilder verwandten als Verbalisierer, während es bei den Texten umgekehrt war (alle Effektstärken $d > 0.75$). Heat-Maps typischer Betrachtungsmuster beider kognitiver Stile beim Lernen visualisieren die Unterschiede zwischen Visualisierern und Verbalisierern sehr anschaulich.

In drei von vier Lernumgebungen unterschieden sich Verbalisierer und Visualisierer zudem in der Menge der erneuten Fixierungen bestimmter Bereiche. Bei der Toilettenspülung, dem Knotenknüpfen und der ATP-Synthase tendierten Verbalisierer dazu, den Text häufiger mehrfach zu lesen als Visualisierer. Visualisierer fixierten in allen Lernumgebungen die Bilder mehrfach und signifikant häufiger als die Verbalisierer (alle Effektstärken $d > 0.76$) und wechselten, wie erwartet, auch häufiger zwischen den Bildern. Allerdings konnte das analoge Verhalten von Verbalisierern bezogen auf den Text nicht festgestellt werden.

Diskussion: Die Untersuchung bestätigt, dass sich Menschen darin unterscheiden, wie sie piktorale und verbale Information wahrnehmen, und dass diese Unterschiede vom kognitiven Stil abhängig sind. Beim Lernen mit Text/Bild-Kombinationen verlassen sich Visualisierer in erheblichem Ausmaß vorrangig auf die Bilder, Verbalisierer auf den Text. Visualisierer scheinen beim Lernen viele Verknüpfungen zwischen einzelnen Bildelementen wie auch zwischen verschiedenen Bildern zu knüpfen. Verbalisierer hingegen konzentrieren sich vorrangig auf den Text, springen dort aber kaum hin und her. Dies kann an der linearen Struktur des Texts liegen, aber auch ein Hinweis auf einen aktiveren Lernstil von Visualisierern sein. Weitere Auswertungen sind geplant.

Mittwoch
11. März

15:45 Uhr

Foyer GC - Ost

Böttcher, Daniela M.; Stahnke, Rebekka; Nitz, Sandra; Upmeier zu Belzen, Annette

Wie sehen sich LehrerInnen Biologieunterricht an? Eine Eyetracking-Studie zur professionellen Unterrichtswahrnehmung

Die professionelle Unterrichtswahrnehmung von LehrerInnen (professional vision) beschreibt ihre Fähigkeit, bedeutsame Aspekte und Ereignisse in einer Unterrichtssituation zu identifizieren (noticing) und diese wissensbasiert zu interpretieren (knowledge-based reasoning) (SEIDEL et al., 2010, BLOMBERG et al., 2011; VAN ES & SHERIN, 2002). Erfahrene Lehrpersonen erkennen beim Betrachten von Unterrichtsszenen mehr bedeutsame Ereignisse (BERLINER et al. 2001; CARTER et al., 1988) und sind in der Lage die gesamte Klasse im Auge zu behalten, während angehende Lehrpersonen (Novizen) eher auf die unterrichtende Person achten als auf die SchülerInnen (SABERS et al., 1991). Niedrig-inferente Maße – z.B. Aufzeichnungen von Blickbewegungen mittels Eyetracking – wurden bisher nur in wenigen Studien zur professionellen Wahrnehmung eingesetzt (VAN DEN BOGERT et al., 2014, YAMAMOTO & IMAI-MATSUMURA, 2013). Eine Verknüpfung der beiden Kompetenzen noticing und knowledge-based reasoning, die für ein tiefergehendes Verständnis der Aufmerksamkeitsverteilung wichtig ist, wurde in diesen Eyetracking-Studien jedoch nicht realisiert. Darüber hinaus wurden in den meisten Studien zur professionellen Unterrichtswahrnehmung die verwendeten Videos nicht unter fachdidaktischen Gesichtspunkten eingesetzt, sondern sollten durch die Probanden auf der Basis ihres pädagogischen Wissens (PK) interpretiert werden. Dadurch bleibt die Frage offen, wie fachspezifische Aspekte wahrgenommen und auf der Basis des fachdidaktischen Wissens (PCK) beurteilt werden. Ein Einfluss des Expertisegrads der Lehrperson auf die professionelle Wahrnehmung von fachspezifischen Aspekten konnte in der Domäne Mathematik gezeigt werden (PHILIPP & SOWDER, 2002). Die Wahrnehmung domänenspezifischer Aspekte des naturwissenschaftlichen Unterrichts wurde bisher nur wenig untersucht. MCDONALD (2008) zeigte, dass angehende Lehrpersonen Schwierigkeiten haben, naturwissenschaftliche Instruktionsstrategien wie z.B. das forschend-entdeckende Lernen angemessen wahrzunehmen. Zudem wurde bisher der fachliche Hintergrund der Probanden nur wenig berücksichtigt, obwohl BLOMBERG und Kollegen (2011) aufzeigen konnten, dass Lehramtsstudierende verschiedener Fächer facheigene und fachfremde Unterrichtsszenen unterschiedlich wahrnehmen.

In diesem Projekt möchten wir exemplarisch für den Biologieunterricht untersuchen, wie angehende und erfahrene Biologielehrerinnen und -lehrer fachspezifische Handlungsmuster im Biologieunterricht identifizieren und wissensbasiert interpretieren. Dazu werden aus gefilmten Biologiestunden Vignetten ausgewählt, die dem zentralen Ansatz des forschend-entdeckendes Lernens folgen und verschiedene Stufen der Schülerorientierung zeigen (COLBURN, 2000). Diese Ausschnitte werden BiologielehrerInnen unterschiedlicher Expertise gezeigt, um herauszufinden, inwieweit diese sich in ihrer Identifikation (noticing) und wissensbasierten Interpretation (knowledge-based reasoning) fachspezifischer Handlungsmuster unterscheiden (Erste Teilstudie).

Andererseits wird untersucht wie angehende und erfahrene LehrerInnen Klassenmanagement im Biologieunterricht wahrnehmen und beschreiben. Dazu werden Szenen ausgewählt, die Aspekte des Klassenmanagements (KÖNIG & LEBENS, 2012) enthalten. Diese Vignetten werden angehenden als auch erfahrenen LehrerInnen gezeigt, welche entweder Biologie oder ein nicht-naturwissenschaftliches Fach unterrichten. Somit soll die Domänenspezifität des allgemein-pädagogischen Klassenmanagements erforscht werden (Zweite Teilstudie).

Ziel dieser beiden Teilstudien ist es, sowohl durch eine Kombination von quantitativen Blickdaten mit qualitativen verbalen Daten, als auch durch eine Verbindung biologiedidaktischer, erziehungs- und kognitionswissenschaftlicher Ansätze einen tieferen Einblick in diejenigen Prozesse zu gewinnen, die der professionellen Unterrichtswahrnehmung zugrunde liegen. Die Probanden werden unter unterschiedlichen Instruktionen mehrere Vignetten betrachten, die zuvor durch Expertenratings als typisch für bestimmte Unterrichtsaspekte bewertet wurden. Dabei werden die Blickbewegungen der Probanden bei erstem Betrachten der Vignetten aufgezeichnet. Anschließend wird ihnen der Videoausschnitt erneut gezeigt, diesmal ohne Ton und mit ihrem eigenen eingblendeten Blickpfad. Die Kommentare der Probanden werden dabei als cued retrospective report aufgenommen (VAN GOG et al., 2005). Die Auswertung der niedrig-inferenten Blickbewegungsdaten erfolgt quantitativ. Für die qualitativen verbalen Daten werden für die Schwerpunkte des forschend-entdeckenden Lernens respektive Klassenmanagements verschiedene Kodierschemata entwickelt. Das gemeinsame Ziel ist es, festzustellen, inwieweit sich Novizen und Experten in ihrer professionellen Unterrichtswahrnehmung unterscheiden. Langfristig können Expertenlösungen in der Lehrerbildung als Beispiele für die professionelle Wahrnehmung sowohl des Klassenmanagements als auch des forschend-entdeckenden Lernens eingesetzt werden, um Studierende darin zu fördern, relevante Situationen im Unterricht zu identifizieren und sowohl auf der Basis ihres domänenübergreifenden, pädagogischen Wissens als auch ihres domänenspezifischen, fachdidaktischen Wissens angemessen zu interpretieren.

Mittwoch
11. März

15:45 Uhr

Foyer GC - Ost

Leber, Jasmin; Skuballa, Irene T.; Renkl, Alexander

Können Blickbewegungsdaten genutzt werden, um Wissenslücken effizienter zu schließen?

Theoretischer Hintergrund: Adaptive Lernsysteme bieten die Möglichkeit, möglichst zeitnah auf das Verhalten eines Lernenden reagieren zu können. So wird z. B. durch Online-Analysen von Prozessdaten ein schnelles und zielgenaues Eingreifen und damit die individuelle Förderung eines Lernenden möglich (Shute & Zapata-Rivera, 2008). Um im Verlauf der Wissensaneignung Informationen über den aktuellen Wissensstand eines Lernenden zu erhalten, haben sich Rapid-Assessment Tasks (RATs) als geeignete Technik bewährt (Kalyuga, 2006a; 2006b). RATs sind kurze in die Lernumgebung eingestreute Fragen, welche den Wissensstand bzw. Wissenslücken eines Lernenden erfassen.

Eine weitere Möglichkeit zur Erfassung von Prozessdaten bietet Eye-Tracking. In den vergangenen Jahren konnten Blickbewegungsindikatoren identifiziert werden, die Aufschluss über Lernprozesse erlauben. So werden beispielsweise Wechsel zwischen Text und Bild als Integration von Text- und Bildinformationen interpretiert, wobei die Anzahl von Wechseln mit dem Lernerfolg assoziiert wird (Mason, Pluchino, Tornatora & Ariasi, 2013). Insbesondere diese Text-Bild-Integration bereitet Lernenden jedoch vielfach Schwierigkeiten (z.B. Ainsworth, 2006; Cierniak, Scheiter & Gerjets, 2009). Zur Förderung des Lernerfolgs erscheint daher Unterstützung bei der Verarbeitung von multimedialen Lernmaterialien sinnvoll und erforderlich (z.B. Kombartzky, Plötzner, Schlag & Metz, 2010). Diese kann mittels eines adaptiven Lernsystems angepasst an die individuellen Bedürfnisse Lernender erfolgen.

Wir entwickelten für computerbasierte, multimediale Lernsysteme einen zweifach adaptiven Mechanismus. Dabei werden die Blickbewegungen eines Lernenden und RATs genutzt, um Wissenslücken zu identifizieren. Blickbewegungsanalysen liefern "Verdachtsmomente" für Wissenslücken, woraufhin entsprechende RATs ausgelöst werden (Adaptivität I). RATs verifizieren oder entkräften den Verdacht. Bei verifizierten Wissenslücken werden Prompts und instruktionale Erklärungen präsentiert (Adaptivität II).

Fragestellung: In einer experimentellen Studie wurden zwei Varianten des Systems untersucht: In einer Variante wurden sowohl RATs, als auch Prompts adaptiv präsentiert (Adaptivität I&II), in einer zweiten Variante wurden alle RATs vorgegeben und nur die Prompts adaptiv ausgelöst (nur Adaptivität II). Es wurde angenommen, dass die adaptive Präsentation von RATs (Adaptivität I&II) eine höhere Sensitivität bei der Detektion von Wissenslücken aufweist als die Vorgabe aller RATs (Hypothese 1). Zudem wurde eine höhere Lerneffizienz erwartet, da eine reduzierte Zahl von RATs Zeit sparen sollte (Hypothese 2).

Methode: Studierende unterschiedlicher Fachrichtungen (N = 56; Alter: M = 22.96, SD = 3.11) wurden randomisiert zu zwei Versuchsgruppen zugeteilt (Adaptivität I&II vs. Adaptivität I). Die Probanden unterschieden sich nicht hinsichtlich Vorwissen, $t(54) = 0.880$, $p = .383$, und Arbeitsgedächtniskapazität, $t(54) = -0.363$, $p = .718$.

Nach der Erhebung von Vorwissen und Arbeitsgedächtniskapazität bearbeiteten die Probanden ein multimediales Lernmaterial (Domäne: Strategisches Management), während ihre Blickbewegungen aufgezeichnet und RATs präsentiert wurden. Abschließend wiederholten alle Teilnehmenden alle RATs und bearbeiteten einen Nachtest mit weiterführenden Fragen zum Lerninhalt.

Ergebnisse: Unsere Annahme bezüglich der Sensitivität (Hypothese 1) konnte bestätigt werden. Die adaptive Präsentation von RATs (Adaptivität I&II) wies eine signifikant höhere Sensitivität bei der Identifikation von Wissenslücken auf als die Vorgabe aller RATs (nur Adaptivität II), $t(53) = 2.217$, $p = .031$, $d = 0.605$. Das bedeutet, dass in der doppelt-adaptiven Bedingung ein größerer Anteil der RATs zur Identifikation von Wissenslücken führte. In der Bedingung, in der alle RATs vorgegeben wurden, erhielten die Lernenden substantiell mehr Fragen, als nötig gewesen wäre, $t(35) = -8.643$, $p < .001$, $d = -2.339$.

Mit Blick auf die Lerneffizienz (Hypothese 2) konnten keine signifikanten Unterschiede zwischen den Versuchsgruppen gefunden werden, $t(55) = 1.097$, $p = .277$.

Diskussion: Zusammenfassend kann festgehalten werden, dass Blickbewegungsdaten geeignet sind, um adaptiv Ereignisse in Lernumgebungen auszulösen und Wissenslücken zu identifizieren. Wir nehmen an, dass zur Steigerung der Lerneffizienz die Auslösemechanismen, die der adaptiven Präsentation der RATs zu Grunde liegen, weiter differenziert werden müssen. So kann sichergestellt werden, dass die adaptive Präsentation von RATs zeitsparend und effizient umgesetzt wird. Hierfür wird das Blickverhalten der Lernenden während der Lernphase weiteren umfassenden Analysen unterzogen werden.

Mittwoch
11. März

15:45 Uhr

Foyer GC - Ost

3

Pawelzik, Janina; Leuchter, Miriam; Möller, Kornelia

Wie wirkt sich ein Einführungspraktikum auf die Entwicklung der Selbstwirksamkeitserwartungen von Sachunterrichtsstudierenden zum naturwissenschaftlichen Unterrichten aus?

Viele Sachunterrichtslehrkräfte vermeiden das Unterrichten naturwissenschaftlicher Themen in der Grundschule (Appleton, 2003); sie haben wenig Interesse, sich mit naturwissenschaftlichen und technischen Themen auseinander zu setzen (Merkens, 2012) sowie nur geringes Vertrauen in ihre eigenen Fähigkeiten (Möller, 2004; Tosun, 2000). Geringes Vertrauen in die eigenen Fähigkeiten kann jedoch das eigene Lehrerverhalten im naturwissenschaftlichen Unterricht wiederum negativ beeinflussen (Appleton & Kindt, 2002). Daher kommt der Ermöglichung von Praxiserfahrungen sowie der Veränderung der negativen Einstellungen und dem Aufbau positiver Selbstwirksamkeitserwartungen bezüglich des eigenen naturwissenschaftlichen Unterrichts im Rahmen der Lehrerausbildung schon zu Beginn des Studiums besondere Bedeutung zu.

Forschungsbefunde zeigen, dass positive Veränderungen der eigenen Selbstwirksamkeitserwartungen von Studierenden im naturwissenschaftlichen Unterricht zum Beispiel durch Praxisphasen erreicht werden können (Velthuis, Fisser & Pieters, 2014). Unklar ist bisher jedoch, wie Praxisphasen gestaltet sein sollten, um einen positiven Einfluss auf die Selbstwirksamkeitserwartung der Studierenden in diesem Bereich zu haben. Studien zeigen, dass besonders die betreuenden Lehrpersonen mit ihren Voraussetzungen und ihrem Wissen einen Einfluss auf die Wirksamkeit der Praxisphase haben (z.B. Hascher, 2011).

An dieser Stelle setzt das ITPP-Projekt („Integration von Theorie und Praxis – Partnerschulen“) an. Grundidee des Projektes ist die Kooperation von Dozierenden der Universität Münster und Lehrkräften an sogenannten Partnerschulen im Regierungsbezirk Münster, um die Betreuung von Studierenden in Praxisphasen zu verbessern. Dazu werden die teilnehmenden Lehrpersonen (N=24) durch mehrjährige Weiterbildungsangebote der Universität Münster im Bereich des naturwissenschaftlichen Sachunterrichts qualifiziert. Ein Ziel des ITPP-Projektes ist es, die Selbstwirksamkeitserwartungen der Studierenden bezüglich naturwissenschaftlichen Unterrichts schon zu Beginn des Studiums durch ein ITPP-Praktikum zu fördern, das von den weitergebildeten Lehrkräften betreut wird. Den Studierenden wird deshalb die Möglichkeit geboten, ihr erstes Praktikum im Studium alternativ zu dem „normalen“ Regelpraktikum als ITPP-Praktikum im Bereich des naturwissenschaftlichen Sachunterrichts in Kleingruppen (2 – 4 Studierende) durchzuführen.

In der vorgestellten Studie untersuchen wir, inwiefern das ITPP-Praktikum mit einer Betreuung durch geschulte Lehrkräfte die Selbstwirksamkeitserwartungen von Studierenden in Bezug auf naturwissenschaftliches Unterrichten beeinflusst. Dazu wurde ein Onlinefragebogen jeweils nach dem ersten Praktikum im Studium bei Sachunterrichtsstudierenden der Westfälischen Wilhelms-Universität Münster mit ITPP-Praktikum (n=55) bzw. mit Regelpraktikum (n=70) eingesetzt. In diesem schätzten die Studierenden auf einer vierstufigen Ratingskala (1=„stimmt gar nicht“ – 4=„stimmt genau“) ein, wie sich ihre Selbstwirksamkeit in Bezug auf das Unterrichten naturwissenschaftlicher Themen durch das Orientierungspraktikum verändert hat (3 Items, $\alpha=.90$). Zusätzlich wurde erfasst, ob die Studierenden in ihrem Praktikum selbst unterrichtet haben und ob sie von einer Lehrkraft mit Bezug zum Sachunterricht betreut wurden (1=Ja, 0=Nein). Zudem gaben die Studierenden die Anzahl der Praktikanten pro Lehrkraft an. Mittels eines standardisierten Fragebogens (Kleickmann, 2008) wurden vor Beginn des Praktikums folgende Kontrollvariablen erhoben: das Interesse der Studierenden an Physik (3 Items, $\alpha=.80$), die Einschätzungen ihrer Fähigkeiten in Physik (4 Items, $\alpha=.76$), das Interesse der Studierenden am Unterrichten physikbezogenen Sachunterrichts (3 Items, $\alpha=.91$) sowie die Einschätzung ihrer Kompetenzen im Unterrichten physikbezogenen Sachunterrichts (3 Items, $\alpha=.80$).

Die Ergebnisse der Varianzanalysen zeigen, dass sich die Studierenden mit ITPP-Praktikum und die Studierenden mit Regelpraktikum bezüglich der motivationalen Eingangsvoraussetzungen vor ihrem Praktikum nicht voneinander unterscheiden ($V=0.061$, $F(4,106)=1.730$, $p=.149$, $\eta_p^2=.061$). Vor dem Hintergrund dieses Ergebnisses erfolgte der Vergleich der beiden Praktikumsgruppen bezüglich der Einschätzung ihrer Selbstwirksamkeitserwartungen nach dem Praktikum: Studierende, die das ITPP-Praktikum absolviert haben, schätzen ihre Selbstwirksamkeitserwartungen bezüglich naturwissenschaftlichen Unterrichts signifikant höher ein als Studierende mit Regelpraktikum ($F(1,122)=9.968$, $p=.002$, $\eta_p^2=.076$). Der positive Effekt des ITPP-Praktikums auf die Einschätzung der eigenen Selbstwirksamkeitserwartungen wird teilweise dadurch beeinflusst, ob die Studierenden selbst unterrichtet haben, ob die Lehrkraft einen Bezug zum Sachunterricht hatte und wie viele Studierende von einer Lehrkraft betreut wurden. Im Vortrag werden diese Zusammenhänge mit Hilfe eines Mediatormodells dargestellt. Zudem werden das Konzept des ITPP-Projektes sowie die Ergebnisse der Befragung vorgestellt und diskutiert.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Rest, Miriam; Minkley, Nina; Kirchner, Wolfgang H.

Lehramtsstudierende im Praktikum: Der Einfluss fachwissenschaftlicher Leistung und pädagogischer Erfahrung auf die Kompetenzentwicklung

Da Lehramtsstudierende im Schulunterricht mit zahlreichen neuen Situationen und vielfältigen Aufgaben konfrontiert sind, nehmen Praktika einen zentralen Stellenwert innerhalb der Lehrerbildung ein. Obwohl ihnen eine besondere Bedeutung und Wirksamkeit zugeschrieben wird, sind empirische Befunde zu diesem Themenbereich heterogen. So belegen z.B. verschiedene Untersuchungen zur Wirksamkeit schulischer Praxisphasen, dass sich sowohl erwünschte (z.B. Identifikation mit dem Lehrberuf) als auch unerwünschte (z.B. autoritäres Handeln) Entwicklungen während des Praktikums vollziehen (Gemsa und Wendland 2011; Hascher 2012; Jäger und Milbach 1994).

Aufgrund dieser z.T. widersprüchlichen Befunde ist bislang noch ungewiss, was den Professionalisierungsprozess angehender Lehrer ausmacht und von welchen Faktoren dieser beeinflusst wird.

Daher wird in der vorliegenden Studie die Frage untersucht, welche Faktoren die Kompetenzentwicklung angehender Biologielehrer in Schulpraktika beeinflussen. Hierbei werden sowohl das biologische Fachwissen als auch pädagogische Vorerfahrungen betrachtet.

Die Probandengruppe der ersten Untersuchungskohorte setzt sich aus 17 Lehramtsstudierenden zusammen, die im Master of Education an der Ruhr-Universität Bochum sowohl Biologie als auch ein weiteres Fach für die Sekundarstufe II Gymnasium/Gesamtschule studieren. Die Praktikumsvorbereitung fand im Rahmen eines fachdidaktischen Seminars statt. Die Dauer des Schulpraktikums betrug vier Wochen, in denen die Studierende wöchentlich mind. 15 Stunden im Biologieunterricht hospitierten sowie mind. eine Unterrichtseinheit eigenständig planten, durchführten und reflektierten.

Um eventuelle Veränderungen hinsichtlich der selbsteingeschätzten berufsbezogenen Kompetenzen in den Bereichen Unterrichten und Erziehen zu erheben, wurden diese jeweils zu drei Messzeitpunkten – vor Beginn, während und nach Abschluss der Praxisphase – mittels eines Onlinefragebogens (EvaStud) erhoben. Die Kompetenzniveaus wurden mithilfe der von Gröschner (2008) entwickelten Standards ermittelt. Diese Standards orientieren sich an den von der Kultusministerkonferenz entwickelten Standards (Standards für die Lehrerbildung: Bildungswissenschaften 2004). Hierbei geben die Studierenden ihre jeweiligen Kompetenzniveaus anhand einer siebenstufigen bipolaren Skala („gar nicht kompetent“ bis „voll kompetent“) an (Gröschner 2008).

Zudem wurde das biologische Fachwissen (Leistung) anhand der Biologie-Fachnote im Bachelor erfasst. Diese setzt sich aus den Noten der Grundmodulprüfungen in den Bereichen Zoologie und Zellbiologie, Botanik und Biodiversität sowie Physiologie und molekulare Biologie zusammen. Außerdem wurden die pädagogischen Vorerfahrungen der Studierenden erfasst. Hierbei handelt es sich um außeruniversitäre Vorerfahrungen in den Bereichen: Betreuung von Kindern, Gestaltung von Freizeitaktivitäten, Nachhilfeunterricht/Hausaufgabenbetreuung als Einzel- oder Gruppenunterricht, Unterrichtstätigkeit in der Schule, Ausbildung, Zivildienst oder Freiwilliges soziales Jahr im pädagogischen Bereich.

Die Ergebnisse der ersten Untersuchungskohorte zeigen, dass das selbsteingeschätzte Kompetenzniveau der Studierenden in den Bereichen Unterrichten und Erziehen vom Zeitpunkt „vor Beginn des Praktikums“ zum Zeitpunkt „während des Praktikums“ signifikant ansteigt ($F(1, 15) = 7,47$; $p = 0,01$; $\eta^2 = 0,33$; ANOVA). Zudem besteht ein Zusammenhang zwischen den beiden Kompetenzbereichen: Studierende, die ihr Kompetenzniveau im Bereich Unterrichten hoch einstufen, stuften dieses auch im Bereich Erziehen hoch ein ($p < 0,001$; $r_p = 0,81$, für Unterrichten Zeitpunkt 1 und Erziehen Zeitpunkt 1; $p < 0,01$; $r_p = 0,65$ für Unterrichten Zeitpunkt 2 und Erziehen Zeitpunkt 2).

Es lässt sich kein signifikanter Zusammenhang zwischen dem Kompetenzniveau und dem biologischen Fachwissen erfassen. Demnach bewerten fachwissenschaftlich leistungsstärkere und leistungsschwächere Studierende ihre Kompetenzniveaus ähnlich.

In Bezug auf die pädagogischen Vorerfahrungen zeigen die Untersuchungsergebnisse, dass Studierende, die Unterrichtserfahrungen in der Schule außerhalb des Studiums aufweisen (bspw. als Vertretungslehrer), ihre Kompetenzniveaus tendenziell höher einstufen als Studierende ohne diese Erfahrungen ($F(1, 15) = 3,26$; $p = 0,09$; $\eta^2 = 0,18$; ANOVA). Zwischen den anderen untersuchten pädagogischen Vorerfahrungen und den Kompetenzniveaus besteht kein signifikanter Zusammenhang.

Ob der Anstieg der selbsteingestuftenen Kompetenz innerhalb der Praxisphase nachhaltig ist, wird in dem noch ausstehenden Messzeitpunkt „nach Abschluss des Praktikums“ erfasst. Zudem werden sowohl weitere Kohorten im Rahmen der vierwöchigen Schulpraktika als auch im neu eingeführten (in NRW) fünfmonatigen Praxissemester untersucht.

Windt, Anna; Lenske, Gerlinde

Mit welchen Fähigkeiten zur Reflexion von Sachunterrichten starten angehende Lehrkräfte in den Vorbereitungsdienst?

Ausgangslage: Das Unterrichten ist eine zentrale Kompetenz, die angehende Lehrkräfte im Vorbereitungsdienst entwickeln sollen (KMK 2004). Dabei werden drei Kompetenzbereiche unterschieden: die Planung, Durchführung und Reflexion des Unterrichts (Merkens 2010). Für den Sachunterricht ist ungeklärt, wie die Entwicklung der Kompetenz des Unterrichts verläuft. Das Projekt EuLe (Windt/ Rumann 2014) untersucht deswegen in drei Teilprojekten die Entwicklung der Kompetenzbereiche Planung, Durchführung und Reflexion von Sachunterricht im Vorbereitungsdienst. Dieser Beitrag fokussiert auf das Teilprojekt zur Reflexion.

Theoretischer Hintergrund: Reflexionsfähigkeit ist eine Voraussetzung für die Entwicklung beruflicher Professionalität, da erst die Wechselwirkung zwischen praktischem Handeln und theoretischer Reflexion eine Weiterentwicklung ermöglicht (Dubs 2008, Helmke 2009, Kansanen 2000, Schön 1983). Graff (2010) geht davon aus, dass die Reflexion im Sachunterricht besondere Relevanz hat, da die Lehrkräfte aufgrund der Vielperspektivität in ihrem Studium Inhalte und Methoden der Bezugsdisziplinen nur exemplarisch lernen.

Reflexion ist eine bestimmte Form des Denkens, ein mentaler Prozess, bei dem es darum geht, Probleme, Situationen, Erfahrungen oder Wissen kognitiv zu strukturieren (Dewey 1933, Hatton/ Smith 1995, Korthagen 2002, Roters 2012). Bei einer Reflexion geht es sowohl um die Rückschau auf erfolgte Handlungen, um sie retrospektiv zu analysieren, als auch um die Vorausschau, um prospektiv mögliche Konsequenzen zukünftigen Handelns abzuschätzen (Valli 1997). Es ist davon auszugehen, dass Reflexionsfähigkeit durch praktische Anwendung erlernbar ist (Dewey 1933, Roters 2012).

Die empirische Erfassung des Konstrukts Reflexion ist schwierig (Calderhead/ Shorrock 1997, Korthagen 2002). Sie ist nur über Rekonstruktionen möglich, die sowohl Novizen als auch Experten schwer fällt (Hatton/ Smith 1995, Roters 2012). Deswegen sollte die Erfassung möglichst kurz nach der Handlung stattfinden (Leuchter/ Reusser/ Pauli/ Klieme 2008).

Forschungsfragen & Hypothesen:

F1: Wie verändern sich die Inhalte der Reflexion von Sachunterricht im Verlauf des Vorbereitungsdienstes?

H1: In der ersten Hälfte des Vorbereitungsdienstes stehen Aspekte der Unterrichtsorganisation und -strukturierung im Fokus, in der zweiten Hälfte Aspekte, die die Individualebene der SuS einbeziehen.

F2: Wie verändert sich die Qualität der Reflexion von Sachunterricht im Verlauf des Vorbereitungsdienstes?

H2: Die Qualität der Reflexion von Sachunterricht nimmt im Verlauf des Vorbereitungsdienstes zu.

Methoden & Design: Da es sich um qualitative Daten handelt, ist die Stichprobe mit 15 LehramtsanwärterInnen (LAA) entsprechend klein. Es werden die mündlichen Reflexionen im Rahmen der verpflichtenden Unterrichtsbesuche längsschnittlich am Anfang, in der Mitte und am Ende des Vorbereitungsdienstes genutzt. Die Reflexionen werden audiographiert, transkribiert und anhand eines Kategoriensystems unter Verwendung der strukturierten Inhaltsanalyse nach Mayring (2000) analysiert.

Dabei werden Inhalte, Veridikalität (Wirklichkeitsgetreue der Reflexion), Niveaustufe, Objektivität und Strukturiertheit kodiert. Die Niveaustufen wurden in Anlehnung an die Reflexionstiefe von Hatton/ Smith (1997) entwickelt: LAA auf Stufe 1 beschreiben in ihrer Reflexion ausschließlich Unterrichtssituationen. LAA auf Stufe 2 bewerten die Unterrichtssituationen zusätzlich hinsichtlich ihrer Qualität. LAA auf Stufe 3 begründen zusätzlich ihre Bewertung. LAA auf Stufe 4 entwickeln zusätzlich alternative Handlungsweisen und LAA auf Stufe 5 leiten zusätzlich aus ihrer Reflexion Konsequenzen ab.

Zur Validierung schätzen die FachleiterInnen und MentorInnen die Reflexion bezüglich derselben Kriterien des Kategoriensystems ein.

Ergebnisse: In der Pilotierung konnte ein Kategoriensystem zur Erfassung der Qualität der Reflexion von Sachunterricht entwickelt und validiert werden. Die ersten Daten des ersten Messzeitpunktes zeigen, dass die LAA sich sowohl hinsichtlich der Inhalte der Reflexion als auch hinsichtlich der Qualität der Reflexion voneinander unterscheiden und dass die Reflexion den LAA im Anschluss an ihr Studium zum Teil schwer fällt. Das Poster berichtet über die Daten des ersten Messzeitpunktes und zeigt auf, mit welchen Fähigkeiten zur Reflexion von Sachunterricht die angehenden Lehrkräfte in den Vorbereitungsdienst starten. Abschließend werden die Ergebnisse hinsichtlich der ersten Ausbildungsphase betrachtet und mögliche Interventionen zur gezielteren Schulung der Reflexionsfähigkeit während des Studiums diskutiert.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Funke, Christina; Langolf, Daniela; Clausen, Marten; Abs, Hermann Josef

Lernen durch Engagement im Praxissemester – Eine quasiexperimentelle Wirksamkeitsstudie

Studierende und Lehrerausbildende an Schulen schätzen die Wirksamkeit von Praxisphasen in der Regel als sehr hoch ein. Aktuellere Studien zeigen jedoch, dass sich die Wirkungen und Wahrnehmungen sowohl zwischen Studierenden als auch vor, während und nach den Praxisphasen unterscheiden (z.B. Hascher, 2006). Dies gilt es auch für das neue Praxissemester im Rahmen des Masterstudiengangs Lehramt in Nordrhein-Westfalen zu überprüfen, insbesondere mit Blick auf die professionelle und allgemeine Kompetenzentwicklung sowie Wahrnehmung und Einstellung der Studierenden. Die hier vorgestellte Studie untersucht die Kompetenzwirksamkeit verschiedener Gestaltungsformen des Praxissemesters, wobei insbesondere Service Learning als Lehr-Lernform in den Blick genommen wird. Service Learning (auch Lernen durch Engagement) verbindet als demokratiepädagogischer Ansatz in spezifischer Weise die Kompetenzvermittlung mit gesellschaftlichem Engagement (Backhaus-Maul & Roth, 2013) und erzielt hinsichtlich verschiedener Kompetenzen Wirkungen, die sich grob drei Bereichen zuordnen lassen (Billig, 2000, Seifert, 2011): (1) Akademisches Lernen: Studierende, die Service Learning betrieben, zeigten deutlich bessere Leistungen, als jene, die dies nicht taten (Furco, 2004). (2) Soziale und persönliche Kompetenzen: Die Teilnahme an Service Learning Projekten verbessert soziale Kompetenzen (Furco, 2002) und steigert das Selbstwirksamkeitsgefühl (Weiler et al., 1998). (3) Zivilgesellschaftliche Einstellungen und Bürgerschaftliches Engagement: Service Learning schafft eine höhere Sensibilität für Probleme in der Gemeinde (Yates & Youniss, 1996) und steigert die politische Identität (Yates & Youniss, 1997). Diese Wirkungen beziehen sich auf die Kompetenzentwicklung derjenigen, die selbst den Service durchführen, den Service-Gebern. Kaum untersucht wurde hingegen die Wirksamkeit für diejenigen, die Service Learning Projekte anleiten, den Service-Geber-Anleitern. Zudem wurden hauptsächlich die Wirkungen auf Schülerinnen und Schüler untersucht, während Lehramtsstudierende bislang noch kaum fokussiert wurden.

Als übergeordnete Fragestellung wird in dieser Studie betrachtet, inwiefern sich die professionsbezogenen und allgemeinen Kompetenzen von Studierenden über das Praxissemester unter divergierenden Vorbereitungs- und Durchführungsvorgaben beim Einsatz von Service Learning unterscheiden. Zwei Gruppen von Studierenden werden Service Learning in ihrem Praxissemester in unterschiedlichen pädagogischen Rollen umsetzen: als Service-Geber oder als Service-Geber-Anleiter. Diese Gruppen werden kontrastiert mit Studierenden, die in ihrem Praxissemester kein Service Learning anbieten. Den Rahmen der Untersuchung bildet das in NRW neu in die Lehramtsausbildung aufgenommene Praxissemester. Dieses wird an der Universität Duisburg-Essen erstmals im Wintersemester 2014/2015 anhand der vorbereitenden Veranstaltung, der Forschungswerkstatt BiWi I, eingeführt und im Sommersemester 2015 realisiert. Zur Untersuchung der Wirksamkeit ist ein quasi-experimentelles Design mit drei jeweils sechs Monate auseinanderliegenden Messzeitpunkten vorgesehen (Pretest, Posttest und Follow-Up, jeweils als Onlinebefragung über limesurvey). Dabei bildet die Art der Maßnahme die quasiexperimentelle Unabhängige Variable ((1) Studierende als Service-Geber vs. (2) als Service-Geber-Anleiter). Als Kontrollgruppe dienen (3) Studierende aus anderen Veranstaltungen. Die Studierenden der Service-Learning-Gruppen (jeweils ca. 40 Teilnehmende) werden zunächst allgemein und später in je zwei Forschungswerkstätten speziell auf ihre Rollen als Service-Geber-Anleiter bzw. Service-Geber im Praxissemester vorbereitet, wobei die Gruppenzuweisung durch Randomisierung erfolgt.

Zur Messung der Kompetenzentwicklung der Studierenden (Abhängige Variable) im Praxissemester werden an den KMK-Standards orientierte, etablierte Instrumente (vgl. Gröschner & Schmitt, 2010) zur Erfassung von professionellem Selbstkonzept und Professionsethos eingesetzt werden. Zusätzliche Skalen erfassen die für Service Learning spezifischen Teilkompetenzen, Einstellungen und Werthaltungen. Eine qualitative Follow-Up-Untersuchung mit semistrukturierten themenzentrierten Leitfadenterviews von vier Akteursgruppen (Studierende, anleitende Lehrkräfte, Schülerinnen und Schüler, Schulleitung) soll vertiefende Erkenntnisse bringen.

Die erste Erhebungswelle erfolgt im Januar 2015. Bis zum Zeitpunkt der Tagung können erste Erfahrungen zur Forschungswerkstatt und Ergebnisse der ersten Erhebung präsentiert werden. Das Projekt widmet sich, mit seinem Fokus auf das Praxissemester, einem wenig erforschten, jedoch hochaktuellen Thema (Arnold, Gröschner & Hascher, 2014). Die längsschnittliche Betrachtung der Kompetenzentwicklung leistet einen wichtigen Beitrag zur evidenzbasierten Gestaltung und Optimierung dieser Ausbildungsphase. Gleichzeitig wird der Lernzuwachs auf Seiten der Lehrenden, der Service-Geber-Anleiter, untersucht. Außerdem wird der innovative Service Learning Ansatzes weiter etabliert, der in der erziehungswissenschaftlichen Lehrerausbildung bislang wenig Raum gefunden hat.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Liegmann, Anke B.; Racherbäumer, Kathrin; van Ackeren, Isabell

Lernen im Praxissemester – Kontextabhängige Entwicklung von Reflexionskompetenz im Kompetenzbereich „Umgang mit Heterogenität“

Theoretischer Hintergrund: Mit der Einführung einer mehrmonatigen Praxisphase (Praxissemester) in Nordrhein-Westfalen erfolgt im Rahmen der Umstellung auf Bachelor-Masterstudiengänge eine Reform der Lehrerbildung, die eine Ausweitung der in der ersten Ausbildungsphase verorteten Praxisanteile in der Lehrerbildung vorsieht. Gleichwohl sind stabile Befunde zum Beitrag von Praxisphasen zur Professionsentwicklung angehender Lehrkräfte beispielsweise aufgrund der sehr unterschiedlichen Settings der Praxisphasen sowie ihrer Komplexität noch nicht ausreichend vorhanden (Hascher, 2012). Die Verknüpfung von Theorie und Praxis gilt als das anspruchsvollste Ziel schulischer Praxisphasen. Dieses Ziel bildet den Schwerpunkt des nordrhein-westfälischen Praxissemesters (MSW NRW, 2010, S. 4). Eine Verknüpfung von Theorie und Praxis beispielsweise durch die Parallelisierung von theoretischem Input und praktischen Erfahrungen ist allein noch kein Garant für einen gelungenen Theorie-Praxis-Transfer, vielmehr bedarf es der gezielten Unterstützung der im Praktikum Lernenden (Artmann, Herzmann, Hoffmann & Proske, 2013; Racherbäumer & Liegmann, 2012). Diese Unterstützung zielt vor allem auf die Entwicklung von Reflexionskompetenz, anhand derer erst die Zusammenführung wissenschaftlichen Wissens und praktischen Könnens erfolgen kann (Hatton & Smith, 1995; Neuweg, 2005). Allerdings ernüchtern die Befunde zur Reflexionskompetenz von Studierenden. In Studien, die ein komplexeres Stufenmodell von Reflexionskompetenz zu Grunde legen, wird eine anspruchsvolle reflexive Durchdringung schulischer Praxisphasen kaum erreicht (z.B. Artmann et al., 2013; Racherbäumer & Liegmann, 2012).

Zudem weisen sowohl Hascher (2012), als auch König et al. (2014) darauf hin, dass in der Praxisphasenforschung oftmals die Prozessqualität der Praxis nur ungenügend spezifiziert ist (Hascher, 2012, S. 124). Dabei kann man davon ausgehen, dass sich individuelle Lerngelegenheiten beispielsweise nach Schulform, Schulgröße, Zusammensetzung der Schülerschaft, Qualität der Betreuung und/oder Akzeptanz der Praktikant/inn/en an der Schule unterscheiden. Insbesondere durch die Einbettung von Schulen in divergierende sozialräumliche Kontexte (Racherbäumer & van Ackeren, 2014) können sich mit Blick auf kontextunterschiedlichen Kompositionen der Lerngruppen an Schulen verschiedener Standorttypen (zum Teil) für Studierende differenzielle Lernmilieus ergeben. Es ist davon auszugehen, dass sich dies in der reflexiven Bearbeitung des Handlungsfeldes „Umgang mit Heterogenität“ widerspiegelt.

Forschungsfragen: Im Einzelnen werden im Projekt folgende Fragen bearbeitet: (1) Welche Lerngelegenheiten im Praxissemester lassen sich standortspezifisch unter besonderer Berücksichtigung des Umgangs mit Heterogenität differenzieren und ggf. typisieren? (2) Wie werden standortspezifische Lerngelegenheiten durch die Studierenden genutzt? (3) Wie gelingt es den Studierenden, die standortspezifischen, insbesondere Diversity-bezogenen Lerngelegenheiten im Praxissemester reflexiv auf ihre theoretischen Kenntnisse zu beziehen? (4) Welche Rolle kommt den Akteuren der drei unterschiedlichen Lernorte (Universität, Schule, Zentren für schulpraktischen Lehrerbildung) bei der Reflexion der Erfahrungen im Praxissemester zu? (5) Welche strukturellen und inhaltlichen Entwicklungsoptionen lassen sich für eine kontextsensible Professionsentwicklung im Praxissemester ableiten?

Forschungsdesign: Längsschnittstudie (Beginn Praxissemester im Februar 2015 bis Eintritt in den Vorbereitungsdienst Februar 2017) mit multimethodischem Zugang:

(A) Quantitative standardisierte Befragungen (gesamte erste Kohorte im Praxissemester der Universität Duisburg-Essen, N = ca. 300): Erfassung des Theoriewissens mit Blick auf das Themenfeld „Umgang mit Heterogenität“ (Tests aus dem Projekt BilWiss, vgl. z.B. Terhart et al. 2012); Erfassung von Lerngelegenheiten an Schulen unterschiedlicher Standorttypen im Praxissemester (Instrument von König et al. 2014, Erweiterung auf spezifische Lernmilieus).

Auswertung mittels deskriptiver und multivariater Verfahren (z.B. Clusteranalysen).

(B) Qualitative Erhebungen (12 maximal kontrastierende Einzelfallstudien (Standorttyp I vs. Standorttyp V): Beobachtungen an verschiedenen Lernorten (Schule, ZfsL); Interviews zu mehreren Zeitpunkten

Auswertung rekonstruktiv mittels dokumentarischer Methode (vgl. z.B. Martens & Asbrand 2009, Pfaff & Krüger 2009).

(C) Dokumentenanalyse: Schriftlich angefertigte Reflexionen im Rahmen der Modulprüfung

Auswertung rekonstruktiv in Anlehnung an Artmann et al. (2013) (Stufenmodell Reflexionskompetenz)

Ergebnisse: Das geplante Projekt (Start November 2014) wird mit einem Poster zur Diskussion bei der GEBF vorgestellt. Zum Zeitpunkt der Tagung werden noch keine Ergebnisse vorliegen.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Herde, Christoph Nils; Stegt, Stephan; Preckel, Franzis

Auswahlverfahren für Masterstudiengänge aus Sicht von Bachelorstudierenden: Eine multimethodale Studie zur Erfassung der Akzeptanz

Theoretischer Hintergrund: Die hochschuleigene Studierendenauswahl birgt Potentiale zur Steigerung der Wettbewerbsfähigkeit von Hochschulen. Diverse Metaanalysen untersuchten entsprechend die prognostische Validität verschiedener Verfahren der Studierendenauswahl (z.B. Hell, Trapmann & Schuler, 2007; Hell, Trapmann, Weigand & Schuler, 2007; Trapmann, Hell, Hirn & Schuler, 2007; Trapmann, Hell, Weigand & Schuler, 2007). Befunde aus dem Kontext der Personalauswahl legen jedoch nahe, dass neben der prognostischen Validität auch die Wahrnehmung eignungsdiagnostischer Verfahren durch die Bewerber zu berücksichtigen ist, da diese z.B. mit Leistungen im Auswahlverfahren, Attraktivitätsurteilen über Unternehmen sowie Intentionen zur Annahme einer angebotenen Stelle oder Weiterempfehlung der Organisation zusammenhängt (Chapman, Uggerslev, Carroll, Piasentin & Jones, 2005; Hausknecht, Day & Thomas, 2004; Uggerslev, Fassina & Kraichy, 2012). Metaanalytische Befunde identifizierten v.a. Bezüge zur Zieltätigkeit sowie die wahrgenommene Augenschein- und prognostische Validität als bedeutsame Determinanten der Akzeptanz (Chan & Schmitt, 2004; Hausknecht et al., 2004). Weitere Befunde legen eine negative Assoziation zwischen der Akzeptanz und der Verfälschbarkeit eignungsdiagnostischer Verfahren nahe (Gilliland, 1995; Steinmayr & Kersting, 2008). Insbesondere die Akzeptanz von Auswahlverfahren für Masterstudiengänge sollte für die Hochschulen von besonderem Interesse sein, da die Hochschulen aufgrund der noch jungen Historie von Masterstudiengängen bisher nur über wenig Erfahrung bzgl. der Gestaltung der zugehörigen Auswahlverfahren verfügen. Bisherige Untersuchungen zur Akzeptanz von Verfahren der Studierendenauswahl im deutschsprachigen Raum beziehen sich entweder auf grundständige Studiengänge (Hell & Schuler, 2005; Täger, 2010; Trost, 1993) oder weisen sehr eingeschränkte Stichproben auf (Stegt, 2011). Entsprechend bieten diese Studien nur eingeschränkte Handlungsempfehlungen zur Gestaltung von Auswahlverfahren für Masterstudiengänge an.

Fragestellung: Die vorliegende Studie exploriert die Akzeptanz von elf Verfahren der Studierendenauswahl für Masterstudiengänge. Aufgrund enger Bezüge zum entsprechenden Masterstudium ist eine hohe Akzeptanz von Bachelornoten, studienfachspezifischen Fähigkeitstests, freien sowie strukturierten Auswahlgesprächen, beruflichen Erfahrungen/Praktika und Arbeitsproben/Assessment Center zu erwarten. Begründet durch geringere Bezüge werden ungünstigere Akzeptanzurteile für Abiturnoten, Intelligenztests und Wartezeiten prognostiziert. Gemäß des Potentials zur aktiven Verfälschung wird eine geringe Akzeptanz von Motivations-/bzw. Begründungsschreiben und Persönlichkeitstest angenommen. Als potentielle Einflussfaktoren der Akzeptanz werden neben Vorerfahrungen mit den einzelnen Verfahren erstmals im Kontext der Studierendenauswahl die Big-Five-Persönlichkeitsfaktoren sowie das akademische Selbstkonzept ergründet.

Methoden: Die Fragestellung wurde mit Hilfe einer Online-Umfrage untersucht. Die Stichprobe umfasste 394 Bachelorstudierende (davon 266 Frauen), welche 19 Hochschulen und 29 verschiedene Studienfächer repräsentieren. Die Akzeptanz der einzelnen Verfahren wurde multimethodal über eine Rangreihung, eine Ratingskala sowie einen kompletten Paarvergleich erhoben. Zur Exploration möglicher Dimensionen des Akzeptanzurteils beurteilten die Teilnehmenden ausgewählte Verfahren auf einer Adjektivskala. Zudem wurden Vorerfahrungen mit den Verfahren, die Big-Five-Persönlichkeitsfaktoren und das akademische Selbstkonzept erhoben.

Ergebnisse: Die Ergebnisse der Rangreihung, Ratingskala und des Paarvergleichs zeigten eine hohe Konvergenz. Studienfachspezifische Fähigkeitstests und Bachelornoten erhielten die höchsten Akzeptanzwerte. Auch frei geführte sowie strukturierte Auswahlgespräche und berufliche Erfahrungen wurden positiv beurteilt. Persönlichkeitstests, Wartezeiten, Intelligenztests und Abiturnoten erlangten eher negative Bewertungen. Die Akzeptanzurteile korrelierten positiv mit der selbsteingeschätzten Leistung in Vorerfahrungen ($.23 < r < .43$). Deskriptive und teils statistisch bedeutsame Tendenzen deuten auf günstigere Akzeptanzurteile durch Teilnehmende mit Vorerfahrungen in den einzelnen Verfahren hin. Ein höheres Ausmaß an Neurotizismus war assoziiert mit einer geringeren Akzeptanz für Bachelornoten ($r = -.13$) und Intelligenztests ($r = -.16$). Ein höheres akademisches Selbstkonzept hing zusammen mit günstigeren Urteilen für Bachelornoten ($r = .23$), Intelligenztests ($r = .19$) und studienfachspezifischen Fähigkeitstests ($r = .11$) sowie mit ungünstigeren Urteilen für Wartezeiten ($r = -.20$). Gemäß einer explorativen Faktorenanalyse ließen sich die Bewertungen der Verfahren auf der Adjektivliste auf die drei Faktoren Modernität/Anregungsgehalt/Aussagekraft, generelle Evaluation (Objektivität, Fairness, Transparenz und Augenscheinvalidität) sowie affektive Wirkung zurückführen. Berufliche Erfahrung/Praktika erhielten die positivsten Urteile auf den Faktoren Modernität/Anregungsgehalt/Aussagekraft sowie affektive Wirkung, während studienfachspezifischen Fähigkeitstests bzgl. der generellen Evaluation der höchste Wert zugewiesen wurde.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

3

Jacob, Marita; Gerth, Maria; Weiss, Felix

Mach Bildung sofort gleich? Soziale Ungleichheiten in studentischer Erwerbstätigkeit und ihrer Entwicklung über die Studiendauer

Ungleichheiten in tertiärer Bildung, die über den sozial selektiven Hochschulzugang hinausgehen, können in vielerlei Hinsicht mit der studentischen Erwerbssituation in Verbindung gebracht werden. Der zeitliche Aufwand, der mit studentischer Erwerbstätigkeit einhergeht, erhöht die Arbeitsbelastung der Studierenden und reduziert die verfügbare Zeit für das Studium. Je nach Art der Erwerbstätigkeit, kann diese aber auch als Investition für den Arbeitsmarkteintritt und das spätere Berufsleben gesehen werden. So können sie mit Qualifikationen, die beim späteren Arbeitsmarkteintritt von den Arbeitgebern als wertvoll erachtet werden, oder mit dem Herstellen arbeitsmarktrelevanter Netzwerke und Kontakte einhergehen. Sollte die Aufnahme solcher „nützlicher“ Erwerbstätigkeiten sozial unterschiedlich verteilt sein, hätte das weitreichende Folgen für soziale Ungleichheiten über den Hochschulzugang hinaus. Forschungsarbeiten zur Frage sozialer Ungleichheit in der Jobqualität von Studierenden werden dennoch nur selten angestrengt. In unserem Beitrag wollen wir die Fragen untersuchen, inwieweit sich die Jobqualität Studierender mit in Zusammenhang mit der sozialen Herkunft steht und wie solche möglicherweise existierenden Unterschiede mit der Studiendauer entwickeln. Wir untersuchen mögliche Mechanismen, die erklären könnten, warum sich diese Erwerbsmuster nach sozialer Herkunft unterscheiden.

Während Erwerbstätigkeit in aller Regel mit monetären Erträgen einhergeht, bieten einige Jobs arbeitsmarktrelevante Qualifikationen und Kontakte als zusätzliche, nicht-monetäre, Erträge. Mit der Akkumulation von Humankapital während des Studiums und einer längeren Suchdauer, wird der Zugang zu diesen high quality-jobs erleichtert. Daher sollte für alle Studierenden mit zunehmender Semesteranzahl, der Anteil der Studenten in high quality jobs steigen.

Da Studierende mit niedriger sozialer Herkunft häufiger arbeiten müssen, um sich ihr Studium finanzieren zu können, sollten sie von Beginn an Jobs aufnehmen, ungeachtet ihrer Qualität. Studierende höherer Klassen sollten dagegen nur dann arbeiten, wenn die Erwerbstätigkeit auch den nicht-monetären Nutzen der high-quality jobs bringt. Mit zunehmender Studiendauer sollte jedoch, durch den Erwerb berufsrelevanter Fähigkeiten, die Eignung für high quality jobs für alle Studierenden steigen und die soziale Ungleichheit in der Jobqualität sinken

Diese Entwicklung erwarten wir vor allem für berufsspezifische Studiengänge, da sie den Studierenden mehr berufsrelevante Fähigkeiten vermitteln. Studienrelevante Jobs für Studierende der Geistes- und Naturwissenschaften sollten hingegen nach trainability vergeben werden, also eher nach soft skills. Da sie weniger im Studium vermittelt werden, sondern teilweise durch Sozialisation erlernt werden, haben Studierende der unteren Klassen in diesen Fächern geringere Möglichkeiten, durch Lernfortschritt im Studium aufzuholen. Wir erwarten daher, dass sich die anfängliche Lücke in der Jobqualität nach sozialer Herkunft weniger deutlich schließt.

Ausgehend von diesen theoretischen Überlegungen vermuten wir außerdem, dass die sozialen Unterschiede in der Qualität der Erwerbsarbeit durch die Motive „finanzielle Notwendigkeit“ und „Arbeitserfahrung sammeln“ sowie „Kontakte knüpfen“ erklärt werden können.

Um unsere Hypothesen zu testen, haben wir Daten der 16. Sozialerhebung analysiert. Wir vergleichen die Unterschiede in der Erwerbstätigkeit und ihrer Qualität nach sozialer Herkunft und wie sich diese Ungleichheit nach der Semesterzahl unterscheidet. In einem zweiten Schritt untersuchen wir mögliche Erklärungsmechanismen. Wir operationalisieren Jobqualität in verschiedenen Arten (Bezug von Job und Studienfach, Lohn, Art des Jobs und Intensität). Deskriptive Befunde stützen unsere Hypothesen. Unter Kontrolle verschiedener soziodemographischer, finanzieller und Studiums-bezogener Variablen finden wir allerdings keine signifikanten Interaktionen zwischen sozialer Herkunft und Semesterzahl. Für die zu erwartenden Unterschiede zwischen den Studienfächern finden wir keine Hinweise. Wir können die Ungleichheit und ihre Entwicklung über das Studium daher nicht auf die Erklärung zurückführen, dass Bildung zu einer schnellen und zeitnahen Angleichung der Erwerbssituation führt. Dekompositionen zeigen hingegen, dass Herkunfts- spezifische Verteilungen in den verschiedenen Erwerbsmotivationen einen Erklärungsbeitrag für die soziale Ungleichheit in der Jobqualität liefern.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Storck, Johanna

Berlin Interdisciplinary Education Network (BIEN)

Ziel des Netzwerkes ist es, Nachwuchswissenschaftler/-innen aus den unterschiedlichen Disziplinen der Bildungsforschung die Gelegenheit zu bieten, sich regelmäßig über ihre Einrichtung und Disziplin hinaus auszutauschen. Das Netzwerk setzt an einer Einzigartigkeit der Berliner Region an: In Berlin und Umgebung ist eine Vielzahl an Forschungseinrichtungen angesiedelt, die sich mit Themen der Bildungsforschung befassen. An dem Netzwerk sind 13 Forschungseinrichtungen beteiligt.

Mit diesen Einrichtungen sind für die Bildungsforschung zentrale Disziplinen, wie die Erziehungswissenschaft, Psychologie, Ökonomie und Soziologie sowie die bildungsgeschichtliche Forschung im Netzwerk vertreten. Die Vielzahl der beteiligten Einrichtungen und Disziplinen macht deutlich, dass im Netzwerk Themen der Bildungsforschung über den gesamten Lebensverlauf behandelt werden: BIEN Lunchtime Seminar (monatlich), Jahrestagungen, Methodenworkshops.

Das Lunchtime Seminar findet während der Vorlesungszeit monatlich am DIW Berlin statt. Renommierte Wissenschaftler/-innen und Nachwuchswissenschaftler diskutieren hier aktuelle Forschungsergebnisse. Jeder Vortrag wird von einem Kommentar von einem/einer Wissenschaftler/-in aus einer anderen Disziplin begleitet. So wird jedes Thema aus mindestens zwei Perspektiven betrachtet. Im Anschluss können die Nachwuchswissenschaftler/-innen beim Lunch Kontakte knüpfen. Darüber hinaus sind pro Jahr eine Jahrestagung und zwei Methodenworkshops geplant, die auch einen Beitrag darstellen, um die fachlichen und methodischen Kenntnisse der beteiligten Nachwuchswissenschaftler/-innen weiter zu entwickeln.

Die Koordinationsstelle von BIEN ist am DIW Berlin angesiedelt. Hier werden die Veranstaltungen des Netzwerkes organisiert und stattfinden. Renommierte Bildungsforscher/-innen an den unterschiedlichen Forschungseinrichtungen agieren als Schirmdamen und Schirmherren für das Netzwerk. Auf der Ebene der Nachwuchswissenschaftler/-innen gibt es an allen beteiligten Einrichtungen eine Ansprechperson für das Netzwerk.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Schmidt, Rebekka; Breuer, Meike

„Meine Hobbys sind malen, turnen ...“ – das Projekt ProMotiKuS (Professionelle Kompetenz- und Motivationsentwicklung im Lehramtsstudium Kunst und Sport)

Ein positives domänenspezifisches Selbstkonzept gilt als prädiktiv für akademische und berufliche Leistungen. Ein Ziel des Lehramtsstudiums sollte es daher sein, zur Entwicklung professioneller Kompetenz beizutragen. Diese kann als Zusammensetzung aus professionellem Wissen, Überzeugungen und Werthaltungen, selbstregulativen Fähigkeiten sowie motivationalen Orientierungen definiert werden (vgl. Baumert & Kunter, 2006). Allerdings fehlen bislang Untersuchungen, wie sich diese Aspekte während des Studiums, des Referendariats und im Verlauf der Berufspraxis entwickeln, anpassen oder evtl. verändern.

Die Studie PROMOTIKUS fokussiert daher, neben den allgemeinen Beweggründen zur Aufnahme des Lehramtsstudiums auch das fachliche Vorverständnis sowie die Entwicklung von fachspezifischem Professionswissen und berufsbezogener Motivation der Studierenden. Es soll untersucht werden, ob sich diese Aspekte fachspezifisch oder übergreifend abbilden lassen, wie sie sich entwickeln und auf das Lehrerhandeln auswirken.

2013 wurde der Studiengang für ein Lehramt an Grundschulen u. a. mit den Fächern Kunst und Sport und Bewegungserziehung an der TU Chemnitz implementiert. Die Räumlichkeiten befinden sich dabei im selben Gebäude wie die Ausbildungsstätte für die Referendare. Darüber hinaus ist davon auszugehen, dass viele der Studierenden im Umkreis von Chemnitz ihre Arbeit als Lehrkraft aufnehmen werden. Aus diesem Grund bietet es sich an, diese Studierenden von Anfang an wissenschaftlich zu begleiten und die Entwicklung professioneller Kompetenzen und berufsbezogener Motivation bezüglich der Bereiche Kunst und Sport und Bewegungserziehung zu untersuchen.

Hierzu werden längsschnittlich als auch im Querschnitt anerkannte quantitative Erhebungsinstrumente eingesetzt und durch qualitative Erhebungen ergänzt und erweitert. Zur Untersuchung der motivationalen Grundlagen wird der „Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA)“ (Pohlmann & Möller, 2010) eingesetzt. Die Erfassung der Ausgangslage bezüglich der professionellen Kompetenzen erfolgt u. a. mit Hilfe des Fragebogens zur „Erfassung berufsbezogener Selbstkonzepte von angehenden Lehrkräften (ERBSE-L)“ (Retelsdorf, Bauer, Gebauer, Kauper & Möller, 2014) erhoben. Die qualitative Untersuchung wird mit der Methode eines problemzentrierten Interviews durchgeführt. Im Fokus stehen dabei vertiefende Aspekte der professionellen Kompetenz und Wahl des Lehramtsstudiums sowie und das Gegenstandsverständnis (Kunst oder Sport) der Studierenden. Da ein solches Verständnis Auswirkungen auf den Unterricht haben kann, gilt es hier detailliert nachzuforschen. Zudem besteht die Möglichkeit dieses Design auch mit späteren Kohorten durchzuführen und auch hier ebenfalls sowohl längsschnittlich als auch im Querschnitt mit anderen Jahrgängen zu vergleichen.

Das Poster verdeutlicht die Anlage und das Design der Untersuchung und stellt erste Ergebnisse der Fragebogenerhebung bezüglich der Motivation und der Kompetenzen dar.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Bäumer, Thomas

Verläufe sportlicher Aktivität im Jugendalter – Befunde aus dem Nationalen Bildungspanel

Sport ist nach wie vor die häufigste Freizeitaktivität bei Kindern und Jugendlichen in Deutschland. In den Startkohorten 3 und 4 des Nationalen Bildungspanels (NEPS) werden daher die Schülerinnen und Schüler jährlich nach der Häufigkeit und dem Ort der Ausübung sowie nach der Sportart gefragt. Somit kann die Sportbeteiligung nicht nur querschnittlich, sondern auch in ihrem Verlauf untersucht werden und dabei der Einfluss heterogenitätserzeugender Merkmale (wie etwa Geschlecht, besuchte Schulform, Bildungshintergrund) abgeschätzt werden. In die Analyse einbezogen werden die ersten drei Erhebungswellen der Startkohorte 3 (Jahrgangsstufe 5 bis 7; N = 6112) sowie der Startkohorte 4 (Jahrgangsstufe 9 bis 11; N = 16425) aus den Schuljahren 2010/2011, 2011/2012 und 2012/2013.

Zwei Fragestellungen stehen im Mittelpunkt der Analysen: (1) Wie stellen sich Art und Weise der Sportbeteiligung in diesen beiden Schülergruppen dar? Hierbei kommen deskriptive Statistiken unter Zuhilfenahme von Designgewichten zum Einsatz, um deutschlandweit repräsentative Aussagen zu gewährleisten. (2) Welche Verläufe der Sportbeteiligung zeigen sich bei diesen beiden Schülergruppen? Während bei Startkohorte 3 der generelle Verlauf über drei Schuljahre betrachtet wird, ist bei Startkohorte 4 insbesondere der Einfluss des Verlassens des allgemeinen Schulsystems nach der neunten oder zehnten Klasse bedeutsam.

Da die Scientific Use Files der NEPS Startkohorten 3 und 4 der entsprechenden Wellen erst gegen Ende des Jahres 2014 veröffentlicht werden, können noch keine Ergebnisse berichtet werden. Erste deskriptive Analysen verweisen allerdings beispielsweise darauf, dass – bei insgesamt hoher Beteiligung – die sportlichen Aktivitäten im Laufe der Schulzeit abnehmen. Geschlecht, besuchte Schulform oder Bildungshintergrund nehmen differenziert Einfluss auf Art und Weise der Sportausübung. Somit sind auch für die anstehenden Analysen sehr interessante Befunde zu erwarten.

Mittwoch
11. März

15:45 Uhr

Foyer GC - West

Schmidtner, Stefanie; Schiepe-Tiska, Anja; Prenzel, Manfred

High Performer in Naturwissenschaften: Kognitive und motivational-affektive Schülervoraussetzungen und ihr Zusammenhang mit der Wahrnehmung von Lehr-Lern-Bedingungen

Betrachtet man die Leistungsspitzen in Naturwissenschaften im internationalen Vergleich, so weist Deutschland im Vergleich zu Ländern wie Finnland oder Singapur einen deutlich geringeren Anteil kompetenzstarker Jugendlicher (High Performer) in Naturwissenschaften auf (Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2013). Deutschland konnte in den letzten Jahren zwar den Anteil leistungsschwacher Jugendlicher verringern und trotz gesteigerter Bildungsbeteiligung das durchschnittliche Niveau des Gymnasiums halten, der Anteil leistungsstarker Jugendlicher konnte jedoch nicht verbessert werden (Schiepe-Tiska et al., 2013).

Die Förderung von Jugendlichen mit einem hohen Potential in Naturwissenschaften setzt eine erfolgreiche Identifikation dieser Jugendlichen voraus. Weil Selektions- und Beurteilungsprozesse im deutschen Bildungssystem größtenteils auf Schulnoten beruhen (Fend, 1981; Schuchart, 2006), spielen Noten dabei eine wichtige Rolle. Die Gliederung des Schulsystems und die häufig praktizierte Benotung nach im Voraus festgesetzter Verteilung verhindert jedoch eine Vergleichbarkeit von Noten zwischen Klassen und Schularten (Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006). Daneben vermitteln Noten Einschätzungen von Leistungen und Leistungspotentialen und interagieren deshalb mit motivational-affektiven Schülervoraussetzungen wie dem fachbezogenen akademischen Selbstkonzept und dem naturwissenschaftsbezogenem Interesse (Trautwein et al., 2006). Sie überwiegen dabei die tatsächliche Leistung im Einfluss auf das Selbstkonzept und sind für High Performer insbesondere im Rahmen des als „Big-Fish-Little-Pond Effect“ (Marsh, 1987) bekannten negativen Effekts des Leistungsniveaus der Bezugsgruppe entscheidend.

Schülervoraussetzungen wie Note, Selbstkonzept und Interesse stehen nicht nur in wechselseitigem Zusammenhang, sondern interagieren auch mit Lehr-Lernumgebungen (Snow, Corno & Belmont, 1996). Insbesondere die im Rahmen der Forschung zu selbstbestimmter Lernmotivation und Interesse identifizierten unterstützenden Lehr-Lern-Bedingungen von Prenzel, Kramer und Drechsel (2002) haben sich dabei als prädiktiv für intrinsische Motivation und die Nutzung elaborierter Lernstrategien erwiesen. Umgekehrt beeinflussen Schülervoraussetzungen wie Vorwissen, Interesse und Selbstkonzept Lehr-Lernumgebung. Forschung zu Lehrer-Schüler-Interaktion zeigt beispielsweise, dass Lehrer Jugendliche, an die sie positive Leistungserwartungen haben, häufiger für richtige Unterrichtsbeiträge loben (Brophy & Good, 1970) und mit Schülern verschiedener Voraussetzungen unterschiedlich häufig interagieren (Skinner & Belmont, 1993). Schülerinnen und Schüler unterschiedlicher Schülervoraussetzungen können sich deshalb in der Wahrnehmung der Qualität ihrer (Micro-)Lehr-Lernumgebung unterscheiden (vgl. Seidel, 2006). Nur wenige Studien betrachten jedoch kognitive und motivational-affektive Schülervoraussetzungen in Kombination und analysieren ihre Rolle bei Lernprozessen oder der Unterrichtswahrnehmung (Jurik, Gröschner & Seidel, 2014; Linnenbrink-Garcia et al., 2012). Insbesondere im Hinblick auf leistungsstarke Jugendliche gibt es dazu bisher keine Erkenntnisse.

Die vorliegende Studie liefert hierzu einen Beitrag, indem sie untersucht, welche Schülervoraussetzungsprofile hinsichtlich Noten, Selbstkonzept und Interesse in Naturwissenschaften innerhalb der Gruppe der leistungsstarken Jugendlichen in Naturwissenschaften existieren und inwiefern sie im Zusammenhang mit der Wahrnehmung von Lehr-Lern-Bedingungen des Naturwissenschaftsunterrichts stehen. Anders als in vorherigen Studien werden dabei sowohl kognitive als auch motivational-affektive Merkmale berücksichtigt und in einer repräsentativen Stichprobe betrachtet. Dazu werden Large-Scale-Assessment Daten von PISA 2006 mit einer Stichprobe von N=9577 Jugendlichen der neunten Klasse verwendet und leistungsstarke Jugendliche über Testwerte in einem standardisierten Kompetenztest in Naturwissenschaften definiert.

Die Analysen ergeben, dass sich leistungsstarke Jugendliche (N=1572) verschiedenen Profilen zuordnen lassen. Dabei ergeben sich insbesondere Gruppen der von außen und sich selbst unterschätzten High Performern (1), der leistungsstarken aber uninteressierten (2) sowie der von den eigenen Fähigkeiten überzeugten und interessierten High Performern (3). Es zeigt sich, dass unterschätzte High Performer mit niedrigem Selbstkonzept im Vergleich zu unterschätzten High Performern mit hohem Selbstkonzept weniger Kompetenzunterstützung und soziale Eingebundenheit im Naturwissenschaftsunterricht berichten. Höheres Interesse geht bei sich selbst unterschätzenden High Performern mit mehr Kompetenz- und Autonomieunterstützung einher, wohingegen es bei von der eigenen Fähigkeit überzeugten High Performern mit einem höher wahrgenommenen Interesse der Lehrperson und stärkerer sozialer Eingebundenheit einhergeht. Die Ergebnisse zeigen wichtige Zusammenhänge von individuellen Voraussetzungen und Unterrichtswahrnehmung und geben Hinweise für eine zukünftige Förderung von High Performern im Naturwissenschaftsunterricht.

Mittwoch
11. März

16:15 Uhr

GA 03/49

Chair(s)

Schmidtner, Stefanie

Mathematisch-naturwissenschaftlicher Unterricht in der Sekundarstufe

Zander, Simon; Krabbe, Heiko; Fischer, Hans E.

Umgang mit Heterogenität im Fachwissen von Schülerinnen und Schülern

Theoretischer Hintergrund: Das Projekt „Ganz In - Mit Ganzttag mehr Zukunft. Das neue Ganztagsgymnasium NRW" zielt darauf ab, unterschiedliche Lernbedingungen von Schüler/-innen am Gymnasium zur Kenntnis zu nehmen und durch unterstützende Förderangebote sowie einer Verbesserung der Unterrichtsqualität die Heterogenität der Schülerschaft fördernd zu berücksichtigen.

Im Rahmen eines Projekts zur Fortbildung von Physikunterricht wurde eine empirisch überprüfte Lehrerfortbildung durchgeführt, die die Lernprozesse der Schülerinnen und Schüler stärker in den Fokus der Unterrichtsstrukturierung rückt und die zumeist verlängerte Unterrichtstaktung im Ganzttag besser ausnutzt. Es wurde erwartet, dass dies insbesondere schwächeren Schülerinnen und Schülern zu Gute kommt (Helmke, 2009). Für die vorliegende Lehrerfortbildung konnten positive Auswirkungen bereits aufgezeigt werden (Beitrag GEBF Tagung 2013). So zeigte sich auf Ebene der Lehrkräfte, dass die Stunden klarere Ziele und eine transparente Struktur bekamen, außerdem wurde mehr Zeit zur Reflexion von Erfahrungen genutzt. Auf Seiten der Schülerinnen und Schüler konnte gezeigt werden, dass zu Beginn des Schuljahres die Vergleichs- und Fortbildungsgruppe in einem Test zur Mechanik gleich gut abschnitten, am Ende des Schuljahres wies die Fortbildungsgruppe doppelt so große Lernzuwächse auf. Schließlich konnten Zusammenhänge zwischen Lernzuwächsen und Qualität der Umsetzung der Basismodelle gezeigt werden.

Fragestellung: Im vorliegenden Beitrag wird aufgezeigt, dass die Fortbildung zudem positive Auswirkungen für schwächere und durchschnittliche Schülerinnen und Schüler hat. Die Forschungsfrage lautete: Profitieren bestimmte Schülergruppen besonders von der verbesserten Strukturierung, z.B. schwache Schülerinnen und Schüler?

Methode: In 30 achten Klassen des Gymnasiums in Nordrhein-Westfalen wurde zu Beginn und am Ende des Schuljahres 2011/ 12 ein Vor- und Nachtest des Fachwissens bei 1610 Schülerinnen und Schüler in Mechanik durchgeführt. 15 Klassen gehörten zur Fortbildungsgruppe, deren Lehrkräfte eine Lehrerfortbildung zu Basismodellen erhielten (Zander, Krabbe & Fischer, 2013), die anderen Klassen bildeten die Vergleichsgruppe und erhielten gewöhnlichen Unterricht. Der Fachwissenstest wurde zu Beginn und am Ende des Schuljahres eingesetzt, er enthielt 27 geschlossene Items und wies eine wle-Reliabilität von .61 auf. Mittels einer latenten Klassenanalyse über die Aufgaben des Vor- und Nachtests wurden Schülergruppen auf Grund ihres Lösungsmusters bezüglich ihrer Fähigkeiten klassifiziert. Kriterium für den besten Modellfit war der kleinste BIC-Wert. Diese Gruppen wurden anhand der durchschnittlichen wle-Personenfähigkeiten im Vor- und Nachtest hinsichtlich ihres Leistungsstands charakterisiert.

Ergebnisse: Die Gesamtstichprobe (Fortbildungs- und Vergleichsgruppe) wird am besten durch drei Gruppen beschrieben, die als schwache ($n = 642$), durchschnittliche ($n = 823$) und starke Schülerinnen und Schüler ($n = 145$) bezeichnet werden können. Die Personenfähigkeit liegt bei den Schwachen bei -0.6 (entspricht einer Lösungswahrscheinlichkeit von 39 %), mit einer Standardabweichung von 0.4, beim Durchschnitt bei 0.2 (entspricht einer Lösungswahrscheinlichkeit von 54 %), mit einer Standardabweichung von 0.4 und bei den Starken bei 1.4 (entspricht einer Lösungswahrscheinlichkeit von 74 %) mit einer Standardabweichung von 0.5. Vergleicht man die schwachen mit den durchschnittlichen Schülerinnen und Schülern liegt ein signifikanter Unterschied vor, $t(1463) = 38.4$, $p < .001$, $d = 2.0$, gleiches gilt für den Vergleich von durchschnittlichen mit starken Schülerinnen und Schülern, $t(966) = 31.4$, $p < .001$, $d = 2.8$.

Die Verteilungen bezüglich der drei Schülergruppen ist für die Vergleichs- und Fortbildungsgruppe im Vortest gleich (Chi-Quadrat-Test, $p = .95$), im Nachtest hingegen unterschiedlich, $p < .001$. Im Vortest zählen 51% der Schülerinnen und Schüler in Vergleichs- bzw. 48 % in der Fortbildungsgruppe zu den schwächeren Schülerinnen und Schülern. Im Nachtest zählen nur noch 39 % der Schülerinnen und Schüler in der Vergleichs- 21 % in der Fortbildungsgruppe zu den zu den schwächeren Schülerinnen und Schülern. Gerade schwächere und durchschnittliche Schülerinnen und Schüler können sich in der Fortbildungsgruppe häufiger verbessern als in der Vergleichsgruppe. Im Vortrag wird der Zusammenhang mit den Basismodellen diskutiert und aufgezeigt, welchen Beitrag die Lehrerfortbildung zum Umgang mit Heterogenität bezüglich des Fachwissens im Physikunterricht leisten kann.

Chair(s)

Schmidtner, Stefanie

Blankenburg, Janet S.; Höffler, Tim N.; Parchmann, Ilka

Was bewegt junge Schülerinnen und Schüler zur Teilnahme an naturwissenschaftlichen Wettbewerben?

Theoretischer Hintergrund: Schülerwettbewerbe sind ein wirkungsvolles Instrument zur Förderung begabter Schülerinnen und Schüler (Campbell, Wagner & Walberg, 2000), weshalb eine Erhöhung der Teilnehmendenzahlen und eine frühzeitige Heranführung an Schülerwettbewerbe ein bildungspolitisch sinnvolles Ziel darstellt.

Aus der Forschung zu leistungsbezogenen Entscheidungen, als welche die Bereitschaft an einem Schülerwettbewerb teilzunehmen angesehen werden kann, ist bekannt, dass diese vor allem über das Interesse und das Fähigkeitsselbstkonzept vermittelt werden (Dionne et al., 2012; Köller et al., 2000). Leistungsbezogene Entscheidungen werden im Erwartungs-mal-Wert-Modell der Leistungsmotivation (Eccles, 2005) über eine Erwartungs- und eine Wertkomponente vermittelt. Die Wertkomponente lässt sich in die Aspekte positives emotionales Erleben, Wichtigkeit, Nützlichkeit und relative Kosten gliedern. Informationen über zu erwartende Erfolgsaussichten in Hinblick auf eine Aufgabe beziehen Individuen vor allem über vorherige Leistungen und eigene Fähigkeitsselbsteinschätzungen.

Das Modell findet auch in der Forschung zu Schülerwettbewerben Anwendung, hier vor allem mit dem Ziel der Erfolgsprädiktion. Beispielsweise identifizierten Urhahne et al. (2012) als stärksten Prädiktor für das Erreichen der dritten Wettbewerbsrunde der Internationalen ChemieOlympiade eine vorherige Teilnahme am Wettbewerb.

Fragestellung: Es wird die Frage untersucht, welche Einflussfaktoren im Rahmen des Erwartungs-mal-Wert-Modells (Eccles, 2005) auf das Teilnahmeinteresse an naturwissenschaftlichen Wettbewerben von Schülerinnen und Schülern der 6. Klasse wirken. Dabei werden aus der Literatur bekannte, mögliche Einfluss nehmende Faktoren wie beispielsweise das akademische und das naturwissenschaftliche Selbstkonzept sowie Zielorientierungen berücksichtigt. Besonderes Augenmerk wird auf geschlechtsspezifische Unterschiede gerichtet. Weiterhin soll geklärt werden, welche über das Interesse an Naturwissenschaften hinausgehenden Gründe und Hindernisse Schülerinnen und Schüler hinsichtlich einer naturwissenschaftlichen Wettbewerbsteilnahme anführen, auf deren Basis längerfristig forschungsbasierte Konzeptionen von Fördermaßnahmen möglich sind.

Methode: Die Daten wurden im Rahmen der Begleitforschung zum Wettbewerbsförderprojekt NaWigator in der IJSO erhoben. An der Untersuchung nahmen insgesamt $N = 474$ Schülerinnen und Schüler der 6. Klasse (44.4 % weiblich; Alter: $M = 11.90$ Jahre, $SD = 0.42$) teil.

Die von den Schülerinnen und Schülern wahrgenommene Nützlichkeit und Wichtigkeit des Naturwissenschaftsunterrichts für ihr Leben (in Anlehnung an Steinmayr & Spinath, 2010) wurde mit jeweils drei Items erhoben. Die emotionale Facette der Wertkomponente des erweiterten Erwartungs-mal-Wert-Modells wurde mit insgesamt vier Items erfasst.

Zudem wurden das absolute akademische (Schöne et al., 2002) und das naturwissenschaftliche Selbstkonzept (Marsh et al., 2005), die Wettbewerbsteilnahmebereitschaft sowie Zielorientierungen im Naturwissenschaftsunterricht (Spinath et al., 2002) erfasst.

Sieben Skalen zur Erfassung der Beweggründe für oder gegen eine naturwissenschaftliche Wettbewerbsteilnahme wurden theoriebasiert (Deci & Ryan, 1993; Prenzel & Drechsel, 1996) entwickelt und durch konfirmatorische Faktorenanalysen abgesichert.

Die Reliabilitäten waren durchgängig zufriedenstellend bis gut (Cronbachs α .72-.86).

Ergebnisse: Jungen und Mädchen unterschieden sich signifikant hinsichtlich ihres naturwissenschaftlichen Selbstkonzepts ($t(462) = 3.59$, $p < .001$, $d = 0.34$) und ihrer Zielorientierungen, etwa zur Arbeitsvermeidung ($t(460) = 2.44$, $p < .05$, $d = 0.23$), bei denen jeweils die Jungen höhere Werte aufwiesen. Weitere Geschlechterunterschiede fanden sich bei den Teilnahmegründen und -hindernissen.

In einer logistischen Regression wurde ermittelt, dass sich unter anderem das Interesse an naturwissenschaftlichen Fächern, der Umstand einer vormaligen naturwissenschaftlichen Wettbewerbsteilnahme, das Selbstkonzept in Bezug auf Wettbewerbsteilnahmen sowie der Teilnahmegrund der intrinsischen Motivation (teilnehmen, um etwas Neues zu lernen) signifikante Prädiktoren für die Teilnahmebereitschaft darstellen ($R^2 = .34$ (Hosmer & Lemeshow), $.35$ (Cox & Snell), $.49$ (Nagelkerke). $X^2(8) = 129.02$, $p < .001$). Bei Mädchen spielen insbesondere eine vorherige Wettbewerbsteilnahme, hohes Interesse an naturwissenschaftlichen Fächern und das akademische Selbstkonzept eine signifikante Rolle, während bei Jungen gute Noten, hohe Nützlichkeitseinschätzung der Naturwissenschaften und geringe Vermeidungs-Leistungsziele prädiktiv wirken.

Die vorliegende Studie charakterisiert damit erfolgreich junge Schülerinnen und Schüler bezüglich ihrer Teilnahmebereitschaft an naturwissenschaftlichen Wettbewerben und liefert so wichtige Informationen für die Gestaltung von Interventionen, wie sie etwa für das Projekt NaWigator in der IJSO bereits genutzt wurden.

Mittwoch
11. März

16:15 Uhr

GA 03/49

Chair(s)

Schmidtner, Stefanie

Wackermann, Rainer

Der Einfluss der Stundenlänge (45 vs. 60 Minuten) auf ausgewählte Aspekte der Unterrichtsqualität im Physikunterricht am Gymnasium

Obwohl bislang kaum Untersuchungen zur optimalen zeitlichen Strukturierung von Unterricht vorliegen, verlängern zurzeit einige Schulen in Deutschland die Schulstundenlänge auf 60 (auch 65 oder 67,5 usw.) Minuten. Damit wird unter anderem das Ziel verfolgt, die Unterrichtsqualität zu verbessern. Die vorliegende, kleine, DFG-geförderte, empirische Studie (WA 2162/2-1) untersucht exemplarisch im Physikunterricht die Auswirkungen dieser Stundenverlängerung auf die Unterrichtsqualität im Rahmen einer prä-post-Untersuchung mit zwei Physiklehrkräften (n=14 Videos).

Die Herkunft der in Deutschland vorherrschenden 45-Minuten-Unterrichtsstunde ist auf die „Überbürdungsdebatte“ zu Beginn des 20. Jahrhunderts zurückzuführen, in deren Folge Unterrichtsstunden in einem Erlass des preußischen Schulministers von 60 auf 45 Minuten verkürzt wurden. Die Einführung der 45 Minuten-Stunde wurde damals weder didaktisch noch pädagogisch-psychologisch begründet. Obwohl bislang kaum Untersuchungen zur optimalen zeitlichen Strukturierung von Unterricht vorliegen, verlängern zurzeit einige Schulen in Deutschland die Schulstundenlänge auf 60 (auch 65 oder 67,5 usw.) Minuten. Damit wird unter anderem das Ziel verfolgt, die Unterrichtsqualität zu verbessern. Aus fachdidaktischer Sicht ist mit einer Unterrichtsstundenverlängerung auch die Frage nach dem Abschluss von Lernprozessen verknüpft, denn Videountersuchungen im Physikunterricht der letzten Jahre haben gezeigt, dass nur ein verschwindend kleiner Teil der Unterrichtszeit auf Transfer, Reflexion oder Vernetzung verwendet wird (Widodo & Duit, 2005; Wackermann, Trendel & Fischer, 2010; Stender, Geller, Neumann & Fischer, 2013). Möglicherweise bieten längere Unterrichtsstunden mehr Gelegenheit für lernprozessabschließende Unterrichtsphasen (Borowski, Fischer, Trendel & Wackermann, 2010).

Für diese Studie wurde der 60-Minuten-Unterricht von zwei Physiklehrkräften aufgezeichnet und ausgewertet und mit dem 45-Minuten-Unterricht derselben Lehrkräfte aus einer früheren Studie verglichen (n=14 Videos). Die Schule der beiden Physiklehrkräfte hatte zwischenzeitlich von traditionellen 45 Minuten- auf 60 Minuten-Stunden gewechselt. Die Analyse erfolgt unter dem Blickwinkel der Basismodelle nach Oser (2001) in der Operationalisierung für Physikunterricht von Wackermann et al. (2010) mit ergänzenden Sichtweisen aus Befragungen von Schülerinnen und Schülern sowie Lehrkräften und einem Expertenrating für Unterrichtsqualität nach Clausen (2002). Alle Untersuchungsinstrumente wurden umfassend und erfolgreich auf ihre Güte geprüft (z.B. Reliabilität Basismodellkodierung Cohens $\kappa = 0,9$).

Zentrales Ergebnis für die untersuchten Fälle ist, dass die zusätzliche Zeit für eine erste Sicherung verwendet wird, die unter der Bedingung 45 Minuten praktisch nicht stattfindet, und dass die Schülerinnen und Schüler bei den längeren Stunden mehr selbst tätig werden können. Jedoch bleibt die kognitive Aktivität der Schülerinnen und Schüler gleich. Außerdem gibt es Hinweise auf eine zeitliche Ausdehnung der Wiederholungsphase bei den längeren Stunden. Die Befragung der Lehrkräfte ergibt subjektiv wahrgenommene qualitative Verbesserungen für den Unterricht sowie für den Schulalltag, allerdings stehen die Befragungsergebnisse teilweise im Widerspruch zu den Ergebnissen der Unterrichtsvideographie. Lernprozessabschließende Unterrichtsphasen konnten weiterhin nicht in nennenswertem Umfang festgestellt werden.

Zusammengefasst bieten die längeren Schulstunden das Potenzial für eine Qualitätsverbesserung. Möglicherweise sind allerdings Lehrkräftefortbildungen notwendig, um das fachdidaktische Potenzial noch mehr auszureizen.

Auf der GEBF-Tagung 2015 in Bochum werden zusätzlich zu den komplett vorliegenden Ergebnissen des 45 vs 60 Minuten-Stunden-Vergleichs auch schon erste Auswertungen eines analog durchgeführten, ebenfalls experimentellen 45 vs 90 Minuten-Doppelstundenvergleichs mit zwei weiteren Physiklehrkräften präsentiert werden, wo die Datenerhebung zum heutigen Zeitpunkt nahezu abgeschlossen ist (Stand Oktober 2014).

Mittwoch
11. März

16:15 Uhr

GA 03/49

Chair(s)

Schmidtner, Stefanie

Wecker, Christof; Ufer, Stefan; Mahl, Christina

Vom Vormachen zum Selbermachen: Fading von Lösungsschritten bei der Demonstration von Strategien im Mathematikunterricht

Professionelles Handeln von Lehrkräften im Sinne einer evidenzbasierten Praxis im Bildungsbereich setzt voraus, dass Forschungsergebnisse zu lernförderlichen Gestaltungsmerkmalen des Unterrichts vorliegen, die Lehrkräfte im eigenen Handeln umsetzen können. Die Grundlagenforschung im Bereich der Lehr-Lernforschung erhebt den Anspruch, unterrichtsrelevante Erkenntnisse bereitzustellen, gewinnt diese jedoch häufig in experimentellen Laborstudien. Ein Beispiel dafür stellt die Forschung zu ausgearbeiteten Lösungsbeispielen dar. So wurde etwa nachgewiesen, dass die schrittweise Erhöhung der Anzahl von Lösungsschritten, die Lernende bei einer Aufgabe selbst bearbeiten („Fading“), im Vergleich zu vollständigen Lösungsbeispielen einen positiven Effekt auf den Lernerfolg hat (Renkl & Atkinson, 2003). Zunehmend wird die Auffassung vertreten, dass die Demonstration einer Strategie („Modeling“) im Unterricht der Darbietung eines Lösungsbeispiels entspricht (van Gog & Rummel, 2010) und die Erkenntnisse zu Lösungsbeispielen auf Demonstrationen übertragbar sind. In der vorliegenden Studie wurde daher untersucht, welchen Effekt das Fading von Lösungsschritten bei Demonstrationen im Unterricht auf den Lernerfolg hat.

Diese Fragestellung wurde mit einem between-subjects-Design untersucht, in der das Fading von Lösungsschritten in Demonstrationen (kein Fading vs. Fading) innerhalb von Schulklassen manipuliert wurde. In die Analyse gingen Daten von 76 Schülerinnen und Schülern aus drei siebten Gymnasialklassen ein (34 % weiblich; Alter: $M = 12.8$ Jahre; $SD = 0.68$). In einer Doppelstunde zu Modellierungsaufgaben aus dem Bereich der Prozentrechnung wurde nach einer individuellen Wiederholungsphase, die zugleich als Vortest diente, anhand von fünf Aufgaben eine vierschrittige Strategie zur Lösung von Prozentrechnungsaufgaben demonstriert. In der Bedingung ohne Fading wurden alle fünf Demonstrationsaufgaben gemeinsam mit den Schülerinnen und Schülern im Plenum gelöst und Schritt für Schritt mit Hilfe einer Dokumentenkamera an der Wand gezeigt. In der Bedingung mit Fading sollten die Schülerinnen und Schüler bei jeder der fünf Aufgaben einen zusätzlichen Lösungsschritt selbst bearbeiten, sodass die fünfte Aufgabe vollständig selbst zu bearbeiten war. Es wurde codiert, ob die Schülerinnen und Schüler die gefadeten Lösungsschritte selbst bearbeitet hatten. Auf diese Demonstrationsphase folgte eine Übungsphase, die zugleich als Nachtest diente.

Die Schülerinnen und Schüler in der Bedingung mit Fading ($M = .71$; $SD = .26$) erzielten signifikant höhere Werte im Nachtest als die Schülerinnen und Schüler in der Bedingung ohne Fading ($M = .57$; $SD = .33$), $F(1; 70) = 4.23$; $p = .04$; $\eta_p^2 = .06$). Während sich Lernende, die die gefadeten Lösungsschritte nicht eigenständig bearbeiteten, im Nachtest nur unwesentlich von Lernenden in der Bedingung ohne Fading unterschieden, $F(1; 41) = 0.62$; $p = .44$; $\eta_p^2 = .02$, erzielten Lernende, die die gefadeten Lösungsschritte eigenständig bearbeiteten, deutlich höhere Werte als Lernende in der Bedingung ohne Fading, $F(1; 60) = 7.07$; $p = .01$; $\eta_p^2 = .11$.

Wie die Ergebnisse zeigen, lässt sich durch das Fading von Lösungsschritten bei der Demonstration einer Strategie im Unterricht deren Beherrschung fördern. Die Wirksamkeit des Fadings erstreckt sich jedoch nur auf Lernende, die die gefadeten Lösungsschritte tatsächlich selbst bearbeiten und somit die angebotene Lerngelegenheit sinnvoll nutzen (Helmke, 2010). Grenzen der Studie liegen unter anderem im unterschiedlichen Ausmaß an Übungsgelegenheiten zwischen den beiden Bedingungen.

Die vorliegende Studie stützt die Annahme, dass in der Grundlagenforschung gewonnene Prinzipien auf schulische Lernsituationen übertragbar sind, sofern dort dieselben kognitiven Mechanismen wirksam sind. Das Paradigma einer experimentellen Unterrichtsforschung stellt eine geeignete Forschungsstrategie zur Untersuchung der Übertragbarkeit laborexperimentell überprüfter Lernprinzipien auf Unterrichtssituationen dar, in denen die entsprechenden Gestaltungsmerkmale durch das Handeln von Lehrkräften umgesetzt werden.

Mittwoch
11. März

16:15 Uhr

GA 03/49

Chair(s)

Schmidtner, Stefanie

Stang, Justine; Urhahne, Detlef

Wie verändert sich die diagnostische Kompetenz von Lehrkräften im Bereich von Leistung, Arbeits- und Sozialverhalten?

Der Schulalltag erfordert es, dass Lehrkräfte nicht nur die Leistung, sondern auch das Arbeits- und Sozialverhalten von Schülern beurteilen. Bemerkungen oder Bewertungen über Mitarbeit und Verhalten der Schüler können in das Zeugnis mit aufgenommen werden (siehe z.B. BayEUG, 2000). Deshalb ist eine akkurate Einschätzung nicht-kognitiver Schülermerkmale von besonderer Bedeutung.

Die Einschätzung nicht-kognitiver Schülermerkmale ist im Gegensatz zu kognitiven Merkmalen jedoch selten untersucht worden. Zu kognitiven Merkmalen wie Leistungen in verschiedenen Schulfächern liegen bereits metaanalytische Ergebnisse vor (z.B. Südkamp, Kaiser & Möller, 2012). Zu nicht-kognitiven Merkmalen existieren vereinzelt Untersuchungen im Bereich von Motivation und Emotion (Urhahne & Zhu, 2014). Befunde zu Arbeits- und Sozialverhalten liegen bislang nicht vor.

Darüber hinaus fehlen Erkenntnisse, wie sich die diagnostische Kompetenz von Lehrkräften über die Zeit hinweg entwickelt. Die Genauigkeit von Lehrkräfteeinschätzungen wird in der Regel nur zu einem Messzeitpunkt erfasst. Durch einen weiteren Messzeitpunkt kann die Veränderbarkeit von Lehrkräften sichtbar gemacht werden. Es ist zu erwarten, dass Lehrkräfte Leistungen, Arbeits- und Sozialverhalten ihrer Schüler im Laufe eines Schuljahres besser beurteilen lernen.

Allerdings ist anzunehmen, dass zwischen Lehrkräften erhebliche Unterschiede in der diagnostischen Kompetenz bestehen. Als Ursache für interindividuelle Differenzen in der Genauigkeit wurden hauptsächlich Lehrkraftmerkmale wie die Berufserfahrung diskutiert. Andere Merkmale wie Alter oder Geschlecht wurden weniger stark beforscht und genauere Schlussfolgerungen sind noch nicht möglich (Südkamp et al., 2012).

Bei der Genauigkeitsbestimmung wird zwischen drei Komponenten unterschieden (Schrader & Helmke, 1987): Rang-, Niveau- und Differenzierungs-komponente. Die Rangkomponente bildet den Zusammenhang von Lehrkräften und Schülermerkmal ab. Die Niveauelemente bezieht sich auf die Differenz und die Differenzierungs-komponente auf die Variabilität von Lehrkräften und Schülermerkmal.

Folgende Forschungsfragen wurden in einer Studie zu Leistung, Arbeits- und Sozialverhalten von Schülern genauer untersucht: (a) Wie gut können Lehrkräfte die verschiedenen Merkmale einschätzen? (b) Ist eine Verbesserung der diagnostischen Kompetenz über die Zeit festzustellen? (c) Können Lehrkraftmerkmale interindividuelle Unterschiede in der diagnostischen Kompetenz erklären?

An der Untersuchung nahmen fünf Realschulen mit 17 fünften Klassen teil. Daten von beiden Messzeitpunkten lagen für 294 Fünftklässler (44.6% weiblich) vor. Die Schüler waren zum zweiten Messzeitpunkt im Schnitt 10.82 Jahre alt (SD = 0.56). Die 17 Mathematiklehrkräfte (64.7% weiblich) waren im Mittel 32.41 Jahre alt (SD = 7.92) und verfügten über eine Berufserfahrung von durchschnittlich 6.50 Jahren (SD = 5.27).

Das mathematische Können wurde mit dem Deutschen Mathematiktest für vierte und fünfte Klassen (Gölitz, Roick & Hasselhorn, 2006; Götz, Lingel & Schneider, 2013) erfasst. Das Arbeitsverhalten mittels des Lern- und Arbeitsverhaltensinventars (Keller & Thiel, 1998), die soziale Kompetenz anhand des Instruments zur Messung sozialer Kompetenzen in Schülerselbstsicht (Frey, 2013). Die Reliabilitäten der eingesetzten Maße waren zufriedenstellend.

Die Untersuchung wurde von geschulten Testleitern im Klassenverband durchgeführt und dauerte zwei Schulstunden. Gegen Ende des ersten und zweiten Schulhalbjahres bearbeiteten die Schüler zuerst den standardisierten Mathematikleistungstest und füllten dann die Selbstbeschreibungsbögen zum Arbeits- und Sozialverhalten aus. Lehrkräfte schätzten parallel dazu Merkmale ihrer Schüler über Einzelitems und Ratingskalen ein.

Die Rangkomponente zeigte einen mittleren Zusammenhang für die Mathematikleistung, aber nur kleine Zusammenhänge für Arbeits- und Sozialverhalten. Die Niveauelemente erbrachte eine Überschätzung der Schülerleistung zum ersten Messzeitpunkt und eine Unterschätzung zum zweiten Messzeitpunkt. Die Differenzierungs-komponente ergab zu beiden Messzeitpunkten eine Überschätzung der Leistungsunterschiede zwischen den Schülern. Eine signifikante Verbesserung des Lehrkräftenurteils war nur für die Mathematikleistung, nicht aber für das Arbeits- und Sozialverhalten zu verzeichnen. Mehrebenenanalysen mit HLM 7.01 (Raudenbush, Bryk & Congdon, 2013) ergaben zum ersten Messzeitpunkt keinen Einfluss von Alter, Geschlecht und Berufserfahrung der Lehrkräfte auf die Genauigkeit der Einschätzung. Zum zweiten Messzeitpunkt zeigte sich ein Einfluss des Geschlechts auf die Beurteilung des Sozialverhaltens. Weibliche Lehrkräfte schnitten signifikant besser ab.

Die Ergebnisse werden auf inhaltlicher und methodischer Ebene diskutiert sowie Implikationen für Forschung und Praxis vorgestellt.

Mittwoch
11. März

16:15 Uhr

GA 03/142

Chair(s)

Stang, Justine

Machts, Nils; Kaiser, Johanna; Möller, Jens

Eine Meta-Analyse zu der Beurteilungsgenauigkeit von Lehrern bei der Beurteilung der kognitiven Fähigkeiten ihrer Schüler

Lehrerbeurteilungen spielen eine entscheidende Rolle für die akademische Laufbahn eines Schülers. Formale Aspekte der Beurteilung finden sich sowohl in Zeugnissen, als auch in Schullaufbahnpfehlungen. In den meisten Studien mit quantitativen Ergebnissen zur Validität von Lehrerurteilen wird eine Schätzung der Beurteilungsgenauigkeit über den Vergleich zwischen einem gemessenem Schülermerkmale und dessen Beurteilung durch die Lehrkraft angegeben. Das übliche Genauigkeitsmaß ist die Korrelation zwischen den Ergebnissen der Merkmalsmessung und den Beurteilungen durch die Lehrer. Die Beurteilungsgenauigkeit der Schulleistung wurde bereits in zwei Meta-Analysen systematisch integriert. Die Analysen von Hoge und Coladarci (1989) ergaben einen Median der Beurteilungsgenauigkeit über 16 Studien von $r = .66$ und Südkamp, Kaiser und Möller (2012) lieferten vergleichbare Ergebnisse mit einer mittleren Beurteilungsgenauigkeit über 75 Studien von $r = .63$.

Die Bedeutung von Lehrerbeurteilungen beschränkt sich jedoch nicht auf die formalen Aspekte der Schullaufbahn ihrer Schüler. Diese Studie beschäftigt sich spezifisch mit der Beurteilung der kognitiven Fähigkeiten der Schüler. Die kognitiven Fähigkeiten stellen den mitunter größten prädiktiven Faktor für schulischen und beruflichen Erfolg dar (Rost, 2009). Dennoch sollten die kognitiven Fähigkeiten nicht als feste und unausweichliche Vorhersage der Entwicklung eines Schülers betrachtet werden, sondern vielmehr als die Ausrüstung, zügig zu lernen und zu verstehen (Carrol, 1997). Um das darin enthaltene Entwicklungspotential zur Entfaltung zu bringen, bedarf es der Aufmerksamkeit und der Kompetenz der Lehrkräfte. Nur wenn eine Lehrkraft Kenntnis über die Lernvoraussetzungen seiner Schüler hat, kann sie für jeden ihrer Schüler adäquate Lernumgebungen schaffen.

Niedrigere Korrelationen für die Beurteilungsgenauigkeit von Schülermerkmalen sind im Vergleich zu curricularer Leistung zu erwarten, da die beobachtbaren Informationen über die Schüler sich weitgehend auf curriculare Leistungen begrenzen. Mit der Beurteilung der den Leistungen vorausgehenden Merkmale ist dagegen eine Abstraktionsleistung zwischengeschaltet. Follman (1991) gibt in einem kurz gefassten Review eine mittlere Korrelation von $r = .55$ für 32 Studien an, in denen IQ-Tests mit IQ-Ratings verglichen wurden.

Südkamp et al. (2012) konnten z.B. zeigen, dass Lehrkräfte eine höhere Beurteilungsgenauigkeit erzielten, wenn sie Kenntnis über das Testverfahren hatten, das zum Vergleich mit ihrer Einschätzung herangezogen wurde. Weiterhin können Schüler- und Testmerkmale Einfluss auf die zum Vergleich herangezogene Merkmalsausprägung haben. Lehrer- und Einschätzungsmerkmalen können die Beurteilungsergebnisse ebenfalls beeinflussen.

Methode: Relevante Studien wurden durch eine systematische Datenbankrecherche in einschlägigen Online Datenbanken identifiziert. Darüber hinaus wurden alle so gewonnen Studien cross-reference-checks unterzogen. Zwei Rater kodierten unabhängig voneinander alle vorliegenden Studien. Zusätzlich zu der Beurteilungsgenauigkeit als Korrelation (oder in eine Korrelation überführbares Maß) zwischen einer Messung der kognitiver Fähigkeiten der Schüler und deren Beurteilung durch die Lehrkräfte, wurden eine Reihe von möglichen Einflussfaktoren auf die Beurteilungsgenauigkeit kodiert. Zu diesen möglichen Einflussfaktoren gehörten Studienmerkmale, Merkmale des Einschätzungsverfahrens, Testmerkmale und Merkmale der Schülerstichprobe. Als Auswertungsverfahren wurde ein meta-analytischer Mehrebenenansatz mit der Software HLM 6 ausgewählt (Raudenbush, Bryk, Cheong, & Congdon, 2004).

Ergebnisse: Es wurden 106 Effektgrößen aus 33 Studien integriert. Die mittlere Beurteilungsgenauigkeit kognitiver Fähigkeit lag bei $r = .45$. Mittels Moderatoranalysen konnten Einflussfaktoren auf die Beurteilungsgenauigkeit festgestellt werden. Es konnte eine niedrigere Beurteilungsgenauigkeit bei bezüglich der kognitiven Fähigkeiten vorselegierten Schülerstichproben gezeigt werden und es konnte eine niedrigere Beurteilungsgenauigkeit für Einschätzungsverfahren gezeigt werden, deren Formulierung keine adäquate Bezugsnorm für die Beurteilung anbot. Darüber hinaus konnte gezeigt werden, dass Lehrereinschätzungen der Intelligenz ihrer Schüler einen höheren Zusammenhang mit der gemessenen Schulleistung der Schüler aufweisen, als mit der gemessenen Intelligenz.

Fazit: Die Ergebnisse dieser Meta-Analyse zeigen eine mittlere Beurteilungsgenauigkeit der kognitiven Fähigkeiten der Schüler durch ihre Lehrkräfte an. Diese scheint jedoch in hohem Maße durch die manifeste Schulleistung der Schüler bedingt zu sein. Zwar können Lehrkräfte die generelle Leistungsfähigkeit ihrer Schüler ordentlich einschätzen, differenzierte Einschätzungen bezüglich der Unterscheidung zwischen kognitivem Fähigkeitenpotential und manifester Schulperformanz, scheinen jedoch im Durchschnitt die Kompetenzen der Lehrkräfte zu übersteigen.

Mittwoch
11. März

16:15 Uhr

GA 03/142

Chair(s)

Stang, Justine

Hellmann, Katharina; Nückles, Matthias

Ich sehe es, wie du es siehst? Der Einfluss von Lehrerfahrung, Spezifität der Urteilsobjekte und Aufgabendesign auf die Güte diagnostischer Urteile

Die Qualität diagnostischer Urteile ist wesentliche Determinante effektiven Lehrerhandelns. Eine spezielle Facette diagnostischer Kompetenz betrifft die Auswahl Schülerangemessener Lernmaterialien (Schrader, 2006). Lehrkräfte müssen befähigt sein, Verständnisschwierigkeiten von Novizen adäquat einzuschätzen und Lernaufgaben entsprechend anzupassen. Der Forschung zum „Expert Blindspot“ (Nathan & Petrosino, 2003;) zufolge besteht jedoch die Gefahr, dass Lehrkräfte ihr Expertenwissen als Anker nutzen (Nickerson, 1999), wenn sie Schwierigkeiten von Aufgaben für Schüler einschätzen. Auf einen „Expert Blind Spot“ kann geschlossen werden, wenn Lehrkräfte ausgehend von ihrer eigenen wahrgenommen Aufgabenschwierigkeit die Schwierigkeit dieser Aufgaben für Schüler unterschätzen. Lehrerfahrung sollte jedoch zu adäquateren Urteilen führen. Auch die Spezifität der Urteilsobjekte sollte Einfluss auf die Urteilsgüte haben. Je bekannter die einzuschätzende Zielgruppe, desto akkurater sollte das Urteil ausfallen (Nickerson, 1999).

Der Cognitive Load Theory (CLT; Sweller, 2005) zufolge setzt sich die Schwierigkeit einer Lernaufgabe hauptsächlich aus intrinsischer (Komplexität) und extrinsischer (Aufgabengestaltung) mentaler Belastung zusammen, wobei letztere durch geeignetes Instruktionsdesign zu reduzieren ist, da sie nachweislich effektives Lernen verhindert. Novizen sollten daher Lernaufgaben erhalten, welche nach CLT-Prinzipien „optimiert“ sind (z. B. durch ausgearbeitete Lösungsbeispiele, Integriertes-Format-Aufgaben, Schritt-für-Schritt-Aufgaben). Bei Experten kann ein solches Instruktionsdesign jedoch gegenteilige Effekte verursachen. „Optimierte“ Aufgaben interferieren mit bereits ausgebildeten, hochvernetzten kognitiven Strukturen und verursachen Redundanzeffekte, welche die extrinsische kognitive Belastung erhöhen. Solch „optimierte“ Aufgaben, die für Novizen lernförderlich sind, werden von Experten tendenziell als schwieriger wahrgenommen (Expertise-Umkehr-Effekt; Kalyuga, 2007). Lehrer als Fachexperten und Pädagogen sollten befähigt sein, Aufgabenschwierigkeiten für Schüler unabhängig von der eigenen mentalen Belastung einzuschätzen.

In zwei quasi-experimentellen Studien verglichen wir Schwierigkeitseinschätzungen von Studierenden, erfahrenen Lehrkräften sowie Lehrkräften, die ihre eigene Klasse beurteilten. Folgende Fragestellungen wurden untersucht: Aufgrund eines generellen Ankereffekts sollten Studierende und Lehrende die Anzahl der Aufgaben überschätzen, die Schüler korrekt lösen können (Überschätzungs-Hypothese). Weiterhin sollten signifikant größere Überschätzungen für nicht-optimierte als für „optimierte“ Lernaufgaben resultieren (Expertise-Umkehr-Hypothese). Die Überschätzungen sollten zudem für erfahrene Lehrkräfte geringer ausfallen als für Studierende (Lehrerfahrungen-Hypothese). Lehrkräfte, die eigene Schulklassen einschätzen, sollten schließlich keine signifikanten Leistungsüberschätzungen mehr zeigen (Spezifitäts-Hypothese).

In Studie 1 nahmen 34 Lehramtsstudierende (Mathematik, durchschnittliche Semesterzahl: 6.12 (SD=3.18)) sowie 22 erfahrene Mathematiklehrkräfte (durchschnittliche Berufserfahrung: 15.5 Jahre (SD=11.71) teil; 54 Schüler der 9. Jahrgangsstufe dienten als Vergleichsgruppe (Durchschnittsalter: 14.26 Jahre (SD= .52)). In Studie 2 nahmen 12 Mathematiklehrkräfte (durchschnittliche Berufserfahrung: 12 Jahre, SD=7.73) mit ihrer jeweiligen Mathematikklasse (239 Schüler der 10. Jahrgangsstufe) teil (Durchschnittsalter: 15.49 Jahre, SD=.59). Das Untersuchungsmaterial bestand aus zehn curricular validen Aufgaben, welche von zwei Mathematikexperten jeweils in einem nicht-optimierten und optimierten CLT-Format gestaltet und den Probanden randomisiert zugewiesen wurden. Der Grad professioneller Expertise (Studierende, Lehrkräfte) sowie das Instruktionsdesign (nicht-optimiert, optimiert) dienten als unabhängige Variablen. Die Einschätzung der Schülerperformanz durch die Expertengruppen und die tatsächliche Schülerperformanz umfassten die abhängigen Variablen.

In Übereinstimmung mit der Überschätzungs-Hypothese überschätzten Studierende, $F(1,86) = 5.71$; $p < .05$, $\eta^2 = .06$, sowie Lehrkräfte, $F(1,74) = 2.88$; $p = .09$, $\eta^2 = .04$, die Anzahl von korrekt gelösten Aufgaben signifikant. Dabei wurde gemäß der Expertise-Umkehr-Hypothese sowohl bei Studierenden, $F(1,86) = 6.17$; $p < .05$, $\eta^2 = .07$, als auch Lehrkräften, $F(1,74) = 2.93$; $p < .09$, $\eta^2 = .04$, die Überschätzung der Schülerleistung signifikant durch das Instruktionsdesign moderiert: obwohl die Schüler gemäß der CLT signifikant mehr optimierte Aufgaben lösten, unterschieden sich die Einschätzungen der Expertengruppen nicht bezüglich beider Aufgabendesigns. Einschätzungen der Expertengruppen für beide Item-Typen unterschieden sich entgegen der Lehrerfahrungen-Hypothese nicht signifikant voneinander, $F(1,54) = .21$, ns, $\eta^2 = .00$. Die Spezifitäts-Hypothese konnte bestätigt werden: Lehrkräfte, die ihre eigene Klasse einschätzten, überschätzten die Schülerleistung für keinen der beiden Item-Typen signifikant, $F(1,11) = .45$, $p = ns$, $\eta^2 = .04$. Zusammengefasst zeigen die Ergebnisse, dass Studierende wie Lehrkräfte in Hinblick auf CLT-Prinzipien und Ankereffekte sensibilisiert werden sollten, um die Güte diagnostischer Urteile zu erhöhen.

Mittwoch
11. März

16:15 Uhr

GA 03/142

Chair(s)

Stang, Justine

Förster, Natalie; Souvignier, Elmar

Urteilsgenauigkeit von Lehrkräften: Schätzen Lehrkräfte schwächere Leser ungenauer ein?

Theoretischer Hintergrund: Lehrkräfte müssen Lernvoraussetzungen, Lernprozesse und Lernergebnisse von Schülerinnen und Schülern einschätzen, um Unterricht angemessen anpassen und Schülerleistungen fair bewerten zu können. Die diagnostische Kompetenz gilt daher als zentraler Aspekt professioneller Lehrerkompetenz (Baumert & Kunter, 2006). Studien zur Akkuratheit von Lehrereinschätzungen zeigen, dass Lehrkräfte die Leistungen ihrer Schüler im Mittel gut einschätzen können (Südkamp, Kaiser & Möller, 2012). Gleichzeitig finden sich jedoch große interindividuelle Unterschiede bezüglich der Akkuratheit von Lehrerurteilen (Helmke & Schrader, 1987; Hoge & Coladarsi, 1989; Lorenz & Artelt, 2009), sowie Hinweise auf größere Schwierigkeiten bei der Einschätzung leistungsschwacher Schülerinnen und Schüler (Feinberg & Shapiro, 2009; Madelaine & Wheldall, 2005). Studien zur Akkuratheit von Lehrkräften basieren dabei häufig auf Korrelationen zwischen Urteil und Schülermerkmal, wobei das Ausmaß einer Über- oder Unterschätzung nicht berücksichtigt wird.

Fragestellung: In der vorliegenden Studie wurde untersucht, ob es einen Zusammenhang zwischen der Lesekompetenz von Schülerinnen und Schülern sowie der Akkuratheit der Lehrereinschätzung gibt. Vor dem Hintergrund vorliegender Befunde wurde dabei angenommen, dass Lehrereinschätzungen bei leistungsschwächeren Schülerinnen und Schülern weniger akkurat ausfallen.

Methode: Insgesamt nahmen 110 Lehrkräfte von dritten und vierten Klassen mit insgesamt 2160 Schülerinnen und Schülern teil. Zu Beginn und zum Ende des Schuljahres wurde die Leseflüssigkeit und das Leseverständnis der Schüler anhand standardisierter Lesetests (SLS 1-4, Hamlet 3-4) erhoben. Die Lehrkräfte wurden gebeten einzuschätzen, wie viele Aufgaben ihre Schülerinnen und Schüler in den jeweiligen Tests lösen würden. Neben der Rangkomponente wurde zusätzlich die Niveauelemente analysiert. Die Genauigkeit der Rangkomponente wurde dabei über die Differenz der Rangplätze der eingeschätzten Leistung durch die Lehrkraft sowie der gemessenen Schülerleistung auf Klassenebene bestimmt. Im Hinblick auf die Niveauelemente wurde der Betrag der Differenz zwischen Lehrereinschätzung und Schülerleistung herangezogen.

Ergebnisse: Die Ergebnisse zeigen, dass die Lehrkräfte die Leistungen ihrer Schülerinnen und Schüler im Mittel sehr gut einschätzen ($r = .63 / .75$) und die Leistungen nur tendenziell überschätzen. Gleichzeitig findet sich jedoch für die Rangkomponente, dass leistungsschwächere Schülerinnen und Schüler ungenauer eingeschätzt werden, als leistungsstärkere Schüler. Dies gilt sowohl für die Leseflüssigkeit, als auch für das Leseverständnis. Im Hinblick auf die Niveauelemente zeigt sich ein weniger einheitliches Bild. Während auch hier das Leseverständnis schwächerer Schüler ungenauer eingeschätzt wird, wird die Leseflüssigkeit leistungsstärkerer Schüler weniger genau eingeschätzt. Das Ergebnismuster ist dabei jeweils zum Schuljahresbeginn und zum Schuljahresende identisch.

Fazit: Zunächst lässt sich festhalten, dass die enge Übereinstimmung der Befunde zum Schuljahresbeginn und Schuljahresende für ein robustes Ergebnismuster spricht. Insbesondere bei der weniger leicht zu beobachtenden Fähigkeit Texte verstehend zu Lesen zeigt sich für beide Maße der Urteilsgenauigkeit wie erwartet, dass Lehrkräfte bei leistungsschwächeren Schülerinnen und Schülern weniger akkurate Leistungseinschätzungen vornehmen. Obwohl die Lehrkräfte das Leistungsniveau und die Leistungsunterschiede ihrer Schüler im Mittel gut einschätzen, erscheint eine ungenauere Einschätzung insbesondere leistungsschwächerer Schülerinnen und Schüler problematisch. Die unterschiedlichen Befunde für die zwei Maße der Urteilsgenauigkeit stellen zudem in Frage, ob es sich bei der diagnostischen Kompetenz von Lehrkräften um ein einheitliches Konstrukt handelt. Neben einer weiteren Klärung des Konstrukts der diagnostischen Kompetenz (vgl. Spinath, 2005) erscheint daher vor allem die Bereitstellung zusätzlicher diagnostischer Informationen im Schulalltag sinnvoll.

Mittwoch
11. März

16:15 Uhr

GA 03/142

Chair(s)

Stang, Justine

Schulz, Andreas; Leuders, Timo

Entwicklung und Validierung eines kognitiven Diagnosemodells zur Eingangsdia­gnose und -förderung in Klasse 5 – Teilmodell zu arithmetischen Basiskompetenzen

Die empirische Modellierung mathematischer Kompetenzen beruht auf der kohärenten Abstimmung vier grundlegender Arbeitsschritte: (i) Entwicklung eines theoretischen Modells, (ii) Beschreibung durch ein psychometrisches Modell, (iii) Operationalisierung mit geeigneten Messkonzepten und Messverfahren sowie (iv) Nutzung im Rahmen von Diagnostik und Assessment (Pellegrino et al., 2001). Vor dem Hintergrund des im schulischen Kontext wichtigen Anliegens, kriteriale individualdiagnostische Information bereit zu stellen, spielen kognitive Diagnosemodelle (Rupp & Mislevy, 2007) eine zunehmend bedeutsame Rolle (Leuders, 2014): Diese fokussieren auf einen engen Kompetenzbereich, welcher über kognitive Theorien und Modelle definiert wird, und zu dem individuell unterschiedliche Merkmalsausprägungen zurückgemeldet werden können. Eine solche empirische Informationsgrundlage eignet sich für die Anbahnung spezifischer Fördermaßnahmen.

Die kohärente Abstimmung dieser vier Arbeitsschritte für die Implementierung einer Eingangsdia­gnose in Klasse 5, die in Baden-Württemberg im Schuljahr 2015-16 landesweit und schulartenübergreifend zum Einsatz kommen soll, wird am Beispiel der Kompetenzbereiche Operations- und Zahlverständnis mit Daten aus einer großflächigen Pilotierungsstudie aus dem Sommer 2014 (ca. 5100 SuS, 260 Klassen) berichtet. Zu beiden Kompetenzbereichen wurden a-priori theoretische Stufenmodelle zur Anforderungsstruktur bzw. zur Vorhersage der Itemschwierigkeiten entwickelt. Diese werden im Vortrag vorgestellt und ihre Geltung mittels der Pilotierungsdaten überprüft.

Zahlverständnis am Ende der Grundschulzeit beinhaltet u.a. (vgl. Padberg & Benz, 2011; Radatz et al. 1999) die Verwendung ordinaler und kardinaler Zahlaspekte, einen verständigen Umgang mit dem dekadischen Stellenwertsystem, Zahlraumvorstellungen und Wissen über die Versprachlichung von Zahlen. Zahlverständnis kommt u.a. dann zum Ausdruck, wenn Kinder Zahlen nicht als starre Einheiten sehen, sondern mit diesen ein ganzes Netz von operativen Beziehungen verbinden, die es ermöglichen, sich flexibel im Zahlenraum bis zu 1 Million zu bewegen. Operationsverständnis bezeichnet die Fähigkeit zu wechselseitigen Übersetzungen von Rechenoperationen zwischen außermathematischen Darstellungen, wie bspw. Texten, Bildern oder realen Situationen einerseits, und mathematisch-symbolischen Darstellungen in Form von Zahlen, Rechenzeichen und Termen andererseits (vgl. Padberg & Benz, 2011; Radatz et al., 1999). Als Grundlage hierfür wird das Verfügen über entsprechende Grundvorstellungen darüber angesehen, welche Rechenoperationen zu einer bestimmten Sachsituation passen könnten oder umgekehrt, welche Situationen sich mit welchen Rechenoperationen modellieren lassen (vgl. vom Hofe, 2003). In die Konstruktion des Stufenmodells zum Operationsverständnis eingeflossen sind Befunde über die Anforderungsstruktur und schwierigkeitsgenerierenden Merkmale von Textaufgaben (u.a. Kintsch & Greeno, 1985).

Die aktuell laufende Datenauswertung orientiert sich an den Analyseschritten, die bereits bei der Überprüfung und Validierung der entwickelten kognitiven Diagnosemodelle zu schriftlichen Rechenverfahren (Subtraktion, Multiplikation, Division) in einer Prepilotierung im Herbst 2013 zur Anwendung gekommen ist (Schulz & Leuders, 2014): (1) Sicherstellung der Passung der Items und Skalen zum Raschmodell sowie Überprüfung der Dimensionalität der Kompetenzstruktur (1-dimensional versus 2-dimensional; vgl. Briggs & Wilson, 2003). (2) Überprüfung der Passung der empirischen Itemschwierigkeiten (Logits entsprechend dem Raschmodell) zu den theoretischen Vorhersagen der a-priori entwickelten Stufenmodelle (linear logistisches Testmodell; vgl. Kubinger, 2008). (3) Bestimmung von Kompetenzstufen auf Personenebene: Hierbei dient als Kriterium für die Identifikation von Schwellenwerten auf der Logit-Skala, dass im arithmetischen Mittel alle Items einer Stufe mit einer erwarteten Wahrscheinlichkeit von 0,65 gelöst werden. (4) Analyse der Zuverlässigkeit dieser Personen-Kompetenzstufen-Zuordnung über den Vergleich der entsprechend der Stufenzugehörigkeit modellkonform richtig bzw. falsch beantworteter Aufgaben (vgl. Reproduzierbarkeitsmaß: Rost, 2004).

Im Vortrag diskutiert wird weitergehend die Eignung des vorgestellten kognitiven Diagnosemodells zum Zahlverständnis und Operationsverständnis für die Generierung kriterialer Leistungsrückmeldungen, auf derer Grundlage Mathematiklehrkräfte zu Beginn von Klasse 5 fokussierte und niveauspezifische Förderentscheidungen treffen können. Dazu wird ein Bezug zu bereits bestehenden und aktuell entwickelten sowie evaluierten Fördermaterialien hergestellt.

Mittwoch
11. März

16:15 Uhr

GA 03/142

Chair(s)

Stang, Justine

Latzko, Brigitte; Dietrich, Sandra

Wertschätzung der Heterogenität von Lehramtsstudierenden durch adaptive Beratungssysteme

In Zusammenhang mit Fragen der Eignungsabklärung in der Lehrerinnen- und Lehrerbildung zeichnet sich der Trend ab, Eignungsabklärung nicht nur unter dem Selektionsaspekt, sondern vielmehr zur diagnosegeleiteten Entwicklung professionsbezogener Kompetenzen einzusetzen (vgl. bspw. Nieskens, Mayr & Meyerdieks, 2011). Folgt man zudem Müller-Fohrbrodt (2012), die dafür sensibilisiert, dass es nicht nur einen erfolgreichen Lehrertypus gibt, greifen Typologien, die zwischen geeigneten vs. weniger geeigneten Bewerbern unterschieden, zu kurz, um die Heterogenität unter den Studienanfängern abbilden und adaptive Beratungsinstrumente entwickeln zu können.

Vor diesem Hintergrund zielen wir im Rahmen des Projekts EULe (Eingangsdagnostik für Lehramtsstudierende) in einem ersten Schritt auf dem Weg zur Implementierung verbindlicher differenzierter Beratungsinstrumente darauf ab, die unterschiedlichen Potenziale der Lehramtsstudierenden für Erfolg im Lehramtsstudium bzw. im Lehrerberuf abzubilden. Dazu gehen wir der Frage nach, in welchen Aspekten sich die Vorstellungen und Motive der Studienanfänger bezogen auf Studium und Beruf unterscheiden?

Im WS2014/2015 wurde die komplette Immatrikulationskohorte der Lehramtsstudierenden der Universität Leipzig (N=1029) eingeladen, an einer Onlineerhebung teilzunehmen (Rücklauf: n = 190). Das Durchschnittsalter betrug 21 Jahre (SD = 3,54). Der Fragebogen erfasste Berufswahlmotive, berufsbezogene Überzeugungen, berufsbezogene Persönlichkeitsmerkmale und Integration ins Studium über etablierte Verfahren wie Factors Influencing Teaching Choice (FIT-Choice, Watt et al. 2012; dt. Adaption von König & Rothland, 2012), Lehrer-Persönlichkeits-Adjektivskalen (LPA, Brandstätter & Mayr, 1994) und die Lehrer-Interessen-Skalen (LIS, Mayr, 1998).

Um Antwortmuster ausfindig zu machen und damit Studierende mit ähnlichen Antwortprofilen zu identifizieren, wurde mittels explorativer Faktorenanalyse eine Dimensionsreduktion auf sieben Faktoren (Pädagogische Motive, Pragmatik, Engagement, Intrinsische Motivation, Zufriedenheit, Lehrerexpertise, Kontaktbereitschaft) durchgeführt. Anhand der Ladungen auf diese Faktoren wurde per hierarchischer Clusteranalyse (Ward-Methode) eine Sechs-Clusterlösung für die Stichprobe ermittelt. Die Clusterlösung wurde so gewählt, dass sich für alle Faktoren signifikante Mittelwertunterschiede zwischen den Clustern ergaben.

Ergebnisse zeigen in Übereinstimmung mit den in der Literatur berichteten Befunden, dass 56% der untersuchten Studierenden ihr Lehramtsstudium mit ungünstigen Vorstellungen und Motiven beginnen. Daraus lässt sich ableiten, dass bei einem Großteil der Erstsemesterstudierenden Beratungsbedarf besteht. Allerdings sensibilisieren die Ergebnisse der Clusteranalysen dafür, dass es sich nicht um eine homogene Risikogruppe handelt, sondern der Beratungsbedarf sehr differenziert ausfällt.

Die Befunde werden mit besonderem Fokus auf die Möglichkeit der Umsetzung adaptiver Beratungsangebote diskutiert, um die professionelle Entwicklung der Lehramtsstudierenden derart zu begleiten, dass professionelle Heterogenität ermöglicht werden kann ohne professionelle Standards in der Lehrerbildung vernachlässigen zu müssen.

Mittwoch
11. März

16:15 Uhr

GA 03/149

Chair(s)

Dietrich, Sandra

Dahm, Gunther

Der Einfluss von Brückenkursen auf die Studienleistungen von nicht-traditionellen Studierenden – zur Wirksamkeit propädeutischer Angebote beim Umgang von Hochschulen mit heterogenen Lernausgangslagen

Spätestens seit der Neuregelung des Hochschulzugangs für beruflich qualifizierte Bewerber ohne schulische Studienberechtigung durch die KMK-Vereinbarung im Jahr 2009 (KMK, 2009) erfährt eine Gruppe Studieninteressierter, die sich in vielerlei Hinsicht von „Normalstudierenden“ unterscheidet, zunehmende wissenschaftliche und bildungspolitische Aufmerksamkeit (Dahm et al., 2013; Freitag, 2012; Wolter, 2012). Bislang vorliegende Untersuchungen zum Studienerfolg dieser sogenannten nicht-traditionellen Studierenden kamen zwar zu dem Ergebnis, dass diese – trotz anfänglicher fachlicher Probleme – letztlich ähnlich erfolgreich studieren wie traditionelle Studierende (Richter, 1995; Scholz, 2006; Schulenberg et al., 1986). Dennoch ist in letzter Zeit eine Verschiebung im Diskurs über diese Studierendengruppe erkennbar – weg von bloßen Fragen der Liberalisierung des Hochschulzugangs hin zu den Anforderungen an eine angemessene Ausgestaltung von Studium und Studienbedingungen (Cusumano, 2012; Hanft & Brinkmann, 2013; Kerres et al., 2012; Orr, 2012). Dabei wird vor allem die Notwendigkeit einer adäquaten Ausgestaltung von Studienformaten sowie der Bereitstellung zielgruppenspezifischer Beratungs- und Unterstützungsangebote betont. Im Hinblick auf die fachlichen Anforderungen eines Studiums sollen propädeutische Angebote in Form sogenannter Brückenkurse etwaige Wissensdefizite am Beginn des Studiums ausgleichen und so den Studienübergang nicht-traditioneller Studierender erleichtern (Banscherus & Pickert, 2013; Kattmann & Wieschowski, 2014; Kerres et al., 2012b, S. 40).

Vor dem theoretischen Hintergrund, dass das Fehlen von (schulischem) Vorwissen spätere Lernprozesse beeinträchtigt (Helmke & Weinert, 1997) und im Falle eines Hochschulstudiums – vor allem in den MINT-Fächern – ein Problem für die Bewältigung von Leistungsanforderungen und den Erwerb des Studienabschlusses darstellen kann (Heublein et al., 2010, S. 21ff., 65ff.), erscheint die Bereitstellung propädeutischer Angebote durch die Hochschulen als geeignete Maßnahme, um das Risiko von Leistungsproblemen und damit auch von Studienabbrüchen zu senken. Allerdings ist zur Wirksamkeit solcher Kursangebote in Bezug auf Kompetenzerwerb, Studienleistungen und Studienabbrüche bislang nur wenig bekannt (vgl. bspw. Bettinger & Long, 2007; Biehler et al., 2014). Während etwa Gensch und Kliegl (2011) hinsichtlich des Studienabbruchs von risikomindernden Effekten propädeutischer Angebote berichten, scheint das Abbruchrisiko gemäß Heublein et al. (2010, S. 75ff.) vom Besuch von Brückenkursen nicht beeinflusst zu werden, wobei jedoch (1) unklar ist, wie dieser Befund bei Kontrolle der Schulabschlussnoten der Studierenden aussähe und (2) ob auch hinsichtlich der Studienleistungen bzw. –noten ein Effekt ausbleibt. Forschungsbefunde aus einem anderen Bereich, nämlich zum Effekt von Bildungsangeboten für Migranti(nn)en deuten darauf hin, dass spät einsetzende Fördermaßnahmen kaum Wirkung auf die Bildungsleistungen der Maßnahmenteilnehmenden entfalten (Becker & Beck, 2011; Becker & Tremel, 2011). Unter dem Vorbehalt der Übertragbarkeit derartiger Ergebnisse auf den vorliegenden Kontext stimmen diese Befunde skeptisch hinsichtlich der potentiellen Wirksamkeit der erst am Studienbeginn ansetzenden und meist nur wenige Wochen dauernden propädeutischen Maßnahmen der Hochschulen (vgl. jedoch die Ergebnisse von Greefrath et al., 2014).

Angesichts der vielzähligen Forderungen nach der Einrichtung von Brückenkursen für nicht-traditionelle Studierende bei zugleich sehr dürftigem Forschungsstand zur Wirksamkeit solcher Angebote soll in diesem Beitrag daher die Frage aufgegriffen werden, inwiefern hochschulische Brückenkurse den ihnen zugewiesenen Zweck erfüllen und dazu beitragen, über den Ausgleich fachlicher Defizite die Studienleistungen von nicht-traditionellen Studierenden positiv zu beeinflussen. Beantwortet werden soll diese Frage mit Hilfe von Auswertungen der Daten der Studienanfängerkohorte des Nationalen Bildungspanels (NEPS). Vorläufige Ergebnisse aus multivariaten Regressionsanalysen zeigen, dass – bei Kontrolle der Schulabschlussnoten, des Hochschultyps und der Fächergruppe – die Studiennoten der Studierenden ohne Hochschulreife im 3. Semester weder mit der Teilnahme oder Nicht-Teilnahme an einem Brückenkurs noch mit der Bewertung solcher Kurse als hilfreich oder nicht hilfreich in systematischem Zusammenhang stehen. Stabilität der Ergebnisse vorausgesetzt, bedeutet dies, dass die aktuell von den Hochschulen bereitgestellten Angebote zum Umgang mit heterogenen Lernausgangslagen nicht geeignet sind, die Studienleistungen von Studienanfänger(innen) ohne schulische Studienberechtigung positiv zu beeinflussen. Mögliche Gründe für dieses Ergebnis werden im Vortrag angesprochen und können als Ausgangspunkt für die Diskussion dienen.

Mittwoch
11. März

16:15 Uhr

GA 03/149

Chair(s)

Dietrich, Sandra

Bosse, Elke; Trautwein, Caroline

Studienbezogene Heterogenität als Zusammenspiel individueller Unterschiede und sozialstruktureller Differenzen

Theoretischer Hintergrund: Im aktuellen hochschulpolitischen Diskurs genießt die Heterogenität der Studierendenschaft wachsende Aufmerksamkeit (Wild & Esdar, 2014). Im Vordergrund stehen sozialstrukturelle Differenzsetzungen (z.B. Bildungsherkunft, Migrationshintergrund), während individuelle Unterschiede (z.B. Studienstrategien, Studienziele) nur vereinzelt Berücksichtigung finden. Welchen Einfluss diese unterschiedlichen Aspekte bzw. ihr Zusammenspiel auf das Hochschulstudium haben, ist bislang nur ansatzweise untersucht (Berthold & Leichsenring, 2012; Bülow-Schramm, 2013; Schulmeister et al., 2012). Somit mangelt es an einer studienbezogenen Klärung des Begriffs „Heterogenität“ sowie an empirischen Befunden für eine wissenschaftlich fundierte Lehr- und Hochschulentwicklung.

Der Vortrag setzt hier insofern an, als Heterogenität mit Blick auf individuelle Studienziele und Sichtweisen kritischer Anforderungen der Studieneingangsphase betrachtet wird. Zur weitergehenden Charakterisierung von Studierendenprofilen werden soziale Kategorien herangezogen. Als Basis dient eine qualitative Studie zur Studieneingangsphase, die Forschungsarbeiten zu Transition und Hochschulsozialisation (Gale & Parker, 2012; Huber, 1991) sowie zum Studienerfolg (Richardson et al., 2012) als theoretische Bezugspunkte nutzt. Die Untersuchung folgt einem theoretischen Modell, das individuelle Voraussetzungen, Studienziele sowie die Bewältigung von institutionellen Studienanforderungen als zentrale Komponenten der Entwicklung von Studierfähigkeit beschreibt (Bosse et al., 2014).

Fragestellung: Die explorativ angelegte Untersuchung orientiert sich an den folgenden Forschungsfragen: (1) Worin unterscheiden sich die individuellen Studienziele der befragten Studierenden? (2) Worin bestehen kritische Anforderungen der Studieneingangsphase aus Sicht der Befragten? (3) Inwiefern lassen sich Zusammenhänge zwischen individuellen Unterschieden und sozialen Differenzen aufdecken?

Insgesamt zielen diese Fragen darauf ab, Verbindungen zwischen individuellen und sozialstrukturellen Kategorisierungen nachzugehen, um Hinweise auf typische Studierendenprofile zu gewinnen.

Methode: In Orientierung an Anforderungsanalyse (Hell et al., 2007) und Critical Incident Technique (Flanagan, 1954) wurden qualitative Leitfadeninterviews mit 25 Studierenden aller Fakultäten der Universität Hamburg durchgeführt. Im Sample sind Studienanfänger/innen sowie fortgeschrittene Studierende vertreten, die sich durch unterschiedliche Eingangsvoraussetzungen und Studiensituationen auszeichnen. Zur Ermittlung individueller Studienziele wurden die Studierenden nach ihrem Verständnis gelingenden Studierens befragt, während kritische Studienanforderungen mittels Fragen nach besonderen Herausforderungen der Studieneingangsphase erhoben wurden. Die Interviewauswertung folgte der qualitativen Inhaltsanalyse, ihre Qualität wurde durch konsensuelles Validieren in einem Auswertungsteam sichergestellt (Kuckartz, 2012; Schreier, 2012).

Ergebnisse: Im Hinblick auf die erste Forschungsfrage nach individuellen Studienzielen lassen sich auf Basis der Interviews 14 Kriterien gelingenden Studierens unterscheiden, die sich entweder auf das Studienergebnis oder den Studienverlauf beziehen. Zugleich sind die ermittelten Kriterien danach zu systematisieren, ob sie für formal vorgegebene Studienziele (z.B. Studienabschluss) stehen oder selbst gesetzte subjektive Maßstäbe (z.B. Persönlichkeitsentwicklung) betreffen. Einzelfallanalyse und Fallvergleich zeigen, dass ein Teil der Befragten formale Kriterien fokussiert, während andere über ein facettenreiches Verständnis gelingenden Studierens verfügen, das sowohl formale als auch subjektive Kriterien umfasst.

Hinsichtlich der zweiten Forschungsfrage wurden 32 kritische Studienanforderungen ermittelt, die sich in inhaltliche, personale, soziale und organisatorische Herausforderungen gliedern lassen. Die Auswertung zeigt, dass Anzahl und Bereich der thematisierten kritischen Anforderungen stark variieren. In einigen Fällen stehen einzelne, vornehmlich personale und organisatorische Herausforderungen im Vordergrund. Andere Sichtweisen zeichnen sich dadurch aus, dass zahlreiche, das gesamte thematische Spektrum umfassende Anforderungen als kritisch beschrieben werden.

Mit Blick auf die dritte Frage lässt sich das Zusammenspiel individueller und sozialer Heterogenitätsaspekte anhand ausgewählter Studierendenprofile veranschaulichen. Der Fall eines „Normalstudenten“ (Middendorf et al., 2013) zeichnet sich z.B. durch Orientierung an formalen Studienzielen sowie die Wahrnehmung personaler und organisatorischer Herausforderungen aus. Für eine „nicht-traditionelle“ Studentin (Teichler & Wolter, 2004) mit einer auf dem 2. Bildungsweg erworbenen Hochschulzugangsberechtigung sowie einer zeitintensiven Nebentätigkeit ist dagegen die Fokussierung subjektiver Studienziele und kritischer sozialer Studienanforderungen charakteristisch. Weitere Profile legen allerdings nahe, die dichotome Unterscheidung zwischen traditionellen und nicht-traditionellen Studierenden weiter auszudifferenzieren, um dem komplexen Zusammenspiel sozialer und individueller Unterschiede gerecht zu werden.

Neben Implikationen für Lehr- und Hochschulentwicklung soll zum Abschluss diskutiert werden, wie die Geltungreichweite der Ergebnisse in einer als BMBF-Begleitforschung zum Qualitätspakt Lehre geplanten Mixed-Methods-Studie zu erweitern ist.

Chair(s)
Dietrich, Sandra

Janke, Stefan; Rudert, Selma Carolin; Marksteiner, Tamara; Dickhäuser, Oliver

Bin ich hier richtig? - Effekte des elterlichen Bildungsniveaus auf Zugehörigkeitsgefühl und Wohlbefinden an der Universität

Die Befriedigung des menschlichen Grundbedürfnisses nach Zugehörigkeit ist von integraler Bedeutung für die Entwicklung des subjektiven Wohlbefindens (Baumeister & Leary, 1995). Dies drückt sich unter anderem darin aus, dass Menschen unzufriedener sind und mehr negativen Affekt erleben, wenn sie sich von ihrem sozialen Umfeld ausgeschlossen fühlen (u.a. Reis, Sheldon, Gable, Roscoe & Ryan, 2000; Williams, 2009). Dieser Effekt kann sich auch in Institutionen des Bildungssystems auswirken, welche bei Angehörigen verschiedener Gruppen ein unterschiedliches Ausmaß an Zugehörigkeitsgefühl hervorrufen können. Dies scheint unter anderem bei Hochschulen der Fall, deren Studierendenschaft sich in einigen Ländern stark aus bestimmten ethnischen Gruppen mit hohem sozio-ökonomischen Status rekrutiert. So konnten Walton und Cohen (2007) beispielhaft an einer US-amerikanischen Hochschule zeigen, dass afroamerikanische Erstsemesterstudierende sensibler auf Hinweise möglichen sozialen Ausschlusses in ihrem akademischen Umfeld reagieren als vergleichbare europäisch-stämmige US-Studierende. Als Folge entwickeln sie zu Studienbeginn ein geringeres Ausmaß an Zugehörigkeitsgefühl und im Studienverlauf ein geringeres Ausmaß an subjektivem Wohlbefinden (Walton & Cohen, 2011). Auch Studierende aus bildungsfernen Schichten erleben ein geringeres Ausmaß an Zugehörigkeitsgefühl an der Universität als Studierende aus bildungsnahen Schichten (Ostrove & Long, 2007). Aufbauend auf diesem Befund wollen wir zeigen, dass Kinder aus Akademikerhaushalten schon zu Studienbeginn ein höheres Zugehörigkeitsgefühl zur Statusgruppe der AkademikerInnen schildern als Kinder, deren Eltern nicht über einen akademischen Abschluss verfügen. Weitergehend nehmen wir an, dass sich diese Unterschiede auf die Studienzufriedenheit und die Prüfungsangst (als Maß negativen Affektes) auswirken. Dabei sollte ein stärkeres Zugehörigkeitsgefühl zum Studienbeginn die Entwicklung eines höheren Ausmaßes an Studienzufriedenheit und eines geringeren Ausmaßes an Prüfungsangst bedingen. Um dieser Fragestellung nachzugehen, befragten wir 319 deutsche Studierende (M = 19,9 Jahre; 67,2 % weiblich) längsschnittlich jeweils zu Anfang des ersten und des zweiten Hochschulsemesters mit einem Online-Fragebogen. Wir konnten zeigen, dass Studierende mit mindestens einem Elternteil mit akademischem Abschluss zu Studienbeginn ein signifikant höheres Zugehörigkeitsgefühl erleben als Studierende, deren Eltern nicht studiert haben. Dieser Effekt blieb auch nach Kontrolle der Abiturnote robust. Es zeigten sich keine Eingangsunterschiede zwischen den Gruppen hinsichtlich Studienzufriedenheit und Prüfungsangst. Zum zweiten Messzeitpunkt schilderten Studierende aus Akademikerhaushalten allerdings signifikant mehr Studienzufriedenheit und weniger Prüfungsangst als Studierende aus Nicht-Akademikerhaushalten. Mit Cross-Lagged Panel Analysen ließ sich außerdem hypothesengemäß nachweisen, dass das Ausmaß des Zugehörigkeitsgefühls zu Studienbeginn einen positiven Effekt auf die Studienzufriedenheit und einen negativen Effekt auf die Prüfungsangst zu Beginn des zweiten Semesters hat. Analysen zur Modellinvarianz zeigten weiterhin, dass dieses Pfadmodell für beide betrachteten Gruppen Gültigkeit hat. Dies bedeutet, dass Akademikerkinder auf Grund ihres zu Studieneingang höher ausgeprägten Zugehörigkeitsgefühls im Laufe des Studiums auch ein höheres Ausmaß an Studienzufriedenheit und ein geringeres Ausmaß an Prüfungsangst entwickeln als Kinder aus bildungsferneren Schichten. Der Grund für diesen Effekt könnte, ähnlich wie bei afroamerikanischen Studierenden an Hochschulen in den USA, in der Selbststigmatisierung von Kindern aus bildungsfernen Haushalten liegen. Es ist aber auch denkbar, dass Kinder aus bildungsnahen Schichten im Umfeld ihres Elternhauses mehr Umgang mit anderen Akademikern (ihre Eltern, Freunde der Eltern, etc.) haben und somit schon vor Studienbeginn einen akademischen Habitus und Klassenbewusstsein aufbauen, was wiederum für die Entstehung des Zugehörigkeitsgefühls zuträglich ist. Welcher dieser beiden möglichen Pfade in welchem Ausmaß als Wirkmechanismus hinter den Zusammenhängen steht, sollte von zukünftiger Forschung betrachtet werden. Für den Moment ergibt sich unabhängig vom Wirkmechanismus allerdings das Bild, dass auch nach Absolvieren vorangehender Bildungshürden soziale Unterschiede zwischen Studierenden mit unterschiedlichem sozialem Hintergrund persistieren. Um diese Effekte abzufedern, wäre es erstrebenswert neue Strategien zu entwickeln, um das Zugehörigkeitsgefühl von Kindern mit bildungsfernem Hintergrund zu ihrer neuen Statusgruppe zu steigern. Hier könnten die Vorarbeiten von Walton und Cohen (2011) zur Förderung sozialer Verbundenheit bei Erstsemester-Studierenden als Inspirationsquelle dienen.

Mittwoch
11. März

16:15 Uhr

GA 03/149

Chair(s)
Dietrich, Sandra

Böhringer, Daniela; Korff, Svea; Gundlach, Julia; Roman, Navina

Heterogenität wahrnehmen? Förderprogramme für Postdocs an deutschen Hochschulen und Forschungseinrichtungen

Ausgangspunkt: Universitäten und Fachhochschulen sind einerseits Bildungsinstitutionen, feste Einrichtungen, die auf einem geteilten Habitus und Regelsystemen beruhen, andererseits aber auch Arbeitsorganisationen, in denen unterschiedliche Statusgruppen beschäftigt sind (Kehm 2012). Ausgangspunkt für den Beitrag ist der empirische Befund, dass es für Frauen – vor allem nach der Promotion - offensichtlich „Aufstiegshemmnisse“ (vgl. Hirschauer 2004) in Wissenschaftsorganisationen gibt. Diesen Aufstiegshemmnissen sollen unter anderem Förderprogramme entgegen wirken. Der Beitrag geht deshalb der Frage nach, wie Universitäten, Fachhochschulen und außeruniversitäre Forschungseinrichtungen ihre Förderangebote für die Gruppe der Postdoktorandinnen und Postdoktoranden (kurz Postdocs) in ihren Internetauftritten nach außen darstellen. Die Gruppe der Postdocs ist unter Gendergesichtspunkten besonders interessant, da auf dieser Ebene der wissenschaftlichen Qualifikation nach wie vor Prozesse zu greifen scheinen, die zu einem Ausschluss von Frauen führen (vgl. Kompetenzzentrum 2006, Metz-Göckel u. a. 2010).

Forschungsfrage und Datengrundlage: Im Beitrag werden folgende Fragen behandelt: (a) In welchen Förderprogrammen werden Postdocs angesprochen und welche Typen von Programmen lassen sich identifizieren? (b) In welcher Form wird Gender in den Förderprogrammen für Postdocs thematisiert?

Der Vortrag basiert auf Erhebungen, die im Rahmen des Forschungsvorhabens „Chancengleichheit in der Postdoc-Phase in Deutschland - Gender und Diversity“ durchgeführt wurden. Datengrundlage bildet eine repräsentative, geschichtete Stichprobe von Homepages deutscher Hochschulen und außeruniversitärer Forschungseinrichtungen, die standardisiert und nicht-standardisiert auf ihre Förderprogramme für Postdocs hin untersucht wurden.

Methodisch ist die Studie in einem mixed methods Design angelegt: (a) Standardisierte Analyse von Universitätshomepages, daran anschließend Clusteranalysen zur Typenbildung von Förderprogrammen für Postdocs, (b) Qualitative Analyse (Membership Categorization Analyse) zur Verwendung von Genderkategorien in Ausschreibungstexten.

Theoretischer Hintergrund: Den theoretischen Rahmen des Beitrags bilden zwei Ansätze:

Zum einen wird auf die Konzeption der „gendered organization“ rekurriert. Acker (1990) geht davon aus, dass gegenwärtige gesellschaftliche Strukturen, in die bereits Geschlechterungleichheiten eingebettet sind, auch auf und in Organisationen wirken. Für Acker sind Organisationen vergeschlechtlich, da die unterschiedlichen Verteilungen von z. B. Einkommen, Aufgaben und Positionen zwischen Männern und Frauen nicht zufällig, sondern systematisch seien. Universitäten gelten dagegen gemeinhin als geschlechtsneutrale Organisationen, herrscht in ihnen doch der Glaube an ein meritokratisches System vor, welches Leistung als einziges Selektionskriterium für eine erfolgreiche Karriere definiert. Doch es sind zumeist Frauen, die nicht von einer Statusposition in die nächst Höhere übergehen, sondern vermehrt aus der Wissenschaft aussteigen („akademisches Frauensterben“).

Zum anderen gehen wir auf der Interaktionsebene aus der Perspektive von „doing gender“ davon aus, dass Geschlecht eher ein Merkmal sozialer Situationen und Arrangements darstellt und von Akteuren hervorgebracht, dargestellt und benutzt werden kann (vgl. West/Zimmerman 1987). Dazu ziehen wir eine Weiterentwicklung des Konzepts von Hirschauer (2001) heran, welches mit „undoing gender“ die Möglichkeit ergänzt, Geschlecht auszublenden sowie für die jeweilige Situation zu neutralisieren (Hirschauer 2001; 2014).

Ergebnisse: Die Typenbildung im quantitativen Untersuchungsteil macht deutlich, dass Förderprogramme für Postdocs vor allem eine „Komm-Struktur“ haben. Sie sind überwiegend für „alle“ Postdocs da, unabhängig von Geschlecht, Herkunft oder anderen Diversity-Kategorien. Insgesamt lassen sich vier Typen von Förderprogrammen ausmachen, wobei sich zeigt, dass das Angebot größtenteils aus frei zugänglichen Informationen sowie finanziellen Förderprogrammen besteht. Ein geringerer Anteil der Programme zielt hingegen auf eine eher kurzfristige Qualifizierung und Beratung von Postdocs ab, oder aber auf eine im Vergleich längerfristige Qualifizierung und Finanzierung von Postdocs hin zur Professur. Die qualitative Analyse zeigt darüber hinaus, dass der Eindruck von „für alle“ oder „Neutralität“ in den Programmausschreibungen durch die Verwendung von Kategorisierungen aktiv hergestellt wird (Böhringer, Gundlach & Korff 2014).

Mittwoch
11. März

16:15 Uhr

GA 03/149

Chair(s)

Dietrich, Sandra

Martens, Thomas

Zur Heterogenität der Lernmotivation im Studium

Theoretischer Hintergrund: Das Rubikon-Modell der Handlungsphasen (Heckhausen, 1989) wurde auf Lernprozesse und Lernmotivation übertragen und um eine vorausgehende Motivierungsphase erweitert, die eine erste Einschätzung der künftigen Lernsituation vornimmt. Dazu wurden Prozessmerkmale der Schutzmotivationstheorie (Rippetoe & Rogers, 1987) sowie der Hilfetheorie (Schwartz & Howard, 1981) in die erste Motivierungsphase integriert (Martens, 2000, 2012; Martens & Rost, 1998).

Im Integrierten Lern- und Handlungsmodell (Martens, 2012) werden deshalb drei Prozessphasen einer vollständigen Lernhandlung unterschieden, die für tiefgreifende Lernprozesse notwendig sind: die Motivierungsphase, die Intentionsphase und die Volitionsphase.

In der Motivierungsphase wird eine Lernmotivation ausgebildet, d.h. es entsteht ein Bedürfnis, eine lernbezogene Soll-Ist-Diskrepanz zu reduzieren. In der Intentionsphase wird eine Lernintention gebildet, die die Lernmotivation erfüllen kann. In der Volitionsphase wird schließlich eine Lernintention in eine tatsächliche Lernhandlung umgesetzt.

Desweiteren wurden die volitionalen Prozesse des Lernens genauer identifiziert (siehe Kuhl, 2000), die ihren Ursprung in den theoretischen Ableitungen der PSI-Theorie (Kuhl, 2001) haben. Hierauf aufbauend wurde zu allen Prozesskomponenten in den drei Phasen der Bezug zu den Makrosystemen der PSI-Theorie spezifiziert. Für die motivationale Regulation sind insbesondere die Assoziationen wichtig, die im Makrosystem Extensionsgedächtnis gebildet werden: hier werden die Bezüge zum Selbst hergestellt. So kann etwa das von Bandura (1977) eingeführte Konstrukt self efficacy theoretisch genauer eingeordnet werden. Die in diesem Konstrukt angelegten Selbstbezüge können als Assoziationen im Makrosystem Extensionsgedächtnis aufgefasst werden, damit wird self efficacy zu einer Handlungs-Selbst-Kongruenz: passen die avisierten Lernhandlungen wirklich zu mir?

Fragestellung: Auf der Grundlage des Integrierten Lern- und Handlungsmodells sollen folgende Fragestellungen beantwortet werden: Können bedeutungsvolle Prozessmuster und Typen der Lernmotivation identifiziert werden? Welchen subjektiven Empfindungen entsprechen die unterschiedlichen Prozessmuster der Lernmotivation? Wie verändern sich die Prozessmuster der Lernmotivation im Laufe der schulischen und universitären Ausbildung? Wie werden Rückmeldungen zum eigenen Lernen und zu den eigenen Lernprozessen verarbeitet und nachfolgend genutzt, um die eigenen Lernprozesse zu verändern und zu regulieren?

Methode: Mit Hilfe eines Mixed-Methods Sequential Explanatory Designs (vgl. Ivankova, Creswell, & Stick, 2006) wurden zunächst 523 Studierende der Erziehungswissenschaften und 516 Studierende der Informatik mit einem Fragebogen befragt. Davon wurden dann 12 ausgewählte Studierende interviewt.

Der Fragebogen umfasste die Konstrukte des Integrierten Lern- und Handlungsmodells (Motivationsphase: Mißerfolgsangst, Repressives Coping, Verantwortungsübernahme; Intentionsphase: Handlungs-Ergebnis-Erwartung, Kompetenzerwartung, Volitionsphase: Abschirmung, Selbstkongruente Zielverfolgung, Lernen mit Peers, Erfolgserleben, Emotionsregulation, Prokastination, Durchhalten, Lernstrategien). In einem ersten Schritt wurde jede Skala mit dem ordinalen Rasch-Modell analysiert. Die Rasch-Reliabilität für die Skalen lag in fast allen Fällen über 0,8 bei 4 oder 5 Items je Skala. In einem zweiten Schritt wurde eine Latent-Class-Analyse angewendet, um die motivationalen, intentionalen und volitionalen Muster der Lerngenese als entsprechende Subpopulationen zu identifizieren. Als Ausgangswerte dafür wurden die gerundeten Personenparameter aus den Rasch-Analysen verwendet. Die resultierende Anzahl der Subpopulationen wurde vor allem mit den mittleren Zuordnungswahrscheinlichkeiten ($>0,9$) und dem Bootstrapping (Prüfgröße $\text{Pearson } X^2 > 0,5$) bestimmt. Die Interviews wurden nach den Regeln des problemzentrierten Interviews durchgeführt (Witzel, 2000).

Ergebnisse:

Für jede Stichprobe konnten jeweils 5 Subpopulationen identifiziert werden. Die Profile der Subpopulationen und die Verteilung der Subpopulationen auf die Studiengänge konnten eine vorausgehende Studien bestätigen (vgl. Metzger, Schulmeister, & Martens, 2012) und gleichzeitig zusätzliche populationspezifische Typen der Lernmotivation identifizieren: eine extrinsische Lernmotivation in der Informatik und eine verunsicherte Lernmotivation in der Erziehungswissenschaft. Die Ergebnisse der Interviews zeigen, dass die Lernmotivation oft mit prägenden Erfahrungen einher geht und die Lernprozessmuster dann über eine lange Zeit nicht mehr verändert werden, gerade wenn die Erfahrung mit negativen Emotionen verbunden ist. Dieser Befund stützt die These, dass die Motivationsregulationsprozesse blockiert werden können, wenn der Abgleich mit dem Makrosystem Extensionsgedächtnis gestört ist.

Mit dieser Untersuchung kann die Heterogenität der Lernmotivation im Studium gezeigt werden. Hieraus lassen sich wertvolle Hinweise für die Gestaltung von Lehr-Lernprozessen an Hochschulen gewinnen.

Mittwoch
11. März

16:15 Uhr

GA 03/149

Chair(s)

Dietrich, Sandra

Griese, Birgit; Bettina, Rösken-Winter

Lernstrategien im Studium: Faktorenstruktur und Reduktion des LIST-Fragebogens

Die Ruhr-Universität Bochum bietet das Projekt MP²-Mathe/Plus zur Unterstützung von Studierenden der Ingenieurwissenschaften in Mathematik an, das auch die Beforschung des Lernverhaltens umfasst. Da der LIST-Fragebogen (Lernstrategien im Studium, Wild & Schiefele, 1994) als bewährtes und veränderungssensibles Instrument zur Datenerhebung durch seine große Zahl von Items viel Zeit beansprucht, wurde das über mehrere Kohorten gesammelte Datenmaterial zur Untersuchung der postulierten Faktorenstruktur bei gleichzeitiger Kürzung der Skalen verwendet.

Theoretischer Hintergrund und Forschungsfrage: Um die Probleme zu verstehen, die beim Übergang zur Universität im Fach Mathematik auftreten, ist nach traditioneller Betonung kognitiver Aspekte nun die Rolle von Affekt, Motivation und Beliefs stärker in den Fokus der Forschung gerückt (vgl. Gómez-Chacón, García-Madruga, Vila, Elosúa, & Rodríguez, 2014), was ein tieferes Verstehen von Lernprozessen ermöglicht. Da Lernstrategien sowohl kognitive als auch affektive Dispositionen spiegeln, kann ihre Untersuchung (vgl. Artelt, 2005; Rach & Heinze, 2011) helfen, Interdependenzen zu erfassen und zielgerichtet zu beeinflussen. Für die entsprechende Datenerhebung stehen verschiedene Fragebögen zur Verfügung. Oft genutzte Messinstrumente wie MSLQ (Pintrich, Smith, Garcia, & McKeachie, 1993) oder LASSI (Weinstein & Palmer, 2002) unterscheiden theoriekonform zwischen (meta-)kognitiven, motivationalen und selbstregulativen Aspekten. Der für MP²-Mathe/Plus eingesetzte LIST-Fragebogen trennt zwischen kognitiven, metakognitiven und ressourcenorientierten Lernstrategien. Die Forschungsfrage zielt daher darauf ab, inwiefern diese Kategorisierung auch mit einem verkürzten Instrument erreicht werden kann:

Kann durch Kürzung des LIST-Fragebogens bei Erhaltung der Faktorenstruktur ein reliables Instrument gewonnen werden?

Methodologie: Im laufenden Projekt MP²-Mathe/Plus wurden Lernstrategien über LIST mit 69 Items (LIST69) erhoben (N=2375). Dieses Sample wurde zufällig geteilt (Random Split); die erste Hälfte (N=1188) wurde für die Reduktion des Fragebogens herangezogen, indem zunächst Cronbach's alpha für die Originalskalen berechnet wurden. Alle Skalen wurden reduziert, indem die Items schrittweise so eliminiert wurden, dass die alpha-Werte möglichst hoch blieben. Dieser Prozess wurde (solange $\alpha > 0.7$) wiederholt, bis die Skala die Hälfte der bisherigen Items enthielt. Dies wurde dreimal für verschiedene Random Splits wiederholt. Weitere Tests umfassten Hauptkomponentenanalysen (PCA, Varimax Rotation, paarweiser Fallausschluss) der verkürzten Fragebogenversion der zweiten Hälfte (N=1187), die kleinere Veränderungen der Itemlöschung induzierten. Konfirmatorische Faktorenanalysen vervollständigten die Untersuchungen.

Ergebnisse: Die postulierten 10 Skalen ließen sich in dem Datensatz zu LIST69 nicht treffgenau wiederfinden. Die Extraktion von 10 Faktoren erklärte 45.65% der Varianz in LIST69. Dass sieben von den zehn Skalen die exakt gleichen Reduzierungen in allen drei Random Splits aufwiesen und die übrigen in zwei Reduktionen übereinstimmten, bewirkte im ersten Reduktionsschritt eine Version mit 45 Items, LIST45, für den 10 Faktoren sogar 59.03% der Varianz erklären, eine signifikante Verbesserung. Einige Items blieben jedoch in beiden Versionen problematisch, z.B. formte Metakognition keine separate Skala. Das passt zu den Ergebnissen anderer Untersuchungen (vgl. Wild, 2000). Die Eliminierung dieser Items führte zu LIST36, dessen Items nicht nur passgenau auf die neun verbleibenden Skalen (alle außer Metakognition) laden, sondern auch gute interne Reliabilitäten ($\alpha > 0.7$) aufweisen und 56.17% der Varianz erklären. Außerdem löst die Reduktion das Problem der Multilinearität in LIST69. Eine abschließende Konfirmatorische Faktorenanalyse bestätigt die Struktur von LIST36 durch gute Passwerte (RMSEA=0.044, zum Vergleich RMSEA=0.042 für LIST69). Zudem zeigen sich die Skalen von LIST36 auch in den einzelnen Umfragen der Kohorten durchgängig reliabel.

Eine kritische inhaltliche Betrachtung des verkürzten Fragebogens LIST36 im Hinblick auf Vernachlässigung bestimmter Aspekte bzw. Verschiebung der Skalenschwerpunkte liefert keine Kontraindikation. Jedoch weist LIST36 keine metakognitiven Items auf; diese Aspekte finden sich auch nicht in vergleichbarer Tiefe in anderen Items wieder. Um valide Daten hierfür zu erhalten, könnten andere Formen der Datenaufnahme, z.B. Interviews oder schriftliche Reflexionen, sinnvoll sein.

Insgesamt konnte ein reliables Instrument generiert werden, das Lernstrategien in eindeutigen Dimensionen erfasst und gleichzeitig den Zeitaufwand der Datenerhebung deutlich reduziert und so eine erhöhte Bereitschaft zur aktiven Teilnahme erwarten lässt.

Mittwoch
11. März

16:15 Uhr

GB 03/49

Chair(s)

Griese, Birgit

Jonas, Katharina; Mägdefrau, Jutta; Michler, Andreas

Aufgabenspezifischer Lernstrategieeinsatz und die Qualität historischen Argumentierens in Schülertexten

Das an der Universität Passau in Kooperation von Geschichtsdidaktik und Erziehungswissenschaft durchgeführte ALGe-Projekt (Adaptive Lernaufgaben im Fach Geschichte) untersucht den Zusammenhang von Personmerkmalen, Aufgabenwahrnehmungen, Lernstrategieeinsatz und Lernerfolg. Theoretisch basiert die Studie zur Aufklärung der von Schülern bei der Aufgabenbearbeitung eingesetzten Lernstrategien auf dem Drei-Phasen-Modell der Lernstrategienutzung (Wild 2000), das einen Einfluss der Wahrnehmung und Interpretation von Lernaufgaben und Lernmaterial durch den Lernenden auf Prozesse der motivationalen Regulation und der Handlungssteuerung während der Lernphase postuliert. Das Modell beansprucht, Lernstrategienutzung, nicht jedoch Lernerfolg erklären zu können. Der Zusammenhang zwischen Lernstrategienutzung und Lernerfolg ist in vorliegenden Studien nicht immer nachweisbar, scheint sich jedoch bei aufgabenspezifischer Erfassung des Lernstrategieeinsatzes zu erhöhen (vgl. Leutner & Leopold 2006).

Ein Ziel des Geschichtsunterrichts ist, bei Lernenden die Fähigkeit auszubilden, historische Urteile zu fällen (Becker 2012). Die Diskussion um die Messung und Graduierung von Kompetenzen historischen Lernens generell und der Urteilskompetenz im Speziellen ist jedoch noch unentschieden. Bisher liegt noch kein Instrument vor, das historisches Argumentieren und Urteilen reliabel messen kann. Die Variable „Leistung“ oder „Lernerfolg“ als abhängige Variable der komplexen Interaktionen von Individuum, Aufgabenwahrnehmung und Lernstrategieeinsatz lässt sich im Fach Geschichte – sofern nicht deklaratives Wissen, sondern historisches Argumentieren gemessen werden soll – wegen des Fehlens standardisierter Testverfahren nur auf Basis von Textanalysen ermitteln. Auf Basis der von Biggs und Collis bereits 1982 vorgeschlagenen und in Deutschland im Bereich der auf historisches Lernen bezogenen empirischen Lehr-Lernforschung praktisch nicht rezipierten SOLO-Taxonomie (Structure of the Observed Learning Outcome) wurden in der ALGe-Studie Schülertexte analysiert. Die SOLO-Taxonomie beruht auf Piagets Modell kognitiver Entwicklungsstufen; mit fortschreitender kognitiver Entwicklung ist danach das Erreichen höherer SOLO-Niveaustufen erwartbar (insgesamt 5 Stufen). Gleichzeitig lassen die theoretischen Überlegungen die Annahme zu, dass sich die Qualität der Schülertexte erhöht, wenn tiefenverarbeitende Lernstrategien bei der Lösung der Aufgabe eingesetzt wurden.

Die Daten des ALGe-Projekts basieren auf Schülerbefragungen aus 30 neunten Klassen bayerischer Realschulen. Zur Prüfung der Annahme, dass die höheren Niveaustufen beim Argumentieren erst bei älteren Schüler/innen erreichbar sind, wurde eine kleine gymnasiale Teilstichprobe erhoben, um schließlich folgenden Fragestellungen nachgehen zu können: (1) Lassen sich aufgrund der SOLO-Taxonomie argumentierende Schülertexte reliabel hinsichtlich des erreichten Argumentationsniveaus einstufen? (2) Zeigen sich höhere Niveaustufen in Abhängigkeit von der Klassenstufe und der besuchten Schulart der Schüler/innen (Realschule/Gymnasium)? (3) Lässt sich ein Zusammenhang zwischen dem von den Schüler/innen berichteten Lernstrategieeinsatz und der erreichten Niveaustufe zeigen?

Eine Stichprobe von N = 265 Realschüler/innen (9. Klasse, 48% männlich) sowie N = 46 Gymnasialschüler/innen (11. Klasse, 59% männlich) bearbeitete eine Lernaufgabe, die schriftliches historisches Argumentieren und Urteilen verlangte. Die Texte wurden von geschulten Ratern mittels quantitativer Inhaltsanalysen auf Basis der SOLO-Taxonomie eingestuft. Nach der Bearbeitung der Aufgabe erhielten die Probanden einen für aufgabenspezifische Erfassung modifizierten Fragebogen zu ihrem gerade vollzogenen Lernstrategieeinsatz (adaptierte Fassung des LIST-R; Wild 2013).

Ergebnisse: (1) Es zeigten sich bereits bei den Realschuldaten akzeptable Übereinstimmungswerte zwischen den Ratern ($\kappa = .65$), nach Überarbeitung des Codebuchs wurden die Kappawerte in der gymnasialen Stichprobe deutlich besser ($\kappa = .87$). (2) Im Vergleich zeigten sich in den gymnasialen Schülertexten häufiger höhere Niveaus der Argumentation als bei den Realschüler/innen: 88% der Gymnasialtexte hatten multistructural oder relational Niveau (Stufen 3 und 4 der Taxonomie) gegenüber 18,1% bei den Realschulertexten. Die Stufe extended abstract (höchste Stufe) wurde nur im Gymnasium erreicht. (3) Die Auswertungen zum Zusammenhang zwischen Niveaustufe und Lernstrategieeinsatz waren zum Zeitpunkt der Einreichung des proposals noch nicht abgeschlossen und werden auf der Tagung berichtet.

Die Ergebnisse der Studie bestätigen grundsätzlich die SOLO-Taxonomie als adäquates Mittel zur Messung von Kompetenzniveaus bezüglich des historischen Argumentierens. Methodische Einschränkungen, die Befunde zum Lernstrategieeinsatz sowie Möglichkeiten der individuellen Förderung von Argumentationsfähigkeit auf Basis der SOLO-Diagnostik werden diskutiert.

Mittwoch
11. März

16:15 Uhr

GB 03/49

Chair(s)
Griese, Birgit

Ryssel, Dipl.-Hdl. Jeannine

Die Lernwirksamkeit von Feedback zu Concept Maps – eine Studie im betriebswirtschaftlichen Planspielunterricht

Komplexe Lehr-Lernarrangements wie z.B. Planspiele werden im Unterricht eingesetzt, um Schüler frühzeitig auf die komplexen Anforderungen des beruflichen Alltags vorzubereiten. Planspiele sind jedoch keine Selbstläufer, sondern erfordern instruktionale Unterstützung. Hierzu können Lernstrategien, wie beispielsweise Concept Maps, genutzt werden. Dabei stellt sich die Frage, inwieweit die Lernstrategie, beispielsweise das Erstellen von Concept Maps, allein eine Tiefenverarbeitung der erlernten Inhalte fördert oder ob es eines zusätzlichen Feedbacks zu dieser Lernstrategie bedarf. Dass Feedback grundsätzlich eine lernförderliche instruktionale Maßnahme ist, hat umfangreiche Forschung gezeigt (z.B. Hattie & Timperley, 2007). Ob bzw. welche Art von Feedback zu Lernstrategien, mit deren Hilfe komplexe Sachverhalte erarbeitet werden sollen, wirksam ist, fand jedoch bisher kaum Beachtung.

Vor diesem Hintergrund wurde eine Studie durchgeführt, welche untersucht, inwieweit Feedback zur Lernstrategie Concept Mapping das Lernen betriebswirtschaftlicher Zusammenhänge unterstützt und welche Art von Feedback lernwirksamer ist.

Folgende Hypothesen wurden formuliert: H1: Feedback ist lernwirksamer als kein Feedback bezogen auf den Wissenszuwachs (H1a) und auf die Qualität der Concept Maps (H1b). H2: Elaboriertes Feedback ist lernwirksamer als Einfaches Feedback bezogen auf den Wissenszuwachs (H2a) und auf die Qualität der Concept Maps (H2b).

Insgesamt nahmen 56 Schüler einer neunten Klasse sowie dreier achter Klassen an der Studie teil. Die Schüler wurden einer von drei Gruppen (kein Feedback, einfaches Feedback, elaboriertes Feedback) zugeordnet. Der Ablauf lässt sich dabei als Dreischritt auffassen, bestehend aus einer Planspielphase, einer darauffolgenden Concept Mapping-Aktivität und ggf. einer Feedbackmaßnahme. Alle Schüler spielten zunächst das Planspiel „Easy Business™“. Die Kontrollgruppe (kein Feedback) erstellte zusätzlich, eingebettet in das Planspiel, insgesamt drei Concept Maps. Die erste Treatmentgruppe (einfaches Feedback) erstellte ebenfalls Concept Maps und erhielt zusätzlich einfaches Feedback in Form einer richtigen Lösung (Referenzmap). In der zweiten Treatmentgruppe (elaboriertes Feedback) wurden darüber hinaus die Fehler sowie das Referenzmap besprochen.

Vor und nach der Intervention wurde ein Wissenstest als Paralleltest durchgeführt. Die Fragen des Wissenstests wurden anhand der Taxonomie von Anderson und Krathwohl (2001) entwickelt. Die Lernwirksamkeit durch das Feedback wurde zum einen durch den Wissenszuwachs (Paralleltest) und zum anderen durch die Qualität der Maps operationalisiert. Die Qualität der drei selbsterstellten Concept Maps wurde mittels eines Scoringverfahrens, bewertet.

Es lassen sich insgesamt keine signifikanten Ergebnisse ableiten, welche eindeutig für eine allgemeine Vorteilhaftigkeit des Feedbacks im Vergleich zu keinem Feedback hinsichtlich des Wissenszuwachses sprechen. Hypothese 1a muss folglich verworfen werden.

Es lässt sich ebenfalls kein signifikanter Unterschied in Bezug auf den Wissenszuwachs für die Gruppen Elaboriertes Feedback und Einfaches Feedback ableiten. Hypothese 2a muss folglich verworfen werden.

Hinsichtlich der Qualität der Concept Maps zeigt sich ein signifikanter Unterschied zwischen Feedback und keinem Feedback für das dritte Mapping. Für das erste und zweite Mapping sind die Unterschiede nicht signifikant. Die Analyse der Veränderung der Map-Qualität über die drei Mappings hinweg zeigt einen signifikanten Vorteil des Feedbacks gegenüber keinem Feedback. Hypothese 1b kann demnach gestützt werden.

Die Gruppe Elaboriertes Feedback ist der Gruppe Einfaches Feedback hinsichtlich der Qualität der Concept Maps für das erste und dritte Mapping signifikant überlegen. Hinsichtlich des zweiten Mappings sind die Unterschiede nicht signifikant. Die Analyse der Veränderung der Map-Qualität zeigt keine signifikanten Gruppenunterschiede. Insgesamt lässt sich Hypothese 2b bezogen auf die Qualität des dritten Mappings bestätigen.

Es stellt sich insgesamt die Frage, weshalb die Ergebnisse für den Wissenszuwachs und die Qualität der Concept Maps differieren. Der Wissenstest enthält einfache und komplexe Aufgaben. Die Selbstkonstruktion von Concept Maps ist ein komplexes Aufgabenformat. Daher wird in einem nächsten Schritt geprüft, inwieweit die Komplexität der Aufgaben des Wissenstests einen Einfluss auf den Wissenszuwachs hat. Profitieren die Schüler bei komplexen Aufgaben eventuell eher vom Feedback als bei einfachen Aufgaben?

Darüber hinaus gilt es, den Einfluss individueller Eingangsvoraussetzungen und weiterer Faktoren auf das Feedback zu analysieren.

Mittwoch
11. März

16:15 Uhr

GB 03/49

Chair(s)
Griese, Birgit

Pehmer, Ann-Kathrin; Gröschner, Alexander; Seidel, Tina

„Ja“, „Nein“, „Vielleicht“ - Die Qualität von Feedback als Komponente einer videobasierten Fortbildung zum produktiven Klassengespräch

Theoretischer Hintergrund: Im Deutschen MINT Unterricht ist das Klassengespräch die dominierende Unterrichtsform (Seidel, 2011). Studien belegen in diesem Zusammenhang die wesentliche Bedeutung von Feedback, um Schülerlernprozesse zu unterstützen (Hattie, 2008). Empirische Befunde zeigen allerdings, dass Feedback oftmals nur Auskunft über die Korrektheit einer Antwort gibt, anstatt Schülern Hinweise für weiteres Lernen zu geben (Hattie & Timperley, 2007). Dies wird von Schülern als wenig hilfreich angesehen (van den Bergh et al., 2014) und hindert sie, weitere Erklärungen abzugeben, welche elementar für ein tiefergehendes Verständnis und eine positive Lernerfahrung in MINT Fächern sind (Duschl & Osborne, 2002).

Der dargestellte Ist-Zustand verdeutlicht den Bedarf einer Lehrerfortbildung zur Förderung des produktiven Klassengesprächs sowie die systematische, ganzheitliche Untersuchung der Effektivität eines solchen Angebots (Fishman et al., 2003). Diesbezüglich wurde ein videobasiertes Interventionsangebot („Dialogic Video Cycle“, DVC; = Interventionsgruppe IG) mit einem erweiterten herkömmlichen Angebot verglichen („Runde Tische“ = Kontrollgruppe KG). Beide Angebote wurden auf aktuelle Forschungsbefunde zu effektiven Lehrerfortbildungen gestützt (Desimone, 2009) und unterschieden sich nicht zuletzt durch das Tool Video systematisch voneinander (Gröschner et al., 2014).

Der DVC wurde zweimal im Schuljahr 2011/12 für die 9. Klasse MINT durchgeführt. Zunächst wurden Elemente des produktiven Klassengesprächs in einen eigenen Unterrichtsplan integriert. In zwei weiteren Sitzungen reflektierten die Lehrkräfte (n=6) im kollegialen Austausch die produktive Klassengesprächsführung (Schüleraktivierung; Begleitung von Schülerlernprozessen) am eigenen Video. Die Lehrkräfte in der KG (n=4) reflektierten ihre Erfahrungen aus herkömmlichen Fortbildungen zum Thema Lehrer-Schüler-Kommunikation zweimal in sog. „Runden Tischen“.

Forschungsfragen: Die Forschungsfragen untersuchten die Wirkungskette (Lehrerlernen, Lehrerhandeln, Schülerwahrnehmung) einer neu entwickelten Lehrerfortbildung zum produktiven Klassengespräch am Beispiel des Elements „Feedback“: (1) Lehrerhandeln: Verändern IG Lehrkräfte die Qualität des Feedbacks im Vergleich zu KG Lehrkräften im Verlauf eines Schuljahres? (2) Schülerwahrnehmungen: Verändert sich die Wahrnehmung tiefenorientierter Lernprozesse von Schülern der IG im Vergleich zu Schülern der KG? (3) Lehrerlernen: Inwiefern thematisieren die IG Lehrkräfte das Element Feedback im Rahmen des DVCs?

Methode: (1) Das Unterrichtshandeln von IG und KG wurde zu Beginn (Pre) und Ende (Post) des Schuljahres videographiert. Die Videodaten wurden nach der Event-Sampling Methode in Sprecherturns (Lehrer, Schüler) unterteilt (Raterübereinstimmung 98.1%). Zur Klärung der ersten Forschungsfrage wurde jeder Lehrerturn nach der Qualität des Feedbacks bewertet (aufgabenbezogenes; lernprozessbezogenes; selbstregulatives; kein Feedback) (Kappa .68). Non-parametrische Varianzanalysen mit R dienten zur Ermittlung systematischer Interaktionseffekte von Angebot und Messzeitpunkt. (2) In beiden Gruppen wurden die Schüler (IG: n=126; KG: n=90) direkt nach der videographierten Unterrichtsstunde zur Wahrnehmung ihrer tiefenorientierten Lernprozesse mit einer 4-stufigen Likert-Skala (0= „trifft nicht zu“ bis 3= „trifft zu“) befragt (apre = .82; apost = .85.) Diese diente als abhängige Variable für ANCOVAS. (3) Zur Stützung der Befunde aus Forschungsfrage 1 und 2 wurden qualitative Statements aus den videographierten Fortbildungsworkshops transkribiert.

Ergebnisse: (1) Non-parametrische Varianzanalysen zeigten systematische Interaktionseffekte von Angebot und Messzeitpunkt zum Vorteil der IG: Aufgabenbezogenes Feedback ($F = 9.72$, $df = 1$, $p = .00$), lernprozessbezogenes Feedback ($F = 6.04$, $df = 1$, $p = .01$) und selbstregulatives Feedback ($F = 3.94$, $df = 1$, $p = .05$). (2) Es kann ebenfalls ein Treatmenteffekt berichtet werden ($F(1,174) = 4.14$, $p = .04$, $\eta^2 = .02$), was auf einen Abfall in der Wahrnehmung der KG zurückzuführen ist. (3) Die gezeigten Ergebnisse konnten durch qualitative Aussagen der Lehrkräfte im Rahmen der Fortbildungsworkshops im DVC gestützt werden, welche die Rolle von Feedback als zentral einschätzen.

Diskussion: Die gezeigten Befunde liefern Implikationen für weitere Untersuchungen zur evidenzbasierten Fortbildung. Dabei ist Feedback ein Element des produktiven Klassengesprächs, das positiv verändert werden kann. Zweitens, Veränderungen im Feedbackverhalten werden von Schülern hinsichtlich ihres tiefenorientierten Lernens wahrgenommen. Drittens, die ganzheitliche Untersuchung der Wirkungskette einer Lehrerfortbildung liefert vielseitige Hinweise bezüglich deren Effektivität.

Mittwoch
11. März

16:15 Uhr

GB 03/49

Chair(s)
Griese, Birgit

3

Murböck, Julia; Stegmann, Karsten

Der Effekt von Scaffolding und des Timings kognitiver Modellierung auf den Fertigkeitserwerb in empirischen Forschungsmethoden

Im Rahmen ihres Studiums müssen Studierende zahlreiche Fertigkeiten erwerben. Zum Fertigkeitserwerb hat sich u.a. die Kombination von kognitiver Modellierung und Bearbeitung von Beispiel-Problempaaren bewährt (Renkl, 2011). Ein bislang nur wenig untersuchter Aspekt dabei ist die zeitliche Anordnung: Wird die Fertigkeit (1) in einem Block vor der Bearbeitung von Beispiel-Problempaaren vollständig modelliert, (2) abschnittsweise verteilt oder (3) erst nach der Bearbeitung modelliert? Für die Modellierung vor der Bearbeitung von Beispiel-Problempaaren (1 & 2) spricht, dass davon ausgegangen werden kann, dass damit eine (überflüssige) zusätzliche kognitive Belastung vermieden werden kann (van Gog et al., 2011). Vor dem Hintergrund einer geringeren kognitiven Belastung sollte die verteilte Modellierung sowohl der vorherigen als auch der nachfolgenden Modellierung überlegen sein. Befunde von Schwartz et al. (2011) liefern jedoch auch einen Beleg, dass Modellieren nach entdeckenden Lernaktivitäten erfolgreich sein kann. Die Lernenden sollten sich so den Lerninhalt aktiv erschließen und bestehende Wissenslücken besser überwachen können, wodurch sie die Instruktion besser mit ihrem Vorwissen verknüpfen können. Diese aktive Auseinandersetzung und Überwachung muss jedoch in der Regel gezielt durch Scaffolding unterstützt werden (Reiser & Tabak, in press). In Kombination mit Scaffolding in Form von Selbsterklärungs- und Monitoring-Prompts sollte die nachfolgende Modellierung zu besseren Lernergebnissen führen. Diese Studie bearbeitet daher die Fragestellung, (RQ1) inwieweit der Zeitpunkt der Modellierung, Scaffolding und deren Interaktion einen Effekt auf den Fertigkeitserwerb haben und (RQ2) inwieweit kognitive Belastung den Effekt mediiert.

Zur Bearbeitung der Fragestellungen wurden die UVs Zeitpunkt der Modellierung (vorher vs. verteilt vs. nachfolgend) und Scaffolding in Form von Selbsterklärungs- und Monitoring-Prompts (ohne vs. mit) in einem 3x2-Design variiert. Im Rahmen der Studie sollten die Studierenden mit Hilfe von kognitiver Modellierung und vier Beispiel-Problempaaren lernen, wie man verschiedene experimentelle Designs darstellt. Pädagogik-Studenten (N = 126) im ersten Semester wurden einer der sechs Bedingungen zufällig zugewiesen. Der Lernerfolg wurde mit Hilfe eines Wissenstests gemessen, bei dem die Lernenden die einzelnen Teilschritte beim Darstellen eines Designs umsetzen mussten (Cohen's Kappa = .82). Die kognitive Belastung wurde während der Übung acht Mal per Selbsteinschätzung erfasst (Cronbach's Alpha = .91).

In den Bedingungen mit verteilter Modellierung wurde signifikant mehr Wissen erworben als in den Bedingungen der nachfolgenden Modellierung ($p = .009$; $d = 0,77$). Der Effekt der verteilten Modellierung wurde durch eine geringere kognitive Belastung mediiert.

Mittwoch
11. März

16:15 Uhr

GB 03/49

Chair(s)
Griese, Birgit

3

Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen II

In dem vom BMBF geförderten Programm „Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor (KoKoHs)“ (www.kompetenzen-im-hochschulsektor.de) forschen seit 2011 bundesweit 24 Projektverbände mit ca. 70 Einzelprojekten zur theoretischen Modellierung von akademischen fachbezogenen und fachübergreifenden Kompetenzen, die in verschiedenen Studiendisziplinen erworben werden sowie an deren Operationalisierung und validen Erfassung. In diesem Symposium werden wichtige Ergebnisse aus diesem neuen Forschungsprogramm präsentiert.

Die Beiträge umfassen Arbeitsergebnisse aus den beiden zentralen Strängen fachspezifische und generische Kompetenzen. Sie decken alle im Forschungsprogramm repräsentierten Studiendomänen ab (Bildungs- bzw. Erziehungswissenschaft, Psychologie, Lehrerbildung, Wirtschafts- und Ingenieurwissenschaften) und liefern Resultate für die Kompetenzmodellierung und -messung in den jeweiligen, bislang nur wenig erforschten Fachdisziplinen. Der inhaltliche Fokus aller Beiträge liegt zum einen auf der theoretischen Modellierung des zu erfassenden Konstrukts und den dazu entwickelten Instrumenten. Zum anderen werden dazu jeweils empirische Ergebnisse v.a. aus Validierungsstudien (z.B. zur theoretisch angenommenen Kompetenzstruktur, zur konvergenten bzw. diskriminanten Validität in Relation zu anderen Konstrukten wie Intelligenz etc.) sowie Analysen zu individuellen (z.B. Geschlecht) und institutionellen Merkmalen (z.B. Lerngelegenheiten) präsentiert, mit denen sich der erfasste Kompetenzstand vorhersagen lässt. Die präsentierten theoretischen Modellierungen, Messinstrumente und empirische Befunde werden im Rahmen des Symposiums kritisch diskutiert und Perspektiven für die weitere Forschung herausgearbeitet.

Mittwoch
11. März

16:15 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen II

Förster, Manuel; Brückner, Sebastian; Happ, Roland; Zlatkin-Troitschanskaia, Olga

Geschlechtereffekte bei der standardisierten Erfassung ökonomischen Fachwissens – eine ländervergleichende Perspektive zwischen US-amerikanischen, deutschen und japanischen Studierenden der Wirtschaftswissenschaften

In mehreren US-amerikanischen Studien werden verschiedene Einflussfaktoren auf das ökonomische Fachwissen untersucht und dabei u.a. geschlechtsspezifische Effekte festgestellt (z. B. Parker, 2006). Die oftmals replizierten und auch kontrovers diskutierten Forschungsergebnisse zeigen, dass Männer einen höheren Testwert als Frauen erzielen (Owen, 2012). Als mögliche Erklärungsansätze für die Befunde werden häufig die Verwendung standardisierter Testformate (wie Multiple Choice-Tests), seltener die Besonderheiten des volkswirtschaftlichen Fachwissens (wie z. B. eine Differenzierung in „economic numeracy“ und „economic literacy“) genannt (z. B. Yamaoka et al., 2010b). Einige Studien zeigen, dass männliche Testteilnehmer bei „numeracy“ Aufgaben besser abschneiden (z. B. Williams et al., 1992). Inwieweit die geschlechtsspezifischen Unterschiede auch in europäischen und asiatischen Ländern und neben dem Schulbereich auch im Hochschulbereich zu beobachten sind, wurde bislang wenig erforscht. Dies liegt mitunter daran, dass bisher kaum Testinstrumente vorliegen, die eine länderübergreifende Untersuchung des „Geschlechtereffekts“ ermöglichen. Gerade die Volkswirtschaftslehre eröffnet aufgrund eines international vergleichbaren Kerncurriculums die Möglichkeit der komparativen Analyse von Testergebnissen. Während in Deutschland amerikanische Lehrbücher im Hochschulbereich etabliert sind (z. B. Mankiw; Samuelson & Nordhaus), ist auch in Japan eine hohe Identität des Curriculum zu den beiden Ländern festzustellen: „Japanese economics majors learn mainstream economics“ (Yamaoka et al., 2010a, S. 451).

Der Vortrag verfolgt zwei Zielstellungen. Im ersten Schritt wird analysiert, ob geschlechtsspezifische Unterschiede im volkswirtschaftlichen Fachwissen von Studierenden länderübergreifend auf der Basis von Testdaten aus Deutschland, den USA und Japan festgestellt werden. Im zweiten Schritt werden die Unterschiede im Fachwissen vor dem Hintergrund konstruktsspezifischer Eigenschaften in den Subskalen „numeracy“ und „literacy“ näher untersucht.

Im Rahmen des Projekts WiwiKom wurde mit dem Test of Understanding in College Economics (TUCE) ein valides, international verbreitetes Messinstrument für den deutschen Hochschulbereich aus dem Amerikanischen adaptiert und umfassend validiert (Zlatkin-Troitschanskaia et al., 2014). Zu diesem Zweck wurden neben den Test Adaption Guidelines (ITC, 2005) auch die Standards for Educational and Psychological Testing (AERA, APA & NCME, 2014) eingesetzt, um das Testinstrument u.a. mittels Curriculaanalysen, Experteninterviews, Onlineratings, kognitiven Interviews mit Studierenden zu validieren und einem Vergleich zugänglich zu machen. Auch für den japanischen Hochschulsektor wurde eine Adaption dieses Testinstruments, das ebenso den internationalen Validierungskriterien (s. Yamaoka et al., 2010b) entspricht, von Projektpartnern entwickelt. Die Vergleichbarkeit der drei Testversionen wurde u.a. mittels Rückübersetzungen und Interviews mit den Testentwicklern und (Rück)Übersetzern überprüft.

Die der Analyse zugrundeliegende (Teil)Stichprobe umfasst 4.400 Studierende aus den USA, 1.629 Studierende aus Deutschland sowie 529 Studierende aus Japan (s. Förster et al., im Druck). In Deutschland kam ein Testheftdesign (nested Youden-Square-Design) (Frey et al., 2009) zum Einsatz. Aufgrund der komplexen Datenstruktur wurde daher eine latente Skalierung der Personenfähigkeit für jedes Land auf Basis von IRT-Modellen notwendig. Hierzu wurden zunächst die fehlenden Werte in der Variable Geschlecht imputiert und die Passung von IRT-Modellen (wie dem Raschmodell) für jedes Land überprüft. Im Anschluss erfolgte die Schätzung von fünf Plausible Values für jedes Land und die zugrundeliegenden Subskalen (Mislevy, 1991). Zwecks der besseren Vergleichbarkeit des Geschlechtereffekts in den Subskalen innerhalb eines Landes und länderübergreifend, wurden die Personenparameter auf einen Mittelwert von 500 und eine Standardabweichung von 100 normiert.

Erste Befunde zeigen, dass für die USA und Deutschland erwartungskonform geschlechtsspezifische Unterschiede in den Testwerten festzustellen sind. Hingegen zeigt sich in Japan kein Unterschied in den Testwerten zwischen weiblichen und männlichen Studierenden. Bei einer Analyse der Subskalen „numeracy“ und „literacy“ wird deutlich, dass die Effekte zwischen den Geschlechtern in der Subskala „literacy“ abnehmen (zugunsten der weiblichen Studierenden), jedoch in der Subskala „numeracy“ zunehmen, und hierbei auch in Japan z.T. signifikante Unterschiede zugunsten der männlichen Studierenden deutlich werden. Dies kann als ein erstes Indiz gewertet werden, dass geschlechtsspezifische Unterschiede auch auf die Besonderheit des Konstruktes zurückgeführt werden können.

Mittwoch
11. März

16:15 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen II

Hammer, Svenja; Ehmke, Timo; Koch-Priewe, Barbara; Köker, Anne; Ohm, Udo

DaZ-Kompetenz von angehenden Lehrkräften: Welche Rolle spielen Lerngelegenheiten im Studium?

In zahlreichen Untersuchungen der letzten Jahre wurde deutlich, dass die Gestaltung angemessener Lernumgebungen für Schülerinnen und Schüler mit nichtdeutscher Herkunftssprache besonderen Bedingungen unterliegt.

Fachlehrkräfte sind allerdings für diese Aufgabe häufig nicht angemessen ausgebildet und sehen sich oft für Maßnahmen zur sprachlichen Förderung auch nicht zuständig (Becker-Mrotzek, Hentschel, Hippmann, & Linnemann, 2012; de Jong, 2013; Li & Zhang, 2004; Simmons, 2009; Zhang & Stephens, 2013).

In der Lehrerbildung sollten demnach entsprechende Lerngelegenheiten angeboten werden, um Kompetenzen im Bereich Deutsch als Zweitsprache (DaZ) zu erwerben und um die spezifischen Lernvoraussetzungen von Schülerinnen und Schülern nicht-deutscher Herkunftssprache im Unterricht berücksichtigen zu können.

Inwieweit angehende Lehrkräfte über Kompetenzen im Bereich DaZ verfügen, steht im Fokus des BMBF-Projektes „DaZKom“. Basierend auf einem theoretischen Kompetenzstrukturmodell wurde ein Testinstrument entwickelt und erprobt, mit dem sich DaZ-Kompetenz auch empirisch erheben und mit IRT-Methoden auswerten lässt.

Das Kompetenzstrukturmodell differenziert DaZ-Kompetenz in drei Inhaltsbereiche, die jeweils noch in zwei Teildimensionen unterteilt sind: (1) „Fachregister“ mit beiden Teildimensionen „grammatische Strukturen und Wortschatz“ und „semiotische Systeme“, (2) „Mehrsprachigkeit“ mit den Teildimensionen „Zweitspracherwerb“ und „Migration“ sowie (3) „Didaktik“ mit den Teildimensionen „Diagnostik“ und „Förderung“.

In dieser Studie soll nun untersucht werden, inwieweit die DaZ-Kompetenz von angehenden Lehrkräften mit der Art und den Inhalten von Lerngelegenheiten im Bereich DaZ zusammenhängen. Konkret wird dabei folgenden Fragestellungen nachgegangen: (1) Welche Zusammenhänge bestehen zwischen der DaZ-Kompetenz von Lehramtsstudierenden und studiengangsspezifischen Merkmalen wie: Studienfach (Deutsch als Zweitsprache vs. Lehramt), Lehramtsunterrichtsfach (z.B. Deutsch vs. Mathematik), Bachelor- vs. Masterniveau?; (2) Inwieweit gibt es Zusammenhänge zwischen der DaZ-Kompetenz von Lehramtsstudierenden und der Art und Modus der Lerngelegenheit z.B.: Lehrveranstaltung (z.B. Seminar vs. Vorlesung), Prüfungsleistung (z.B. Klausur vs. Hausarbeit), praktischen Erfahrung im Bereich DaZ (z.B. Nachhilfe vs. Unterrichtspraktikum)?; (3) Welche Zusammenhänge bestehen zwischen der DaZ-Kompetenz von Lehramtsstudierenden und der Behandlung konkreter Inhalte und Thematiken (z.B. Mehrsprachigkeit), der Häufigkeit der Thematisierung in Lehrveranstaltungen (z.B. in einer Vorlesung einmal gehört vs. in mehreren Seminaren behandelt)?

Die Datenbasis bildet eine Stichprobe von N = 252 Lehramtsstudierenden von den Universitäten Bielefeld, Lüneburg, Dortmund, Münster und Bamberg. Das DaZ-Testinstrument besteht aus 68 Testitems, die sich jeweils einer der drei Inhaltsbereiche zuordnen lassen. Zusätzlich wurde ein Fragebogen zu DaZ-bezogenen Lerngelegenheiten mit 44 Items eingesetzt. Die Bearbeitungszeit beträgt insgesamt 70 Minuten.

Erste vorläufige Ergebnisse der Studie zeigen, dass Lehramtsstudierende mit dem Unterrichtsfach Deutsch eine höhere DaZ-Kompetenz aufweisen als Studierende anderer Unterrichtsfächer. Zudem steigt die DaZ-Kompetenz mit zunehmender Anzahl an Studiensemestern an. Lehramtsstudierende, die häufiger DaZ-spezifische Lerngelegenheiten im Studium wahrgenommen haben, weisen ebenfalls eine höhere DaZ-Kompetenz auf.

Aus den Ergebnissen sollen Schlussfolgerungen für die Gestaltung der universitären Lehre im Bereich DaZ abgeleitet werden, um ein höheres professionelles Niveau von Lehramtsstudierenden zu erreichen.

Mittwoch
11. März

16:15 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen II

Brauch, Nicola; Wäschle, Kristin; Lehmann, Thomas; Nückles, Matthias

Projekt SOSCIE Modellierung und Erfassung geschichtsdidaktischer Kompetenz künftiger Geschichtslehrkräfte vor der Praxis

Theorie: Durch die Erhöhung der Anteile fachdidaktischer und bildungswissenschaftlicher Studienelemente und die Integration von Praxisphasen in das reformierte Lehramtstudium ergibt sich der Bedarf nach geeigneten Messinstrumenten zur Bestimmung domänenspezifischer und berufsbiographisch angemessenen Kompetenzüberprüfungen. Die Geschichtsdidaktik befindet sich im Unterschied zu anderen Domänen dabei in der Phase der Grundlagenforschung. Zu modellieren gilt es zunächst die domänenspezifischen Ausprägungen individuellen fachdidaktischen Denkens, das wir nach Shulman (1986) als Integration der Konzepte des Faches (ck), der Fachdidaktik (pck) und der Bildungswissenschaften (pk) zu erfassen suchen. Künftige Lehrkräfte können sich die dazu notwendige kognitive Fähigkeit der Wissensintegration (Anmarkrud, Braten & Stromsö 2014) bereits vor der Praxis in curricular verankerten fachdidaktischen Lerngelegenheiten aneignen (vgl. auch KMK 2005/2013). Durch den Fokus auf das Sample künftiger Lehrkräfte vor deren Erstbegegnung mit der Praxis legt sich für den inhaltlichen Rahmen der Testentwicklung der Teilaspekt professioneller Lehrerkompetenz „Aufgaben“ nahe (vgl. die Kompetenzmodelle bei COACTIV und TEDS-M). Es wird der Kompetenzbegriff im Sinne „kognitiver Leistungsdispositionen“ (Koeppen et al. 2013) zugrundegelegt.

Fragestellung: Erstens: In welchem Ausmaß sind Studierende vor der Praxis in der Lage, Wissen aus den drei Domänen der Geschichtswissenschaft, der Geschichtsdidaktik und der Lernpsychologie zu integrieren, wenn sie angeregt werden, über geeignete Aufgabenstellungen für den Geschichtsunterricht nachdenken? Zweitens: Welche Hinweise für die Konstruktion von Items für ein standardisiertes Test-Instrument lassen sich aus den qualitativen Beobachtungen in ökologisch validen Settings theoretisch ableiten?

Methode: Zunächst führten wir ein experimentelles Cognitive Lab (N = 52) bei Studierenden aus curricularen geschichtsdidaktischen Veranstaltungen durch. Die Studierenden wurden randomisiert der Kontrollgruppe oder der Experimentgruppe zugeteilt. Die Studie begann mit dem Lesen dreier disziplinär distinkter Texte (ck, pck, pk), über die im Anschluss ein Lerntagebuch verfasst wurde (Glogger et al. 2012). In einem Posttest wurde daraufhin Faktenwissen über die drei Texte und geschichtsdidaktische Begründungskompetenz bei der Entwicklung und Bewertung von Aufgaben für 15jährige SchülerInnen zum Thema Nationalsozialismus geprüft. Die Experimentgruppe erhielt Prompts zur Abfassung des Lerntagebuches, die die Studierenden zur Wissensintegration anregen sollten.

Die Studierenden sollten die drei Texte zur Vorbereitung einer Aufgabenkonstruktion für das Thema „Nationalsozialismus und Holocaust“ lesen. Dabei ging es inhaltlich um die geschichtswissenschaftliche Kontroverse zwischen denjenigen, die von einem schon bei der Machtübernahme gefestigten Plan Hitlers ausgehen und denjenigen, die den Massenmord an den europäischen Juden als eine multikausale Entwicklung argumentieren.

Ergebnisse: Den Studierenden in der Experimentgruppe gelang es besser als denjenigen in der Kontrollgruppe die Texte nicht nur zu paraphrasieren, sondern deren zentrale Konzepte mit dem Fokus auf Aufgabenkonstruktion für den Geschichtsunterricht zu integrieren (Brauch et al. 2014, Wäschle et al., subm.). Aus fachspezifischer Perspektive ist weiterhin die Beobachtung bemerkenswert, dass der fachdidaktische Text (Henke-Bockschatz 2004) häufig missverstanden wurde. Während dieser eine analytische statt moralische didaktische Umsetzung des Themas Holocaust im Geschichtsunterricht beinhaltete, argumentierten viele der Studierenden in entgegengesetzter Richtung. Dies lässt auf Effekte der Art des Themas und des Textes auf das Antwortverhalten der Studierenden schließen, wobei sich stabile Überzeugungen gegen angeleitetes Wissen über kompetenzorientiertes Unterrichten durchzusetzen scheinen. Darüber hinaus bildete sich die Vertrautheit der Studierenden im Umgang mit dem fachwissenschaftlichen Text in den Integrationsleistungen ab, bei denen der Umgang mit dem bildungswissenschaftlichen Text am schwächsten vertreten war.

Für die Konstruktion von Test-Items bedeutet dies, dass die Integration disziplinärer Konzepte generell eine schwierige aber durch Unterstützung bewältigbare Aufgabe darstellt. Wir ziehen daraus die Konsequenz, dass die Aufgabenstämme der zu konstruierenden Items deutlicher Kennzeichnung des integrativen Charakters fachdidaktischer Entscheidungen, und halb-offene Aufgaben gezielten Promptings bedürfen. Basales Wissen über die Konzeption lernprozessanregender Aufgaben (Tulodziecki, Herzig & Blömeke 2009) lässt sich zur Vereinfachung der Integration in die Aufgabenstämme integrieren.

Mittwoch
11. März

16:15 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen II

Hammann, Marcus; Joachim, Cora; Hasse, Sascha; Carstensen, Claus H.; Bögeholz, Susanne

Modellierung und Erfassung von Vermittlungs- und Beurteilungskompetenzen angehender Biologielehrkräfte zum Experimentieren

Vermittlungskompetenzen – d.h. die Kompetenzen der Analyse und Planung von Fachunterricht – sowie Beurteilungskompetenzen von fachlichen Schülerleistungen gelten als wenig beforscht (Benner 2009). Zwar werden allgemeine Modelle zur Planung und Analyse von Unterricht häufig genutzt (z.B. Klafki 1996), doch fehlen empirische Studien zu Vermittlungs- und Beurteilungskompetenzen im Fachunterricht. Angenommen wird, dass fachdidaktische Spezifika bestehen (Arnold & Piewe 2010). Anknüpfend an Studien zur professionellen Wahrnehmung (Van Es & Sherin 2002) und zur Erfassung von Kompetenzen mit Hilfe von Unterrichtsvignetten und -videos (Seidel & Prenzel 2007), fokussiert die vorliegende Studie auf die Nutzung von biologiedidaktischem Wissen (PCK) bei Vermittlungs- und Beurteilungskompetenzen angehender Biologielehrkräfte. Zur Aufbereitung des Theoriehintergrundes wurde das PCK-Modell von Magnusson et al. (1999) für den experimentellen Biologieunterricht konkretisiert. Da Lehrkräfte z.B. über die Kompetenz verfügen müssen, den inquiry cycle im Unterricht umzusetzen, wurde die PCK-Teildimension „knowledge about instructional strategies“ (Magnusson et al. 1999) durch fachdidaktische Kompetenzen zur Vermittlung von Experimentalunterricht präzisiert. Ähnlich wurde bei Beurteilungskompetenzen für den Biologieunterricht vorgegangen (vgl. confirmation bias siehe „knowledge of students' understanding of science“).

Fragestellungen: Die Fragestellungen des Projekts umfassen die theoretische und empirische Fundierung eines Kompetenzmodells für Vermittlungs- und Beurteilungskompetenzen angehender Biologielehrkräfte zum Experimentalunterricht und die Entwicklung von reliablen und validen Messinstrumenten. Ziel ist die Klärung der Dimensionalität des postulierten dreidimensionalen Kompetenzmodells, das Analysekompetenz, Planungskompetenz und Beurteilungskompetenz von Schülerleistungen fasst.

Methode: Zunächst wurde ein Modell zu Vermittlungs- und Beurteilungskompetenz aus der Theorie hergeleitet. Gemäß dem Modell für Schülerkompetenzen beim Experimentieren (Hammann 2004), sollen angehende Biologielehrkräfte, die Kompetenzen von Lernenden bei der Hypothesenbildung, Experimentplanung und Datenauswertung schulen und beurteilen können.

Für die empirische Überprüfung wurden drei Fragebögen entwickelt (Analyse-, Planung- und Beurteilungskompetenzen). Entwickelt wurden offene Aufgaben zu curricular validen fachlichen Konzepten der Sekundarstufe I (z.B. Samenkeimung, Photosynthese), die realistische Beschreibungen konkreter Unterrichtssituationen enthalten. Die Aufgabenentwicklung erfolgte nach dem „betweenitem-multidimensionality“-Ansatz (Wu et al. 1998). Die Aufgaben erlauben das Scoring von poly- und dichotomen Items. Die Entwicklung des Messinstruments erfolgte gemäß dem Entwicklungszyklus nach Wilson (2005; Eggert & Bögeholz 2014).

Erprobt wurden die Aufgaben zunächst mit der Methode des Lauten Denkens (n = 14 „Beurteilungskompetenzen“, 16 „Vermittlungskompetenzen“; Ericsson & Simon 1993). Die Erkenntnisse dienen der Weiterentwicklung der Aufgaben. Anschließend wurde eine Prä-Pilotierung (n = 24 „Beurteilungskompetenzen“, 27 „Vermittlungskompetenzen“) zur Optimierung der Aufgaben und zur Weiterentwicklung des Scoring Guides durchgeführt. Zur inhaltlichen Validierung wurden Expertinnen und Experten (n = 6 „Beurteilungskompetenzen“) aus Praxis und Wissenschaft zur Repräsentativität, Relevanz und Realitätsnähe der Aufgaben befragt.

Die Pilotierung für die Teilkompetenz „Beurteilungskompetenzen“ (N = 145 Studierende für das Lehramt Biologie) erfolgte über 27 Aufgaben, die über 9 Testhefte mit jeweils 9 Aufgaben anhand eines incomplete block designs (Frey, Hartig & Rupp 2009) getestet wurden. Jedes Testheft enthielt 9 verschiedene Aufgabentypen, die über 3 Bearbeitungskontexte realisiert wurden und von denen jeder Bearbeitungskontext sowohl das Beurteilen der Hypothesenbildung, der Experimentplanung als auch der Datenanalyse umfasste.

Ergebnisse: Die Vortestung erbrachte erste empirische Hinweise zu den postulierten drei Teilkompetenzen: Für die Teilkompetenz „Beurteilungskompetenzen“ lag in der Pilotierung bei allen eingesetzten neun Testheften der Cronbachs' Alpha zwischen .62 und .80 (bei 5 Testheften >.71, 1 Testheft > .80). Studierende, die am Ende des Masterstudiums stehen, beurteilten die Experimentierkompetenzen von Schülerinnen und Schülern signifikant besser als Studierende aus den ersten Bachelorsemestern. Für die Erfassung von Analyse- und Planungskompetenzen wurden Reliabilitäten um .55 pro Testheft ermittelt. Die Ergebnisse der Pilotierung wurden zur Optimierung der Aufgabe bzw. zur Aufgabenauswahl für die Hauptstudie verwendet sowie zur Weiterentwicklung des Scoring Guides.

Ausblick: In der laufenden Hauptstudie erfolgen IRT-Modellierungen. Die resultierenden Teilkompetenzen werden dann mit Instrumenten für (angehende) Biologielehrkräfte zur Diagnosekompetenz, zu Selbstwirksamkeitserwartungen und zu wissenschaftlichem Denken validiert.

Mittwoch
11. März

16:15 Uhr

HGB 20

Chair(s)

Pant, Hans Anand
Zlatkin-Troitschanskaia, Olga

Stand, Herausforderungen und Perspektiven der Messung von fachspezifischen und fachübergreifenden akademischen Kompetenzen in verschiedenen Studiendomänen II

Gerick, Julia

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 2

Mit der International Computer and Information Literacy Study (ICILS 2013) werden erstmalig computer- und informationsbezogene Kompetenzen im Sinne von ICT Literacy von Schülerinnen und Schülern der Sekundarstufe I international vergleichend, repräsentativ und computerbasiert, mittels Live-Software-Umgebung und eigens für die Studie entwickelter Aufgabenformaten untersucht und ein empirisch geprüftes Kompetenzstufenmodell auf der Datengrundlage der Testergebnisse von rund 60 000 Schülerinnen und Schülern aus weltweit 21 Bildungssystemen vorgelegt (vgl. Fraillon, Schulz & Ainley, 2013). Die Studie stellt für Deutschland höchst relevante Informationen über den Kompetenzstand von Achtklässlerinnen und Achtklässlern in dieser fächerübergreifenden Schlüsselkompetenz, medial vermittelte Informationen zu verstehen, zu nutzen und zu kommunizieren zu Verfügung und untersucht die Rahmenbedingungen des Kompetenzerwerbs zur Verfügung (vgl. Bos, Eickelmann, Gerick et al., 2014, 2014, in Druck). ICILS 2013 erweitert damit den Kanon der regelmäßig in Studien international vergleichend erfassten Kompetenzen in den Bereichen Mathematik, Lesen und Naturwissenschaften um den im 21. Jahrhundert immer relevanter werdenden Bereich der computer- und informationsbezogenen Kompetenzen.

Im Rahmen dieses zweiteiligen Symposiums werden zentrale Befunde von ICILS 2013 sowie vertiefende Analysen vorgestellt. Die Inhalte des Symposiums gehen dabei über die nationale Berichtslegung (vgl. Bos, Eickelmann, Gerick et al., in Druck, 2014) hinaus und nehmen thematisch besonderen Bezug zum Umgang mit Heterogenität in diesem Kompetenzbereich auf. Die Beiträge nutzen unterschiedliche quantitative methodische Zugänge, die der Mehrebenenstruktur und Stichprobengröße der Studie Rechnung tragen.

Anknüpfend an den ersten Teil des Symposiums stehen im zweiten Teil Analysen zur Nutzung neuer Technologien und ihr Zusammenhang mit dem Kompetenzerwerb sowie die differenzierte Betrachtung der computer- und informationsbezogenen Kompetenzen nach den verschiedenen Differenzlinien im Mittelpunkt. Schaumburg, Eickelmann, Drossel und Lorenz untersuchen, welche Lehr- und Lernbedingungen für die Nutzung digitaler Medien durch Lehrpersonen relevant sind. Im Beitrag von Lorenz, Gerick, Schulz-Zander und Eickelmann werden mögliche geschlechtsspezifische Unterschiede in den computer- und informationsbezogenen Kompetenzen betrachtet und die Bedeutung der computerbezogenen Selbstwirksamkeitserwartung betrachtet. Drossel, Vennemann, Wendt und Schwippert legen den Fokus auf die Analyse herkunftsbedingter Disparitäten und stellen Analysen zu verschiedenen Indikatoren sozialer Herkunft und ihrer Relevanz – international wie national – für die computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern vor. Abschließend betrachten Senkbeil, Eickelmann, Schwippert und Vennemann die Relevanz des Migrationshintergrunds für die Erklärung von Unterschieden in den computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich wie auch im nationalen Kontext. Die Beiträge, auch aus dem ersten Teil des Symposiums, werden von Naumann vor dem Hintergrund methodischer, wissenschaftlich-analytischer und prozeduraler Aspekte aus der Perspektive der empirischen Bildungsforschung diskutiert.

Mittwoch
11. März

16:15 Uhr

HGB 30

Chair(s)

Eickelmann, Birgit
Bos, Wilfried
Gerick, Julia

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 2

Schaumburg, Heike; Eickelmann, Birgit; Drossel, Kerstin; Lorenz, Ramona

Die Nutzung digitaler Medien von Lehrkräften im Unterricht und der Zusammenhang mit Lehr- und Lernbedingungen im internationalen Vergleich

National und international wird seit mehr als 40 Jahren die Nutzung von Computern und digitalen Medien in schulischen Lehr- und Lernkontexten untersucht. Diesbezüglich belegen zahlreiche bisherige Studien sowohl für den Primar- als auch für den Sekundarbereich, dass in Deutschland im internationalen Vergleich digitale Medien unterdurchschnittlich häufig in der Schule und im Unterricht eingesetzt werden (vgl. dazu u.a. Eickelmann, Schaumburg, Drossel & Lorenz, 2014; OECD, 2011; Schulz-Zander, Eickelmann & Goy, 2010; Senkbeil & Wittwer, 2008). Dabei geben die vorliegenden Untersuchungen Hinweise darauf, dass sowohl schulische Lehr- und Lernbedingungen als auch Faktoren auf der Individualebene der Lehrpersonen wie das Alter und das Geschlecht sowie die wahrgenommene Nützlichkeit des unterrichtlichen Einsatzes digitaler Medien und die selbsteingeschätzte Fähigkeit der Verwendung digitaler Medien im Unterricht sowie computerbezogene Einstellungen relevante Faktoren sind, die die Nutzung neuer Technologien in schulische Lehr- und Lernprozesse beeinflussen (vgl. Eickelmann, 2010; Knezek, Christensen & Fluke, 2003; Legris, Ingham & Collette, 2003; Schaumburg, Prasse, Tschackert & Blömeke, 2007; Venkatesh, Morris, Davis & Davis, 2003). Die bisherigen Untersuchungen fokussieren oft nur auf eine Auswahl von Faktoren und untersuchen diese vielfach nur in kleineren Stichproben.

Mit dem vorliegenden Beitrag wird der Effekt der schulischen Lehr- und Lernbedingungen und der vorgenannten anderen Voraussetzungen auf den schulischen Einsatz digitaler Medien auf der Grundlage der Daten der Studie ICILS 2013 umfassend für Deutschland und im internationalen Vergleich untersucht (vgl. Eickelmann et al., 2014). Dazu wird entlang des theoretischen Rahmenmodells der Studie (Fraillon, Schulz & Ainley, 2013; Eickelmann, Bos, Gerick & Kahnert, 2014), in dem die schulische Mediennutzung durch Lehrpersonen als Prozess auf der Ebene der Schule und der Klasse verortet wird, durch Voraussetzungen auf unterschiedlichen verschiedenen Ebenen beeinflusst, deren empirische Relevanz mit dem vorliegenden Beitrag untersucht wird.

Auf Grundlage der Daten von ICILS 2013 wird für Deutschland sowie für weitere ausgewählte ICILS-2013-Teilnehmerländer untersucht, (1) welche schulischen Lehr- und Lernbedingungen, (2) welche einstellungsbezogenen Merkmale der Lehrpersonen sowie (3) welche weiteren Merkmale der Lehrkräfte Effekte auf die unterrichtliche Computernutzung haben.

Die Frage, welche Indikatoren zur Vorhersage des Computereinsatzes durch Lehrkräfte im Unterricht bedeutsam sind, wird auf der Grundlage von Regressionsanalysen für die einzelnen betrachteten Länder untersucht. Um schulische Lehr- und Lernbedingungen zu analysieren, werden neben ausstattungsbezogenen Variablen, wie das Computer-Schüler-Verhältnis, die Einschätzungen schulischer Akteure bezüglich technischer Probleme und die Verfügbarkeit von IT-Ressourcen einbezogen. Die wahrgenommenen Potenziale, die Lehrkräfte dem Einsatz digitaler Medien im Unterricht zuschreiben sowie ihre selbsteingeschätzten didaktischen und technischen Kompetenzen gehen als einstellungsbezogene Merkmale in das Modell ein. Als weitere Merkmale werden das Alter und das Geschlecht der Lehrkräfte aufgenommen.

Die Ergebnisse für die betrachteten Teilnehmerländer der Studie werden vor dem Hintergrund der Kontextfaktoren auf der Ebene der Bildungssysteme, die mit dem sogenannten National Context Survey vorliegen, diskutiert und so vertiefende Hinweise zur Identifikation und Interpretation förderlicher Lehr- und Lernbedingungen für den schulischen Einsatz neuer Technologien entwickelt.

Mittwoch
11. März

16:15 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 2

Chair(s)
Eickelmann, Birgit
Bos, Wilfried
Gerick, Julia

Lorenz, Ramona; Gerick, Julia; Schulz-Zander, Renate; Eickelmann, Birgit

Geschlechtsspezifische Disparitäten im Kontext von computer- und informationsbezogenen Kompetenzen in der Sekundarstufe I : Welche Rolle spielt die computerbezogene Selbstwirksamkeitserwartung?

Mit ICILS 2013 können erstmalig auf der Grundlage von computerbasierten Tests international vergleichend geschlechtsspezifische Disparitäten in den computer- und informationsbezogenen Kompetenzen im internationalen Vergleich untersucht werden (vgl. Lorenz, Gerick, Schulz-Zander & Eickelmann, 2014). Im Rahmen dieses Beitrags wird der Fokus auf mögliche Kompetenzunterschiede zwischen Jungen und Mädchen gelegt und vertiefend der Frage nachgegangen, welche Bedeutung die computerbezogene Selbstwirksamkeitserwartung für die Erklärung möglicher Leistungsunterschiede hat. Bisherige nationale und internationale Studien liefern vielfach Hinweise auf Geschlechterdisparitäten in bereichsspezifischen Kompetenzen (vgl. u.a. Bos et al., 2007; Brehl, Wendt & Bos, 2012; Pant et al., 2013; Prenzel, Sälzer, Klieme & Köller, 2013). Im Zuge des gesellschaftlichen Wandels hin zu einer Informations- und Wissensgesellschaft steigt die Bedeutung von fächerübergreifenden Kompetenzen wie computer- und informationsbezogenen Kompetenzen und macht auch die Untersuchung von Disparitäten im Kontext dieses Kompetenzbereiches notwendig, da damit auch Fragen nach Chancengerechtigkeit und Bildungspartizipation von Mädchen und Jungen einhergehen. Wie sich die Kompetenzen von Mädchen und Jungen tatsächlich abbilden, ist bisher nur vereinzelt testbasiert hinsichtlich ähnlicher Konstrukte dieses Kompetenzbereiches erfasst worden, wobei die Ergebnisse nicht konsistent zugunsten einer Gruppe ausfallen (u.a. ACARA, 2012; Pelgrum et al., 1993).

Zur Untersuchung von möglichen Leistungsunterschieden in den computer- und informationsbezogenen Kompetenzen zwischen Mädchen und Jungen wird für den vorliegenden Beitrag die neben den in ICILS 2013 eingesetzten Leistungstest erfasste computerbezogene Selbstwirksamkeitserwartung als ein Merkmal fokussiert, von dem angenommen wird, dass es geschlechtsspezifische Kompetenzunterschiede erklären kann. Bereits vorliegende Studien haben gezeigt, dass Jungen in der Regel über eine höhere computerbezogene Selbstwirksamkeit verfügen als Mädchen (u.a. Moos & Azevedo, 2009; Senkbeil & Wittwer, 2007; Tømte & Hatlevik, 2011). Zudem gilt die Selbstwirksamkeitserwartung generell als eine der größten Determinanten von Leistung (u.a. Bandura, 1993; Helmke & Weinert, 1997).

Der vorliegende Beitrag greift im Rahmen von ICILS 2013 die Frage nach geschlechtsspezifischen Disparitäten in den computer- und informationsbezogenen Kompetenzen von Schülerinnen und Schülern in der achten Jahrgangsstufe auf. Vertiefend wird für Deutschland und zusätzlich ausgewählte ICILS-2013-Teilnehmerländer die Relevanz der computerbezogenen Selbstwirksamkeitserwartung für die Erklärung von möglichen geschlechtsspezifischen Unterschieden untersucht. Die computerbezogene Selbstwirksamkeit wird dabei differenziert in Hinsicht auf basale und fortgeschrittene Fähigkeiten betrachtet.

Folgenden Fragestellungen wird in diesem Beitrag nachgegangen: (1) Lassen sich Unterschiede zwischen Mädchen und Jungen hinsichtlich der computer- und informationsbezogenen Kompetenzen in Deutschland und im internationalen Vergleich identifizieren? (2) Inwieweit können mögliche geschlechtsbedingte Disparitäten im computer- und informationsbezogenen Kompetenzen durch die computerbezogene Selbstwirksamkeit hinsichtlich basaler und fortgeschrittener Fähigkeiten erklärt werden? Die erste Fragestellung wird mittels deskriptiver Analysen beantwortet (vgl. Lorenz et al., 2014). Die Analysen im Rahmen der zweiten Fragestellung fokussieren vertiefend auf die Erklärung möglicher geschlechtsspezifischer Leistungsunterschiede in Deutschland und ausgewählten weiteren Teilnehmerländern. Nach Prüfung des Konstrukts der computerbezogenen Selbstwirksamkeitserwartung auf Messinvarianz nach Geschlecht (u.a. van de Schoot, Lugtig & Hox, 2012), wird die computerbezogene Selbstwirksamkeitserwartung im Rahmen eines latenten Regressionsmodells mit Mplus (vgl. Muthen & Muthen, 2012) auf ihre Bedeutung für die Erklärung von möglichen geschlechtsspezifischen Unterschieden in den computer- und informationsbezogenen Kompetenzen analysiert. Dabei werden weitere Kontrollvariablen, wie z.B. die Dauer der Computererfahrung berücksichtigt.

Als Ergebnis werden im Rahmen des internationalen Vergleichs die mittleren Kompetenzwerte der Schülerinnen und Schüler sowie deren Verteilung auf die Kompetenzstufen für die in ICILS 2013 beteiligten Bildungssysteme nach Geschlecht berichtet, um mögliche geschlechtsspezifische Disparitäten in computer- und informationsbezogenen Kompetenzen in Deutschland und in den anderen beteiligten Bildungssystemen zu identifizieren. Darüber hinaus wird dargestellt, ob und in welcher Weise mögliche Leistungsunterschiede zwischen Mädchen und Jungen durch geschlechtsspezifische computerbezogene Selbstwirksamkeitserwartung erklärt werden können.

Mittwoch
11. März

16:15 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 2

Chair(s)
Eickelmann, Birgit
Bos, Wilfried
Gerick, Julia

Drossel, Kerstin; Vennemann, Mario; Wendt, Heike; Schwippert, Knut

Computer- und informationsbezogene Kompetenzen von Jugendlichen mit unterschiedlicher sozialer Herkunft

Befunde der empirischen Bildungsforschung belegen seit Jahren die enge Kopplung zwischen dem sozioökonomischen Status von Familien und dem Bildungserfolg der Kinder und Jugendlichen (vgl. Müller & Ehmke, 2013; OECD, 2013; Stubbe, Tarelli & Wendt, 2012; Wendt, Stubbe & Schwippert, 2012). Unterschiede, die sich in Abhängigkeit der sozialen Herkunft einerseits im Zugang zu digitalen Medien und andererseits in ihrer Nutzung ergeben, werden in diesem Zusammenhang konzeptionell vor dem Hintergrund des digital divide (oder auch der digitalen Spaltung) und damit unter dem Aspekt der Chancengerechtigkeit diskutiert (Hilbert, 2011). Die Messung sozialer Disparitäten in der empirischen Bildungsforschung lehnt sich oftmals an die Vorarbeiten von Pierre Bourdieu (1983, 1986) und James Coleman (1988) an, die zur Beschreibung des Status einer Person bzw. einer Familie drei Kapitalformen unterscheiden: ökonomisches, kulturelles und soziales Kapital (vgl. Müller & Ehmke, 2013; OECD, 2013; Stubbe et al., 2012; Wendt et al., 2012). Weitgehend ungeklärt ist bislang, inwieweit sich ein digital divide bei Schülergruppen verschiedener sozialer Herkunft ausmachen lässt. In Deutschland ist dieses Desiderat vor allem im Hinblick auf verschiedene Schulformen ersichtlich.

Auf Grundlage der repräsentativen Schülerdaten von ICILS 2013 wird für Deutschland untersucht, (1) ob und inwieweit sich Disparitäten in den computer- und informationsbezogenen Kompetenzen in Abhängigkeit vor sozialen Lagen der Schülerinnen und Schüler abbilden lassen, wobei zwischen Schülerinnen und Schülern an Gymnasien und anderen Schulformen der Sekundarstufe I, die nicht oder nicht ausschließlich einen gymnasialen Bildungsgang anbieten, differenziert wird. (2) Darüber hinaus wird der Frage nachgegangen, welche Indikatoren (z.B. Computernutzung in der Schule und zu Hause sowie die schulische Förderung lern- und prozessbezogener Computernutzung) zur Vorhersage computer- und informationsbezogener Kompetenzen für Schülerinnen und Schüler an Gymnasien und anderen Schulformen bedeutsam sind und welche Unterschiede sich identifizieren lassen.

Für die differenzierte Erfassung der sozialen Herkunft werden als Indikator für das kulturelle Kapital der Buchbesitz im Haushalt und der Bildungsabschluss der Eltern herangezogen, für das ökonomische Kapital stehen EGP-Klassen sowie der HISEI zur Verfügung. (1) Die Darstellung von Unterschieden in den computer- und informationsbezogenen Kompetenzen von Achtklässlerinnen und Achtklässlern erfolgt deskriptiv für die Gesamtskala und differenziert nach Schulform und sozialer Lage der Schülerschaft. (2) Die Frage welche Indikatoren zur Vorhersage computer- und informationsbezogener Kompetenzen relevant sind, wird im Rahmen eines sequenziellen Regressionsansatzes untersucht. Dabei werden sowohl verschiedene Indikatoren für die Beschreibung der sozialen Lage der Schülerschaft einbezogen als auch Variablen der schulischen und außerschulischen Computernutzung und das computerbezogene Fähigkeitsselbstkonzept sowie Hintergrundmerkmale, wie das Geschlecht der Schüler bzw. des Schülers oder die kognitiven Fähigkeiten.

Mittwoch
11. März

16:15 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 2

Chair(s)
Eickelmann, Birgit
Bos, Wilfried
Gerick, Julia

Senkbeil, Martin; Eickelmann, Birgit; Schaumburg, Heike; Schwippert, Knut; Vennemann, Mario
Computer- und informationsbezogene Kompetenzen von Jugendlichen mit Migrationshintergrund

Sozial- und migrationsbedingte Disparitäten werden in Bezug auf Computer und Internet als digital divide (digitale Ungleichheit; vgl. Zilien, 2009) bezeichnet. Dabei wird der digital divide als mehrdimensionales Konzept verstanden und Benachteiligungen (a) im Zugang zu Computer und Internet, (b) in den Nutzungsarten des Computers, (c) in den motivationalen Orientierungen (z.B. Selbstkonzept, Interesse) sowie (d) im computer- und informationsbezogenen Wissen als zu differenzierende Facetten berücksichtigt (vgl. van Dijk & Hacker, 2003). Benachteiligte Personengruppen werden vornehmlich anhand soziodemographischer Indikatoren wie Einkommen, Bildung, Alter und Geschlecht identifiziert (vgl. Yu, 2006). Weitgehend ungeklärt ist bislang, inwieweit auch die Gruppe der Migranten – aufgrund sozioökonomischer und bildungsbezogener Disparitäten und aufgrund sozio- und ethnisch-kultureller Merkmale (z.B. häuslicher Sprachgebrauch) – einem digital divide in einer oder mehrerer der oben angeführten Dimensionen unterworfen ist (vgl. z.B. Bonfadelli, Bucher & Piga, 2007).

Diese Frage greift dieser Beitrag im Rahmen von ICILS 2013 auf und untersucht, ob und inwieweit sich Disparitäten in den computer- und informationsbezogenen Kompetenzen in Abhängigkeit vom Migrationsstatus abbilden lassen. Dabei kann vermutet werden, dass die Ergebnisse der in ICILS 2013 beteiligten Bildungssysteme – ähnlich wie in schulischen Kernkompetenzen – variieren. Beispielsweise zeigte sich in der jüngsten PISA-Erhebung für den Bereich Mathematik, dass in ausgewählten Ländern wie Australien Schülerinnen und Schüler mit Migrationsstatus besser abschneiden als solche ohne Migrationshintergrund (vgl. Gebhardt, Rauch, Mang, Sälzer & Stanat, 2013).

Für die differenzierte Erfassung des Migrationshintergrunds wird die Zuwanderergeneration (kein Migrationshintergrund vs. erste, zweite oder 2,5. Zuwanderergeneration) anhand des Geburtslands der Eltern des Jugendlichen sowie des Jugendlichen selbst bestimmt. Da die Beherrschung der Sprache in einem Land als Indikator für die kulturelle Assimilation einer Person gilt (vgl. Schwippert, Wendt & Tarelli, 2012) und diese einen bedeutsamen Faktor im Kompetenzerwerb darstellt (vgl. Stanat, Rauch & Segeritz, 2010), wird zusätzlich die Sprache berücksichtigt, die zuhause am häufigsten gesprochen wird (vgl. Eickelmann, Schaumburg, Senkbeil, Schwippert & Vennemann, 2014, in Druck).

Für die in Deutschland untersuchte Stichprobe wird ergänzend im Rahmen eines sequenziellen Regressionsansatzes untersucht, inwieweit mögliche migrationsbedingte Disparitäten im computer- und informationsbedingten Wissen durch sozioökonomische und bildungsbezogene Disparitäten sowie ethnisch-kulturelle Merkmale erklärt werden können. Dabei werden die Nutzungshäufigkeit von Computer und Internet, das Geschlecht und die Schulform als Kontrollvariablen berücksichtigt.

Im Rahmen des internationalen Vergleichs werden die mittleren Kompetenzwerte der Schülerinnen und Schüler in Abhängigkeit ihres Zuwanderungshintergrunds und der überwiegend zu Hause gesprochenen Sprache für die in ICILS 2013 beteiligten Bildungssysteme berichtet. Anhand dieser Ergebnisse kann geprüft werden, ob migrationsbedingte Disparitäten im computer- und informationsbezogenen Wissen, wie von der digital divide-Forschung vermutet, generell, d.h. über alle untersuchten Bildungssystemen anzutreffen sind oder ob diese vornehmlich vom Grad der kulturellen und strukturellen Assimilation der Zuwandererfamilien (Sprachgebrauch in der Familie, sozioökonomischer Status, schulische und außerschulische Nutzung) abhängen. Für die in Deutschland untersuchten Schülerinnen und Schüler wird auch nach Schulformen getrennt dargestellt, ob über die genannten Assimilationsmerkmale und berücksichtigten Kontrollvariablen (Geschlecht, Schulart, Nutzungshäufigkeit, kognitive Fähigkeiten) hinaus Disparitäten im computer- und informationsbezogenen Wissen bestehen.

Mittwoch
11. März

16:15 Uhr

HGB 30

Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Zentrale Befunde und vertiefende Analysen zur IEA-Studie ICILS 2013 (International Computer and Information Literacy Study). Teil 2

Chair(s)
Eickelmann, Birgit
Bos, Wilfried
Gerick, Julia

Ulferts, Hannah

Wie gut ist frühkindliche mathematische Bildung in Deutschland? – Befunde und Perspektiven der Qualitäts-, Effektivitäts- und Professionalisierungsforschung

Der Forschungsstand zu mathematischer Bildung in frühkindlichen Einrichtungen weist auf die Komplexität des Untersuchungsgegenstandes hin. Studien belegen das entwicklungsförderliche Potenzial institutioneller Erfahrungen für den mathematischen Bereich, jedoch scheint die Qualität in den Einrichtungen für die Auswirkungen auf die kindliche Entwicklung von zentraler Bedeutung zu sein (Anders, 2013). Die Qualitätsforschung untersucht in diesem Zusammenhang Struktur-, Prozess- und Orientierungsmerkmale von Einrichtungen (Kuger & Kluczniok, 2008). Hierbei scheint es zudem wichtig, die Heterogenität kindlicher Hintergrundmerkmale zu berücksichtigen: Studien zeigen, dass verschiedene Merkmale der Kindes (z.B. Geschlecht und Alter) und des familiären Hintergrundes (z.B. sozioökonomischer Hintergrund oder häusliche Lernumwelt) differenzielle Befunde zu Auswirkungen institutioneller Erfahrungen erklären können (Anders et al., 2012; Sammons et al., 2002). Als zentraler Einflussfaktor auf die Qualität und auf das entwicklungsförderliche Potenzial von Einrichtungen geraten die professionellen Kompetenzen der frühpädagogischen Fachkräfte zunehmend in das Blickfeld der Forschung und der Praxis (Gasteiger, 2010). Das Symposium möchte der Komplexität des Untersuchungsgegenstandes durch einen facettenreichen Blick auf elementare mathematische Bildung Rechnung tragen.

Die ersten beiden Beiträge beschäftigen sich mit Aspekten der Qualität in den Einrichtungen. Duchhardt und Jordan untersuchen mithilfe von Daten des Nationalen Bildungspanels die Auswirkung von Struktur- und Prozessmerkmalen auf die mathematische Kompetenz von Kindern ein halbes Jahr vor der Einschulung. Der Beitrag von Ulferts, Hachfeld und Anders thematisiert Bildungseffektivität als Facette der Qualität. Am empirischen Beispiel der BiKS-3-10-Studie wird diskutiert, inwiefern sich das Bildungsergebnis selbst, d.h. die kindliche Rechenfertigkeit, als Qualitätsmerkmal eignet. Beide Vorträge berücksichtigen die Rolle unterschiedlicher Merkmale der Kinder und des familiären Hintergrunds für ihre Fragestellungen.

Die weiteren zwei Beiträge berichten Ergebnisse des KomMa-Projekts und fokussieren unterschiedliche Aspekte professioneller Kompetenzen. Der Beitrag von Dunekacke, Jenßen, Grassmann und Blömeke betrachtet mathematikbezogene Überzeugungen von Fachkräften und damit Merkmale der in der Forschung wenig berücksichtigten Orientierungsqualität (Kuger & Kluczniok, 2008). Präsentiert werden Befunde zu Zusammenhängen dieser Überzeugungen mit der Situationswahrnehmung und Handlungsplanung. Jenßen, Dunekacke und Blömeke untersuchen mathematisches Fachwissen und mathematikdidaktisches Wissen von Fachkräften und die Zusammenhänge zwischen den beiden Wissenskomponenten unter Kontrolle von Intelligenz.

Im Anschluss stellt die DiskutantIn Miriam Leuchter die vier Beiträge einander gegenüber. Thematisiert wird wie sich die Beiträge mit ihren unterschiedlichen Untersuchungsebenen (z.B. Einrichtung, Fachkräfte und Kinder), methodischem Vorgehen und Forschungsperspektiven (z.B. Forschung zu Qualität, Effektivität und Professionalisierung) ergänzen und so insgesamt zu einem umfassenderen Verständnis mathematischer Bildung in frühkindlichen Einrichtungen beitragen.

Mittwoch
11. März

16:15 Uhr

HGB 40

Chair(s)
Anders, Yvonne

Wie gut ist frühkindliche mathematische Bildung in Deutschland? – Befunde und Perspektiven der Qualitäts-, Effektivitäts- und Professionalisierungsforschung

Duchhardt, Christoph; Jordan, Anne-Katrin

Auswirkungen von Struktur- und Prozessmerkmalen von Kindergärten auf die mathematische Kompetenz von Kindern: Ergebnisse des Nationalen Bildungspanels

Relevanz und Theoretischer Hintergrund: Längsschnittliche Studien – wie das EPPSE-Projekt (Effective Pre-School, Primary and Secondary Education) in Großbritannien oder die BiKS-Studie (Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vor- und Grundschulalter) in Deutschland – haben in den letzten Jahren wiederholt die positiven Einflüsse guter häuslicher Lernumwelt und hoher Qualität vorschulischer Einrichtungen auf mathematische Kompetenzen von Kindergartenkindern sowie die weitere Kompetenzentwicklung während der Grundschulzeit nachgewiesen (vgl. Anders et al., 2012, Anders et al., 2013, Sammons et al., 2004, Sylva et al., 2013). Dabei kann man institutionelle Qualität als Komposition von allgemeinen strukturellen Charakteristika (z.B. Gruppengrößen, Ausbildung der Erzieherinnen und Erzieher, Verfügbarkeit von mathematikspezifischen Materialien) und Prozessmerkmalen (z.B. Dynamik des pädagogischen Geschehens, entwicklungsangemessener auf die Bedürfnisse abgestimmter Umgang mit dem Kind, positives Interaktionsklima...) auffassen (vgl. Anders et al., 2013), die auch mathematikspezifische Komponenten beinhalten (z.B. pädagogische Interaktionen mit mathematischem Gehalt). Weiterhin konnten Blevins-Knabe et al. (1996) einen Einfluss eigenständiger mathematikbezogener Aktivitäten von Kindern auf deren Mathematikleistungen zeigen.

Dabei fokussieren diese Studien auf den – in dieser Altersspanne zentralen – mathematischen Bereich von „Mengen und Zahlen“, numeracy. Im Nationalen Bildungspanel (NEPS, vgl. Blossfeld et al., 2011) sollen Kompetenzverläufe über die gesamte Lebensspanne abgebildet werden. Daher liegt der Erfassung mathematischer Kompetenz ein Rahmenkonzept zugrunde, das für alle Altersgruppen umgesetzt wird (vgl. Ehmke et al., 2009). Insbesondere werden also auch schon im Kindergarten Bereiche wie „Daten und Zufall“ und „Raum und Form“ einbezogen.

Fragestellung: In diesem Beitrag untersuchen wir daher, welche Faktoren die – inhaltlich breit aufgefasste – mathematische Kompetenz von Kindergartenkindern beeinflussen. Dabei liegt ein Fokus auf strukturellen und prozessbezogenen Merkmalen von Kindergärten. Indirekt führt uns das dazu, die Instrumente, die im NEPS zur Erfassung von häuslicher Lernumgebung, Prozessqualität und mathematikbezogenen Aktivitäten eingesetzt werden, den entsprechenden Instrumenten aus EPPSE bzw. BiKS gegenüberzustellen.

Methode: Diesem Beitrag liegen Daten der Startkohorte 2 des NEPS zugrunde. Die mathematische Kompetenz der teilnehmenden Kinder (N ≈ 3000 Kinder, Durchschnittsalter = 5;10 Jahre, n ≈ 700 Kindergartengruppen) wurde Anfang 2012, etwa ein halbes Jahr vor ihrer Einschulung, in standardisierten, materialbasierten Einzelinterviews erhoben. Zusätzlich wurden Eltern sowie Erzieherinnen und Erzieher zu sich selbst, dem Kind und der Familie bzw. der Kindergartengruppe des Kindes befragt (u.a. zu Prozess- und Strukturmerkmalen der Einrichtungen).

Da die Kinder in Gruppen geschachtelt sind, nutzen wir Mehrebenenanalysen, um mathematische Kompetenz auf Individual- bzw. Gruppenebene zu beschreiben. Auf Individualebene dienen hierbei neben Alter, Geschlecht, kognitiven und sprachlichen Fähigkeiten auch familiäre Hintergrundvariablen wie Migrationshintergrund, Bildungsstand der Mutter, häusliche Lernumwelt und mathematische Aktivitäten des Kindes als Prädiktoren. Auf Gruppenebene beziehen wir Gruppengröße, aggregierten Bildungsstand der Mütter, Verfügbarkeit von mathematikbezogenen Materialien, mathematische Aktivitäten sowie Aus- und Weiterbildungshintergrund der Erzieherinnen und Erzieher mit ein.

Ergebnisse: Erste Ergebnisse zeigen, dass 18,7% der Varianz mathematischer Kompetenz auf Unterschiede zwischen den Gruppen zurückgeführt werden können. Auf beiden Ebenen zeichnet sich der Bildungsstand der Mutter als besonders relevanter Einflussfaktor ab. Auf Individualebene bilden z.B. Geschlecht, Alter und kognitive Fähigkeiten weitere wichtige Prädiktoren für die mathematische Kompetenz der Kinder. Der Einfluss mathematikspezifischer Gruppenvariablen wie mathematische Aktivitäten im Kindergarten wird u.a. unter Einbezug der Gruppengröße untersucht. Die Ergebnisse werden abschließend denen der EPPSE-/BiKS-Studie gegenübergestellt und Unterschiede und Zusammenhänge in der Erfassung diskutiert.

Mittwoch
11. März

16:15 Uhr

HGB 40

Ulferts, Hannah; Hachfeld, Axinja; Anders, Yvonne

Bildungseffektivität frühkindlicher institutioneller Bildung in Deutschland im Bildungsbereich Mathematik

Theoretischer Hintergrund: Mathematische Bildung ist in den Bildungsplänen der meisten Bundesländer verankert (Royar, 2007). Dennoch ist anzunehmen, dass sich Einrichtungen darin unterscheiden, wie gut sie den Bildungsauftrag umsetzen. In der Forschung lassen sich zwei Ansätze unterscheiden, welche dieser Frage nachgehen (Melhuish, 2001). Der Ansatz der Qualitätsforschung verwendet hierfür Merkmale der Einrichtungen, die als Indikatoren von Qualität herangezogen werden. Dieser Ansatz dominiert die frühkindliche Forschung, allerdings stellen sich einige messtheoretische und empirische Herausforderungen, beispielsweise ist Qualität ein normatives und perspektivabhängiges Konstrukt und die Erfassung der Indikatoren ist teilweise sehr aufwendig. Der Ansatz der Effektivitätsforschung vergleicht Einrichtungen anhand des Leistungsstandes bzw. der Leistungsentwicklung der Kinder in den Einrichtungen. Obschon der Ansatz in der Schulforschung auch in Deutschland angewendet wird, ist die EPPSE-Studie (Effective Pre-School, Primary and Secondary Education) bislang die einzige Studie, welche die Effektivität frühkindlicher, institutioneller Bildung untersucht hat (Sylva, Melhuish, Sammons, Siraj-Blatchford, & Taggart, 2010).

Fragestellung: Der Beitrag diskutiert die Anwendbarkeit der Effektivitätsforschung auf die frühkindliche mathematische Bildung in Deutschland. Es werden verschiedene Herangehensweisen zur Bestimmung der Effektivität gegenübergestellt (Sammons et al., 2002). Darüber hinaus werden Kriterien angesprochen, welchen Studien genügen müssen, damit Effektivität ermittelbar ist. Für die Daten der BiKS-Studie (Bildungsprozesse, Kompetenzentwicklung und Selektionsentscheidungen im Vorschul- und Schulalter) wird die Anwendbarkeit empirisch geprüft (Maurice et al., 2007).

Dabei werden vier verschiedene Herangehensweisen verglichen: Es wird untersucht, ob sich Einrichtungen im mittlerem Leistungsniveau ihrer Kinder unterscheiden (Statusvergleich) und ob Unterschiede auch im mittleren Leistungszuwachs feststellbar sind (Zuwachsvvergleich). Des Weiteren wird untersucht, welche Rolle weitere Merkmale der Kinder (z.B. Geschlecht) und des familiären Hintergrundes (z.B. mütterlicher Bildungsabschluss) bei den Vergleichen spielen (kontextualisierter Status- bzw. Zuwachsvvergleich).

Methode: Datengrundlage der Mehrebenenanalysen bildet die Längsschnittstudie BiKS-3-10, welche 544 Kinder aus 97 Einrichtungen in Bayern und Hessen umfasst. Verwendet wurden als Maß früher Rechenfertigkeit Ergebnisse der Skala Rechnen des K-ABC zu drei Messzeitpunkten, an denen sich die Kinder (Alter 3-5 Jahre) in frühkindlichen Einrichtungen befanden. Für die Zuwachsvvergleiche und kontextualisierten Vergleiche werden Merkmale auf Individualebene kontrolliert und dann die verbleibende Varianz auf Gruppenebene berichtet. Für die kontextualisierten Modelle werden verschiedene Merkmale der Kinder (z.B. Geschlecht, Alter zum Testzeitpunkt) und des familiären Hintergrundes (z.B. sozioökonomischer Status, mütterlicher Bildungsabschluss) hinzugezogen.

Ergebnisse: Bei Vergleich des mittleren Leistungsniveaus früher Rechenfertigkeiten ließen sich signifikante Unterschiede zwischen den Einrichtungen zu allen Messzeitpunkten finden (ICCs = .08-.21). Keine Unterschiede ergaben sich beim Vergleich des Leistungszuwachs über ein bzw. zwei Jahre (ICCs < .04). Für die kontextualisierten Vergleiche, bei denen Hintergrundmerkmale kontrolliert wurden, unterschieden sich Einrichtungen weder beim Status- noch beim Zuwachsvvergleich (ICCs < .05). Zu allen Messzeitpunkten, erwiesen sich Merkmale der Kinder und des familiären Hintergrundes als signifikante Prädiktoren des Leistungsstandes bzw. -zuwachs. Mit zunehmendem Alter ließ sich jedoch eine Abnahme im Determinationskoeffizient auf Individualebene beobachten, was insgesamt auf eine Abnahme des Zusammenhangs mit den Hintergrundvariablen hindeutet. Stärkste Prädiktoren waren die Leistungen zu früheren Messzeitpunkten.

Während sich bei Statusvergleichen Effektivitätsunterschiede beobachten ließen, ergaben Zuwachsvvergleiche oder kontextualisierte Vergleiche (sogenannte „faire Vergleiche“, Scheerens, 1992), keine Unterschiede für die Effektivität frühkindlicher Einrichtungen in mathematischer Bildung. Eine mögliche Erklärung für die Befunde stellen Selektionseffekte dar: vermutlich unterscheiden sich Einrichtungen bereits von Anfang an hinsichtlich des Ausgangsniveaus früher Rechenfertigkeiten und relevanter Hintergrundmerkmale ihrer Kinder.

Mittwoch
11. März

16:15 Uhr

HGB 40

Chair(s)

Anders, Yvonne

Wie gut ist frühkindliche mathematische Bildung in Deutschland? – Befunde und Perspektiven der Qualitäts-, Effektivitäts- und Professionalisierungsforschung

Dunekacke, Simone; Jenßen, Lars; Grassmann, Marianne; Blömeke, Sigrid

Mathematikbezogene Überzeugungen von angehenden Erzieher/innen und ihr Einfluss auf die Fähigkeit zur mathematikbezogenen Situationswahrnehmung und Handlungsplanung

Relevanz und Fragestellung: Kinder begegnen bereits vor dem Schuleintritt der Mathematik und erwerben dabei für den Schulerfolg in Mathematik und anderen Fächern bedeutsame Fähigkeiten (Krajewski & Schneider, 2009). Die Entwicklung dieser Fähigkeiten ist auf die Begleitung durch Eltern und Erzieher/innen angewiesen (van Oers, 2009). Ob Erzieher/innen im Rahmen ihrer Ausbildung die dafür erforderlichen Kompetenzen erwerben ist bislang kaum untersucht (National Advisory Panel, 2008). Überzeugungen von Erzieher/innen zur Mathematik sowie ihr Zusammenhang zur Performanz wurden dagegen schon in unterschiedlichen Studien untersucht.

Unsere Studie untersucht den Zusammenhang von mathematikbezogenen Überzeugungen angehender Erzieher/innen mit ihrer Fähigkeit zur Handlungsplanung. Zusätzlich wird dabei die Fähigkeit der angehenden Erzieher/innen zum Erkennen mathematischer Lernsituationen mit einbezogen.

Theoretischer Hintergrund: Situationswahrnehmung (PERC) wird im Kontext der Lehrerbildungsforschung immer wieder als eine wichtige Fähigkeit beschrieben, da sie von zentraler Bedeutung ist, um in hochkomplexen pädagogischen Situationen adaptiv handeln zu können (Star & Strickland, 2008; van Es & Sherin, 2006). Unsere Studie fokussiert vor allem auf das Identifizieren wichtiger Kontextbedingungen (z.B. welche mathematikbezogenen Materialien werden genutzt) der Situation (van Es & Sherin, 2006) aus mathematikbezogener Perspektive.

Darüber hinaus werden die Zusammenhänge zur Handlungsplanung (ACT) angehender Erzieher/innen untersucht. Handlungsplanung wird als bedeutsamer Teilschritt des komplexen Handlungsprozesses angesehen, der für Erzieher/innen wichtig ist, um auch in alltäglichen Situationen die Aufmerksamkeit auf die Interessen der Kinder zu richten (Schäfer, 2005). Handlungsplanung beinhaltet das Abwägen verschiedener situationsspezifischer Alternativen und die Entscheidung über eine der Handlungsalternativen (Widulle, 2009).

Lee und Ginsburg (2007) zeigen, dass fachbezogene Überzeugungen mit der Qualität der Lernangebote zusammenhängen. Benz (2012) zeigt außerdem, dass fachbezogene Überzeugungen auch mit lehr-lernbezogenen Überzeugungen zusammenhängen. In unserer Studie erheben wir fachbezogene Überzeugungen in Anlehnung an Grigutsch et al. (1998) und unterscheiden die formalistisch-schematische Perspektive (FSP), die prozessbezogene Perspektive (PP) und die anwendungsbezogenen Perspektive (AP).

Methode: Mit einem Fragebogen werden die mathematikbezogenen Überzeugungen erfasst. Die Skalen stellen gekürzte Fassungen der Skalen von Grigutsch et al. (1998) dar. Mit einem videobasierten Test werden Situationswahrnehmung und Handlungsplanung erfasst. Die Teilnehmer/innen sehen drei Videos, die mathematikhaltige Situationen in einer Kindertagesstätte zeigen. Anschließend müssen sie innerhalb einer bestimmten Zeitspanne die mathematikhaltigen Aspekte der Situation beschreiben und mehrere Möglichkeiten nennen als Erzieher/in hier zu handeln. Die Aussagen wurden dann anhand eines Codierschemas dichotom (richtig/falsch) codiert. Die Inhaltsvalidität wurde durch eine Expertenbefragung nachgewiesen.

Beide Instrumente wurden bei 354 angehenden Erzieher/innen an zwei unterschiedlichen Messzeitpunkten eingesetzt. Das Durchschnittsalter betrug 23 Jahre (SD = 4.1) und 83% waren weiblich. Die Datenanalysen wurden mit der Software MPlus 5.2 durchgeführt. Im ersten Schritt wurden für alle Konstrukte konfirmatorische Faktorenanalysen durchgeführt und die identifizierten Messmodelle im zweiten Schritt in einem linearen Strukturgleichungsmodell zusammengeführt.

Ergebnisse: Wir konnten ein Strukturgleichungsmodell mit guten Modellfitwerten ($\chi^2(55) = 71.341$, $p = 0.684$, RMSEA = 0.030 [0.00; 0.05], SRMR = 0.05, CFI = 0.98) identifizieren. Im Gegensatz zu Grigutsch et al. (1998) finden wir niedrige bis hohe Korrelationen zwischen allen drei Perspektiven mathematikbezogener Überzeugungen. Allerdings ist nur die anwendungsbezogene Perspektive ein Prädiktor für die Fähigkeit zur Situationswahrnehmung, die wiederum ein Prädiktor für die Fähigkeit zur Handlungsplanung ist.

Eine Erklärung für das von den Lehrkräften abweichende Korrelationsmuster könnte in den widersprüchlichen Erfahrungen mit Mathematik, die die Erzieher/innen im Rahmen ihrer Ausbildung machen liegen (eher negative in der Schule vs. (wenn überhaupt) eher positive, auf die Arbeit mit Kindern bezogene in der Ausbildung). Eine grundsätzliche Limitation der Studie ergibt sich aus einem sehr hohen Zusammenhang zwischen PERC und ACT. Eine Ursache könnte sein, dass beide Konstrukte zum selben Zeitpunkt mit denselben Videos erhoben wurden.

Mittwoch
11. März

16:15 Uhr

HGB 40

Chair(s)
Anders, Yvonne

Wie gut ist frühkindliche mathematische Bildung in Deutschland? – Befunde und Perspektiven der Qualitäts-, Effektivitäts- und Professionalisierungsforschung

Jenßen, Lars; Dunekacke, Simone; Blömeke, Sigrid

Zusammenhänge von mathematikdidaktischem Wissen und mathematischem Fachwissen angehender Erzieher/-innen unter Berücksichtigung von Intelligenz

Relevanz und Fragestellung: Die Frühpädagogik im Bereich Mathematik fokussiert vor allem das mathematikdidaktische Wissen (MPCK) frühpädagogischer Fachpersonen als bedeutsame Facette professioneller Kompetenz (z.B. Sarama & Clements, 2009). In der Lehrerbildungsforschung wird jedoch davon ausgegangen, dass auch mathematisches Fachwissen (MCK) eine bedeutsame Facette professioneller Kompetenz im Bereich Mathematik darstellt (Shulman, 1986) und hohe Zusammenhänge zu MPCK aufweist (vgl. Blömeke, Kaiser & Lehmann, 2010 für Primarstufenlehrkräfte). Empirische Evidenzen für den Elementarbereich stehen bislang jedoch aus.

Aus der Perspektive der Intelligenzforschung wird hingegen postuliert, dass dieser Zusammenhang auf allgemein-kognitive Fähigkeiten, die hierarchisch über den Kompetenzfacetten stehen (z.B. Gustafsson, 1994), zurückzuführen sei. Kernpunkt der Diskussion ist dabei vor allem die Frage, ob Kompetenzfacetten von Intelligenz hinreichend differenziert werden können.

In dem vorliegenden Beitrag soll daher zwei Fragen nachgegangen werden: Wie hängen MCK und MPCK bei angehenden Erzieher/-innen zusammen? Welcher Zusammenhang ergibt sich, wenn Intelligenz auf Konstruktebene ebenfalls modelliert wird?

Theoretischer Hintergrund: In den letzten Jahren wird in der Frühpädagogik zunehmend deutlicher, dass mathematisches Fachwissen aufseiten frühpädagogischer Fachpersonen wichtig ist, um mathematikhaltige Situationen im Kita-Alltag zu erkennen und um mathematische Kompetenzen der Kinder anschließend fördern zu können (Ginsburg & Ertle, 2008; Dunekacke, Jenßen & Blömeke, submitted). Theoretisch kann daher ein hoher Zusammenhang zu MPCK angenommen werden.

In der Kompetenzforschung besteht seit Jahren die Diskussion, inwieweit sich Kompetenzen von Intelligenz abgrenzen lassen bzw. verbunden damit auch die Frage, inwieweit Intelligenz die Zusammenhänge zwischen einzelnen Kompetenzfacetten erklären kann. Theoretisch lassen sich Kompetenzen und Intelligenz dahingehend unterscheiden, dass angenommen wird, dass Kompetenzen im Gegensatz zu Intelligenz erlernbar sind und domänenspezifischer sind (vgl. Hartig & Klieme, 2006). In der Bildungsforschung werden beide Konstrukte daher getrennt modelliert. Es zeigen sich jedoch hohe Zusammenhänge (vgl. Rindermann, 2006). Zugrunde liegt dabei die theoretische Annahme, dass eine allgemein-kognitive Fähigkeit g hierarchisch über allen Kompetenzen steht.

Methode: Im Projekt KomMa wurden Kompetenztests entwickelt, die MCK und MPCK angehender Erzieher/-innen erfassen sollen. Der MCK-Test umfasst 24 Items, die Kombinationen aus verschiedenen mathematischen Bildungsinhalten und mathematischen Prozessen darstellen. Der MPCK-Test umfasst 12 Items, die relevante Themen der Frühpädagogik aus mathematischer Perspektive (z.B. Gestaltung mathematischer Situationen) beinhalten. Beide Tests weisen akzeptable Reliabilitäten auf (Cronbachs Alpha = .89 für MCK und .79 für MPCK) und können als inhaltsvalide angesehen werden (vgl. Jenßen et al., 2013). Mithilfe des IST-Screening wurden numerische, verbale und figurale Intelligenz anhand von 60 Items reliabel (Cronbachs α = .89) erfasst (Liepmann et al., 2012).

Um die Forschungsfragen zu beantworten, wurden 354 angehende Erzieher/-innen aus 16 Fachschulen in Niedersachsen und Berlin getestet. Anhand von CFA-Modellen unter Berücksichtigung der geclusterten Datenstruktur wurden die Korrelationen auf latenter Ebene zwischen den Konstrukten ermittelt.

Ergebnisse: Die Ergebnisse deuten auf einen signifikanten hohen Zusammenhang zwischen den beiden Kompetenzfacetten von .67 ($p < .001$) hin, bei gutem Modell-Fit ($\chi^2(df = 8, N = 340) = 6.067, p = .6398, CFI = 1.00, RMSEA = .00, SRMR = .02$). Das Modell, in dem Intelligenz gemeinsam mit den Kompetenzfacetten modelliert wurde, wies ebenfalls einen guten Fit auf ($\chi^2(df = 26, N = 353) = 38.225, p = .0577, CFI = .98, RMSEA = .03, SRMR = .04$). Der Zusammenhang zwischen MCK und MPCK lag hier bei .65 ($p < .001$). Erwartungsgemäß zeigte sich ein stärkerer Zusammenhang zwischen Intelligenz und MCK (.81, $p < .001$) als zwischen Intelligenz und MPCK (.69, $p < .001$).

Der Beitrag liefert erstmals einen empirischen Hinweis, dass MCK bedeutsam und hoch mit MPCK bei angehenden Erzieher/-innen zusammenhängt. Unter Kontrolle von Intelligenz bleibt ein ähnlich starker Zusammenhang bestehen. Die Korrelationen zwischen Intelligenz und den Kompetenzfacetten sind zwar erwartungsgemäß hoch, erlauben aber noch eine inhaltliche Trennung der Konstrukte. Diese Befunde belegen die Konstruktvalidität der konstruierten Testverfahren in KomMa.

Mittwoch
11. März

16:15 Uhr

HGB 40

Wie gut ist frühkindliche mathematische Bildung in Deutschland? – Befunde und Perspektiven der Qualitäts-, Effektivitäts- und Professionalisierungsforschung

Chair(s)
Anders, Yvonne

Spörer, Nadine

Soziale Integration und Sozialverhalten in inklusiven Schulklassen

Schulische Inklusion bedeutet, dass alle Kinder eine gemeinsame Schule besuchen und gemeinsam in einer Klasse lernen. In diesen Klassen sollen die Kinder in ihrer individuellen akademischen und persönlichen Entwicklung bestmöglich durch ein multiprofessionelles Team unterstützt werden. Das Gelingen einer inklusiven Beschulung zeigt sich aber nicht nur am akademischen Lernerfolg der Kinder, sondern auch darin, wie sehr alle Kinder in den Klassenverband und dessen soziale Strukturen integriert werden können. Aktuelle Studien weisen nach, dass in inklusiv arbeitenden Schulen leistungsschwächere Kinder und Kinder mit einem sonderpädagogischem Förderbedarf häufiger von sozialer Ausgrenzung bedroht sind als ihre leistungsstärkeren Peers. Unklar bleibt derzeit jedoch, welche Wirkungsgefüge diesen Ausgrenzungsprozessen zugrunde liegen und wie diese vermieden bzw. abgemildert werden können.

Daran anknüpfend sollen im Symposium mit vier Originalbeiträgen und einem Diskussionsbeitrag aus der Sonderpädagogik, der Pädagogischen Psychologie und der Grundschulpädagogik die Heterogenität von sozialer Integration und Sozialverhalten in inklusiven Lernsettings aus einer interdisziplinären Perspektive betrachtet werden. Verbindendes Element ist die Analyse längsschnittlicher Daten. Johanna Krull et al. (Universität Potsdam) analysieren im ersten Beitrag, inwiefern ein erhöhter Förderbedarf als Ursache von sozialer Ausgrenzung zu sehen ist oder ob er vielmehr das Resultat sozialer Ausgrenzung ist. Dazu wurden 56 erste Klassen über insgesamt zwei Schuljahre hinweg untersucht. Yvonne Blumenthal et al. (Universität Rostock) thematisieren im zweiten Beitrag intra- und interpersonale Schutzfaktoren im Zusammenhang mit sozialer Integration. Anhand der Daten von 23 Grundschulklassen untersuchen sie, ob die Integration leistungsschwacher Grundschulkindern in den Klassenverband von der Anzahl der Freundschaften und/oder dem sozial-emotionalen Entwicklungsstand eines Kindes abhängt. Der dritte Beitrag von Carmen Zurbriggen et al. (Universität Fribourg, Schweiz) betrachtet, inwiefern das Ausmaß der sozialen Integration beeinflusst, wie Grundschüler Unterrichtssituationen wahrnehmen. Hierzu wurden Grundschüler aus 40 Klassen längsschnittlich begleitet und es wurde erfasst, ob sich eine geringe soziale Eingebundenheit auf positive und negative Aktivierungen der Schüler in verschiedenen sozialen Lernkontexten auswirkt. Im vierten und letzten Beitrag von Thorsten Luka et al. (Universität Potsdam) wird anhand von Daten des Projekts „Inklusive Grundschule“ der Frage nachgegangen, wie stabil die sozialen Netzwerke von Grundschulkindern mit und ohne sonderpädagogischen Förderbedarf sind und welche individuellen Merkmale der Schüler ihre soziale Integration beeinflussen. Die Studie analysiert dabei die Angaben von Grundschulkindern aus 72 Klassen, die über zwei Schuljahre hinweg begleitet wurden.

Zusammen sollen die Studien dazu beitragen, Theorien zur Ausbildung von Freundschaften bzw. zur sozialen Ausgrenzung im Rahmen inklusiver Lernkontexte zu prüfen und über die Analyse von Prädiktoren sozialer Integration Ableitungen für die Unterrichtsgestaltung zu diskutieren.

Mittwoch
11. März

16:15 Uhr

HGB 50

Chair(s)

Spörer, Nadine

Krull, Johanna; Wilbert, Jürgen; Hennemann, Thomas

Der Zusammenhang zwischen schulischen Lern- und Verhaltensproblemen und sozialer Ausgrenzung in inklusiven Klassen

Das Bedürfnis nach sozialer Eingebundenheit gehört zu den psychologischen Grundbedürfnissen eines jeden Menschen (Krapp & Ryan, 2002). Dem folgend liegt nach Martschinke, Kopp und Ratz (2012) das angestrebte und erklärte Ziel jeden Unterrichts darin, dass alle Schülerinnen und Schüler einer Klasse sozial integriert sind und sich gleichermaßen als integriert wahrnehmen. Dennoch deuten sowohl nationale wie auch internationale Forschungsbefunde zur sozialen Partizipation von Schülerinnen und Schülern in inklusiven Klassen auf eine überwiegend ungünstige soziale Integration für Kinder mit sonderpädagogischem Förderbedarf (z.B. Bless & Mohr, 2007; Kavale & Forness, 1996; Pijl & Frostad, 2010; Huber & Wilbert, 2012). Die Mehrheit dieser Studien stützt sich dabei auf querschnittliche Untersuchungsdesigns, deren Ergebnisse keine kausale Interpretation zulassen. Somit stellt sich die Frage, ob einerseits ein erhöhter Förderbedarf als Ursache von sozialer Ausgrenzung zu sehen ist oder andererseits, ob er das Resultat sozialer Ausgrenzung ist.

Zur Beantwortung unserer Forschungsfragen wurden in den Jahren 2012 und 2013 sowohl Einzel- als auch Gruppenbefragungen bei 972 Schülerinnen und Schülern mit und ohne erhöhten Förderbedarf aus 56 ersten sowie zweiten Klassen aus 28 Grundschulen im Kreis Mettmann (NRW) durchgeführt. Dabei gaben die Kinder Auskunft über ihre subjektiv empfundene sozio-emotionalen Schulsituation, gemessen u.a. anhand der Subskala Gefühl des Angenommenseins des FEES 1-2 (Rauer & Schuck, 2004). Zusätzlich erfolgte die Erfassung der sozialen Integration auf Basis einer soziometrischen Peer-Befragung nach Moreno (1967). Die Operationalisierung des erhöhten schulischen Förderbedarfs beruhte auf den Angaben der verantwortlichen Klassenlehrkräfte, die diejenigen Kinder benennen sollten, die im Unterricht Verhaltens- und/oder Lernprobleme aufwiesen.

Die Analyse der Daten erfolgte mithilfe dreier cross-lagged-panel Designs (Lazarsfeld & Fiske, 1938) durch Strukturgleichungsmodelle. Dabei wurde dem zentralen Forschungsinteresse nachgegangen, ob und wie sehr sich bei Schülerinnen und Schülern mit schulischen Lern- und Verhaltensproblemen in der ersten Grundschulklasse anhand ihres erhöhten Förderbedarfs die soziale Ausgrenzung in der zweiten Jahrgangsstufe vorhersagen lässt und umgekehrt, wie sehr dieser erhöhte Förderbedarf ihre soziale Ausgrenzung im darauffolgenden Schuljahr prognostiziert. Das erste panel-Design überprüft einen möglichen Ursache-Wirkungs-Zusammenhang zwischen schulischen Verhaltensproblemen und sozialer Ausgrenzung, das zweite Design analysiert diesen Zusammenhang bei Kindern mit schulischen Lernproblemen und im dritten Design wird die Kausalkette von schulischen Verhaltensproblemen und dem Gefühl von der Lehrkraft angenommen zu sein, untersucht.

Für den erhöhten Förderbedarf im Bereich Verhalten in Klasse eins zeigt sich, dass dieser zu einer stärkeren Ausgrenzung in Klasse zwei führt ($\beta = 0.08$). Umgekehrt wird jedoch ebenso deutlich, dass eine soziale Ausgrenzung in der ersten Jahrgangsstufe ein erhöhtes Förderbedarfsrisiko im Folgejahr nach sich zieht ($\beta = 0.19$). Darüber hinaus erweist sich der zweite Pfad als signifikant größer als der erste Pfad. Im Hinblick auf den erhöhten Förderbedarf im Bereich Lernen lassen die Befunde diese Zusammenhänge nicht erkennen. Es findet sich weder ein Effekt des erhöhten Förderbedarfs auf ein höheres Risiko sozialer Ausgrenzung ein Jahr später ($\beta = 0.02$), noch gibt es aufgrund einer sozialen Ausgrenzung in Klasse eins, ein höheres Risiko schulische Lernprobleme über den gleichen Zeitraum zu entwickeln ($\beta = 0.03$). Die Datenauswertung für das dritte panel-Design ergab, dass sich die Zweitklässlerinnen und Zweitklässler weniger gut von Ihrer Lehrkraft angenommen fühlen, wenn sie in der ersten Klasse schulische Verhaltensprobleme aufweisen ($\beta = -0.10$). Andersherum konnte kein signifikanter Effekt festgestellt werden, der von einem schlechten Gefühl des Angenommenseins auf einen erhöhten Förderbedarf im darauffolgenden Jahr, schließen lässt ($\beta = -0.03$). Zum Abschluss des Beitrags werden die Ergebnisse im Hinblick auf ihre pädagogische Bedeutsamkeit für die inklusive Praxis diskutiert.

Mittwoch
11. März

16:15 Uhr

HGB 50

Chair(s)

Spörer, Nadine

Blumenthal, Yvonne; Marten, Katharina; Voß, Stefan; Hartke, Bodo

Ein Freund, ein guter Freund... Freundschaften in der Schulklasse als Schutzfaktor sozialer Ausgrenzung leistungsschwacher Schülerinnen und Schüler

Internationalen Studien zufolge sind schulische Minderleistungen und Auffälligkeiten in der sozial-emotionalen Entwicklung als zentrale Risikofaktoren für eine ungünstige soziale Inklusion anzusehen (u. a. Bless, 2000, Huber & Wilbert, 2012, Kavale & Forness, 1996, Newcomb, Bukowski & Pattee, 1993, Pearl & Donahue, 2008). Dennoch weisen aktuelle Untersuchungen auch auf differenzielle Effekte hin. So nehmen nicht alle leistungsschwachen Schülerinnen und Schüler sozial ungünstige Positionen innerhalb ihrer Klasse ein. Bis zu 40 % der leistungsschwachen Kinder sind dem Integrationsmodell von Coie und Dodge (1988) zufolge als durchschnittlich gut integriert oder als beliebt zu klassifizieren (Blumenthal & Marten, 2014). Es ist daher anzunehmen, dass protektive Faktoren sowohl inter- als auch intrapersonell auf das Bedingungsgefüge des sozialen Eingebundenseins von leistungsschwachen Kindern in ihren Klassen einwirken. So spielen auf interpersoneller Ebene Peerbeziehungen, insbesondere Freundschaften, eine wesentliche Rolle für gelingende soziale Inklusionsprozesse in Schulklassen. Intrapersonell kann in der angemessenen sozialen und emotionalen Entwicklung eines leistungsschwachen Kindes ein Schutzfaktor gesehen werden.

Mit der vorliegenden Studie soll der Einfluss der genannten inter- und intrapersonellen Schutzfaktoren auf die soziale Inklusion von Kindern mit schulischen Minderleistungen näher betrachtet werden. Dabei sind folgende Fragestellungen erkenntnisleitend: Beeinflusst (1) das Vorhandensein von zumindest einer Freundschaft zwischen einem leistungsschwachen Kind und einem Mitschüler oder einer Mitschülerin bzw. (2) eine altersadäquate emotional-soziale Entwicklung eines leistungsschwachen Kindes dessen soziale Stellung positiv?

Zur Beantwortung dieser Fragestellungen werden Leistungs- und Entwicklungsdaten von als leistungsschwach klassifizierten Kindern (eine Leistung entsprechend einem PR < 25 in akademischen Bereichen) aus 23 inklusiv beschulten Grundschulklassen der Insel Rügen analysiert. Die Operationalisierung der Kriteriumsvariablen basiert grundlegend auf der Methode der Soziometrie (Moreno, 1954). Durch ein soziometrisches Nominierungsverfahren erfolgt die Ermittlung der sozialen Statuskategorie (Coie & Dodge, 1988) anhand je einer Wahl- und Abwahlfrage zu den Bereichen Sitznachbar und Geburtstagsgast ohne Begrenzung der Nominierungsmöglichkeiten. Zudem wird der soziale Integrationsstatus (IST, Huber, 2006) als Maß für die soziale Integration eines Kindes berücksichtigt. Die unabhängigen Variablen (Schutzfaktoren) werden zum einen durch die Anzahl reziproker Wahlen im o. g. Nominierungsverfahren (Freundschaften) sowie durch den Strengths and Difficulties Questionnaire (Goodman, 2005) (emotionale und soziale Entwicklung des Kindes) operationalisiert.

Hinweis: Wir bitten um Verständnis dafür, dass zum aktuellen Zeitpunkt noch keine endgültigen Informationen zu den Analyseverfahren und den Ergebnissen berichtet werden können, da sich ein Teil der Daten derzeit in der Eingabephase befindet.

Mittwoch
11. März

16:15 Uhr

HGB 50

Chair(s)

Spörer, Nadine

Zurbriggen, Carmen; Venetz, Martin

Subjektiv erlebte soziale Partizipation im Unterrichtsalltag inklusiver Schulklassen

Soziale Partizipation wird definitionsgemäß als eine der Kernfragen in der Inklusionsdebatte betrachtet (z.B. Bossaert, Colpin, Pijl & Petry, 2013). Soziale Teilhabe bedeutet dabei mehr als die physische Anwesenheit von Kindern oder Jugendlichen mit sonderpädagogischem Förderbedarf in Regelklassen, sondern beinhaltet neben der Akzeptanz der Klassenkolleginnen und –kollegen sowie gegenseitige Freundschaften auch positive soziale Interaktionen im Rahmen des Schulunterrichts (Koster, Nakken, Pijl & van Houten, 2009).

Forschungsüberblicke zur sozialen Dimension von Inklusion berichten mehrheitlich von Schwierigkeiten der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, sich in der Klassengemeinschaft einzufügen respektive von den Mitschülerinnen und Mitschülern akzeptiert zu werden (Bless & Mohr, 2007; Ruijs & Peetsma, 2009). Soziale Ausgrenzung oder Randpositionen waren dabei vorrangig mit tiefen schulischen Leistungen verknüpft. Ergebnisse aus einzelnen Studien verweisen zudem darauf, dass Verhaltensprobleme einen weiteren Risikofaktor darstellen (z.B. Mand, 2007). Die Frage nach der sozialen Eingebundenheit wurde bislang grösstenteils mittels soziometrischer Befragung der Peers untersucht. Studien, in denen die Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf diesbezüglich nach deren subjektiven Einschätzung gefragt wurden, lassen sich im deutschen Sprachraum hingegen nur wenige ausmachen (Haeberlin, Bless, Moser & Klaghofer, 1990; Huber & Wilbert, 2012). Unbeachtet blieb jeweils, ob und wie sich eine (subjektiv erlebte) geringe soziale Eingebundenheit auf das Erleben in verschiedenen sozialen Lernkontexten des konkreten Unterrichts niederschlägt.

Im Fokus dieses Beitrags steht daher die Frage, ob Schülerinnen und Schüler, die sich als vergleichsweise schlecht in die Klasse integriert beschreiben, unterschiedliche Sozialformen im Unterricht anders als ihre Mitschülerinnen und Mitschüler erleben. Vorgängig soll die Frage geklärt werden, wie sich Schülerinnen und Schüler charakterisieren lassen, die sich als sozial vergleichsweise schlecht im Klassenverband integriert einschätzen.

Zur Klärung der Fragen haben 697 Schülerinnen und Schüler aus 40 Regelklassen der sechsten Primarstufe zum einen Auskunft über ihre soziale Eingebundenheit in der Klasse gegeben und zum anderen ihr unterrichtsbezogenes Erleben im Verlaufe einer Schulwoche zu je 14 Zeitpunkten mittels der Experience Sampling Method (ESM; Hektner, Schmidt & Csikszentmihalyi, 2007) protokolliert. Die Erfassung der sozialen Integration erfolgte anhand der gleichnamigen Skala des KFDI (Venetz, Zurbriggen & Eckhart, 2014). Vorliegend waren des Weiteren die Resultate aus standardisierten Schulleistungstests in Mathematik und Deutsch sowie die Einschätzung der Lehrperson zu Stärken und Schwächen der Verhaltenskompetenzen der Schülerin oder des Schülers (SDQ; Goodman, 2001). Die Operationalisierung der Erlebensqualität basiert auf dem Modell der positiven und negativen Aktivierung (Yik, Russel & Feldmann Baret, 1999). Zusätzlich wurden zu allen 14 Zeitpunkten die aktuelle Konzentration sowie die Freude an der Tätigkeit erhoben.

Für die Analysen zur ersten Fragestellung wurden ordinale Regressionsmodelle spezifiziert. Die Ergebnisse zeigen, dass die subjektiv eingeschätzte soziale Integration einerseits mit den Deutschleistungen und andererseits mit einem Förderbedarf im Bereich Verhalten verknüpft ist. Vertiefende Analysen ergeben, dass Deutschleistung und sprachliche Herkunft konfundiert sind, und dass eine vergleichsweise tiefe soziale Integration primär mit Verhaltensproblemen im Umgang mit Peers in Verbindung steht. Die Auswertungen zur zweiten Fragestellung erfolgten anhand von Multilevelmodellen. Hierzu kann als erstes festgehalten werden, dass der größte Anteil der Varianz im Erleben auf situative Unterrichtsgegebenheiten zurückgeht (ca. 60-75%), rund 23-37% liegen auf Schülerebene und 3-5% auf Klassenebene. Schülerinnen und Schüler, deren Einschätzungen zur sozialen Integration im untersten Quartil befinden, sind im Unterricht generell weniger positiv aktiviert (z.B. motivierter) und stärker negativ aktiviert (z. B. gestresster) als ihre Peers. Zudem erleben sie weniger Freude an den Tätigkeiten und sind weniger konzentriert. Unterrichtsaktivitäten, denen gemeinsam mit anderen nachgegangen wird, werden im Allgemeinen positiver erlebt – wobei sich zwischen den Gruppen jeweils Unterschiede in der Ausprägung, jedoch keine Interaktionseffekte zeigen. Zum Abschluss des Beitrages werden die Befunde im Kontext des aktuellen Forschungsstandes zur sozialen Partizipation sowie Implikationen für die inklusive Praxis diskutiert.

Mittwoch
11. März

16:15 Uhr

HGB 50

Chair(s)

Spörer, Nadine

3

Henke, Thorsten; Jäntsich, Christian; Lambrecht, Jennifer; Bosse, Stefanie; Spörer, Nadine

Wollen wir Freunde sein? Eine längsschnittliche Analyse der Freundschaftsdynamiken in inklusiv lernenden Klassen

Der Aufbau freundschaftlicher Beziehungen zu Gleichaltrigen ist ein essentieller Bestandteil der sozial-kognitiven Entwicklung von Kindern und Jugendlichen (Wentzel, 2005). Insbesondere enge Freundschaften wirken protektiv gegen Ausgrenzung und haben eine unterstützende Wirkung auf ängstliche Kinder (Buhs, Ladd, & Herald, 2006; Rubin, Wojslawowicz, Rose-Krasnor, Booth-LaForce, & Burgess, 2006).

Ein wichtiges Ziel einer inklusiven Schule ist die Schaffung eines positiven Sozialklimas und der Einbezug aller Kinder in die Klassengemeinschaft (Prengel, 2013). Insbesondere Kinder mit einem Sonderpädagogischen Förderbedarf (SPF) haben jedoch einen geringeren soziometrischen Status und schätzen das Sozialklima der Klasse schlechter ein (Huber & Wilbert, 2012; Krull, Wilbert, & Hennemann, 2014). Zwar konnten Rubin et al. (2006) zeigen, dass auch diese Kinder nicht gänzlich von der Klassengemeinschaft isoliert sind und überwiegend mindestens einen reziproken Freund haben. Diese Freundschaften wurden jedoch als weniger eng und vertraut beschrieben. Darüber hinaus erfolgt die Bildung von Freundschaften homophil (McPherson, Smith-Lovin, & Cook, 2001), sodass abgelehnte Kinder Freundschaften eher unter ebenfalls abgelehnten Kindern suchen (Frostad & Pijl, 2007), wodurch sich ihre Möglichkeiten zur Teilhabe am Klassengeschehen nur unwesentlich verbessern.

Bislang bleibt jedoch ungeklärt, ob ein kausales Verhältnis des SPF auf eine geringere Peerakzeptanz vorliegt (Huber & Wilbert, 2012). Bezüglich des Zusammenhangs von akademischer Leistung und der Akzeptanz durch die Peers legen Ladd, Buhs und Seid (2000) anhand längsschnittlicher Analysen den Schluss nahe, dass eine seltenere Nennung als Freund geringere akademische Leistungen zu einem späteren Zeitpunkt evoziert; der umgekehrte Fall gilt jedoch nicht. Nimmt man weiterhin zur Kenntnis, dass Schüler mit einem SPF häufiger Verhaltensprobleme aufweisen (Koster, Nakken, Pijl, & van Houten, 2009), seltener Schüler mit ähnlichen Merkmalen in ihrer Klasse finden (Frostad & Pijl, 2007) und ihre Freundschaften weniger stabil sind (Rubin et al., 2006), dann lässt sich vermuten, dass ein SPF einen negativen Effekt auf den Aufbau von Peerbeziehungen hat. Es werden somit folgende Hypothesen formuliert: (1) Ein diagnostizierter SPF hat nach Kontrolle anderer Einflussfaktoren (Leistungs- und Geschlechterdifferenzen) einen negativen Effekt auf die Wahl als Freund durch eine Person ohne einen solchen Förderbedarf. (2) Die Bereitschaft ein Freundschaftsangebot nicht zu erwidern oder ein bestehendes reziprokes Verhältnis aufzulösen wird nach Kontrolle anderer Einflussfaktoren (Leistungs- und Geschlechterdifferenzen) durch den SPF des Gegenübers beeinflusst.

Grundlage der Studie bilden die Untersuchungen zu Schüler-, Lehrer- und Unterrichtsmerkmalen im Rahmen des Pilotprojekts „Inklusive Grundschule“ des Landes Brandenburg, das seit dem Schuljahr 2012/13 von der Universität Potsdam wissenschaftlich begleitet wird. Ziel der als Längsschnittstudie angelegten Begleitforschung ist die Identifizierung von Gelingenbedingungen für gemeinsamen Unterricht von Kindern mit und ohne SPF. Die untersuchte Stichprobe umfasst N = 1465 Schülerinnen (50%) und Schüler (50%) in 72 Klassen (53% 2. Klasse; 47% 3. Klasse). Hiervon haben 85 (6%) Kinder einen SPF. Dieser wurde separat für die Bereiche Lernen, Emotionale & soziale Entwicklung, Körperliche und motorische Entwicklung, Hören, Sehen sowie Geistige Entwicklung von den Klassenlehrern erfragt. Eine Rückmeldung erfolgte für 52 Klassen, im Mittel gab es in jeder Klasse ein Kind mit SPF ($M = 1.31$, $SD = 1.62$). Die Freundschaftsbeziehungen innerhalb einer Klasse wurden mithilfe von Soziogrammen (Frederickson & Furnham, 2001) an zwei Messzeitpunkten am Ende des Schuljahres 2012/13 sowie 2013/14 erhoben. Die Anzahl anzugebender Freunde unterlag keinerlei Restriktionen. Es war somit möglich vollständige Netzwerke für jede Klasse zu extrahieren.

Um die längsschnittliche Entwicklung der multiplen Netzwerke und die durch die Netzwerkstruktur aufgeworfenen Interdependenzen der Schülerinnen und Schüler adäquat kontrollieren zu können, wird ein agentenbasiertes Modell für die Analyse eingesetzt (Snijders, Van de Bunt, & Steglich, 2010). Fehlende Netzwerkdaten werden nach den Vorschlägen von Huisman (Huisman, 2009; Huisman & Snijders, 2003; Huisman & Steglich, 2008) imputiert. Es wird die Stabilität bzw. Variabilität der Netzwerke berichtet und vor dem Hintergrund der beeinflussenden Merkmale diskutiert.

Mittwoch
11. März

16:15 Uhr

HGB 50

Chair(s)

Spörer, Nadine

Steinmayr, Ricarda

Selbst- und Fremdbeurteilung schulischer Leistung und Fähigkeiten: Einflussfaktoren und Effekte

Dieses Symposium vereint aktuelle Forschungsarbeiten aus der Psychologie und Erziehungswissenschaft, deren Ergebnisse über diese Disziplinen hinaus auch für die Lehrerbildung und die Fachdidaktiken Deutsch und Mathematik Relevanz besitzen. Zwei der Beiträge behandeln die unterschiedlichen Bildungserfolge von Jungen und Mädchen, wobei schulischer Erfolg insgesamt sowie domänenspezifische Selbstkonzepte und Leistungen fokussiert werden. Übergreifendes Ziel des Symposiums ist es, aus verschiedenen theoretischen Perspektiven und mit unterschiedlichen methodischen Designs neue Erkenntnisse zu gewinnen, wie die Selbst- und Fremdbewertungen gezeigter schulischer Leistungen zustande kommen und sich nachfolgend auswirken.

Inwiefern Schüler/innen durch Störverhalten gezielt die Attributionen für schulischen Misserfolg beeinflussen können, untersuchen Heyder und Kessels in einer experimentellen Studie. Aus den attributionstheoretischen Arbeiten Heckhausens (1972), Weiners (1994) und Butlers (1994) leiten sie die These ab, dass das bewusste Evozieren eines Tadels durch die Lehrkraft eine Attribution schlechter Leistungen auf fehlende Anstrengung statt auf fehlende Begabung wahrscheinlicher macht sowie das Ansehen in der Peergruppe und die zugeschriebene Maskulinität befördert. Somit wäre bei schulischem Misserfolg Störverhalten zur Aufrechterhaltung eines positiven Fähigkeitsselbstkonzeptes funktional, und dieses vor allem bei Jungen.

Weidinger, Spinath und Steinmayr untersuchen in einer längsschnittlichen Studie, wie im Fach Mathematik das Fähigkeitsselbstkonzept und die intrinsische Motivation vom Notenniveau beeinflusst wird. Aufbauend auf den Theorien von Marsh und der Selbstbestimmungstheorie von Deci und Ryan (1985) nehmen sie an, dass sich in Abhängigkeit vom Notenniveau interindividuelle Unterschiede in der intraindividuellen Entwicklung des Fähigkeitsselbstkonzepts und der intrinsischen Motivation nachweisen lassen.

Hannover und Wolter untersuchen in ihrer längsschnittlichen Studie, inwiefern die Selbstzuschreibung geschlechtstypisierter Merkmale zu einer Erklärung von Geschlechtsunterschieden in der Entwicklung von Leistungen und des akademischen Fähigkeitsselbstkonzepts in der Grundschule beitragen kann. Vor dem Hintergrund kognitiver Entwicklungstheorien zur Aneignung von Geschlechterstereotypen in das eigene Selbstkonzept gehen die Autorinnen davon aus, dass in dem Maße, wie Kinder sich selbst zu Schulbeginn geschlechtstypisiert beschreiben, sie stärker geschlechtstypisierte domänenspezifische Fähigkeitsselbstkonzepte und Leistungsprofile entwickeln.

Im letzten Beitrag befassen sich Meißner, McElvany und Steinmayr mit den Auswirkungen der Wahl des Leistungsindikators (Note versus standardisierter Test) im Bereich der Lesekompetenz auf die prädiktive Validität des Fähigkeitsselbstkonzeptes im Vergleich zur Intelligenz. Basierend auf vorliegenden Arbeiten zur relativen Bedeutsamkeit kognitiver und motivationaler Prädiktoren zur Vorhersage von Leistungen (Steinmayr & Meißner, 2013) untersuchen die Autorinnen mithilfe von latenten Strukturgleichungsmodellen und Kommunalitätenanalysen, inwiefern sich die spezifischen, aber auch gemeinsamen Anteile der beiden Prädiktoren (Fähigkeitsselbstkonzept und Intelligenz) bei der Varianzaufklärung von Lesenoten und Lesekompetenztests unterscheiden.

Malte Schwinger diskutiert abschließend alle Beiträge aus pädagogisch-psychologischer Perspektive.

Mittwoch
11. März

16:15 Uhr

HGC 20

Chair(s)

Steinmayr, Ricarda
Kessels, Ursula**Selbst- und Fremdbeurteilung schulischer Leistung und Fähigkeiten:
Einflussfaktoren und Effekte**

Heyder, Anke; Kessels, Ursula

Faul statt dumm? Wie Störverhalten selbstwertdienliche Attributionen hervorruft

Theoretischer Hintergrund: Eine selbstwertdienliche Attribution bei Misserfolg ist das Zurückführen der schlechten Leistung auf fehlende Anstrengung, wohingegen es den Selbstwert beschädigt, wenn fehlende Begabung als Grund gesehen wird (z.B. Heckhausen, 1972; Skaalvik, 1994; Weiner, 1994). Auch die Reaktionen anderer Personen auf einen Misserfolg unterscheiden sich je nach zugeschriebener Ursache (Weiner, 1994). Ab dem Jugendalter sind Schüler/innen in der Lage, aus der Art der Reaktion der Lehrkraft abzulesen, ob eine schlechte Leistung auf fehlende Anstrengung oder aber auf fehlende Begabung attribuiert wird: Tadel und Ermahnungen seitens der Lehrkraft erfolgen nur, wenn diese die schlechte Leistung auf Faulheit oder Nachlässigkeit zurückführt, nicht aber, wenn sie annimmt, sie sei der fehlenden Begabung geschuldet (Butler, 1994). Unsere Studie basiert auf der Annahme, dass Jugendliche ihr Wissen um diesen Zusammenhang gezielt nutzen, um im Falle schlechter Schulleistungen selbstwertdienliche Attributionen auf fehlende Anstrengung (statt Begabung) nahezulegen. Wir nehmen an, dass ein Grund für Störverhalten in der Schule darin liegt, dass die dadurch hervorgerufenen Ermahnungen der Lehrkraft fehlende Anstrengung (statt fehlende Begabung) des/der Störer/in signalisieren. Dadurch würden sie sowohl den eigenen Selbstwert als auch ihr Ansehen unter Gleichaltrigen schützen.

Fragestellung: Unsere Studie untersucht, ob Störverhalten die Fremdeinschätzung von schulischen Fähigkeiten und sozialem Ansehen beeinflusst. Wir prüfen, ob Jugendliche schlechte Leistungen von Schülerinnen und Schülern, die bewusst den Unterricht stören und dadurch Ermahnungen provozieren, stärker auf mangelnde Anstrengung (und nicht auf fehlende Begabung) attribuieren als die genauso schlechten Leistungen von Schülerinnen und Schülern, die sich unauffällig verhalten. Diese Attribution sollte ihre Beliebtheit in der Klasse fördern. Jungen sollten von diesem Mechanismus stärker profitieren, weil Mädchen Störverhalten seltener zeigen (z.B. Trautwein, Köller & Baumert, 2004) und dieses bei Mädchen als nicht geschlechtsrollenkonform (und damit als deviant) gelten würde.

Methode: In einem 2 (between: Geschlecht Targetperson) x 2 (between: Geschlecht Versuchsperson) x 2 (within: Verhalten: stören versus unauffällig)–faktoriellen Vignettenexperiment wurde die Wirkung von schulischem Störverhalten auf die Fremdeinschätzung schulischer Fähigkeiten bei schlechten Leistungen untersucht. Jugendliche der 9. Jahrgangsstufe zweier Gymnasien bearbeiteten einen Fragebogen, in dem zwei verschiedene fiktive Personen (entweder zwei Mädchen oder zwei Jungen) beschrieben wurden. Beide Personen bekamen häufig schlechte Noten, aber eine/r der beiden verhielt sich unauffällig und der/die andere störte häufig den Unterricht. Nachdem die Jugendlichen sich die Personen kurz vorgestellt hatten, gaben sie an, welche Reaktion der Lehrkraft sie gegenüber der Person erwarten und warum die Person nicht so gute Noten bekommt. Weitere abhängige Variablen waren die zugeschriebene Beliebtheit, Männlichkeit, Weiblichkeit und Ähnlichkeit mit einem typischen Jungen/Mädchen. Die Verhaltensweisen zur Beschreibung der/s auffälligen Schüler/in waren an einer studentischen Stichprobe vorgetestet worden, dass sie eindeutig Störverhalten im Unterricht darstellen.

Ergebnisse: Die Jugendlichen erwarteten, dass die störenden Schülerinnen und Schüler häufiger von der Lehrkraft ermahnt würden und mehr Ärger provozieren würden. In Übereinstimmung mit unseren Hypothesen sahen sie die Ursache der schlechten Leistungen der störenden Schülerinnen und Schüler stärker in einem Mangel an Anstrengung als in einem Mangel an Begabung und schrieben ihnen eine höhere Beliebtheit zu. Störende Schülerinnen und Schüler wurden außerdem als männlicher und weniger weiblich wahrgenommen. Zudem ähnelte der Junge, der stört, aus Perspektive der Jugendlichen einem typischen Jungen am meisten, während das störende Mädchen die geringste Ähnlichkeit mit einem typischen Mädchen aufwies.

Zusammenfassend deuten unsere Ergebnisse darauf hin, dass Jugendliche mit schlechten Leistungen wohlmöglich auch deshalb Störverhalten zeigen, weil ihre Mitschüler/innen ihre schlechten Leistungen dann mit mangelnder Anstrengung erklären. So kann Störverhalten zur Aufrechterhaltung eines positiven Fähigkeitsselbstkonzeptes trotz schulischer Misserfolge beitragen. Die Ergebnisse werden vor dem Hintergrund diskutiert, inwiefern Störverhalten besonders für Jungen als funktional gelten kann.

Mittwoch
11. März

16:15 Uhr

HGC 20

Chair(s)

Steinmayr, Ricarda
Kessels, UrsulaSelbst- und Fremdbeurteilung schulischer Leistung und Fähigkeiten:
Einflussfaktoren und Effekte

Weidinger, Anne; Spinath, Birgit; Steinmayr, Ricarda

Zur Bedeutung von Mathematiknoten für die Veränderung des domänenspezifischen Fähigkeitsselbstkonzepts und der Intrinsischen Motivation

Das Fähigkeitsselbstkonzept (FSK) und die Intrinsische Motivation (IM) gelten als wichtige Determinanten des Lern- und Leistungsverhaltens im Bildungskontext (z.B. Wigfield, Eccles, Schiefele, Roeser & Davis-Kean, 2006). Zu Beginn der Grundschulzeit sind beide Konstrukte zumeist sehr positiv ausgeprägt, sie fallen jedoch im Mittel über die Grundschulzeit hinweg ab (z.B. Gottfried, Fleming & Gottfried, 2001; Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002). Dieser Abfall wird unter anderem als Folge von Leistungsrückmeldungen angesehen, mit denen Kinder in der Grundschule zunehmend konfrontiert werden (z.B. Eccles & Wigfield, 2002; Wigfield et al., 2006). Es wird angenommen, dass vor allem Schulnoten in diesem Zusammenhang von Bedeutung sind, da sie einen unmittelbaren normativen Leistungsvergleich ermöglichen. In einigen Studien konnte tatsächlich nachgewiesen werden, dass Noten und das FSK (z.B. Helmke & van Aken, 1995; Niepel, Brunner & Preckel, 2014) bzw. die IM (z.B. Corpus, McClintic-Gilbert & Hayenga, 2009; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005) über die Zeit hinweg miteinander zusammenhängen. In diesen Studien wurde der längsschnittliche Zusammenhang auf interindividueller Ebene betrachtet, wobei Unterschiede in der intraindividuellen Entwicklung der Konstrukte unberücksichtigt blieben. Deshalb untersuchten wir die Fragestellung, ob sich das FSK und die IM bei allen oder nur bei bestimmten Kindern verändern und welche Bedeutung hierbei den Noten zukommt. Bezüglich des FSKs der Kinder ist zu erwarten, dass es sich in Abhängigkeit der Noten unterschiedlich verändert. Bei der Entwicklung einer realistischen Fähigkeitsselbstwahrnehmung ziehen Schulanfänger/innen vor allem Leistungsrückmeldungen als eine wichtige Informationsquelle heran (vgl. Skaalvik & Skaalvik, 2002). Kinder, die schlechte Noten erhalten, sollten ihr FSK somit stärker nach unten korrigieren als Kinder mit guten Noten. Im Hinblick auf die Veränderung der IM lassen sich vor dem Hintergrund der Selbstbestimmungstheorie (Deci & Ryan, 1985) zwei theoretisch plausible Vorhersagen ableiten: Einerseits ist denkbar, dass die IM bei allen Kindern absinkt, da aufgrund der veränderten schulischen Umwelt (i.e. Einführung von Ziffernnoten) ihre Selbstbestimmtheit eingeschränkt wird („Autonomieannahme“). Andererseits ist plausible, dass die IM von Kindern, die schlechte Noten erhalten, stärker absinkt, da deren Kompetenzbedürfnis unbefriedigt bleibt („Kompetenzannahme“).

Die Daten zur Untersuchung der Fragestellung stammen aus einer Längsschnittstudie mit $N = 542$ Grundschulkindern ($t1$: $M = 7.95$ Jahre, $SD = 0.58$). Das FSK und die IM für Mathematik (Selbstberichtfragebögen) sowie die Mathematiknote (Angabe der Lehrkraft) wurden zu sieben Messezeitpunkten vom Ende der 2. bis zum Ende der 4. Klasse erhoben. Die Daten wurden mit Hilfe latenter Wachstumskurvenmodelle (LGCM) ausgewertet, da diese ermöglichen, die intraindividuellen Veränderungen in der absoluten Ausprägung von Merkmalen, die interindividuellen Unterschiede in diesen Veränderungen sowie die Wechselbeziehungen der Veränderungen zweier Merkmale zu untersuchen (Duncan & Duncan, 2004).

Univariate LGCM zeigten, dass alle drei Variablen im Mittel über die Zeit abfielen, es jedoch signifikante interindividuelle Unterschiede in der intraindividuellen Veränderung gab. Das bivariate LGCM, in dem die Veränderung der Note im Zusammenhang mit der Veränderung des FSKs bzw. der IM betrachtet wurde, zeigte einen signifikanten Zusammenhang zwischen den Veränderungskurven der Note und den Veränderungskurven der jeweiligen motivationalen Variablen. Darüber hinaus beobachteten wir einen moderaten bzw. geringen signifikanten Zusammenhang zwischen der Ausgangsnote und der Veränderung des FSKs bzw. der IM über die Zeit. Je schlechter die Mathematiknote zu $t1$ war, umso stärker veränderten sich die motivationalen Variablen über den Untersuchungszeitraum hinweg. Die Zusammenhänge waren für die IM signifikant geringer als für das FSK.

Unsere Ergebnisse stützen die Annahme, dass Kinder u.a. Noten als Informationsquelle zur Generierung realistischer Fähigkeitsselbstwahrnehmungen nutzen (z.B. Spinath & Spinath, 2005). Im Hinblick auf die Veränderung der IM in der Grundschule scheinen Noten eher von geringerer Bedeutung zu sein, was dafür spricht, dass die Rückmeldung niedriger Kompetenz (z.B. durch schlechte Noten) nicht zwangsläufig damit einhergehen muss, dass man den Spaß an Aufgaben verliert (vgl. Spinath & Steinmayr, 2008, 2012).

Mittwoch
11. März

16:15 Uhr

HGC 20

Chair(s)

Steinmayr, Ricarda
Kessels, Ursula

Selbst- und Fremdbeurteilung schulischer Leistung und Fähigkeiten:
Einflussfaktoren und Effekte

Hannover, Bettina; Wolter, Ilka

Selbstzuschreibung geschlechtstypisierter Merkmale im Schulanfangsalter: Ihr Einfluss auf Fähigkeitsselbstkonzepte und Leistungen in geschlechtskonnotierten Domänen

Mit der Aneignung von Geschlechtsstereotypen und der Herausbildung der Geschlechtsidentität des Kindes ist schon im Vorschulalter die Grundlage dafür geschaffen, dass es auf die Geschlechter bezogenes Wissen in sein Selbst integriert, d.h., sich selbst geschlechtstypisierte Merkmale zuschreibt. So charakterisieren sich Jungen und Mädchen bereits im Schuleintrittsalter auf geschlechtstypisierte Weise; z.B. sagen Jungen über sich selbst eher, dass sie "sich beim Spielen schmutzig machen" oder "beim Spielen kämpfen", Mädchen hingegen, dass sie "sich gerne hübsch anziehen" oder "Rollenspiele spielen, wie Vater-Mutter-Kind" (Wolter, 2011). In unserem Vortrag gehen wir der Frage nach, inwiefern diese Selbstzuschreibung geschlechtstypisierter Merkmale zu einer Erklärung von Geschlechtsunterschieden in der Entwicklung von Leistungen und des akademischen Fähigkeitsselbstkonzepts in der Grundschule beitragen kann.

Geschlechtsunterschiede in Fähigkeitsselbstkonzepten und Leistungen lassen sich teilweise bereits in der Grundschule nachweisen (z.B. Mielke, Goy & Pietsch, 2006; Niklas & Schneider, 2012); in anderen Studien wurden sie für diese Altersphase noch nicht gefunden (z.B. IGLU-E Studie, z.B. Prenzel, Geiser, Langeheine & Lobemeier, 2003). Deutlich sind sie dann zum Ende der Grundschulzeit (z.B. Herbert & Stipek, 2005; Muzzatti & Agnoli, 2007) und in der Sekundarstufe I nachweisbar (z.B. Artelt, Baumert, Klieme et al., 2001; Klieme, Artelt, Hartig et al., 2010).

Wir haben angenommen, dass in dem Maße, wie Kinder sich selbst zu Schulbeginn geschlechtstypisiert beschreiben, sie stärker geschlechtstypische Fähigkeitsselbstkonzepte und domänenspezifische Leistungsprofile entwickeln. Für die Prüfung dieser Hypothese haben wir in einem Interview bei 135 Kindern zu Schulbeginn geschlechtstypisierte Selbstbeschreibungen erhoben und geprüft, inwiefern diese einen Einfluss auf ihre Fähigkeitsselbstkonzepte und Leistungen in den männlich bzw. weiblich konnotierten Domänen Mathematik und Lesen am Ende der ersten und der zweiten Klasse haben.

Zur Messung der geschlechtstypisierten Selbstbeschreibungen entwickelten wir ein Instrument, das verhaltensnahe feminine und maskuline Selbstbeschreibungen vorgibt (Wolter, 2011). Die Fähigkeitsselbstkonzepte wurden mit einer adaptierten Form des Self Description Questionnaire (SDQP, Marsh, Ellis & Craven, 2002) erfasst. Leistungen im Lesen wurden mit der Würzburger Leise Leseprobe (WLLP; Küspert & Schneider, 1998) und in Mathematik mit dem Heidelberger Rechentest (HRT 1-4; Haffner, Baro, Parzer & Resch, 2005) erhoben.

Die Ergebnisse zeigen, dass die Selbstbeschreibungen der Kinder bereits zu Schuleintritt auf geschlechtstypisierte Weise ausgeprägt sind: Mädchen beschrieben sich mit mehr femininen und Jungen beschrieben sich mit mehr maskulinen Attributen als Mädchen. In Pfadanalysen zeigte sich, dass eine stark feminine Selbstbeschreibung mit einem höheren Fähigkeitsselbstkonzept im Lesen am Ende der ersten und zweiten Klasse zusammenhängt. Umgekehrt war eine stark maskuline Selbstbeschreibung mit einem höheren Fähigkeitsselbstkonzept in der Mathematik ein halbes Jahr später am Ende der ersten Klasse verbunden. Weiterhin zeigte sich, dass das Fähigkeitsselbstkonzept im Lesen mit der Leseleistung zusammenhängt. Wir konnten keine Zusammenhänge zwischen dem Fähigkeitsselbstkonzept in der Mathematik und den entsprechenden Leistungen in der ersten Klasse aufzeigen. Leistungen und Fähigkeitsselbstkonzept im Lesen fallen ein Jahr später am Ende der zweiten Klasse stabiler aus als in der männlich-konnotierten Domäne Mathematik.

Es scheint, in dem Maße, wie Kinder die Geschlechtskonnotation von Schulfächern kennen lernen, werden auch ihre Fähigkeitsselbstkonzepte von dem Bestreben geprägt, sich selbst in Übereinstimmung mit derjenigen sozialen Geschlechtsrolle zu beschreiben, die zum eigenen biologischen Geschlecht "passt". Dass in der Mathematik keine Zusammenhänge zu Leistungen nachgewiesen werden konnten, kann möglicherweise damit erklärt werden, dass kausale Effekte von Fähigkeitsselbstkonzepten auf die Leistung erst über einen längeren Zeitraum hinweg zur Wirkung kommen.

Mittwoch
11. März

16:15 Uhr

HGC 20

Chair(s)

Steinmayr, Ricarda
Kessels, Ursula

Selbst- und Fremdbeurteilung schulischer Leistung und Fähigkeiten:
Einflussfaktoren und Effekte

Meißner, Anja; McElvany, Nele; Steinmayr, Ricarda

Spielt der Leistungsindikator eine Rolle? Zur relativen Bedeutung der Intelligenz und des Fähigkeitsselbstkonzeptes bei der Vorhersage der Lesekompetenz

Das Verstehen und die Erklärung von schulischen Leistungsunterschieden zwischen Schülerinnen und Schülern stellt ein zentrales Forschungsanliegen dar, dem interdisziplinär Rechnung getragen wird. Zur Erklärung interindividueller Unterschiede in schulischen Leistungen werden neben bewährten kognitiven Konstrukten wie der Intelligenz zunehmend motivationale Faktoren untersucht. Studien, die kognitive und motivationale Prädiktoren berücksichtigten, zeigen, dass motivationale Variablen sowohl unabhängig von kognitiven Variablen als auch gemeinsam mit kognitiven Variablen einen Teil der Unterschiede in der Schulleistung vorhersagen konnten (z.B. Spinath, Spinath, Harlaar & Plomin, 2006; Steinmayr & Spinath, 2009). Die Anteile der allein durch einen Prädiktor oder durch beide Prädiktoren gemeinsam aufgeklärten Varianz in der Schulleistung variieren jedoch erheblich zwischen den Studien. Diese heterogene Befundlage könnte darauf zurückzuführen sein, dass in den Studien unterschiedliche Schulleistungsindikatoren (z.B. standardisierte Tests, Schulnoten) verwendet wurden. Um Schlussfolgerungen über die relative Bedeutsamkeit kognitiver und motivationaler Prädiktoren für verschiedene Leistungsindikatoren ziehen zu können, bedarf es Studien, die sowohl kognitive und motivationale Prädiktoren als auch verschiedene schulische Leistungsindikatoren an einer Stichprobe untersuchen. Bislang existieren nur sehr wenige Studien, die dies realisiert haben (z.B. Helmke, 1992; Sauer & Gamsjäger, 1996; Steinmayr & Meißner, 2013). Für die Domäne Mathematik konnte gezeigt werden, dass die Intelligenz der beste Prädiktor zur Vorhersage der Leistung in einem standardisierten Test ist, während die Varianz in der Note vergleichbar hoch durch Intelligenz und Fähigkeitsselbstkonzept erklärt wurde (Helmke, 1992; Steinmayr & Meißner, 2013). Die bestehende Forschung begrenzt sich vorwiegend auf die Domäne Mathematik (Helmke, 1992; Steinmayr & Meißner, 2013), sodass eine Replizierbarkeit der Ergebnisse für andere Domänen zu überprüfen ist. Aus diesem Grund soll in der vorliegenden Studie für die Domäne Lesen untersucht werden, ob sich die relative Bedeutsamkeit kognitiver und motivationaler Prädiktoren in Abhängigkeit vom Schulleistungskriterium bei Grundschulkindern unterscheidet.

Zur Beantwortung der Fragestellung wurden die Daten von 458 Schülerinnen und Schülern ($M = 12.0$ Jahre, $SD = 0.6$, 215 Mädchen) der 6. Klasse aus 22 Berliner Grundschulen herangezogen. Erfasst wurden die Leistung in einem standardisierten Lesetest (DTD; Nauck & Otte, 1980) und die Lesenote (Angabe der Lehrkraft) als Kriterien sowie die allgemeine Intelligenz (KFT 4–12+ R; Heller & Perleth, 2000) und das Fähigkeitsselbstkonzept im Lesen (McElvany, 2008) als Prädiktoren. Es wurden latente Strukturgleichungsmodelle und Kommunalitätenanalysen (s. Pedhazur, 1997) berechnet. Die Analysen erfolgten getrennt für Testleistung und Note.

Im Rahmen der latenten multiplen Regressionsanalysen konnten die allgemeine Intelligenz und das Fähigkeitsselbstkonzept im Lesen einen bedeutsamen eigenständigen Beitrag sowohl zur Vorhersage der Leistung im Lesetest als auch zur Lesenote leisten. Bei der Vorhersage der Lesetestleistung war der spezifische Beitrag der Intelligenz bedeutsam größer als der Beitrag des Fähigkeitsselbstkonzeptes, während das Fähigkeitsselbstkonzept einen größeren spezifischen Beitrag als die Intelligenz zur Vorhersage der Lesenote leistete. Ein bedeutsamer Anteil der Varianz im Lesetest und der Note wurde durch beide Prädiktoren gemeinsam erklärt.

Die Ergebnisse sprechen dafür, dass sich die relative Bedeutsamkeit kognitiver und motivationaler Prädiktoren in Abhängigkeit vom untersuchten Leistungsmaß der Lesekompetenz unterscheidet. Die Vorhersagbarkeit schulischer Leistungen hängt demzufolge sowohl vom untersuchten Prädiktor als auch der Art des Kriteriums ab. Die Ergebnisse werden in Hinblick auf die Bedeutsamkeit unterschiedlicher Kriterien bei der Vorhersage von Schulleistung diskutiert.

Mittwoch
11. März

16:15 Uhr

HGC 20

Chair(s)

Steinmayr, Ricarda
Kessels, Ursula

Selbst- und Fremdbeurteilung schulischer Leistung und Fähigkeiten:
Einflussfaktoren und Effekte

Lipowsky, Frank

Qualität und Wirkung pädagogischen Handelns: Zum 60. Geburtstag von Eckhard Klieme

Eckhard Klieme hat mit seinen Forschungsprojekten und seinen konzeptionellen Beiträgen z.B. zu den Bildungsstandards die empirische Bildungsforschung, aber auch die Praxis des Bildungssystems entscheidend beeinflusst. Darüber hinaus nimmt er bis heute zahlreiche Funktionen in Gremien der Bildungsforschung wahr. So gehört er zu den Gründungsmitgliedern der GEBF und ist langjähriges Mitglied der AEPF, deren Vorsitzender er von 2003 bis 2006 war. In vielfältigen internationalen und nationalen Funktionen wirkt Eckhard Klieme seit 2000 an der Durchführung der PISA-Studien mit. Forschungsprojekte wie die „Studie zur Entwicklung von Ganztagschulen“ (STEG), „Deutsch Englisch Schülerleistungen International“ (DESI), „Unterrichtsqualität und mathematisches Verständnis in verschiedenen Unterrichtskulturen“ (sog. Pythagorasstudie), „Demokratie lernen und leben“, „Individuelle Förderung und adaptive Lerngelegenheiten im Grundschulunterricht“ (IGEL) sowie die Studie „Nutzung und Auswirkungen der Kompetenzmessung in mathematischen Lehr-Lern-Prozessen“ (CoCa) in dem gemeinsam mit Detlev Leutner koordinierten DFG Schwerpunktprogramm „Kompetenzmodelle“ sind beeindruckender Ausdruck der wissenschaftlichen Vielfalt und der mannigfaltigen Forschungsinteressen von Eckhard Klieme.

Für seine wissenschaftlichen Leistungen wurde Eckhard Klieme 2010 mit dem Wissenschaftspreis des Stifterverbandes „Gesellschaft braucht Wissenschaft“ und 2014 mit dem Forscherpreis der Deutschen Gesellschaft für Erziehungswissenschaft ausgezeichnet.

Aus Anlass des 60. Geburtstags von Eckhard Klieme im Herbst 2014 stellen in diesem Symposium vier ehemalige Mitarbeiter/-innen Eckhard Kliemes und deren Teams aktuelle Ergebnisse ihrer Forschungsprojekte vor. Diese Projekte wurden direkt oder indirekt durch die Zusammenarbeit mit Eckhard Klieme beeinflusst oder sind aus dieser hervorgegangen.

Die Studie von Maag-Merki und Oerke, welche exemplarisch für Eckard Kliemes Interesse an Forschung zur neuen Steuerungsmechanismen und ihren Wirkungen steht, untersucht die kurz- und längerfristigen Auswirkungen des Zentralabiturs auf motivationale Dimensionen des Lernens.

Die Untersuchung von Hartig und Harsch steht exemplarisch für die Arbeiten von Eckhard Klieme zur Modellierung und Messung von Kompetenzen. Sie geht der Frage nach, wodurch die Schwierigkeiten von Aufgaben im Leseverstehen im Fach Englisch erklärt werden können.

Im Vortrag von Lipowsky, Pauli, Drollinger-Vetter, Rakoczy und Reusser werden die Bedeutung der fachdidaktischen Qualität des Unterrichts für den Lernerfolg der Schüler/-innen und der Zusammenhang mit den Wahrnehmungen der Schüler/-innen untersucht. Dieser Beitrag aus dem Pythagorasprojekt steht stellvertretend für die Projekte zur Unterrichtsqualität, die Eckhard Klieme durchgeföhrt und beeinflusst hat.

Exemplarisch für die Projekte, die Eckhard Klieme im Bereich der außerschulischen Bildungsforschung geleitet und angestoßen hat, geht Stecher in seinem Vortrag der Frage nach, wie die Qualität außerunterrichtlicher Bildungsangebote kontextunabhängig bzw. kontextübergreifend erfasst werden kann. Hierzu werden Daten des Nationalen Bildungspanels (NEPS, Säule Lernumwelten) herangezogen.

Mittwoch
11. März

16:15 Uhr

HGC 30

Chair(s)

Lipowsky, Frank
Wirth, Joachim**Qualität und Wirkung pädagogischen Handelns: Zum 60. Geburtstag von Eckhard Klieme**

Maag Merki, Katharina; Oerke, Britta

Die Wirkungen zentraler Abiturprüfungen auf motivationale Dimensionen des Lernens der Schüler/-innen in Abhängigkeit von Unterrichtserfahrungen

Theoretischer Hintergrund: Empirische Befunde zur Effektivität zentraler Abschlussprüfungen zeigen einzelne positive, teilweise aber auch ambivalente und unklare Effekte. Insbesondere kann gezeigt werden, dass zentrale Abschlussprüfungen in einem Zusammenhang mit negativen motivationalen und emotionalen Dimensionen des Lernens stehen wie beispielsweise erhöhte Leistungsangst oder ein geringeres Fachinteresse (u.a. Jürges et al., 2012; Meyer et al., 2009; Ryan et al., 2007). Ebenso werden ambivalente Befunde hinsichtlich der Veränderung der Unterrichtsqualität sichtbar, wobei neben einzelnen positiven auch negative Effekte zu identifizieren sind (für eine Übersicht siehe Maag Merki, 2010).

Die empirische Befundlage ist allerdings unbefriedigend, da Längsschnittstudien fehlen, die den Wechsel des Prüfungssystems sowie den Effekt der Implementation zentraler Abschlussprüfungen über mehrere Jahre auf zentrale motivationale Dimensionen des Lernens der Schüler/-innen und des Unterrichts untersucht haben. Des Weiteren sind bisherige Studienergebnisse undifferenziert, da fachspezifische und anforderungsspezifische Differenzen zwischen den Lernsettings und die Interdependenz zwischen motivationalen Dimensionen des Lernens der Schüler/-innen und ihrer Unterrichtserfahrungen nicht genügend berücksichtigt wurden.

Fragestellung: Auf der Basis eines komplexen theoretischen Modelles (Maag Merki, 2014) wird der Frage nachgegangen, welche kurz- und längerfristigen Effekte die Einführung zentraler Abiturprüfungen auf motivationale Dimensionen des Lernens in Abhängigkeit der Unterrichtserfahrungen ausüben. Aufgrund bisheriger Ergebnisse der Unterrichts- und Lernforschung (Klieme, 2006; Reusser & Pauli, 2010) ist davon auszugehen, dass der Effekt zentraler Abiturprüfungen auf motivationale Dimensionen des Lernens durch die individuellen Unterrichtserfahrungen der Schüler/-innen mediiert wird. Es ist mit positiven wie auch negativen Effekten zu rechnen. Zudem sind fach- und anforderungsspezifische Unterschiede zu erwarten (Maag Merki, 2012).

Methode: Basis ist eine von der DFG geförderte Längsschnittstudie in den beiden Bundesländern Bremen und Hessen (Maag Merki, 2012). In Hessen wurden in 2007 zentrale Abiturprüfungen in allen drei schriftlichen Prüfungsfächern gleichzeitig, in Bremen hingegen nur in den Grundkursen, in 2008 zusätzlich in den Leistungskursen in einzelnen Fächern eingeführt. Damit ist es für Bremen möglich, den Effekt zentraler Abiturprüfungen in den Leistungskursen kurz- (2007-2008) und langfristig (2007-2011) zu untersuchen. Zudem können in beiden Ländern die längerfristigen Veränderungen nach Einführung des Zentralabiturs untersucht werden. In jedem Jahr (2007-2009, 2011) wurden standardisierte Befragungen bei Schülern/-innen des Abiturjahrganges in 19 Gymnasien in Bremen und in 9 Gymnasien in Hessen (repräsentativ bezüglich Region, Typus, Größe) einen Monat vor den Abiturprüfungen durchgeführt. Ausgewählt wurden für die nachfolgenden Ergebnisse die Schüler/-innen, die Deutsch oder Mathematik als Grund- oder Leistungskurs gewählt haben.

Stichprobe Bremen: Deutsch LK: N=181 – 204 Hessen, N=330-487 Bremen; Mathematik LK: N=150-181 Hessen, N=220-338 Bremen; Deutsch GK: N= 97-113 Hessen, N=153-245 Bremen; Mathematik GK: N=192-259 Hessen, N=132-203 Bremen.

Abhängige Variablen: Fachinteresse, Selbstwirksamkeit, Persistenz (α .69-.86).

Unterrichtserfahrungen als unabhängige Variablen: Motivierungsfähigkeit der Lehrperson, Autonomieunterstützung, Kompetenzunterstützung (Deci & Ryan, 1993) (α .65-.82).

Auswertungsverfahren: Mehrebenenanalyse (HLM 6.06; Level 1: SuS, Level 2: Schule), difference-in-differences Analysen unter Kontrolle von Geschlecht und kogn. Fähigkeit (KFT).

Ergebnisse: Die Analysen zeigen signifikante Effekte der Implementation des Zentralabiturs auf die motivationalen Dimensionen des Lernens in langfristiger Perspektive. Erwartungsgemäß variieren die Resultate in Abhängigkeit des Faches und Anspruchsniveaus. Entsprechend den Hypothesen wird zudem sichtbar, dass die Unterrichtserfahrungen bedeutsam den Zusammenhang zwischen Zentralabitur und Motivationen mediiieren. In den Leistungskursen Deutsch in Bremen werden die positiven Effekte auf das Fachinteresse vollständig durch die über die Jahre positiver erlebte Kompetenzunterstützung der Lehrpersonen erklärt. Diese verstärkte Kompetenzunterstützung kompensiert zudem ausbleibende oder negative Effekte auf die schulische Selbstwirksamkeit und Persistenz. In Hessen zeigen sich positive Interaktionseffekte bezüglich des Fachinteresses und der schulischen Selbstwirksamkeit in Mathematik Leistungskursen. Hingegen kann der positive Effekt auf die schulische Selbstwirksamkeit in Deutsch Leistungskursen nicht mit einer verstärkten Kompetenzunterstützung erklärt werden. In den Grundkursen, in denen die Veränderungen nach Einführung des Zentralabiturs untersucht worden sind, zeigen sich ebenfalls systematische Interaktionseffekte, weniger allerdings für die Grundkurse in Deutsch.

Mittwoch
11. März

16:15 Uhr

HGC 30

Chair(s)

Lipowsky, Frank
Wirth, Joachim

Qualität und Wirkung pädagogischen Handelns: Zum 60. Geburtstag von
Eckhard Klieme

Hartig, Johannes; Harsch, Claudia

Kriterienorientierte Testwertinterpretation mithilfe schwierigkeitsbestimmender Aufgabenmerkmale am Beispiel Englisch als Fremdsprache

Zur kriterienorientierten Interpretation von Testwerten aus Schulleistungstests werden in den meisten großen Studien Kompetenzniveaus oder Kompetenzstufen gebildet. Die ursprünglich kontinuierliche Skala wird hierzu in Abschnitte unterteilt, und diese Abschnitte werden bezogen auf die Anforderungen der Testaufgaben und damit bezogen auf die Fähigkeiten der Testteilnehmer inhaltlich beschrieben. Zur Bildung solcher Kompetenzniveaus gibt es eine Vielzahl von Methoden (z. B. Beaton & Allen, 1992), die oft auch unter dem Begriff „Standard Setting“ beschrieben werden (z. B. Cizek & Bunch, 2007). Eine attraktive Möglichkeit zur modellgeleiteten Bestimmung von Schwellen zwischen Kompetenzniveaus ist die Verwendung von schwierigkeitsbestimmenden Aufgabenmerkmalen, also a priori beschriebener Merkmale der Testitems, die (vermutlich) einen Effekt auf die Anforderungen der Items haben. Kompetenzniveaus können über die Schwierigkeiten von Testaufgaben mit charakteristischen, unterschiedlich schweren Merkmalskonfigurationen beschrieben werden (Hartig, 2007; Hartig, Frey, Nold & Klieme, 2012).

Im vorliegenden Beitrag wird das Vorgehen am Beispiel von Daten zum Leseverstehen im Englischen illustriert. Die Merkmale der Testaufgaben waren bei der Aufgabenentwicklung systematisch variiert worden, um die Effekte verschiedener Faktoren unabhängig voneinander abschätzen zu können. Variiert wurden die Textschwierigkeit, die Anforderungen der zur Beantwortung der Items notwendigen Informationsverarbeitung sowie das Antwortformat (kurze offene Antworten vs. Multiple Choice). In die Analyse gehen Daten von N=2328 Neuntklässlern/-innen ein, die in einem rotierten Design eine Gesamtmenge von 82 Testitems bearbeitet hatten. Neben den inhaltlichen Ergebnissen werden die Effekte unterschiedlicher Analysemethoden auf die Ergebnisse demonstriert. Verwendet wird zu einem ein einfaches zweistufiges Verfahren, in dem die Itemschwierigkeiten in einer eindimensionalen Raschskalierung geschätzt und anschließend in einer Regressionsanalyse mit den Aufgabenmerkmalen vorhergesagt werden. Zum anderen werden erklärende IRT-Modelle (explanatory item response models, Wilson & DeBoeck, 2004) geschätzt, in denen die Aufgabenmerkmale direkt in das Skalierungsmodell einbezogen werden.

Beim Methodenvergleich finden sich mit allen Methoden nahezu identische Ergebnisse. Hinsichtlich der Aufgabenmerkmale findet sich ein starker Effekt der zur Beantwortung der Items notwendigen Informationsverarbeitung sowie des Antwortformates. Die Schwierigkeit der Texte hat wider Erwarten keinen bedeutsamen Einfluss auf die Aufgabenschwierigkeit. Mögliche Schwellen zwischen Kompetenzniveaus werden anhand der Merkmale mit statistisch bedeutsamen Effekten gebildet und exemplarisch beschrieben.

Die Ergebnisse zeigen, wie ein Niveaumodell für Leseverstehen im Englischen für deutsche Sekundarschüler aussehen könnte, und veranschaulichen den Nutzen der Definition von schwierigkeitsbestimmenden Aufgabenmerkmalen. Der Methodenvergleich zeigt, dass derartige Analysen auch mit vergleichsweise einfachen Analysemethoden durchgeführt werden können.

Mittwoch
11. März

16:15 Uhr

HGC 30

Chair(s)

Lipowsky, Frank
Wirth, Joachim

Qualität und Wirkung pädagogischen Handelns: Zum 60. Geburtstag von Eckhard Klieme

Lipowsky, Frank; Pauli, Christine; Drollinger-Vetter, Barbara; Rakoczy, Katrin; Reusser, Kurt

Unterrichtsqualität – fachdidaktisch betrachtet

Theoretischer Hintergrund: In gegenwärtigen Modellen der Unterrichtsqualität wird von drei Qualitätsdimensionen ausgegangen: Kognitive Aktivierung, effektive Klassenführung und unterstützendes Unterrichtsklima gelten als wesentliche Qualitätsmerkmale des Unterrichts, die die Lern- und Motivationsentwicklung von Schüler/-innen positiv beeinflussen (Klieme, Pauli & Reusser, 2009; Kunter & Voss, 2011).

Darüber hinaus werden zur Konzeptualisierung der Qualität von Unterricht zunehmend auch fachspezifische Merkmale des Unterrichts herangezogen (Klieme & Rakoczy, 2008). Im Mathematikunterricht gelten z.B. die Behandlung von relevanten Kernideen und Konzepten („attending to concepts“, Hiebert & Grouws, 2007; „richness and development of the mathematics“, Learning Mathematics for Teaching Project, 2010) und deren Verknüpfungen („making connections“, Hiebert & Carpenter, 1992) als wesentliche Bedingungen eines qualitativ hochwertigen und lernwirksamen Mathematikunterrichts.

Fragestellungen: Im Vortrag wird der Frage nachgegangen, inwiefern die fachdidaktische Qualität von Mathematikunterricht mit den drei Basisdimensionen der Unterrichtsqualität und mit Aspekten der Lehrer-Schüler-Interaktion zusammenhängt und welche Effekte die fachdidaktische Qualität des Unterrichts auf die Lernentwicklung von Schüler/-innen hat. Zudem wird geprüft, inwieweit die fachdidaktische Qualität des Unterrichts mit den auf das eigene Verständnis bezogenen Wahrnehmungen der Schüler/-innen kovariert.

Methode: Die Fragestellungen werden anhand der Stichprobe des Pythagorasprojekts untersucht. Die Stichprobe besteht aus 39 deutschen und schweizerischen Klassen des 8. (Schweiz) bzw. 9. Schuljahres (Deutschland) mit ca. 1.000 Schüler/-innen, welche über ein Schuljahr hinweg untersucht wurden. Im Mittelpunkt des Projekts standen die beiden Unterrichtseinheiten „Einführung in den Satz von Pythagoras“ und „Umgang mit Textaufgaben“, welche videografiert und analysiert wurden. Für den vorliegenden Vortrag wurden nur die Daten aus der Pythagoraseinheit herangezogen.

In dieser Pythagoraseinheit wurde die fachdidaktische Unterrichtsqualität anknüpfend an die internationale Diskussion (Hiebert & Grouws, 2007) und anknüpfend an die Psychologische Didaktik Aebli (Aebli, 1994) als Behandeln, Strukturieren und Verknüpfen sogenannter Verstehenselemente begriffen (Drollinger-Vetter, 2011; Drollinger-Vetter & Lipowsky, 2006). Verstehenselemente werden dabei als Teilkonzepte des zu verstehenden Konzepts (hier des Satzes von Pythagoras) verstanden, die für den Aufbau eines Verständnisses dieses Konzepts grundlegend sind (Drollinger-Vetter, 2011). Diese Verstehenselemente müssen von den Schüler/-innen verknüpft werden, um ein tragfähiges Konzept vom Satz des Pythagoras zu entwickeln. Entsprechend wurde die fachdidaktische Qualität im Rahmen der Unterrichtseinheit zum Satz des Pythagoras anhand der drei 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 hängen eng zusammen (von $r=.72$ bis $r=.82$) und wurden daher zu der Skala „Fachdidaktische Qualität“ zusammengefasst.

Die drei Basisdimensionen der Unterrichtsqualität wurden über hochinferente Ratings und die Facetten der Lehrer-Schüler-Interaktion über eine Skala zur Diskursqualität ($\alpha=.80$ - $\alpha=.99$), welche sich aus niedrig inferenten Kodierungen speiste, erfasst.

Die Schülerwahrnehmungen bezogen sich auf die eigenen kognitiven Lernaktivitäten ($\alpha=.86$), gemessen unmittelbar nach der Unterrichtseinheit zum Satz des Pythagoras, und auf die Verstehensorientierung, gemessen am Ende des Schuljahres ($\alpha=.86$).

Die abhängigen Schülerleistungen wurden unmittelbar nach der Unterrichtseinheit zum Satz des Pythagoras und ca. 3 Wochen danach mit Testaufgaben erfasst, die das konzeptuelle Verständnis und die Anwendungsfähigkeiten der Lernenden erfassten (EAP/PV: .74; .78).

Ergebnisse: Ergebnisse von Mehrebenenanalysen unter Kontrolle verschiedener Lernvoraussetzungen und Kontextvariablen zeigen, dass die fachdidaktische Qualität des Unterrichts den Lernerfolg vorhersagt und zudem mit den – auf das eigene Verständnis bezogenen – Wahrnehmungen der Schüler/-innen in Zusammenhang steht.

Da die Korrelationen mit den drei Basisdimensionen der Unterrichtsqualität und mit den Facetten der Lehrer-Schüler-Interaktion (Pauli & Reusser, submitted) mehrheitlich schwach ausfallen, weisen die Befunde darauf hin, dass das dreidimensionale Modell von Unterrichtsqualität durch stärker fachdidaktisch konzeptualisierte Qualitätsdimensionen ergänzt werden sollte. Zudem deuten die Ergebnisse darauf hin, dass Beobachtereinschätzungen und Schülerwahrnehmungen bezogen auf den Aufbau von Verständnis enger zusammenhängen als wenn sich diese Wahrnehmungen und Einschätzungen auf allgemeine Merkmale von Unterrichtsqualität beziehen (Clausen, 2002).

Mittwoch
11. März

16:15 Uhr

HGC 30

Chair(s)

Lipowsky, Frank
Wirth, Joachim

Qualität und Wirkung pädagogischen Handelns: Zum 60. Geburtstag von
Eckhard Klieme

Stecher, Ludwig

Extended Education – Die Messung pädagogischer Qualität außerunterrichtlicher Angebote auf der Basis der Daten des Deutschen Bildungspanels

Theoretischer Hintergrund: In vielen Ländern lässt sich beobachten, wie lernunterstützende bzw. den Bildungserfolg flankierende Angebote und Programme außerhalb des Unterrichts bzw. außerhalb der klassischen Institutionen Schule, Berufsschule oder Hochschule zunehmend an Bedeutung gewinnen. Hierzu zählen etwa außerunterrichtliche Angebote an Ganztagschulen, Musikschulangebote, Volkshochschulkurse, Praktika, Nachhilfe, um nur einige zu nennen. Analog zum Unterricht als pädagogisch intentionalem Setting stellt sich auch in diesen non-formalen Angeboten die Frage, was ein qualitativ gutes Angebot ausmacht und was es bewirkt.

In der Forschung wurden hierzu in den letzten Jahren Modelle entwickelt. Diese leiten sich in der Regel aus der schulischen Effektivitätsforschung bzw. aus der Forschung zur Qualität guten Unterrichts ab, wie sie unter anderem von Eckhard Klieme und seinem Team vorgelegt wurden (Klieme et al., 2006; Stecher et al., 2007, 2009; Rakoczy et al., 2010).

Fragestellung: Die Modelle guten Unterrichts unterscheiden in der Regel drei Grunddimensionen: die Strukturdimension, die Dimension der kognitiven Aktivierung und die Unterstützungsdimension. Jenseits der Plausibilität der Übertragbarkeit dieser Grunddimensionen auf alle Formen pädagogisch intentionaler Lernsettings und Kontexte auch außerhalb des Unterrichts (siehe Radisch et al., 2008a, b) fehlt es allerdings bislang weitgehend an empirischen Belegen für die tatsächliche Übertragbarkeit der dimensional Struktur entsprechender Messmodelle auf außerunterrichtliche pädagogische Settings.

Dies greift der vorliegende Beitrag auf und geht anhand der Daten über unterschiedliche Lernkontexte der Frage nach, inwieweit die dreifaktorielle Struktur des postulierten Qualitätsmodells auf außerunterrichtliche Lernsettings übertragbar ist. Damit wird nicht nur eine Antwort auf die Frage vergleichbarer Messung pädagogischer Qualität in unterrichtsnahen und unterrichtsfernen Lernkontexten gegeben, sondern die für die Pädagogik sehr viel zentralere Frage bearbeitet, ob die drei Dimensionen guten Unterrichts tatsächlich als Fundamentalkategorien pädagogischer Arbeit gelten können, wie dies von Eckhard Klieme und Kollegen postuliert wird (Radisch et al., 2008b).

Methoden/Datengrundlage: Die Qualität außerunterrichtlicher Angebote der Extended Education wurde in einer Arbeitsgruppe des Nationalen Bildungspanels (NEPS, Säule Lernumwelten) aufgenommen (Bäumer et al., 2011). Zum ersten Mal liegen hier über verschiedene Bereiche der Extended Education vergleichbare Daten zur Messung der pädagogischen Qualität vor. Dabei bezieht sich der Beitrag auf verschiedene Felder der Extended Education: Nachhilfe in der 5./9. Jahrgangsstufe; Schulpraktika 9. Jahrgangsstufe; Praktika im Vorfeld des Studiums (Studierendenkohorte); Fortbildungsveranstaltungen (Erwachsenenkohorte).

Ergebnisse: Die Analysen zeigen, dass die Dimensionen kognitive Aktivierung und Unterstützung sich gut in den verschiedenen Settings abbilden und klare Faktoren bilden. Über beide Dimensionen sind kontextunabhängige Messungen der pädagogischen Qualität möglich. Der dritte Faktor – die Strukturdimension – zeigt in unterschiedlichen Kontexten eine unterschiedliche Struktur und bildet damit einen eher kontextspezifischen Faktor.

Mittwoch
11. März

16:15 Uhr

HGC 30

Chair(s)

Lipowsky, Frank
Wirth, Joachim

Qualität und Wirkung pädagogischen Handelns: Zum 60. Geburtstag von
Eckhard Klieme

Dirk, Judith

Stimmung und Emotion von Schülern und Lehrern im Schulkontext: Neue Erkenntnisse aus Experience Sampling Studien

Schüler wie Lehrer verbringen täglich viele Stunden im Schulkontext. Dabei verfolgen sie persönliche Ziele im Leistungsbereich und in Bezug auf ihre Sozialbeziehungen. Aufgrund dieser subjektiven Relevanz und der intensiven sozialen Interaktionen ist das affektive Erleben im Klassenraum von zentraler Bedeutung (Schutz & Pekrun, 2007). Obwohl bereits seit Langem bekannt ist, dass Stimmung im Alltag sowohl bei Kindern (Larson & Lampman-Petratis, 1989) als auch bei Erwachsenen (Wessman & Ricks, 1966) variiert, gibt es kaum Studien, die Schwankungen von Stimmung und Emotion alltagsnah im Schulkontext untersuchen und dabei sowohl Prädiktoren als auch Konsequenzen von Schwankungen im affektiven Erleben betrachten. Darüber hinaus haben Studien im Schulkontext zumeist interindividuelle Unterschiede betrachtet, wenngleich die theoretischen Modelle zur Rolle von Affekt im Erleben und Handeln allgemein (Carver & Scheier, 1998) sowie spezifisch im Schulkontext (Frenzel, 2014; Pekrun, Elliot & Maier, 2009) intraindividuelle Prozesse postulieren. Dies ist problematisch, da interindividuelle Befunde keine Schlussfolgerungen über intraindividuelle Prozesse zulassen (Molenaar, 2004; Schmitz, 2006).

In unserem Symposium möchten wir daher aktuelle Studien vorstellen, die Stimmung und Emotion aus Schüler- und Lehrerperspektive im Schulkontext alltagsnah mit Tagebüchern oder elektronischen Geräten erfasst und besonders intraindividuelle Prozesse untersucht haben. Dabei werden sowohl tägliche Einflussfaktoren auf Stimmung und Emotion als auch Konsequenzen von affektivem Erleben betrachtet. Wir ergründen damit die Breite des Forschungsfeldes und beantworten zum einen die Frage, wodurch Stimmung und Emotion im Schulkontext beeinflusst werden. Im Beitrag von Könen, Dirk und Schmiedek wird der Frage nachgegangen, welchen Einfluss Schlafqualität und -quantität auf die tägliche Stimmung von Grundschulern haben. Die Frage, welche Rolle Leistungsziele für Leistungsemotionen von Schülern in der Sekundarstufe haben, wird im Beitrag von Götz, Sticca, Murayama und Elliot diskutiert. Aus Sicht der Lehrer fragen Keller und Becker, inwiefern Klassenmerkmale und deren Bewertungen in der Entstehung von Emotionen im Alltag von Lehrern zusammenspielen. Zum anderen untersucht das Symposium auch Konsequenzen affektiven Erlebens im Alltag. Dieser Frage folgend diskutieren Dirk und Schmiedek, wie Stimmung und Motivation gemeinsam die kognitiven Leistungen von Grundschulern beeinflussen.

Damit kombiniert unser Symposium die Schüler- und Lehrerperspektive in Bezug auf das affektive Erleben im Schulkontext und zeigt die Vorteile des experience sampling Ansatzes für die Bildungsforschung. Wir betrachten Heterogenität im affektiven Erleben innerhalb von Schülern und Lehrern über die Zeit. Diese alltagsnahen Untersuchungen der affektiven Prozesse im Schulkontext können als Grundlage für spezifische Interventionen nutzbar gemacht werden, welche auf das Wohlbefinden und die Unterstützung bei der Zielerreichung von Schülern und Lehrern gerichtet sind.

Mittwoch
11. März

16:15 Uhr

HGC 40

Chair(s)
Dirk, Judith**Stimmung und Emotion von Schülern und Lehrern im Schulkontext: Neue Erkenntnisse aus Experience Sampling Studien**

Könen, Tanja; Dirk, Judith; Schmiedek, Florian

Das Zusammenspiel von Stimmung und Schlaf im Alltag von Grundschulkindern

Die Stimmung ist ein zentraler Aspekt im alltäglichen Erleben und Verhalten von Kindern. Obwohl schon lange bekannt ist, dass die Stimmung von gesunden Kindern im Alltag deutlich variiert (Larson & Lampman-Petratis, 1989), sind die Prädiktoren und Konsequenzen dieser Schwankungen weitgehend unerforscht. Forschung mit Erwachsenen legt nahe, dass Stimmung in einem bidirektionalen Verhältnis zu Schlafverhalten steht (z.B. Deliens, Gilson & Peigneux, 2014; Kahn, Sheppes & Sadeh 2013). Da Kinder sich jedoch in vielen zentralen Aspekten von Erwachsenen unterscheiden, z.B. in der Schlafdauer, der Schlafarchitektur, der kognitiven Entwicklung und den alltäglichen Erfahrungen, stellt sich die Frage des Zusammenhangs von Stimmung und Schlafverhalten für sie neu.

In einer Studie aus dem Projekt FLUX beantworteten 110 Kinder (8-11 Jahre alt) vier Wochen lang viermal täglich innerhalb und außerhalb der Schule Fragen zu ihrer aktuellen Stimmung auf Smartphones. Dabei wurden die Stimmungsfaktoren positiver und negativer Affekt sowie Aktivierung und Deaktivierung anhand einer kindgerechten und alltagsnahen Kurzsкала erfasst. Weiterhin berichteten die Kinder in der ersten Sitzung am Morgen ihr Schlafverhalten der letzten Nacht, wobei Quantität und wahrgenommene Qualität des Schlafes berücksichtigt wurden.

Die vierwöchige mikrolängsschnittliche Erhebung ermöglichte uns, interindividuelle Unterschiede zwischen Kindern und intraindividuelle Variabilität innerhalb von Kindern zu untersuchen. Im Gegensatz zu den meisten experimentellen Studien, in denen jeweils ein Aspekt von Schlaf oder Stimmung manipuliert wurde, betrachteten wir die alltägliche und natürlich auftretende Variabilität von mehreren Stimmungsfaktoren und Schlafvariablen. Die Schlafquantität und -qualität variierten deutlich, sowohl zwischen den Kindern als auch innerhalb der Kinder über die Zeit (Könen, Dirk & Schmiedek, 2014). Auch die Stimmung schwankte systematisch zwischen Kindern, zwischen Tagen und innerhalb von Tagen.

Wir nutzten die Stimmung zum letzten täglichen Messzeitpunkt (19:00 Uhr) um das Schlafverhalten der folgenden Nacht vorherzusagen. Das Schlafverhalten wiederum nutzten wir zur Vorhersage der Stimmung zu den ersten drei Messzeitpunkten am nächsten Tag (8:50 Uhr, 11:25 Uhr und 15:00 Uhr). Es zeigte sich, dass die Kinder besser schliefen, wenn sie zuvor guter Stimmung waren und länger schliefen, wenn sie zuvor aktiver waren. Bei der Vorsage des täglichen Affektes zeigt sich, dass die Schlafqualität im Gegensatz zur Schlafquantität für alle Stimmungsfaktoren relevant war, dieser Zusammenhang aber über den Tag hinweg schwächer wurde. Beispielweise waren die Kinder am Morgen und Vormittag besser gelaunt, wenn sie zuvor gut geschlafen hatten, doch für den Nachmittag zeigt sich dieser Zusammenhang nicht mehr.

Insgesamt konnten wir das Zusammenspiel zwischen Schlafverhalten und Stimmung direkt im Alltag von Grundschulkindern untersuchen. Das Schlafverhalten hing einerseits mit der Stimmung des Vorabends zusammen und war andererseits prädiktiv für die Stimmung am nächsten Morgen und Vormittag. Möglicherweise ist das Schlafverhalten ein relevanter Faktor für den Start in den Tag, während der restliche Verlauf der Stimmung über den Tag hinweg vom Tagesgeschehen abhängt. Es ist wichtig, dass Kinder am Morgen und Vormittag gut gelaunt und aktiviert sind, denn dies ermöglicht ihnen, die Schule und den Schulkontext als eine positive Erfahrung wahrzunehmen.

Mittwoch
11. März

16:15 Uhr

HGC 40

Chair(s)
Dirk, Judith

Stimmung und Emotion von Schülern und Lehrern im Schulkontext: Neue Erkenntnisse aus Experience Sampling Studien

Dirk, Judith; Schmiedek, Florian

Zusammenhänge von Stimmung, Motivation und Arbeitsgedächtnisleistung im Alltag von Grundschulern

Die Stimmung und die Motivation sind zentrale Aspekte im täglichen Erleben und Verhalten von Kindern. Stimmung beeinflusst kognitive Leistungen im Schulkontext (Linnenbrink, 2007). Motivationale Aspekte wie Interesse und Leistungsorientierung sagen Unterschiede in kognitiven Leistungen vorher (Eccles & Wigfield, 2002). Diese Ergebnisse beruhen wesentlich auf interindividuellen Befunden obwohl theoretische Annahmen zum Zusammenhang von Stimmung, Motivation und Leistung intraindividuelle Prozesse postulieren. So wird zum einen angenommen, dass das Regulieren von Emotionen kognitive Ressourcen benötigt (Ellis & Ashbrook, 1988), so dass Personen schlechtere Leistungen zeigen, wenn sie in schlechter Stimmung sind. Zum anderen hilft Stimmung das Verhalten zu regulieren, indem sie die motivationalen Ressourcen darauf richtet, unerwünschte Zustände zu vermeiden und erwünschte Zustände zu verstärken (Carver & Scheier, 1998). Zusammen betrachtet lässt sich daher annehmen, dass Stimmung und Motivation einzeln aber auch gemeinsam die kognitive Leistung beeinträchtigen. Während neuere mikrolängsschnittliche Studien mit Erwachsenen bereits intraindividuelle Zusammenhänge zwischen Stimmung, Motivation und Leistung gezeigt haben (z.B., Brose, Schmiedek, & Lövdén, & Lindenberger, 2012), fehlen vergleichbare Studien mit Kindern im Schulalltag.

Daher wurden im aktuellen Beitrag tagtägliche Schwankungen in Stimmung, Motivation und kognitiven Leistungen bei Grundschulkindern im Alltagskontext anhand einer Studie des Projektes FLUX mit ambulanten Messungen auf Smartphones betrachtet. Dabei untersuchten wir affektive, motivationale und kognitive Prozesse innerhalb von Kindern über die Zeit im Schulkontext. Wir nahmen zum einen an, dass die kognitive Leistung der Kinder niedriger ist in Situationen, in denen sie mehr negative und deaktivierende Stimmung und weniger positive und aktivierende Stimmung berichten. Zum anderen postulierten wir, dass die kognitive Leistung der Kinder niedriger ist in Situationen, in denen sie weniger motiviert sind. Schließlich untersuchten wir gemeinsame Effekte von Stimmung und Motivation auf Leistung basierend auf Modellen der Ressourcen-Bereitstellung (Ellis & Ashbrook, 1988) und motivationalen Theorien (Carver & Scheier, 1998). Dabei testeten wir, inwiefern der Einfluss von Stimmung auf die tagtägliche kognitive Leistung mit tagtäglichen Unterschieden in der Motivation zusammenhängt.

In der FLUX-Studie beantworteten 110 Kinder (8-11 Jahre) vier Wochen lang zu Beginn (8:50 Uhr) und am Ende des Unterrichts (11:25 Uhr) sowie am Nachmittag (ca. 15:00 Uhr) Fragen zu ihrer Stimmung und Motivation und lösten Arbeitsgedächtnisaufgaben. Dabei wurden die Stimmungsfaktoren positiver und negativer Affekt sowie Aktivierung und Deaktivierung anhand einer kindgerechten und alltagsnahen Kurzsкала mit jeweils 3 Fragen erfasst. Die Motivation wurde mit sechs kurzen Fragen zu aktueller Anstrengung und Freude bei der Aufgabenbearbeitung erfasst. Die Arbeitsgedächtnisaufgaben wurden für die Studie entwickelt und basieren auf Aufgaben, die bereits mit Jugendlichen und Erwachsenen erfolgreich in Mikrolängsschnittstudien eingesetzt wurden (Schmiedek, Lövdén, & Lindenberger, 2013; Riediger, Wrzus, Schmiedek, & Lindenberger, 2011). Aufgrund der empirischen und theoretischen Zusammenhänge zwischen Arbeitsgedächtnis und Schulleistung stellen sie ein optimales Maß für Leistungsschwankungen im Schulkontext dar.

Die Stimmung, die Motivation, und die Arbeitsgedächtnisleistung zeigten systematische tagtägliche Schwankungen. An Tagen, an denen Kinder mehr positive und aktivierende Stimmung erlebten, sich mehr anstrengten und mehr Spaß bei der Aufgabenbearbeitung hatten, zeigten sie im Mittel auch bessere Arbeitsgedächtnisleistungen. Im Gegensatz dazu waren die Arbeitsgedächtnisleistungen schlechter an Tagen, an denen Kinder mehr negative und deaktivierende Stimmung berichteten. Diese Effekte zeigten signifikante Zufallseffekte, was darauf hinweist, dass es Heterogenität in diesen Befunden gibt und die Zusammenhänge nicht für alle Kinder gleich sind. Des Weiteren teilten Anstrengung und Stimmung sowie Freude und Stimmung gemeinsame Varianz in der Vorhersage tagtäglicher kognitiver Leistungen.

Diese Ergebnisse replizieren Befunde mit Erwachsenen und bestätigen theoretische Annahmen über intraindividuelle Prozesse im Schulalltag von Kindern zwischen 8 und 11 Jahren. Stimmung und Motivation tragen gemeinsam zum Erbringen von kognitiven Leistungen im Schulkontext bei. Die systematischen Schwankungen in allen drei Bereichen und die Heterogenität in den gefundenen Zusammenhängen bildet eine wichtige Grundlage für die Untersuchung von interindividuellen Unterschieden in tagtäglichen Lernprozessen.

Mittwoch
11. März

16:15 Uhr

HGC 40

Chair(s)
Dirk, Judith**Stimmung und Emotion von Schülern und Lehrern im Schulkontext: Neue Erkenntnisse aus Experience Sampling Studien**

Götz, Thomas; Sticca, Fabio; Murayama, Kou; Andrew J., Elliot

Die Effekte von Leistungszielen auf Schüleremotionen: Eine Experience-Sampling-Studie

Im Rahmen der vorliegenden Studie wurde die Wirkung von Leistungszielen auf Lern- und Leistungsempfindungen von Schülerinnen und Schülern untersucht. Die theoretischen Modelle zu diesen Wirkungen beziehen sich ausschließlich auf intraindividuelle Mechanismen, d.h. auf Mechanismen innerhalb von Personen – unter Bezugnahme auf intraindividuelle Heterogenität (z.B. Linnenbrink & Pintrich, 2002; Pekrun, Elliot & Maier, 2009). Fast alle empirischen Studien zu diesem Thema haben jedoch interindividuelle Beziehungen, d.h. Beziehungen zwischen Personen untersucht – unter Bezugnahme auf interindividuelle Heterogenität. Da inter- und intraindividuelle Zusammenhänge jedoch statistisch völlig unabhängig voneinander sind (basierend auf intra- vs. interindividuellen Verteilungen; Molenaar, 2004) müssen Analysen auf interindividueller Ebene mit dem Ziel einer Testung von Hypothesen, die sich auf intraindividuelle Mechanismen beziehen, als höchst problematisch bezeichnet werden.

Die vorliegende Studie greift diese Unverbundenheit von Theorien und empirischer Forschung auf. Hierzu wurden sowohl inter- als auch intraindividuelle Zusammenhänge zwischen den im trichotomen Modell der Zielorientierung formulierten Konstrukten (Mastery Goals, Performance-Approach Goals, Performance-Avoidance Goals; Quelle: Achievement Goals Questionnaire (AGQ); Elliot & McGregor, 2001) und sechs Lern- und Leistungsempfindungen (Freude, Stolz, Angst, Scham, Ärger, Langeweile; Quelle: Achievement Emotions Questionnaire (AEQ), Pekrun, Goetz, Frenzel, Barchfeld & Perry, 2011) untersucht. Die Annahmen zu den Zusammenhängen basierten auf einem von Pekrun, Elliot und Maier (2006, 2009) entwickelten theoretischen Modell, in welchem Lern- und Leistungsempfindungen im Hinblick auf ihre Valenz (positiv vs. negativ), als auch im Hinblick auf ihren Fokus (Handlung vs. Ergebnis) unterschieden werden. Die Stichprobe bestand aus N = 120 Schweizer Schülerinnen und Schülern der 10. Jahrgangsstufe (37% weiblich). Um interindividuelle Mechanismen zu untersuchen (Trait-Ansatz), wurden Ziele und Emotionen anhand von Selbstbericht-Fragebögen erfasst. Für die Analyse intraindividuelle Beziehungen (State-Ansatz) wurden Ziele und Emotionen anhand der Experience-Sampling-Methode (Hektner, Schmidt & Csikszentmihalyi, 2007) für einen Zeitraum von 10 Schultagen im regulären Schulunterricht erfasst (insgesamt 1409 Real-Time-Erhebungen; Hardware: iPod Touch 4G; Experience-Sampling-Software: iDialog Pad; Kubiak & Krog, 2012). Sowohl beim State- als auch beim Trait-Ansatz wurden die Daten in vier unterschiedlichen Schulfächern erhoben, nämlich in Mathematik, Deutsch, Englisch und Französisch.

Vorausgehende Analysen zeigten einige bemerkenswerte Befunde. Das Muster an Korrelationen zwischen den untersuchten Variablen unterschied sich zum Teil gravierend bei den inter- versus intraindividuellen Analysen. So konnten beispielsweise Performance-Approach Goals und Performance-Avoidance Goals auf intraindividuelle Ebene deutlich besser voneinander getrennt werden ($r = .48$), als auf interindividueller Ebene ($r = .86$). Mehrebenenanalysen zur Heterogenität der Daten auf intraindividuelle Ebene zeigten für den State-Ansatz (Erhebungen innerhalb von Personen), dass die Intraklassenkorrelationen (ICCs) für Leistungsziele und Lern- und Leistungsempfindungen sehr ähnlich und relativ schwach waren ($[.257; .332]$ für Ziele; $[.204; .325]$ für Emotionen). D.h., der größte Anteil der Varianz resultierte für beide Konstrukte aus situationsbedingten Fluktuationen innerhalb von Personen. Bezüglich der zentralen Hypothesen waren die Ergebnisse der State-Analysen (intraindividuelle Mechanismen) im Einklang mit den theoretischen Annahmen: Mastery-Goals zeigten positive Effekte auf Freude und negative Effekte auf Langeweile und Ärger; Performance-Approach Goals waren positive Prädiktoren von Stolz; und Performance-Avoidance Goals zeigten positive Effekte auf Angst und Scham. Ein zentrales Ergebnis der Studie war auch, dass die Ergebnisse der Trait-Analysen (interindividuelle Beziehungen) mit denen der State-Analysen völlig übereinstimmten, d.h. ebenfalls die Hypothesen bestätigten. Die Ergebnisse beider Ansätze waren auch bei statistischer Kontrolle von Geschlecht, Alter, Noten und untersuchtem Unterrichtsfach stabil.

Möglichkeiten einer Integration inter- und intraindividuelle Forschungsparadigmen zu Leistungszielen und Emotionen sowie Implikationen für die unterrichtliche Praxis werden diskutiert. Das Wissen zu intraindividuellen Mechanismen ist eine Voraussetzung für die Entwicklung von Interventionsprogrammen (Goetz et al., 2013; Voelkle, Brose, Schmiedek und Lindenberger, 2014). Somit kann die vorliegende Studie einen Beitrag zur Entwicklung von Lernumgebungen leisten, welche einerseits Emotionen begünstigen, die Lernprozesse von Schülerinnen und Schülern zielgerichtet fördern und andererseits Emotionen reduzieren, die solchen Prozessen hinderlich sind.

Mittwoch
11. März

16:15 Uhr

HGC 40

Chair(s)
Dirk, Judith

Stimmung und Emotion von Schülern und Lehrern im Schulkontext: Neue Erkenntnisse aus Experience Sampling Studien

Keller, Melanie M.; Becker, Eva S.

Kognitive Appraisals medieren den Zusammenhang zwischen Klassenmerkmalen und emotionalem Erleben bei Lehrkräften: Ergebnisse einer Tagebuchstudie

Es wird davon ausgegangen, dass Lehreremotionen nicht nur für Gesundheit und Wohlbefinden, sondern auch für die Unterrichtsgestaltung und endgültig auch Schüleroutcomes von hoher Relevanz sind (Keller, Frenzel, Goetz, Pekrun, & Hensley, 2014). Bezüglich der Frage, wie Emotionen bei Lehrkräften entstehen, ist aber über einige qualitative Interviewstudien (siehe Sutton & Wheatley, 2003) hinaus wenig bekannt. Aus theoretischer Sicht ist davon auszugehen, dass nicht die Situation selbst (z.B. Leistungsniveau der Klasse), sondern die kognitive Interpretation und Bewertung der Situation durch die Lehrperson (sog. Appraisals) deren Emotionen bestimmen (Ellsworth & Scherer, 2003). Derartige Appraisaltheorien postulieren demnach Zusammenhänge zwischen Appraisals und Emotionen auf situativer Ebene (state), d.h. innerhalb von Personen, wohingegen die Mehrheit der bisherigen Emotionsforschung Emotionen auf einer generalisierten Ebene (trait) untersucht; dass jedoch trait- und state-Angaben von Emotionen auch für Lehrkräfte nicht unbedingt übereinstimmen, wurde bereits gezeigt (Keller et al., 2014). In einem Rahmenmodell für Lehreremotionen von Frenzel (2014) werden verschiedene Appraisals unterschieden (z.B. Zielzutraglichkeit, Zielrelevanz, Kontrolle) und als Mediator zwischen Klassenmerkmalen und Lehreremotionen postuliert; eine empirische Überprüfung der Zusammenhänge insbesondere auf Basis momentanen Emotionserlebens steht allerdings noch aus.

Ausgehend von Frenzels (2014) Modell wird untersucht, inwieweit Klassenmerkmale (Leistung, Motivation, Disziplin) und Appraisals (Zielzutraglichkeit, Zielrelevanz, Kontrolle) das momentane emotionale Erleben (Freude, Stolz, Ärger, Angst) bei Lehrpersonen im Unterricht vorhersagen. Es wird davon ausgegangen, dass Appraisals den Zusammenhang zwischen Klassenmerkmalen und Emotionen medieren.

Um momentanes Emotionserleben von Lehrpersonen gleichzeitig mit von Schüler/innen wahrgenommenen Klassenmerkmalen alltagsnah erfassen zu können, wurde die Methode des Tagebuchs gewählt. Hierbei wurden N = 28 Mathematiklehrkräfte der 9./10. Klasse an Gymnasien und deren Schüler/innen jeweils im Anschluss an eine Unterrichtsstunde (N = 282) mittels eines Kurzfragebogens über einen Zeitraum von 2-3 Wochen befragt. Klassenmerkmale (Leistung, Motivation, Disziplin) wurden dabei durch Schülerwahrnehmung erfasst und jeweils pro Unterrichtsstunde aggregiert. Die Appraisals Zielzutraglichkeit, Zielrelevanz, und Kontrolle wurden von Lehrpersonen jeweils mittels Einzelitems eingeschätzt (z.B. „In dieser Stunde war das Schülerverhalten meinen Unterrichtszielen zuträglich.“), die Emotionen Freude, Stolz, Ärger, und Angst jeweils mittels zweier Items (z.B. „In dieser Stunde machte mir das Unterrichten Freude“). Alle Items konnten auf einer fünfstufigen Skala von (1) stimmt gar nicht bis (5) stimmt genau bewertet werden. Unter Berücksichtigung der genesteten Datenstruktur wurden Mehrebenenpfadmodelle (Mplus 6.1; Muthén & Muthén, 1998-2012) getrennt für die Emotionen gerechnet und die Zusammenhänge zunächst nur auf within-level modelliert.

Freude ist im Mittel die am stärksten auftretende unterrichtsbezogene Emotion von Lehrpersonen, gefolgt von Stolz; negative Emotionen sind deutlich geringer ausgeprägt. Emotionales Erleben zeigt sich stark kontextualisiert, d.h. die Mehrheit der Varianz ist innerhalb von Personen zu finden ($ICC_{Freude/Stolz/Ärger/Angst} = .16/.33/.22/.22$). Ähnliches zeigt sich auch für die Appraisals ($ICC_{Zielzutraglichkeit/Zielrelevanz/Kontrolle} = .30/.28/.36$).

Analysen der Zusammenhänge zwischen Klassenmerkmalen und Appraisals mit Emotionen zeigen mehrheitlich signifikante Effekte. Beispielsweise ist Kontrolle insbesondere hinsichtlich der Entstehung aller Emotionen bedeutsam ($\beta_{Freude/Stolz/Ärger/Angst} = .39^{***}/.20^{*}/-.34^{***}/-.31^{***}$), wohingegen Zielzutraglichkeit insbesondere in Zusammenhang mit dem Erleben von Freude ($\beta = .38^{***}$) steht. Die Analyse der Pfadmodelle zeigt, dass die Appraisals den Zusammenhang zwischen Klassenmerkmalen und Emotionen teilweise medieren, wobei sich Unterschiede je nach Klassenmerkmal, Appraisal, und Emotion zeigen. So mediert beispielsweise Kontrolle den Effekt von Disziplin auf Freude, Ärger und Angst (Freude: $\beta_{indirekt} = .20^{**}/-.12^{**}/-.14^{**}$), nicht jedoch auf Stolz. Ergebnisse für alle Zusammenhänge und Modelle werden im Rahmen des Vortrags dargestellt.

Die Ergebnisse der vorliegenden Studie unterstützen die theoretische Annahme zum einen der unterschiedlichen Bedeutsamkeit der Appraisals je nach Emotion und zum anderen der medierenden Funktion von Appraisals. Dies ist insbesondere deshalb von Bedeutung, da über die Appraisals Einflussnahme auf das Emotionserleben von Lehrkräften möglich scheint, beispielsweise mittels sog. Reappraisal-Strategien (z.B. Gross & John, 1998) und diese somit grundlegend sind im Kontext der Optimierung von emotionalem Erleben und Wohlbefinden von Lehrpersonen.

Mittwoch
11. März

16:15 Uhr

HGC 40

Chair(s)
Dirk, Judith

Stimmung und Emotion von Schülern und Lehrern im Schulkontext: Neue Erkenntnisse aus Experience Sampling Studien

Winther, Esther

Validität authentischer Testformate – Befunde aus der ASCOT-Initiative

Das Symposium bedient aus einer inhaltlichen Perspektive drei Kernbereiche der beruflichen Bildung: die gewerblich-technische Domäne, die kaufmännische Domäne sowie den Bereich der Gesundheits- und Pflegeberufe. Aus einer methodischen Perspektive greifen die Einzelbeiträge ein zentrales Desiderat der Testkonstruktion auf: die Sicherstellung von Validität. Alle Beiträge sind im Rahmen der ASCOT-Forschungsinitiative des BMBF entstanden, deren Zielsetzung die Modellierung und Messung beruflicher Kompetenzen ist, wobei die Messung selbst mittels technologisierten Testformaten erfolgt, um (a) der Authentizität beruflicher Kompetenzmessung Rechnung zu tragen und um (b) die Komplexität beruflicher Handlungsabläufe entsprechend zu würdigen. Der ASCOT-Forschungsverbund lebt durch einen hohen Anteil inter- und transdisziplinärer Arbeiten, die über die eingereichten Beiträge abgebildet werden können. Neben (1) wirtschafts- und berufspädagogischen Fragestellungen, die primär auf Aspekte des Kompetenzerwerbs und des Lernens in Berufen und den entsprechenden Modellierungen bezogen sind, werden für die theoretische Fundierung der Modellannahmen beispielsweise (2) Ansätze der soziologischen Arbeitsforschung und (3) der Psychologie herangezogen.

Die vier Einzelbeiträge des Symposiums greifen die Frage nach der Validität der eingesetzten authentischen Testformate unterschiedlich auf: Im Beitrag von Julia Sangmeister, Dagmar Festner & Esther Winther wird eine inferenzstatistische Prüfung der Dimensionalität beruflicher Kompetenz von Auszubildenden im Ausbildungsberuf Industriekaufmann / Industriekauffrau mittels Modellen der IRT angestrebt, wobei verschiedene (und zum Teil konkurrierende) Modellannahmen diskutiert werden. Hiermit wird (a) aufgezeigt, dass die Diskussion um die Modellierung beruflicher Kompetenz ohne explizit inhaltlichen und damit fachdidaktischem Diskurs zu kurz greifen würde und dass (b) berufliche Kompetenzen als mehrdimensionales Konstrukt Aspekte beruflicher Entkulturation, wie beispielsweise berufstypische Urteilsstrukturen und betriebsspezifische Entscheidungs- und Handlungsmuster, berücksichtigen. Susan Seeber und Michelle Liedtke verfolgen das Ziel, berufliche Kompetenzmodelle für verschiedene Berufsfelder zu validieren und entsprechend den Generalisierungsgrad solcher Modelle zu erhöhen. Hierfür werden die Kompetenzprofile zweier Ausbildungsgruppen (Industrie-sowie Speditionskaufleute) verglichen. Es wird geprüft, inwieweit sich differentielle Itemfunktionen auf eine konstrukt-irrelevante Varianz zurückführen lassen, oder ob berufsspezifische Lerngelegenheiten die verschiedenen Kompetenzprofile in den zwei kaufmännischen Berufen erklären. Für die gewerblich-technische Domäne stellen Nico Link & Bernd Geißel Befunde vor, die aufzeigen, in welchem Umfang Leistungen von Auszubildenden im Beruf Elektroniker/innen für Automatisierungstechnik in einer technologisierten Testumgebung als konvergent valide hinsichtlich der erbrachten Leistung in einer realen Arbeitsumgebung interpretiert werden können. Für den Bereich der Gesundheits- und Pflegeberufe wird von Agnes Dietzen,

Moana Monnier, Christian Srbeny & Tanja Tschöpe vorrangig aus Perspektive einer inhalts- und konstruktvaliden Prüfung ein Situational Judgment Test präsentiert, der sozial-kommunikative Kompetenzen von Medizinischen Fachangestellten abbildet.

Mittwoch
11. März

16:15 Uhr

HGC 50

Chair(s)
Winther, Esther

Sangmeister, Julia; Festner, Dagmar; Winther, Esther

Valide und authentisch messen – Beispiele aus kaufmännischen Handlungssituationen

Der Beitrag stellt den im Projekt CoBALIT (Competencies in the field of Business and Administration – Learning, Instruction, and Transition) entwickelten computerbasierten Test zur Messung kaufmännischer Kompetenzen für Auszubildende im Ausbildungsberuf Industriekaufmann / Industriekauffrau vor und beleuchtet hierbei Aspekte der Validitätssicherung. Der Test folgt auf Basis von 71 Testitems einer komplexen Geschäftsprozessmodellierung und ist an ein Modellunternehmen (ALUSIM®) gekoppelt, um weitreichende berufliche Authentizität sicherzustellen.

Theoretischer Hintergrund: Auf Basis einer technologiebasierten Kompetenzmessung bei Auszubildenden im Ausbildungsberuf Industriekaufmann / Industriekauffrau mit der authentischen Testplattform ALUSIM werden Analysen zur Validitätsprüfung der Messung vorgestellt. Dabei liegt der Fokus auf dem Aspekt der Konstruktvalidität (u. a. Messick, 1989) und damit auf einer Prüfung von Kompetenzstrukturannahmen bzw. von Dimensionalität beruflicher Kompetenz. Eine inferenzstatistische Prüfung der Dimensionalität erfolgt dabei mit Modellen der Item-Response-Theorie (IRT). Diese Prüfung ist insbesondere dadurch motiviert, dass in der beruflichen Bildung eine Vielzahl verschiedener Kompetenzmodelle vertreten werden, die entweder einer inhaltlichen Logik folgen (u. a. Geschwendtner, Abele & Nickolaus, 2009; Winther, 2010) oder sich an kognitiven Strukturüberlegungen bzw. Handlungsmodellen ausrichten (u. a. Lehmann & Seeber, 2007). Diese Modelle scheinen bislang unvereinbar, obwohl sie auf vergleichbaren Annahmen der Item- und Testkonstruktion sowie auf sehr ähnlichen Zugängen zu den beruflichen Domänen basieren.

Fragestellungen: Ziel des vorliegenden Beitrags ist es darzustellen, wie sich die psychometrische Güte eines beruflichen Kompetenztests überprüfen lässt und welche Anhaltspunkte zur Interpretation der geschätzten Item- und Personenparameter im Hinblick auf die Testintention und die Konstruktvalidität des Tests gefunden werden können. Hierbei ist in besonderem Maße zu berücksichtigen, dass berufliche Kompetenzmessungen (a) der Authentizität und (b) der Komplexität betrieblicher Arbeitsabläufe und Entscheidungen verpflichtet sind. Mit dieser Verpflichtung steigt gleichsam das Risiko, dass das ursprüngliche Zielkonzept der Testung (hier Abbildung beruflicher Kompetenz) von dem, was mit der Testendform tatsächlich erfasst wird, abweicht. Es sind daher Binnenkriterien bei der Itemanalyse heranzuziehen, mit denen sich zeigen lässt, dass Testintention und Testbefund möglichst wenig differieren. Vor diesem Hintergrund ergeben sich (zum Teil) konkurrierende Modelle der Kompetenzstrukturen mit inhalts- und handlungslogischen Bezügen. Diese Strukturen sind kritisch zu überprüfen und sowohl hinsichtlich ihrer empirischen Passung als auch mit Blick auf die inhaltlichen und damit auch fachdidaktischen Deutungen zu analysieren, da sie letztlich ein Maß für die Konstruktvalidität des beruflichen Kompetenztests darstellen.

Methode und Ergebnisse: An einer umfangreichen Stichprobe von Auszubildenden im Ausbildungsberuf Industriekaufmann / Industriekauffrau (N = 620 aus Pilotierungsdatensatz) werden vier konkurrierende Modelle zur Kompetenzstruktur vorgestellt und mittels Verfahren der IRT überprüft. Es zeigt sich, dass eine 2-faktorielle Lösung eine gute Modellanpassung erzielt. Dieses Modell wird an einer zweiten Stichprobe (N = 1450 aus Hauptstudien datensatz) konfirmatorisch repliziert. Zusammenhänge mit konvergenten und diskriminanten Außenkriterien belegen insbesondere die Konstruktvalidität einer domänenspezifischen Kompetenzstruktur (und damit die Abbildung beruflicher Aktivität und Bewertung), während gleichsam eine domänenverbundene Kompetenz als dispositionale Hilfskategorie für die Bewältigung authentischer Situationen in beruflichen Kontexten abgebildet werden kann (Skalierung mittels MRCMLM; EAP/PV Reliabilität: .712; 71 Testitems im zulässigen Intervall von $0,75 \leq wMNSQ \leq 1,33$ sowie korrespondierende T-values $\leq 1,0$; Item Thresholds (difficulty) von -2,98 bis 2,79).

Mittwoch
11. März

16:15 Uhr

HGC 50

Chair(s)
Winther, Esther

Seeber, Susan; Liedtke, Michelle

Differentielle Itemfunktionen: Testärgernis oder Anhaltspunkt für Stärken und Schwächen im Kompetenzprofil Jugendlicher bei verschiedenen kaufmännischen Berufen?

Ziele/Fragestellung: Für die Messung beruflicher Kompetenz ist es eine besondere Herausforderung, Inhalte zu identifizieren, die nicht nur valide im Hinblick auf das Berufsspezifische sind, sondern auch mit Blick auf das Berufsfeldtypische. Im Projekt CoBALIT-Transfer wird geprüft, inwieweit für verschiedene kaufmännische Berufe relevante Handlungsfelder, deren Aufgaben und Anforderungen in einem simulationsbasierten Assessment in einen berufs- und betriebsspezifischen Handlungskontext eingebunden sind, eine Messung beruflicher Kompetenzen bei verschiedenen kaufmännischen Berufen gestatten. Mit Blick auf die Validität stellen sich Fragen der inhaltlichen Relevanz und Repräsentativität der in den Testaufgaben modellierten beruflichen Anforderungen, aber auch Fragen nach angemessenen Testmodellen für die Auswertung und Interpretation der Ergebnisse.

Theoretischer Rahmen: Bisherige Studien zu einzelnen kaufmännischen Berufen zeigen, dass sich kaufmännische Kompetenz als kontinuierliche latente Variable modellieren lässt (z. B. Seeber, 2008; Winther 2010). Weitgehend offen ist, ob auch vergleichende Aussagen zur beruflichen Fachkompetenz über verschiedene Berufe einer Berufsfamilie möglich sind. Die Diagnose beruflicher Kompetenzen setzt eine Bestimmung der Anforderungsbereiche voraus. Damit verbunden sind Fragen zur Domänenabgrenzung und Reichweite der jeweiligen Kompetenzen. Messick (1989) unterscheidet verschiedene Validitätsaspekte, u. a. Relevanz und Repräsentativität, Mess- und Score-Modelle zur Datenauswertung, aber auch die Validität hinsichtlich der Interpretation der Testergebnisse.

Methodisches Vorgehen: Für die Erfassung der beruflichen Fachkompetenz bei Industriekaufleuten und Speditionskaufleuten wurde ein computerbasierter Test mit authentischen beruflichen Handlungssituationen entwickelt. Dieser besteht aus mehreren kaufmännischen Handlungsbereichen, wobei sich die vorgestellten Analysen auf das Modul „Logistik“ beziehen, das 18 Aufgaben umfasst. Zunächst wurde geprüft, inwiefern Ziele, Inhalte und Arbeitshandlungen in den beiden Berufen strukturell übereinstimmen. Neben Arbeitsplatzanalysen wurden Ausbildungsordnungen, Berichtshefte, Curricula, Lehrbücher und IHK-Prüfungsaufgaben analysiert. Auf der Grundlage von Differential Item Functioning Analysen wurde mit Daten aus der ersten Kohorte (Juni 2014) geprüft, ob sich – bei gleicher Fähigkeit - die Lösungswahrscheinlichkeit der Items zwischen Speditions- (N=423) und Industriekaufleuten (N=179) unterscheidet (Abschluss der Erhebungen der zweiten Kohorte im November 2014).

Ergebnisse: Über den Martin-Löf-Test wurde geprüft, ob das Testmodul eindimensional ist. Dieser wurde nicht signifikant, auch Faktorenanalysen lieferten keine Anhaltspunkte für eine Mehrdimensionalität des Tests. Eine Skalierung über das Partial-Credit-Modell zeigte zufriedenstellende Test- und Itemfitstatistiken (EAP/PV Reliabilität: .788; Weighted MNSQ-Fits: 0.91 bis 1.10; T-values ≤ 1.0 ; Item-Total-Correlations zwischen 0.23 und 0.61; Item Thresholds von -2,23 bis 1,99). Allerdings wurde ein deutlicher Leistungsvorsprung der Speditionskaufleute gegenüber den Industriekaufleuten in der Bearbeitung der Logistikaufgaben festgestellt. Die dabei jeweils festgestellten Effektstärken erlangen zum Teil beachtliche Ausmaße und deuten eine hohe praktische Relevanz der Leistungsunterschiede an. DIF-Prüfungen zur Lösungswahrscheinlichkeit zeigten bei einer Reihe von Items eine Konfundierung mit der beruflichen Gruppenzugehörigkeit.

Dies bedeutet, unabhängig von generellen Unterschieden in der mittleren Testleistung beider Gruppen weisen Items unterschiedliche Messeigenschaften auf, in Bezug auf den vorliegenden Beitrag insbesondere unterschiedliche Schwierigkeitsparameter zwischen den Gruppen. Ziel des Beitrags ist es, zu untersuchen, was Items auszeichnet, die besonders gut von Speditionskaufleuten oder besonders gut von den Industriekaufleuten gelöst werden. Zu prüfen ist ferner, ob berufsspezifische Lerngelegenheiten zu Stärken und Schwächen im Kompetenzprofil führten. Dies hat Bedeutung für die Interpretation und die Rückbindung der Assessmentergebnisse an die Weiterentwicklung von Ausbildungsbedingungen und die Sicherung der Ausbildungsqualität. In einem abschließenden Schritt wird geprüft, welches Messmodell am besten für die vorliegenden Daten geeignet ist, wobei insbesondere Mehrgruppenmodelle in den Analysen einbezogen werden.

Mittwoch
11. März

16:15 Uhr

HGC 50

Chair(s)
Winther, Esther

Dietzen, Agnes; Monnier, Moana; Srbeny, Christian; Tschöpe, Tanja

Messung sozialer Kompetenzen bei Medizinischen Fachangestellten im Projekt CoSMed

Der Beitrag stellt den im Projekt CoSMed (Competence Measurement based on Simulations and adaptive Testing in Medical Settings) entwickelten computerbasierten Situational Judgement Test zur Messung sozialer Kompetenzen von Medizinischen Fachangestellten vor und beleuchtet hierbei Aspekte der Validitätssicherung.

Theoretischer Hintergrund: Dem Testverfahren liegt ein Kompetenzmodell zugrunde, das auf der Basis empirischer Analysen der Domäne sowie theoretischen Erwägungen entwickelt wurde. Das Kompetenzmodell beinhaltet vier Kompetenzdimensionen: Emotionsregulation, Perspektivenkoordination, Kommunikationsstrategien und Sprechen/ Zuhören (näheres zum Kompetenzmodell vgl. Monnier, Srbeny & Tschöpe, 2014; Srbeny, Monnier, Dietzen & Tschöpe, im Druck). Für die theoretische Fundierung der Kompetenzdimensionen und die Ableitung von Bewertungsschemata für die Testantworten wurden Ansätze aus verschiedenen Disziplinen, u. a. der Pädagogik, der soziologischen Arbeitsforschung, Psychologie und den Kommunikationswissenschaften herangezogen. Bei der Operationalisierung der Dimension Emotionsregulation wurden u.a. Ansätze von Gross (2009) herangezogen, die den Einsatz verschiedener Strategien beim Umgang mit Emotionen beschreiben. Die Dimension Perspektivenkoordination beruht maßgeblich auf dem entwicklungspsychologischen Konzept von Selman (2003), das die Fähigkeit zur Koordination verschiedener Perspektiven bei interpersonellen Konflikten beschreibt. Die Dimension Kommunikation bezieht Ansätze von Rosenberg (2006) zur gewaltfreien Kommunikation und das Kommunikationsquadrat von Schulz von Thun (2004) ein.

Fragestellung: Die zentrale Fragestellung des Projekts lautet: Wie können sozial-kommunikative Kompetenzen Medizinischer Fachangestellter mittels eines computergestützten, situationsbezogenen Testinstruments valide, authentisch und realitätsnah gemessen werden? Der vorliegende Beitrag beschreibt, wie das erarbeitete Kompetenzmodell der sozial-kommunikativen Kompetenzen in einen computergestützten Test überführt wurde. Dazu werden Schritte zur Validitätssicherung sowie zur Erhöhung der Realitätsnähe und Akzeptanz dargestellt.

Methode: Bei der empirischen Anforderungsanalyse für das Kompetenzmodell kamen u.a. Interviews und Befragungen nach der Critical Incident Technique (Flanagan, 1954) zum Einsatz. Für die Kompetenzmessung wurde im Projekt ein an Situational Judgement Tests (Lievens, Peeters & Schollaert, 2008) orientiertes Instrument entwickelt, welches anders als klassische SJTs nicht nur Fragen zur Verhaltenstendenz in den dargestellten Situationen, sondern Fragen zu allen für den Beruf als relevant erachteten sozialen Kompetenzbereichen enthält. Der Test arbeitet mit Videosequenzen von kritischen Interaktionssituationen mit Patientinnen/Patienten und Kolleginnen/Kollegen, zu denen offene und geschlossene Fragen am Computer gestellt werden. Im Beitrag werden Auszüge aus dem Test vorgestellt, wobei u.a. auf die Kategorisierung und Systematik der verwendeten Situationen, die emotionsauslösende Wirkung der Videos sowie die Auswahl der Item-Formate eingegangen wird. Der Test wurden zweimal an Stichproben von jeweils ca. 300 Auszubildenden am Ende der Ausbildung pilotiert und in einer Haupterhebung mit weiteren ca. 400 Auszubildenden (insgesamt n ~ 1000) eingesetzt, wobei zusätzliche Daten zur Fachkompetenz und zu weiteren Variablen erhoben wurden (z.B. Lesekompetenz, Ausbildungsbedingungen, Demographie). Zur Validierung des Tests wurden darüber hinaus in den Pilotierungen standardisierte Verfahren zur Messung sozialer und emotionaler Kompetenzen eingesetzt (u.a. ISK - Inventar Sozialer Kompetenzen von Kanning, 2009; ERI - Emotionsregulations-Inventar von König, 2011). Die inhaltliche Validität wurde in allen Schritten der Testkonstruktion durch Interviews, Gruppendiskussionen und Online-Befragungen von Experten immer wieder abgesichert und hierbei Urteile von wissenschaftlichen Experten verschiedener Disziplinen ebenso berücksichtigt wie von Medizinischen Fachangestellten, Ärzten sowie Kammer- und Verbandsvertretern.

Ergebnisse: Aus den Pilotierungen werden Ergebnissen präsentiert, die Auswirkungen auf die Testentwicklung und die Itemgestaltung hatten. So kam es beispielsweise aufgrund der sehr heterogenen Zusammensetzung der Auszubildenden hinsichtlich des Schulabschlusses oder der Sprachkompetenz in ersten Versionen zu Verzerrungen. Bezogen auf die aktuelle Testversion werden Gütekriterien durch Kennwerte der klassischen Testtheorie vorgestellt, beispielsweise Interrater-Reliabilitäten, interne Konsistenzen, Trennschärfen oder Faktorladungen. Insbesondere wird auf bereits erfolgte und noch anstehende Prüfungen der Validität, z.B. zur Inhalts- oder Konstruktvalidität eingegangen. Neben einem Beitrag zur Klärung des Konzepts der berufsbezogenen Sozialkompetenz und seiner Messung werden weitere potenzielle Einsatzmöglichkeiten vorgeschlagen.

Mittwoch
11. März

16:15 Uhr

HGC 50

Chair(s)
Winther, Esther

Link, Nico; Geißel, Bernd

Entwicklung und Validierung eines Messinstruments zu Erfassung berufsspezifischer konstruktiver Problemlösefähigkeit bei Elektroniker/innen für Automatisierungstechnik

Theoretischer Hintergrund: Elektroniker/innen für Automatisierungstechnik arbeiten an komplexen automatisierten Industrieanlagen und sind für deren Programmierung, Inbetriebnahme und Instandhaltung verantwortlich. Das Ziel des Forschungsprojekts aus welchem hier berichtet wird, ist es, arbeitsprozessnahe Messinstrumente zur Erfassung berufsfachlicher Kompetenzen von Elektroniker/innen für Automatisierungstechnik zu entwickeln. Der Fokus dieses Beitrags liegt auf der Entwicklung und Validierung eines Messinstruments zu Erfassung berufsspezifischer konstruktiver Problemlösefähigkeit unter besonderer Berücksichtigung der Programmierung von speicherprogrammierbaren Steuerungen. Die fachspezifische Problemlösefähigkeit wird hier als ein Bündel kognitiver Leistungsdispositionen definiert, das das selbstständige, methodengeleitete und sachgerechte Handeln in berufsfachlichen problemhaltigen Situationen ermöglicht (Hartig & Klieme, 2007; KMK, 2007; Abele, Walker & Nickolaus, 2014). So treten bei Elektroniker/innen für Automatisierungstechnik neben analytischen Anforderungssituationen, wie bei der Fehleranalyse in einer automatisierten Anlage, auch konstruktive Anforderungssituationen im Berufsalltag auf. Hinsichtlich der analytischen Problemlösefähigkeit konnte im Beruf des KFZ-Mechatronikers gezeigt werden, dass die Testleistung in einer Computersimulation, als konvergent valide hinsichtlich der erbrachten Testleistung in einer realen Arbeitsumgebung interpretiert werden können (vgl. Nickolaus, Gschwendtner & Abele, 2009). Vergleichbare Ergebnisse zur Konstruktvalidität der konstruktiven Problemlösefähigkeit liegen bislang nicht vor. Erste Überlegungen zum Konstrukt der konstruktiven Problemlösefähigkeit legen nahe, dass für das Erstellen von Steuerungsprogrammen normierte Symbole und Grafiken verwendet werden und die Erstellung der Steuerungsprogramme in der Realität in computerbasierten Entwicklungsumgebungen und zumeist zeitlich und räumlich getrennt von einem automatisierten System (wie z.B. dem Mensa-Band) erfolgt (Langmann, 2004). Vor dem Hintergrund dieser Annahmen erscheint der Vergleich der erbrachten Testleistung in der computerbasierten Entwicklungsumgebung mit der Testleistung, welche in einem authentischen Paper-Pencil-Test erbracht wurde, hinsichtlich der konstruktiven Problemlösefähigkeit erfolgsversprechend.

Fragestellung: Da die Verfügbarmachung von realen Programmierumgebungen, gegebenenfalls in Kombination mit automatisierten Anlagen, für ein Large-Scale-Assessment aus ökonomischen und Praktikabilitätsgründen ausscheidet, soll geprüft werden, inwiefern die mit Hilfe eines authentischen Paper-Pencil-Tests erfasste Testleistung mit jener in der Entwicklungsumgebung als konvergent valide interpretiert werden kann.

Methode: Im Rahmen des Beitrages wird das Validierungsdesign vorgestellt, das auf einem Cross-over-Design basiert und um Positions- und Carry-Over-Effekte kontrollieren zu können auf lateinische Quadrate zurückgreift. Es wurden acht komplexe konstruktive Problemstellungen, basierend auf curricularen Analysen und in Abstimmung mit Experten beruflicher Ausbildungspraxis, entwickelt. Die Hypothesenprüfung erfolgt auf der Basis von Datensätzen des Ausbildungsberufs Elektroniker/in für Automatisierungstechnik (N=300) am Ende ihrer Ausbildung. Es wird als Kriterium der konvergenten Validität überprüft, ob ein Zusammenhang größer gleich .9 zwischen Entwicklungsumgebung und Paper-Pencil-Test erreicht wird.

Ergebnisse: Aktuell liegen Ergebnisse der ersten Haupterhebungsphase vor. Die untersuchten Auszubildenden (N=145) waren überwiegend männlich (87,8%), im Mittel 18,9 Jahre alt (M=18,9 SD=1,6). Sie besitzen als höchsten Schulabschluss einen mittleren Bildungsabschluss (84,5%). Bei der ersten Gruppe (zuerst Paper-Pencil-Test dann Problemstellungen in der Entwicklungsumgebung bearbeitet) zeigt sich beim 2-Faktormodell eine latente Korrelation von .68***. In einem nächsten Schritt wurde die Korrelation zwischen den beiden latenten Variablen auf .9 fixiert. Beide Modelle besitzen einen guten Modellfit (z.B. $X^2/df=0.93$, CFI=1.00, RMSEA=0.000 [0.000-0.055], WRMR=0.628). Beim Modellvergleich ergab sich $\Delta X^2(1)=3.329$ ($p=0.0681$), $\Delta CFI=.00$. Das heißt, das strengere Modell mit der fixierten Korrelation weicht nicht signifikant von den Daten ab. Bei der zweiten Gruppe (zuerst Entwicklungsumgebung dann Paper-Pencil-Test) ergab sich beim 2-Faktormodell eine latente Korrelation von .79***. Es wurde auch hier in Anlehnung an das Vorgehen bei Gruppe 1 ein Modelltest durchgeführt. Das strengere Modell mit der fixierten Korrelation weicht auch hier nicht signifikant von den Daten ab ($\Delta X^2(1)=2.180$ mit $p=0.1398$, $\Delta CFI=.00$). Wie der Modellvergleich der zwei Gruppen gezeigt hat, kann die im Paper-Pencil-Test erzielte Testleistung mit jener in der Entwicklungsumgebung als konvergent valide interpretiert werden.

Gegenwärtig erfolgt die Dateneingabe der zweiten Haupterhebung vom April 2014. Bis zur Konferenz im Frühjahr 2015 wird die Auswertung weiterer 160 Probanden abgeschlossen sein, sodass der vollständige Datensatz zur Hypothesenprüfung zur Verfügung steht.

Mittwoch
11. März

16:15 Uhr

HGC 50

Chair(s)
Winther, Esther

Schmidt, Juliane; Klusmann, Uta

Was macht den Alltag von Lehrkräften belastend? Eine Tagebuchstudie zu Häufigkeiten und Effekten von täglichen Stressoren und Ressourcen

Der Berufseinstieg von Lehrkräften wird als eine Phase mit hoher Beanspruchung beschrieben (Fives, Hamman & Olivarez, 2007). Ohnehin zählt der Lehrerberuf als Berufsgruppe, in der Burnout und psychische Erkrankungen häufig auftreten (Weber, Weltle & Lederer, 2004). Das „Job Demands-Resources Model“ (JD-R Modell; Demerouti, Bakker, Nachreiner, & Schaufeli, 2001) erklärt die Entstehung von Burnout bzw. das Leitsymptom der emotionalen Erschöpfung durch ein Ungleichgewicht von arbeitsbezogenen Ressourcen und Stressoren. Zahlreiche Forschungsarbeiten haben es sich zum Ziel gesetzt, berufsspezifische Ressourcen und Stressoren zu identifizieren, die das Erleben von Stress und Burnout im Lehrerberuf erklären können. Bisherige Befunde zeigen, dass auf Ebene des beruflichen Umfeldes die soziale Unterstützung (Klusmann, Kunter, Trautwein, Lüdtke & Baumert, 2008), eine positive Beziehung zu den Schülerinnen und Schülern (Klassen, Perry & Frenzel, 2012) sowie ein pädagogischer Konsens im Kollegium (vgl. Rothland & Klusmann, 2012) protektive Faktoren bilden, während Schülerfehlverhalten (Geving, 2007) sowie administrative und organisatorische Aufgaben (vgl. Kyriacou, 2001) potentielle Stressoren darstellen. Bisherige Forschungsarbeiten verstehen Ressourcen und Stressoren dabei allerdings hauptsächlich als stabile und überdauernde Faktoren und betrachten zudem lediglich Ausschnitte aus der beruflichen Tätigkeit der Lehrkraft.

Die aktuelle Studie konzentriert sich auf das tägliche Erleben von Ressourcen, Stressoren und Wohlbefinden von Lehrkräften. Ziel ist es, tägliche Ressourcen und Stressoren in den Tätigkeitsbereichen der Lehrerinnen und Lehrer zu beschreiben, die Variabilität der Ressourcen und Stressoren aufzuzeigen und Unterschiede in der täglichen Beanspruchung der Lehrkräfte durch die täglichen Ressourcen und Stressoren zu erklären.

Die Analysen basieren auf einer Tagebuchstudie, in der 141 Lehrkräfte in den ersten vier Berufsjahren 14 Tage lang täglich via Internet Angaben zu ihren positiven Ereignissen (Ressourcen) und negativen Ereignissen (Stressoren) sowie zur emotionalen Erschöpfung machten. Die täglichen Ressourcen und Stressoren wurden, einem idiographisch-nomothetischen Ansatz folgend, zunächst offen von den Lehrkräften berichtet und anschließend – entsprechend der Tätigkeitsbereiche von Lehrkräften – kategorisiert. Daraus ergaben sich acht Kategorien: Unterrichten, Vor- und Nachbereitung, Erziehen (außerunterrichtliche Interaktion mit Schülerinnen und Schülern), Beratung von Eltern, Interaktion mit Kollegen, Fortbildung und Organisation sowie Sonstiges.

Die Ergebnisse zeigen, dass die Lehrkräfte besonders häufig tägliche Ressourcen und Stressoren im Bereich des Unterrichts erleben, wobei Ereignisse, die die persönliche Beziehung zu den Schülerinnen und Schülern thematisieren, eine wichtige Ressource und Ereignisse, die die Klassenführung betreffen, einen entscheidenden Stressor darstellen. Lediglich im Bereich der Organisation nennen die Lehrerinnen und Lehrer mehr Stressoren als Ressourcen. Die Intraklassenkorrelation zeigt, dass ein substantieller Anteil der Varianz der Ressourcen und Stressoren auf Tagesebene liegt. Die Mehrebenenanalyse zeigt, dass tägliche positive Ereignisse einen negativen Effekt und tägliche negative Ereignisse einen positiven Effekt auf die emotionale Erschöpfung haben.

Mittwoch
11. März

16:15 Uhr

GC 03/142

Chair(s)

Schmidt, Juliane

Rollett, Wolfram; Scharenberg, Katja; Bos, Wilfried

Effekte der Klassenkomposition auf Beanspruchung und Berufszufriedenheit von Grundschullehrkräften

In der empirischen Forschung zum Lehrerberuf ist die schulische Beanspruchung von Lehrkräften einer der aktuellen Arbeitsschwerpunkte (z.B. Rothland, 2012; Terhart, Bennewitz, & Rothland, 2011). Dabei sind die Auswirkungen berufsbezogener Belastungen nicht nur hinsichtlich des persönlichen Leidensdrucks der Betroffenen bedeutsam. Eine geringere Belastungen bzw. Beanspruchung werden auch als wichtige Aspekte der Schulqualität verstanden (Holtappels, Kamski & Schnetzer, 2009): Einerseits können sich Beanspruchungen nachteilig auf wichtige Aspekte der schulischen Unterrichts- und Organisationskultur auswirken (z.B. Klusmann u.a., 2006, Lossen, Rollett & Willems, 2013). Andererseits hängt der Erfolg von Schulentwicklungsmaßnahmen auch davon ab, inwiefern Überlastungen bei den beteiligten Personen vermieden werden können (Böhm-Kasper, u.a. 2001).

In der Lehrerbelastungsforschung wird aus arbeitswissenschaftlicher Perspektive zunehmend auch die Bedeutung von Charakteristika des Arbeitsplatzes, der Arbeitstätigkeit und der organisatorischen Rahmenbedingungen der Arbeit für Gesundheit, Wohlbefinden und Leistungsfähigkeit von Lehrkräften thematisiert (Krause, Dorsewagen, & Alexander, 2011). Bislang weitgehend unbeachtet ist dabei die Frage, welche Bedeutung die Zusammensetzung der jeweils unterrichteten Schülerschaft für die Beanspruchung von Lehrkräften hat. Dies ist überraschend, da die Frage, wie mit der Unterschiedlichkeit von Schülerinnen und Schülern umgegangen werden soll, ein Schlüsselproblem der beruflichen Tätigkeit von Lehrkräften ist, das auch hohe Anforderungen an das professionelle Lehrerhandeln stellt (Trautmann & Wischer, 2011).

Mit Rückgriff auf das Erfurter Modell schulischer Belastung und Beanspruchung (Böhm-Kasper u.a., 2001) geht der vorliegende Beitrag daher den folgenden Forschungsfragen nach: (1) Welche Bedeutung haben verschiedene Dimensionen der Klassenkomposition als objektive Tätigkeitsanforderungen des Lehrerberufs für die berufsbezogene Beanspruchung von Lehrkräften? (2) Werden ggf. nachweisbare Effekte der Schülerkomposition den Annahmen der Belastungs-Beanspruchungssequenz (Rudow, 1994) entsprechend über subjektive Deutungsmuster bzw. Bewertungen der objektiven Belastung durch die Lehrkräfte vermittelt?

Als Analysegrundlage dienen dabei Daten von Hamburger Grundschullehrkräften der Schulleistungsstudie „Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 4“ (KESS4, Bos & Pietsch, 2006). Für 283 Deutschlehrkräfte konnten die Kompositionsmerkmale der von ihnen jeweils unterrichteten Schulklasse (insgesamt 4.624 Schülerinnen und Schüler) bestimmt werden.

Zur Messung der zeitlichen und psychophysischen Beanspruchung wurden Skalen aus dem Erfurter Belastungsinventar (EBI, Böhm-Kasper u.a., 2000) eingesetzt. Dies gilt auch für die Skala Berufszufriedenheit, die als Indikator für subjektive Deutungsprozesse dient (Böhm-Kasper, 2004). Zur Erfassung der Klassenkomposition wurden das mittlere Leistungsniveau und die Leistungsheterogenität im Leseverständnis, die mittlere soziale Zusammensetzung und soziale Heterogenität (HISEI) sowie der Migrantanteil herangezogen. Zudem wurden individuelle Hintergrundvariablen der Lehrkräfte (Geschlecht, Berufsjahre, Beschäftigungsstatus) berücksichtigt. Zur Analyse der Daten wurden schrittweise multiple lineare Regressionsanalysen durchgeführt. Fehlende Werte wurden mit einem EM-Algorithmus in SPSS 22 (IBM Corporation, 2013) imputiert. Allen Analysen wurde ein Signifikanzniveau von $p < .05$ zugrunde gelegt.

In den Analysen zeigen sich für die Leistungsheterogenität der Schulklassen kleinere ungünstige Effekte auf die Berufszufriedenheit der Lehrkräfte ($\beta = -.16$) und ihre psychophysische Beanspruchung ($\beta = .14$). Letztere werden allerdings vollständig über die Berufszufriedenheit vermittelt. Eine ungünstigere soziale Klassenkomposition hat einen nachteiligen Effekt auf die zeitliche Beanspruchung ($\beta = -.22$). Unter Kontrolle der Berufszufriedenheit werden nachteilige Effekte einer ungünstigeren sozialen Schülerkomposition auch auf die zeitliche und psychophysische Beanspruchung aufgedeckt ($\beta = -.25$ und $-.18$). Dies weist darauf hin, dass der Berufszufriedenheit nicht nur eine mediierende, sondern auch eine kompensatorische Rolle zukommt.

Insgesamt zeigen die Befunde, dass es sich bei bestimmten Dimensionen der Klassenkomposition (Leistungsheterogenität, mittlere soziale Lage) um psychologisch wirksame Belastungsfaktoren im Sinne der Belastungs-Beanspruchungssequenz handelt. Die nachgewiesenen Effekte sind erfreulicherweise aber nur von kleiner Größenordnung. Die Befunde unterstützen die Annahme der Belastungs-Beanspruchungssequenz, dass subjektive Deutungsprozesse eine Mediationsfunktion für die Auswirkungen belastender Bedingungen auf die Beanspruchung haben. Bedeutsam sind aber auch die Nullbefunde: So haben sich für das mittlere Leistungsniveau, die soziale Heterogenität und den Migrantanteil keine nachteiligen Effekte auf die Berufszufriedenheit oder die Beanspruchung ergeben. Offensichtlich gehen die Lehrkräfte mit diesen Aspekten der Klassenkomposition in ihrem professionellen Handeln in einer Weise um, dass sie keine beanspruchungsrelevanten Belastungsfaktoren darstellen.

Mittwoch
11. März

16:15 Uhr

GC 03/142

Chair(s)
Schmidt, Juliane

Denger, Corinna; Schüle, Christoph; Besa, Kris-Stephen; Arnold, Karl-Heinz

Berufswahlmotivation, Lehrerselbstwirksamkeit und Belastungserleben von Lehramtsstudierenden

Entsprechend der Theorie der Ressourcenerhaltung (COR-Theorie, Buchwald & Hobfoll, 2004) entsteht Stress, wenn die Ressourcen einer Person schneller verbraucht werden, als sie ersetzt werden können. Mündet dieser Ressourcenverlust in eine Verlustspirale, resultiert der COR-Theorie folgend Burn-Out. Um sich vor einem derartigen Ressourcenverlust schützen zu können, müssen Individuen bereit sein, in ihre Ressourcen zu investieren. Dies setzt bei Lehramtsstudierenden eine Investitionsbereitschaft zu Beginn ihres Studiums voraus, mit denen sie sich notwendige Ressourcen für ihre spätere Lehrtätigkeit aneignen können. Als besondere Ressource zur Bewältigung von Stress und Belastung wird hierbei die individuelle Selbstwirksamkeitserwartung einer Person diskutiert (Schwarzer & Schmitz, 1999). Bandura (1977) postuliert vier Quellen wahrgenommener Selbstwirksamkeit, was wiederum eine Motivation zum Aufsuchen derartiger Quellen voraussetzt (Tönjes & Dickhäuser, 2009).

Nach der COR-Theorie kann vermutet werden, dass Studierende des Lehramtes aufgrund geringer intrinsischer und hoher extrinsischer Motivationsfacetten weniger bereit sind Quellen wahrgenommener Selbstwirksamkeit aufzusuchen, weshalb sie weniger Ressourcen aufbauen und bereits im Schulpraktikum eine erhöhte Belastung erleben. Im vorliegenden Beitrag wird geprüft, inwiefern die Berufswahlmotivation zu Beginn des Studiums die Entwicklung der Selbstwirksamkeitserwartung im Rahmen von Schulpraktika vorhersagt und so das Belastungserleben als Lehrkraft auch in weiteren praxisbezogenen Studienelementen beeinflusst.

An einer Stichprobe von N = 443 Studierenden wurde die Berufswahlmotivation zu Beginn des Studiums im Jahr 2010 (t1) mit dem FEMOLA (Pohlmann & Möller, 2010) erfasst. Anderthalb Jahre später wurde die Lehrerselbstwirksamkeitserwartung der Studierenden im Anschluss an ein Allgemeines Schulpraktikum (t2) mithilfe der von Schwarzer und Jerusalem (1999) entwickelten Skala erhoben. Im letzten Mastersemester wurde innerhalb eines Fachpraktikums darüber hinaus das wahrgenommene Belastungserleben der Studierenden in ihrer Rolle als Lehrkraft (t3) mittels der Skala zur beruflichen Belastung von Enzmann und Kleiber (1989) erfasst.

Im Ergebnis eines latenten Regressionsmodells ($\chi^2_{48} = 80.31$; $p \leq .01$; CFI = .96; RMSEA = .04; SRMR = .05) beeinflussen sowohl die extrinsische Berufswahl zu t1 ($\beta = .37$; $p \leq .01$) als auch die Lehrerwirksamkeitserwartung zu t2 ($\beta = -.55$; $p \leq .01$) das Belastungserleben der Studierenden in einem Fachpraktikum am Ende ihres Studiums (t3). Darüber hinaus kann ein indirekter Effekt der intrinsischen Berufswahlmotivation zu t1 auf das Belastungserleben zu t3 nachgewiesen werden, welcher durch die individuelle Lehrerselbstwirksamkeitserwartung zu t2 vermittelt wird.

Entsprechend der COR-Theorie zeigt die vorliegende Untersuchung, dass eine positive intrinsische Motivation den Aufbau bedeutsamer Ressourcen im Verlauf des Studiums erleichtert und sich positiv auf das Belastungserleben im Lehrerberuf auswirken kann. Eine hohe extrinsische Berufswahlmotivation begünstigt demgegenüber fortschreitende Überlastung bereits im geschützten Raum eines universitären Fachpraktikums des letzten Mastersemesters.

Mittwoch
11. März

16:15 Uhr

GC 03/142

Chair(s)
Schmidt, Juliane

Wach, F.-Sophie; Karbach, Julia; Brünken, Roland; Spinath, Frank M.

Warum sind manche Studierende zufriedener als andere? Ergebnisse einer Längsschnittstudie

Theoretischer Hintergrund: Die Zufriedenheit von Studierenden mit ihrem Studium gilt als wichtiger Bestandteil einer erfolgreichen universitären Ausbildung, auch weil das Konstrukt mit einer Vielzahl anderer wichtiger Erfolgsmaße, wie Studienerfolg (Bean & Bradley, 1986), Verbleib im Studium (Aitken, 1982; Starr, Berth & Menne, 1972) und Belastungstoleranz (Schiefele & Jacob-Ebbinghaus, 2006) zusammenhängt. Bezüglich der strukturellen Konzeption unterscheiden Westermann, Heise, Spies und Trautwein (1996; cf. Schiefele & Jacob-Ebbinghaus, 2006) drei Facetten der Studienzufriedenheit: (1) die Zufriedenheit mit den Studieninhalten, (2) die Zufriedenheit mit den Studienbedingungen und (3) die Zufriedenheit mit der individuellen Bewältigung von Studienbelastungen. Darüber hinaus wird angenommen, dass die Zufriedenheit der Studierenden durch eine Vielzahl von Variablen determiniert wird. So bestätigen Befunde zur Studienzufriedenheit bzw. dem verwandten Konzept der Arbeitszufriedenheit die Wichtigkeit von Persönlichkeitsvariablen (emotionale Stabilität, Gewissenhaftigkeit, Extraversion; Judge, Heller & Mount, 2002; Lounsbury, Saudargas, Gibson & Leong, 2005). Auch Zusammenhänge zu motivationalen Konstrukten wie Interesse (Heise, Westermann, Spies & Schiffler, 1997) und Berufswahlmotiven (Hiemisch, Westermann, & Michael, 2005) konnten in Studien gezeigt werden. Apenburg (1980) diskutiert darüber hinaus die Relevanz der Studienleistungen für die Studienzufriedenheit.

Dennoch ist der Forschungsbereich aufgrund des Mangels an Replikationsstudien, des Fehlens einer allgemein etablierten Definition der Studienzufriedenheit und der großen Vielfalt an Erhebungsinstrumenten (cf. Reed, Lahey & Downey, 1984) noch immer von widersprüchlichen Ergebnissen und mangelnder Vergleichbarkeit der Studien gekennzeichnet.

Zielsetzung: Ziel der vorliegenden Arbeit war es daher, relevante Determinanten der drei Studienzufriedenheitsfacetten zu identifizieren und deren prädiktive Validität zu bestimmen. Dabei wurden gleichzeitig Variablen aus dem Persönlichkeits-, Motivations- und Leistungsbereich berücksichtigt.

Methode: Es wurden längsschnittliche Daten der „Studie zu individuellen und organisationalen Einflüssen auf Studien- und Ausbildungserfolg in der Lehrerbildung (SioS-L)“ analysiert. Im Rahmen des Forschungsprojekts bearbeiteten 620 Lehramtsstudierende in ihrem ersten Semester (t1, Alter: M=20,77; SD=3,22; 62,7% weiblich) Fragebögen zu demografischen Informationen (u.a. Alter, Geschlecht, höchster Bildungsabschluss der Eltern), motivationalen Aspekten (berufliche Interessensorientierungen, akademisches Selbstkonzept, Leistungsmotivation, Selbstregulation, Berufswahlmotive) und ihrer Persönlichkeit (Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit). Darüber hinaus wurden sowohl ihre kognitiven Fähigkeiten als auch ihre Studienleistungen erfasst. Nach durchschnittlich zwei Jahren (t2) erfolgte die Erhebung der Studienzufriedenheit sowie der allgemeinen Lebenszufriedenheit und der individuellen Bedeutsamkeit des Studiums (n=255).

Vor der Datenanalyse wurden Prädiktoren und Kriterien um den Einfluss des Alters, die Kriterien zusätzlich um den Einfluss der Lebenszufriedenheit korrigiert. Signifikante Korrelationen auf manifester Ebene führten zu einer Berücksichtigung des entsprechenden Prädiktors in den sich anschließenden Strukturgleichungsanalysen zur Vorhersage der drei Studienzufriedenheitsfacetten.

Ergebnisse: Das akademische Selbstkonzept, die Leistungsmotivation, das fachliche Interesse als intrinsisches Berufswahlmotiv, Neurotizismus, Gewissenhaftigkeit und die individuelle Bedeutsamkeit des Studiums wiesen auf manifester Ebene signifikante Korrelationen mit der Zufriedenheit mit den Studieninhalten auf. In einem Strukturgleichungsmodell, in dem auch Korrelationen zwischen den Prädiktoren berücksichtigt wurden, leisteten jedoch nur Neurotizismus und das fachliche Interesse einen signifikanten Beitrag zur Vorhersage dieser Facette der Studienzufriedenheit.

Auch die Zufriedenheit mit den Studienbedingungen korrelierte auf manifester Ebene mit einer Vielzahl von Variablen (Geschlecht, künstlerisch-sprachliche und unternehmerische Interessensorientierung, verschiedene Berufswahlmotive [pädagogisches Interesse, Fähigkeitsüberzeugung, geringe Schwierigkeit], Neurotizismus und Studienleistungen). Im Strukturgleichungsmodell konnte jedoch nur der Einfluss von Neurotizismus bestätigt werden.

Neurotizismus erwies sich auch als signifikanter Prädiktor der Zufriedenheit mit der Bewältigung von Studienbelastungen. Ein signifikanter Zusammenhang mit dem Geschlecht zeigte sich nur auf manifester Ebene.

Zusammenfassend betonen die Ergebnisse die Wichtigkeit des Persönlichkeitstraits Neurotizismus für alle drei Facetten der Studienzufriedenheit. Studierende mit einer geringeren emotionalen Stabilität berichteten demnach auch eine geringere Zufriedenheit mit den Inhalten sowie den Bedingungen ihres Studiums und erwiesen sich auch als unzufriedener mit ihrer individuellen Bewältigung von Studienbelastungen.

Mittwoch
11. März

16:15 Uhr

GC 03/142

Chair(s)
Schmidt, Juliane

Gläser-Zikuda, Michaela; Hagenauer, Gerda; Moschner, Barbara; Bernholt, Andrea; Dunker, Nina; Klaß, Susi; Lüschen, Iris; Wolf, Nicole

Individuelle und kontextuelle Prädiktoren der Studienzufriedenheit von Lehramtsstudierenden

Die Studie untersucht Prädiktoren für Studienzufriedenheit von Lehramtsstudierenden auf individueller sowie kontextueller Ebene (vgl. Blüthmann, 2012; Künsting & Lipowsky, 2011).

Studienzufriedenheit wird nach Westermann et al. (1996) als multidimensionales Konstrukt definiert, wobei zwischen (1) der Zufriedenheit mit den Studieninhalten, (2) den Studienbedingungen und (3) der Zufriedenheit mit der Bewältigung von Studienbelastungen unterschieden wird. Schiefele und Jabob-Ebbinghaus (2006) zeigten in einer Studie mit Psychologiestudierenden, dass jede dieser Dimensionen durch verschiedene Faktoren bedingt wird. Allerdings vernachlässigen die meisten Studien diese Multidimensionalität der Studienzufriedenheit (z.B. Garcia-Aracil, 2012).

Die vorliegende Studie trägt zur Weiterentwicklung der Forschung im Bereich Studienzufriedenheit bei. Sie wird der Komplexität dieses Forschungsbereichs besser gerecht und betrachtet zunächst die oben erwähnten drei Dimensionen der Studienzufriedenheit differenziert. Studienzufriedenheit kann in Hinblick auf die zunehmende Kompetitivität zwischen den Hochschulen, im Sinne der Kundenorientierung, auch als ökonomischer Faktor verstanden werden (z.B. Gruber et al., 2010). Die Qualität der in der Lehrerbildung evozierten Lernprozesse bei Lehramtsstudierenden dürfte sich langfristig als eine wesentliche Bedingung für die Lehr- bzw. Unterrichtsqualität an Schulen herausstellen. Schließlich ist davon auszugehen, dass bezogen auf salutogene Faktoren die Förderung von Selbstregulation bei Lehramtsstudierenden zu höherem Wohlbefinden (Mattern & Bauer, in press) und damit letztlich zu einer höheren Studienzufriedenheit führt.

An unserer paper pencil Befragung in regulären Hochschulveranstaltungen nahmen 792 Lehramtsstudierende (davon 71.4% weiblich, durchschnittliche Semesteranzahl = 3.48; mittleres Alter = 22.62 Jahre) von 5 Universitäten (80.7%) bzw. Pädagogischen Hochschulen (19.3%) aus verschiedenen Bundesländern teil. Die Studienzufriedenheit wurde nach Westermann et al. (1996) durch drei Faktoren erfasst: (1) Zufriedenheit mit den Inhalten ($\alpha=.86$), (2) Zufriedenheit mit den Bedingungen ($\alpha=.77$), und (3) Zufriedenheit mit der Bewältigung von Studienbelastungen ($\alpha=.70$). Die Prädiktorvariablen wurden ebenfalls mit bekannten Skalen erhoben und geringfügig angepasst. Auf der individuellen Ebene wurden z.B. als Persönlichkeitsfaktoren die Big5 (Kauper et al., 2012; $\alpha>.62$), Selbstwirksamkeit (Schiefele et al., 2002; $\alpha=.90$), Lernmotivation (Spinath et al., 2002; z.B. Vermeidungsleistungszielorientierung, $\alpha=.91$) und Selbstregulation (Schwarzer & Jerusalem, 1999; $\alpha=.83$) erhoben. Auf der Kontextebene der Lehr- und Studienbedingungen wurden u.a. folgende Variablen erfasst: Qualität des Theorie-Praxis-Bezugs (Schiefele et al., 2002; $\alpha=.79$), studierendenzentrierter (; $\alpha=.73$) und lehrendenzentrierter Lehransatz (Johannes et al., 2011; $\alpha=.59$), Leistungsdruck (Schiefele et al., 2002; $\alpha=.73$), Autonomieunterstützung (Kauper et al., 2012; $\alpha=.80$) und Soziale Eingebundenheit (Kauper et al., 2012; Schiefele et al., 2002; $\alpha=.89$). Schließlich wurden Kontextfaktoren auf Ebene der Institution Hochschule erhoben, wie z.B. die Studierbarkeit (Kauper et al., 2012, $\alpha=.77$), oder die Informations- und Beratungsangebote (Greiner, 2009; $\alpha=.85$).

Hierarchische Regressionsanalysen wurden in 6 Schritten berechnet (Multikollinearität und Autokorrelationen lagen nicht vor). Die drei verschiedenen Dimensionen von Studienzufriedenheit konnten durch verschiedene Indikatoren vorhergesagt werden:

Die Zufriedenheit mit den Studieninhalten wurde mit einer aufgeklärten Varianz von .43 (adj. $\eta^2=.42$) am besten vorausgesagt. Stärkste Prädiktoren waren unter anderem die Qualität des Theorie-Praxis-Bezugs ($\beta=.26$), Lernerfolgsorientierung ($\beta=.17$), Vermeidungsleistungszielorientierung ($\beta=-.14$) und soziale Eingebundenheit ($\beta=.14$).

Die Zufriedenheit mit den Studienbedingungen ($\eta^2=.32$; adj. $\eta^2=.30$) wurde - wie erwartet - durch Faktoren der Studienumgebung vorhergesagt: z.B. durch die Machbarkeit des Studiums ($\beta=.15$), die Qualität der Bibliothek ($\beta=.08$), die Ausstattung der Veranstaltungsräume ($\beta=.10$) und die Qualität des Theorie-Praxis-Bezugs ($\beta=.14$).

Schließlich wurden 38% der Varianz der Zufriedenheit mit der Bewältigung von Studienbelastungen aufgeklärt (adj. $\eta^2=.35$). Prädiktoren sind z.B. Leistungsdruck ($\beta=-.36$), Autonomieunterstützung ($\beta=.10$) und Studierbarkeit ($\beta=.08$).

Ergebnisse der Untersuchung werden in Hinblick auf bisherige Studien dieses Forschungsschwerpunktes im Allgemeinen reflektiert und vor allem mit Blick auf die Bedeutung des Theorie-Praxis Bezugs für die Studienzufriedenheit von Lehramtsstudierenden und deren Rolle bzgl. einer Weiterentwicklung der Lehrerbildung diskutiert.

Mittwoch
11. März

16:15 Uhr

GC 03/142

Chair(s)
Schmidt, Juliane

Vogel, Freydis; Kollar, Ingo; Wecker, Christof; Fischer, Frank

Effekte von computerunterstützten Kollaborationsskripts auf domänenspezifisches Wissen: Eine Metaanalyse

Theoretischer Hintergrund: Für den erwarteten positiven Effekt von kollaborativem Lernen in Kleingruppen auf den domänenspezifischen Wissenserwerb ist es wichtig, dass die Lernenden die domänenspezifischen Inhalte vertieft kognitiv elaborieren (King, 2007). Im kollaborativen Lernprozess zeigen Lernende jedoch eher selten selbstständig die hierfür entscheidenden Aktivitäten (Cohen, 1994). Um kollaboratives Lernen zu fördern, werden daher computerunterstützte Kollaborationsskripts eingesetzt, die Lernende einer Kleingruppe dazu auffordern, in einer bestimmten Sequenz festgelegte Aktivitäten durchzuführen, die meist an Rollen gebunden sind, welche an die Lernenden verteilt werden (Kollar, Fischer, & Hesse, 2006). Zur Sequenzierung der Aktivitäten und Verteilung von Rollen eignet sich der Einsatz computerunterstützter Lernumgebungen. Computerunterstützte Kollaborationsskripts wurden für den Wissenserwerb in unterschiedlichen Domänen entwickelt (z.B. Informatik, Psychologie und Medizin). Insgesamt zeigt die Forschung jedoch ein heterogenes Bild bezüglich der Effekte von Kollaborationsskripts auf den domänenspezifischen Wissenserwerb (z.B. Rummel, Spada, & Hauser, 2009; Stegmann, Weinberger, & Fischer, 2007). Daher stellt sich die Frage, wie gut derartige Skripts generell dazu geeignet sind, domänenspezifisches Wissen zu fördern, und welche kollaborativen Aktivitäten von Kollaborationsskripts angeregt werden sollten, damit sich positive Effekte von Kollaborationsskripts auf den domänenspezifischen Wissenserwerb einstellen. Auf Basis von Arbeiten von Chi (2009) kann etwa vermutet werden, dass Kollaborationsskripts umso effektiver für den domänenspezifischen Wissenserwerb sind, je mehr sie die Verwendung transaktiver Aktivitäten anregen (Fischer, Kollar, Stegmann, & Wecker, 2013), also solcher Aktivitäten, mit denen Lernende auf die Beiträge ihrer Lernpartner Bezug nehmen, indem sie diese etwa kritisieren, überarbeiten oder weiterentwickeln (Teasley, 1997).

Fragestellungen: (1) Welchen Effekt haben computerunterstützte Kollaborationsskripts auf das domänenspezifische Wissen der Lernenden? (2) Inwiefern wird dieser Effekt von der durch das Kollaborationsskript angeregten Transaktivität moderiert?

Methode: Zur Beantwortung der Fragestellungen führten wir eine Metaanalyse zum Lernen mit computerunterstützten Kollaborationsskripts durch. Zur Identifikation relevanter Studien wurde einer Recherche in bibliographischen Datenbanken (ERIC und ISI Web of Science) durchgeführt, die eine Ausgangsstichprobe von 248 Studien ergab. Folgende Inklusionskriterien wurden angelegt: (1) Die Studie musste ein computerunterstütztes Kollaborationsskript experimentell variieren und (2) ausreichende Angaben zur Schätzung der Effektstärke für domänenspezifisches Wissen enthalten. Zur Auswahl der inkludierten Studien kodierten drei unabhängige Kodierer die Studien gemäß den Inklusionskriterien (Fleiss' kappa = 0,72). Diese Kodierung ergab eine endgültige Stichprobe von 21 Studien, in denen 33 Effekte für domänenspezifisches Wissen berichtet werden. Um die Rolle der Transaktivität für die Effekte von Kollaborationsskripts auf den domänenspezifischen Wissenserwerb zu bestimmen, wurde kodiert, ob das Kollaborationsskript darauf ausgelegt war, transaktive Aktivitäten anzuregen oder nicht (Fleiss' kappa = 0,82). Die Kodierung ergab 6 Effekte für Kollaborationsskripts mit Transaktivität und 27 Effekte für Kollaborationsskripts ohne Transaktivität. Zur statistischen Analyse wurde das random-effects-Modell herangezogen (Borenstein et al., 2009). Das Signifikanzniveau wurde für alle Analysen auf 5% festgelegt.

Ergebnisse: Computerunterstütztes Lernen mit Kollaborationsskripts hat im Vergleich zu Lernen ohne Kollaborationsskript insgesamt einen signifikanten kleinen positiven Effekt auf das domänenspezifische Wissen, $d = 0.20$; $SE = 0.10$; $p = 0.02$. Für die Transaktivität von Kooperationsskripts konnte kein statistisch signifikanter Effektivitätsunterschied festgestellt werden, $Q(1) = 0.28$, $p = 0.60$. Deskriptiv wurde für transaktive Kollaborationsskripts ein statistisch nicht signifikanter Effekt geschätzt, $d = 0.05$; $SE = 0.31$; $p = 0.43$, während für nicht-transaktive Kollaborationsskripts ein signifikanter positiver Effekt ermittelt wurde, $d = 0.22$; $SE = 0.10$; $p = 0.02$.

Die Ergebnisse zeigen, dass Kollaborationsskripts domänenspezifisches Wissen fördern können. Sie geben Hinweise darauf, dass die Anregung transaktiver Aktivitäten mittels Kollaborationsskripts die kognitive Verarbeitung und damit den domänenspezifischen Wissenserwerb positiv beeinflussen kann (Fischer et al., 2013; King, 2007). Weitere Analysen sollten klären, inwiefern Kollaborationsskripts durch die Anregung transaktiver Aktivitäten tatsächlich kollaborative Lernaktivitäten beeinflussen und inwiefern die Effekte auch auf andere Lernerfolgsdaten übertragbar sind (z.B. Kollaborationsfertigkeiten). In zukünftigen Studien könnte dies durch eine einheitliche Auswertung und Berichterlegung zu kollaborativen Lernaktivitäten umgesetzt werden.

Mittwoch
11. März

16:15 Uhr

GC 03/149

Chair(s)

Vogel, Freydis

3

Stegmann, Karsten; Weinberger, Armin; Fischer, Frank

Mechanismen der argumentativen Wissenskonstruktion in Online-Diskussionen

Online-Diskussionen werden in der Hochschullehre eingesetzt, um die aktive Beteiligung von Studierenden in Vorlesungen und Seminaren zu verbessern (Petko, 2003). Implizit steht dahinter die Annahme, dass Lernende in Online-Diskussionen ihre Standpunkte formulieren und begründen, die Position ihrer Lernpartner verstehen und kritisieren und Konflikte lösen. Im Rahmenmodell der argumentativen Wissenskonstruktion in Online-Diskussionen von Weinberger und Fischer (2006) werden unterschiedliche Dimensionen von Online-Diskussionen spezifiziert, die den Wissenserwerb durch Online-Diskussionen erklären sollen:

Die Partizipation stellt eine Voraussetzung für argumentative Wissenskonstruktion in Online-Diskussionen dar. Die Partizipation bedeutet hier, wie häufig sich Lernende in einer Lernumgebung anmelden und in welchem Ausmaß sie dort aktiv sind. Die Qualität einzelner Argumente wird anhand des Modells der Argumentation von Toulmin (1958) konzeptualisiert. Die Qualität der Argumentsequenzen wird im vorliegenden Modell anhand des Argument-Gegenargument-Synthese-Schemas bestimmt. Transaktivität beschreibt den Grad, zu dem Lernende auf den Wissensbeiträgen eines anderen Lernenden aufbauen bzw. darauf Bezug nehmen. Die epistemische Aktivität beschreibt, welche Konzepte Lernende anwenden und miteinander verknüpfen, bzw. inwieweit Lernende vom Thema abschweifen, weniger relevante Aspekte der Lernumgebung koordinieren oder welche Fehlkonzepte sie konstruieren.

Vor dem Hintergrund des Modells der argumentativen Wissenskonstruktion wird in der Studie die Frage untersucht, inwiefern Partizipation, Transaktivität, Qualität der Argumentation (einzelne Argumente und Argumentsequenzen), epistemische Aktivitäten und individueller Wissenserwerb zusammenhängen. Es soll überprüft werden, inwiefern das Modell der argumentativen Wissenskonstruktion in Online-Diskussionen in der Lage ist den inhaltlichen Wissenserwerb zu erklären.

Für die Studie wurden 36 Online-Diskussionen von Dreiergruppen (N = 108 Lernende) analysiert. Die Teilnehmerinnen und Teilnehmer waren Erstsemester-Studierende der Pädagogik die im Rahmen einer Lehrveranstaltung an einer virtuellen Lernsitzung teilgenommen haben. Die Aufgabe der Lernenden war es, Problemfälle mit Hilfe einer psychologischen Theorie zu analysieren.

Ausgewertet wurden sowohl die individuellen Analysen nach der Online-Diskussion als auch die Online-Diskussionen der Dreiergruppen. Die individuellen Analysen und Online-Diskussionen wurden von trainierten Kodierern in Segmente unterteilt. Die Segmentierung wurde auf der Ebene von Propositionen durchgeführt. Bezüglich der Segmentierung erreichten die Kodierer eine zufriedenstellende Übereinstimmung von 84%. Diese Segmente wurden auf mehreren Dimensionen kodiert. Auf Grundlage der Kodierung wurden die folgenden Variablen gebildet (Cohens Kappa \geq .70): Partizipation (Anzahl Zeichen), epistemische Aktivitäten (Anzahl der korrekten Propositionen während der Onlinediskussion), domänenspezifisches Wissen (Anzahl der korrekten Propositionen im Nachtest), Transaktivität (Anzahl konflikt- und integrationsorientierter Konsensbildungen), Qualität der einzelnen Argumente (Anzahl der begründeten und/oder relativierten Behauptung) und Qualität der Argumentsequenzen (Anzahl der Gegenargumente).

Getestet wurde ein Pfadmodell in dem die Partizipation die Transaktivität, die Qualität der einzelnen Argumente und die Qualität der Argumentsequenzen vorhersagt. Die Qualität der einzelnen Argumente und die Qualität der Argumentsequenzen wurden wiederum als Prädiktoren für die epistemische Aktivität verwendet. Die epistemische Aktivität wurde schließlich als Prädiktor für den Wissenserwerb verwendet. Das Pfadmodell wurde mit Hilfe von AMOS überprüft. Das beschriebene Modell erreichte eine akzeptable Passung.

Die Partizipation sagt signifikant und mit jeweils großer Effektstärke die Transaktivität (27% Varianzaufklärung), die Qualität der einzelnen Argumente (23% Varianzaufklärung) und die Qualität der Argumentsequenzen (15% Varianzaufklärung) vorher. Die Transaktivität, die Qualität der einzelnen Argumente und die Qualität der Argumentsequenzen klären 36% der Varianz der epistemischen Aktivität auf. Die Qualität der einzelnen Argumente ist ein signifikanter Prädiktor mit einer großen Effektstärke. Die Transaktivität ist ebenfalls ein signifikanter Prädiktor, weist jedoch nur eine kleine Effektstärke auf. Die Qualität der Argumentsequenzen weist ebenfalls eine kleine, nicht signifikante Effektstärke auf. Die epistemische Aktivität klärt signifikant 11% der Varianz des domänenspezifischen Wissens im Nachtest auf.

Das Modell der Argumentativen Wissenskonstruktion erwies sich weitgehend geeignet, um die Zusammenhänge zwischen Prozessen und Ergebnissen der gemeinsamen Wissenskonstruktion zu beschreiben. Auf dieser Basis können instruktionale Ansätze entwickelt werden, die gezielt bestimmte Prozesse und so den Wissenserwerb fördern.

Mittwoch
11. März

16:15 Uhr

GC 03/149

Chair(s)
Vogel, Freydis

Brendel, Nina

Reflexionsprozesse bestimmen und fördern

Schüler/-innen auf eine globalisierte Gesellschaft vorzubereiten, ist ein zentrales Anliegen des Globalen Lernens im Fach Geographie. Ziel ist es dabei, durch Förderung von Systemkompetenz und Bewertungskompetenz die Lernenden zu persönlichem Urteilen und Handeln unter dem Leitbild der nachhaltigen Entwicklung zu befähigen (SCHRÜFER 2013).

Dabei schreibt der Orientierungsrahmen für den Lernbereich Globale Entwicklung der Reflexion im Bereich Bewerten einen hohen Stellenwert zu, jedoch ohne sie zu definieren oder operationalisieren (BMZ, KMK 2007). So postulieren GRYL und KANWISCHER (2011, S. 177): „Im Gegensatz zur herausgehobenen Bedeutung der Reflexion für kompetenzorientiertes Lernen ist die theoriegeleitete Entwicklung eines kompetenzorientierten Reflexionsmodells aus Sicht der Fachdidaktiken bislang ein Desiderat.“

Im Rahmen des vorgestellten Dissertationsprojektes wird daher der Forschungsfrage nachgegangen, welche Ausprägungen reflexiven Denkens bei Schüler/-innen im Rahmen Globalen Lernens bestimmt werden können und welche Faktoren Reflexionsleistungen beeinflussen. Dadurch soll Lehrenden ein Instrument an die Hand gegeben werden, Reflexionsleistungen von Schüler/-innen im Rahmen Globalen Lernens zu bestimmen und heterogene Reflexionsprozesse individuell zu fördern.

In den letzten Jahren haben sich besonders Weblogs als effektives Lernwerkzeug erwiesen, um reflexives Denken bei Lernenden anzuregen (DABBAGH & KITSANTAS 2011). Denn im Gegensatz zu herkömmlichen (analogen) Lerntagebüchern ermöglichen sie neben der individuellen Reflexion des einzelnen Schülers auch Kommunikation und Kollaboration innerhalb des Klassenverbandes oder darüber hinaus (z. B. durch Kommentarfunktionen) (ROSA 2013).

Auf dieser Grundlage wurde 2013 ein Projekt mit 5 Oberstufenkursen aus NRW (10. und 11. Klasse) durchgeführt. Über einen Zeitraum von 4-8 Wochen reflektierten 94 Schüler/-innen in persönlichen Weblogs über die Unterrichtsinhalte zum Globalen Lernen und kommentierten die Blogs ihrer Mitschüler. Um Rahmenbedingungen in den einzelnen Klassen zu erheben, wurden außerdem vorab leitfadengestützte Interviews mit den unterrichtenden Lehrpersonen zu zentralen Lehrkonzepten und Lernkultur erhoben. Zudem waren die Lehrkräfte dazu angehalten, das Unterrichtsgeschehen im Sinne eines reflective practitioners (ABELS 2011) in einem privaten Blog zu reflektieren.

Um verschiedene Stufen von Reflexion im Globalen Lernen unter den Schüler/-innen zu bestimmen, wurden alle rund 500 Blogbeiträge nach der Methode der Skalierenden Strukturierung (MAYRING 2010) unter Verwendung der anerkannten fünfstufigen Reflexionsskala nach BAIN ET AL (1999) codiert. Die Auswertung erfolgte in Synthese mit den Ergebnisse der Qualitativen Inhaltsanalyse (MAYRING 2010) der Lehrerinterviews sowie einer offenen Codierung der Schüler- und Lehrerblogs.

Erste Auswertungsergebnisse zeigen sehr heterogene Leistungsniveaus der Schüler/-innen auf Basis von Faktoren wie u.a. Unterrichtsmethoden, Fachinhalte, Geschlecht oder Lernkultur und lassen Strategien erkennen, wie individuell auf Heterogenität im Bereich Reflexion eingegangen werden kann.

Mittwoch
11. März

16:15 Uhr

GC 03/149

Chair(s)
Vogel, Freydis

Schlag, Myriam

Wie wirken sich computerbezogene Einstellungen von Lehramtsstudierenden auf deren Selbstregulation bei der Erstellung eines E-Portfolios in OneNote aus?

Theoretischer Hintergrund: Ein (E-)Portfolio ermöglicht es Lernenden, ihren Lernprozess zu strukturieren und zu dokumentieren, indem Dokumente gesammelt und reflektiert werden. Das (E-)Portfolio eignet sich zum einem zur Kompetenzförderung (Hornung-Prähauser, et al., 2007), zum anderen zum Umgang mit Heterogenität in Lerngruppen (Stratmann, 2009). Lernende unterscheiden sich in den Voraussetzungen und Lernwegen sowie in ihren computerbezogenen Einstellungen (Richter, Naumann & Horz, 2010). Bisherige Studien haben gezeigt, dass diese computerbezogenen Einstellungen (Erice & Ertaş, 2011) ebenso wie selbstreguliertes Lernen (Alexiou & Paraskeva, 2013; Cheng & Chau, 2013) mit dem Lernprozess beim E-Portfolio in Verbindung stehen. Weiterhin ist Computerängstlichkeit beim Einsatz von E-Portfolios zu berücksichtigen (Erice & Ertaş, 2011). Ziel dieser Untersuchung ist es, den Zusammenhang zwischen computerbezogenen Einstellungen, Computerängstlichkeit und selbstreguliertem Lernprozess der E-Portfolio Erstellung zu erfassen.

Fragestellungen: (a) Wie verläuft das selbstregulierte Lernen bei der Erstellung eines E-Portfolios bei Lehramtsstudierenden? (b) In welchem Zusammenhang stehen computerbezogene Einstellungen und selbstreguliertes Lernen mit einem E-Portfolio sowie dessen Akzeptanz? (c) Wie entwickelt sich die Computerängstlichkeit bei der Erstellung eines E-Portfolios und in welchem Zusammenhang steht sie zu computerbezogenen Einstellungen?

Methode: Die Teilnehmer der Studie besuchten im Wintersemester 2013/2014 ein Seminar im Rahmen der Bildungswissenschaften. Die Studierenden arbeiteten in jeder Seminarsitzung mit dem Programm Microsoft OneNote (Herzog & Kessler, 2014) an ihrem E-Portfolio. Sie erhielten jede Woche von der Dozentin einen vorstrukturierten Abschnitt mit den Folien der Veranstaltung, Arbeitsmaterialien und konnten dabei zwischen verschiedenen Aufgaben auswählen. Am Ende jeder Sitzung reichten die Studierenden ihren aktuellen OneNote-Abschnitt über eine Lernplattform ein.

Insgesamt 32 Lehramtsstudierende des OneNote-Kurses beantworteten einen Online-Fragebogen zu drei Messzeitpunkten. Neben der Abfrage demographischer Daten erfasste der Online-Fragebogen Skalen zum selbstregulierten Lernen. Dazu wurde der Fragebogen zur aktuellen Motivation (Rheinberg, Vollmeyer & Burns, 2001) eingesetzt. Die Computerängstlichkeit und die computerbezogenen Einstellungen wurden mit den entsprechenden Skalen aus dem Inventar zur Computerbildung von Richter, Naumann und Horz (2010) erhoben. Die Akzeptanz des Portfolios wurde mit einer eigenen Skala erfasst.

Ergebnisse: Die Werte des Fragebogens zur aktuellen Motivation zeigen stabile Verläufe bei drei aus vier Subskalen. Allein bei der Subskala Herausforderung zeigen sich signifikante Unterschiede ($F(2,62) = 3.76, p = .03, \eta^2 = .108$), wobei diese zunächst stabil blieb und dann zwischen der Mitte und dem Ende des Semesters abnahm. Die Computerängstlichkeit der Lehramtsstudierenden nahm vom Beginn zur Mitte des Semesters signifikant ab und blieb dann stabil ($F(1,22, 36.72) = 11.51, p = .001, \eta^2 = .277$).

Im Zusammenhang mit den computerbezogenen Einstellungen ergaben sich zu allen drei Messzeitpunkten signifikante Korrelationen mit dem selbstregulierten Lernen. Je mehr der Computer als nützliches Werkzeug für Arbeit und Lernen wahrgenommen wurde, desto höher war die Herausforderung (MZP1: $r = .36, p = .05$, MZP3: $r = .46, p = .01$) und das Interesse (MZP3: $r = .47, p = .01$) und desto geringer war die Computerängstlichkeit (MZP1: $r = -.37, p = .04$). Je stärker Studierende jedoch glaubten, es handle sich beim Computer um eine unbeeinflussbare Maschine zum Lernen und Arbeiten, desto höher waren die Misserfolgserwartungen (MZP1: $r = .50, p = .01$) und die Computerängstlichkeit (MZP2: $r = .78, p < .01$, MZP3: $r = .74, p < .01$) und desto geringer war das Interesse (MZP3: $r = -.49, p = .01$) und die Akzeptanz für das Portfolio (MZP3: $r = .44, p = .01$). Wurde der Computer als unbeeinflussbare Maschine zur Kommunikation und Unterhaltung wahrgenommen, gab es positive Korrelationen zur Computerängstlichkeit (MZP2: $r = .51, p = .01$, MZP3: $r = .51, p = .01$).

Die Ergebnisse untermauern die Annahme, dass auch computerbezogene Einstellungen bei der Arbeit mit E-Portfolios von Lernenden und Lehrenden beachtet werden müssen, damit diese Lern- und Arbeitsform erfolgreich sein kann.

Mittwoch
11. März

16:15 Uhr

GC 03/149

Chair(s)

Vogel, Freydis

Peter, Johannes; Leichner, Nikolas; Mayer, Anne-Kathrin; Krampen, Günter

Einfluss epistemologischer Überzeugungen auf die Förderung von Informationskompetenz

Theoretischer Hintergrund und Hypothesen: Die Bedeutung der Förderung von Informationskompetenz bzw. professionellem Informationssuchverhalten wird verstärkt im Hochschulkontext diskutiert (Hochschulrektorenkonferenz, 2012), da zahlreiche Studien belegen, dass die dafür erforderlichen Fähigkeiten unter Studierenden oftmals unzureichend sind (z. B. Heinze, 2008). Neben technischen Fähigkeiten (z.B. Schlagwortsuche in Fachdatenbanken) sind bei Informationssuchen insbesondere metakognitive Prozesse bedeutsam. Diese werden vor allem durch individuelle Konzepte über die Beschaffenheit von Wissen und Wissensprozessen bzw. epistemologische Überzeugungen beeinflusst (Hofer, 2004).

Die individuelle Entwicklung derartiger Überzeugungen zeichnet sich durch den Wandel einfacher objektivistischer zugunsten komplexer, zunächst ebenfalls undifferenzierter, dann jedoch zunehmend differenzierter subjektivistischer Positionen aus (Kuhn, 2000). Verschiedene Studien zeigen, dass Individuen mit komplexeren epistemologischen Überzeugungen bei der Informationsbewertung eher dazu neigen, Inhalte differenziert zu vergleichen und die Glaubwürdigkeit von Quellen kritisch zu hinterfragen (z. B. Mason, Boldrin & Ariasi, 2010). Über den Einfluss epistemologischer Überzeugungen auf die Förderung professionellen Informationssuchverhaltens ist hingegen vergleichsweise wenig bekannt. Es ist jedoch zu erwarten (Hypothese 1), dass Individuen mit undifferenzierten objektivistischen Positionen den Nutzen des Vergleichs verschiedener Quellen weniger gut erkennen und daher weniger geneigt sind, Lerninhalte zu akzeptieren, die sich mit dem Auffinden und dem Bewerten von Quellen beschäftigen (vgl. Hofer, 2004). Darüber hinaus ist anzunehmen (Hypothese 2), dass undifferenzierte subjektivistische Positionen weniger mit der deterministischen Natur des für Informationssuchen erforderlichen technischen Wissens vereinbar sind (vgl. Elby & Hammer, 2001).

Methode: Die beiden Hypothesen wurden im Kontext zweier Evaluationsstudien zu einem Blended-Learning Training der Informationskompetenz geprüft, an denen $N = 67$ bzw. $N = 64$ Studierende der Psychologie an der Universität Trier teilnahmen. In den Trainings wurden theoretische Grundlagen zu Informationssuchen vermittelt und mit praktischen Übungen verbunden. In Training 1 wurden in den praktischen Übungen standardisierte Aufgaben verwendet und der Lerngewinn (prä-post) durch einen Informationskompetenz-Wissenstest (Leichner, Peter, Mayer & Krampen, 2013; $k = 35$ Items; $\alpha = .63/.55$) erhoben. In Training 2 recherchierten die Teilnehmenden in den praktischen Übungen zu eigenen Themen (z. B. BSc-/MSc-Arbeit). Der Lerngewinn wurde hier (prä-post) mittels einer subjektiven Skala zur Erfassung Informationskompetenz-bezogenen thematischen Wissens erhoben (z. B. „Ich könnte spontan mehrere Zeitschriften nennen, in denen regelmäßig Artikel zu meinem Themengebiet publiziert werden.“; $k = 7$ Items; $\alpha = .82/.87$).

Epistemologische Überzeugungen wurden jeweils vor dem Training mittels eines Fragebogens erhoben, der mit zwei Subskalen undifferenzierte objektivistische ($k = 12$; $\alpha = .70$) und undifferenzierte subjektivistische Überzeugungen ($k = 11$; $\alpha = .64$) erfasst (Peter, Leichner, Mayer & Krampen, 2014).

Ergebnisse: Um Maße für den vom Ausgangswert unabhängigen Lerngewinn zu erhalten, wurde jeweils eine Regression mit dem Prä-Wert als Prädiktor und dem Post-Wert als Kriterium durchgeführt. Die Residuen dieser Regressionen wurden als Veränderungsmaße verwendet (vgl. Vögele, 2008).

Für beide Trainings wurden separate multiple Regressionsanalysen mit den beiden epistemologischen Skalen als Prädiktoren und dem jeweiligen Veränderungsmaß als Kriterium durchgeführt. In Training 1 lässt sich der mittels des Wissenstests erfasste objektive Lerngewinn signifikant durch epistemologische Überzeugungen vorhersagen. Während objektivistische Überzeugungen erwartungswidrig nicht bedeutsam für den Lerngewinn sind ($\beta = -.14$, $p = ns$), fällt dieser hypothesenkonform bei hoch ausgeprägten subjektivistischen Überzeugungen geringer aus ($\beta = -.24$, $p < .05$). In Training 2 erweist sich die objektivistische Skala wiederum nicht als signifikanter Prädiktor des subjektiven Lerngewinns ($\beta = -.07$, $p = ns$), während stärkere subjektivistische Überzeugungen mit einem geringeren subjektiven Erkenntniszuwachs einhergehen ($\beta = -.27$, $p < .05$).

Die Ergebnisse werden mit Blick auf die Gestaltung von Informationskompetenz-Trainings diskutiert. Eine zentrale Frage ist es, wie die deterministischen Lerninhalte zur Informationskompetenz aufbereitet werden können, um kompatibler zu subjektivistischen Überzeugungen zu sein. Darüber hinaus nimmt die Diskussion Bezug auf die aktuelle Forschung zu epistemologischen Überzeugungen und betont die Bedeutsamkeit der separaten Erfassung objektivistischer und subjektivistischer Positionen.

Mittwoch
11. März

16:15 Uhr

GC 03/149

Chair(s)

Vogel, Freydis

Rösken-Winter, Bettina; Lehmann, Malte; Weißenrieder, Jochen; Blömeke, Sigrid
Qualitätsmerkmale von Lehrerfortbildungen

Einführung und Theorie: Von Lehrerfortbildungen wird angenommen, dass sie sich positiv auf die Kompetenzentwicklung von Lehrkräften und dadurch letztlich auf die Unterrichtsqualität auswirken (Hattie, 2008). Die Angebote reichen von Ein-Tages-Veranstaltungen bis hin zu umfangreichen Professionalisierungskonzepten (Goldsmith, Doerr & Lewis, 2014; Guskey, 2004; Loucks-Horsley, Love, Stiles, Mundry & Hewson, 2003). Allerdings müssen wirksame Veranstaltungen bestimmten inhaltlichen und methodischen Qualitätsmerkmalen genügen (Timperley et al., 2007; Lipowsky & Rzejak, 2010). Die Arbeiten von Desimone, Porter, Garet, Yoon and Birman (2002) sowie Garet et al. (2008) zeigen auf, dass die Merkmale Content focus, Active learning, Coherence, Duration und Collective Participation zielführend für die Gestaltung effektiver Fortbildungen sind. Für die über das DZLM (Deutsches Zentrum für Lehrerbildung Mathematik) angebotenen Professionalisierungsmaßnahmen sind entsprechend sechs Gestaltungsprinzipien abgeleitet worden: Kompetenzorientierung, Teilnehmerorientierung, Kooperationsanregung, Fallbezogenheit, Methodenvielfalt und Reflexionsförderung (vgl. DZLM, 2014). Die Qualitätsmerkmale von Lehrerfortbildungen lassen sich in zwei Gruppen ausdifferenzieren: Vier Merkmale beschreiben den Prozess, wie die Fortbildungen durchgeführt werden (teilnehmerorientiert, kooperationsanregend, methodisch vielfältig und fallbezogen) und zwei Merkmale sind produktbezogen (kompetenzorientiert und reflexionsfördernd). Während die prozessorientierten Merkmale als sogenannte „Sichtstrukturen“ weniger Einfluss auf die kognitiven Entwicklungen haben (vgl. Oser & Patry, 1990; Seidel, 2003), sind die produktorientierten Gestaltungsprinzipien eng mit einer hohen Kompetenzentwicklung verbunden. Zudem ist die Umsetzung der Fortbildungsinhalte im eigenen Unterricht von entscheidender Bedeutung. In der explorativen Studie werden die folgenden Forschungsfragen verfolgt: (1) Welchen Einfluss nehmen produktorientierte Gestaltungsprinzipien auf den Kompetenzzuwachs der TeilnehmerInnen? (2) Gibt es einen Zusammenhang zwischen der Umsetzung der Gestaltungsprinzipien in Fortbildungen und dem möglichen Einsatz von Fortbildungsinhalten in der Praxis?

Methodologie: In der Studie wurden die LeiterInnen (DZLM-MitarbeiterInnen, assoziierte ProfessorInnen und externe FortbildnerInnen) von 42 DZLM-Fortbildungen zu den sechs Gestaltungsprinzipien auf einer sechsstufigen Likertskala (keine Bedeutung – zentrale Bedeutung) befragt. Außerdem schätzten 161 TeilnehmerInnen die Umsetzung der Gestaltungsprinzipien in den jeweiligen Fortbildungen anhand von zwölf Items auf einer sechsstufigen Likertskala (trifft nicht zu – trifft zu) ein. Die Kompetenzentwicklung der TeilnehmerInnen wurde durch ein retrospektives Pre-Posttest-Design erfasst. Dabei werden Selbsteinschätzungen der eigenen Kompetenz nicht vorab erfragt, sondern über einen Vorher-Nachher-Vergleich im Anschluss an eine Intervention. Retrospektive Selbsteinschätzungen entgehen dem Problem, dass sich im Laufe einer Intervention der Bewertungsmaßstab verändert vor dessen Hintergrund die Einschätzungen vorgenommen werden (Coulter, 2012; Nimon, Zigarmi & Allen, 2011). Abschließend wurde mit einem Fragebogen, (Evaluation von Online-Fortbildungskursen für Mathematiklehrkräfte, Szymanski & Bruder, 2012) erfasst, inwieweit die TeilnehmerInnen planen, das Gelernte in ihrem eigenen Unterricht einzusetzen.

In einem ersten Schritt wurde die Qualität aller Skalen mittels klassischer Reliabilitätsanalysen in SPSS (Cronbachs Alpha) und auf robusten Maximum-Likelihood-Verfahren basierenden konfirmatorischen Faktorenanalysen (CFA) in MPlus geprüft. Die Zusammenhänge zwischen der Umsetzung der Qualitätsmerkmale von Lehrerfortbildungen in den Fortbildungen, ihre Auswirkungen auf die subjektiv eingeschätzte Kompetenzentwicklung und die Bereitschaft, das Gelernte in die Praxis umzusetzen, wurden zunächst mittels bivariater Korrelationen geprüft, um einfache Zusammenhänge festzustellen.

Ergebnisse: Den Gestaltungsprinzipien wird von den LeiterInnen der hier untersuchten Fortbildungen hohe Bedeutung zugesprochen. Allerdings zeigt sich zugleich, dass nicht alle Qualitätsmerkmale den Fortbildnern gleich wichtig für die Umsetzung ihrer Fortbildung sind. Die Teilnehmerorientierung ist mit einer durchschnittlichen Bewertung von 5,7 (SD 0,47) das bedeutendste Gestaltungsprinzip. Zwischen der Umsetzung der produktbezogenen Gestaltungsprinzipien und der Einschätzung, wie sich die eigenen Kompetenzen im Laufe der Fortbildung verändert haben, besteht ein signifikant positiver Zusammenhang ($r=0,2$; $p<0,01$). Je höher die TeilnehmerInnen die Umsetzung der Gestaltungsprinzipien einschätzen, desto geringer ist ihre Sorge, diese in ihrem Unterricht einzusetzen (produktbezogen: $r=-0,3$; $p<0,001$; prozessbezogen: $r=-0,2$; $p<0,05$).

Die vorliegende Studie gibt erste Hinweise darauf, wie Qualitätsmerkmale von Fortbildungen auf die Kompetenzentwicklung der TeilnehmerInnen und die Einschätzungen zur möglichen Umsetzung wirken. Um vertiefte Aussagen tätigen zu können, sind weitergehende Analysen nötig, die auf dem vorliegenden heuristischen Vorgehen aufbauen können.

Martin, Andreas; Schoemann, Klaus

Zum Einfluss der regionalen Heterogenität auf die Weiterbildungsbeteiligung in Deutschland

Regionale Heterogenität wird in Deutschland seit den Ergebnissen der PISA-Studien thematisiert. Im Vordergrund des Interesses steht meistens die regionale Gliederung nach Bundesländern. Diese Form der institutionell bedingten Heterogenität von Schulsystemen ist zunächst plausibel da die Bundesländer jeweils spezifische Formen des primären und besonders des sekundären Bildungssystems in der Nachkriegszeit entwickelt haben. Die Varianz von Bildungspolitik ist auf der Ebene der Bundesländer hinreichend gegeben. Aus regionaltheoretischer Sicht ist aber der Kultur- und Bildungsföderalismus nur eine der möglichen Ebenen auf denen sich Heterogenität äußern kann. Unser Beitrag befasst sich daher mit der Fragestellung welche weiteren regionalen Ebenen spielen bei Bildungsprozessen eine wichtige Rolle. Wir befassen uns in diesem Beitrag mit der Beteiligung an allgemeiner und beruflicher Weiterbildung.

Neben der Ebene der Bundesländer lassen sich mit Daten des Mikrozensus auch die sozialen und ökonomischen Prozesse der Beteiligung sowie Benachteiligung bei Weiterbildung mit Raumordnungsregionen als zweite Ebene und die Kreise als dritte räumliche Ebene untersuchen. Es kann somit erstmals empirisch überprüft werden welche regionale Gliederung den größten Beitrag zur Erklärung von Weiterbildungsbeteiligung leistet. Dazu werden zwei und drei Ebenen umfassende Mehrebenenmodelle geschätzt.

Um der Fragestellung nach der Relevanz einzelner Raumgliederungen nachzugehen, wurden Daten des Mikrozensus 2012 verwendet. Mit dieser Datengrundlage ist es zum ersten mal möglich, Weiterbildungsbeteiligung tiefregionalisiert zu erfassen. Mittels eines 4-Ebenen-logitmodells wurden die Varianzkomponenten auf der Ebene der Individuen, der Kreise, der Raumordnungsregionen und der Bundesländer berechnet. Aus den Varianzanteilen lässt sich die Intraclasscorrelation berechnen, die Aufschluss über die Varianzanteile der einzelnen Ebenen gibt.

Die Ergebnisse zeigen, dass die individuelle Ebene mit weitem Abstand die höchste Relevanz aufweist. Die regionalen Kontexte weisen sehr unterschiedliche Erklärungskraft auf. Die wichtigste regionale Ebene ist die Kreisebene. Entgegen der Ausgangshypothese zeigt sich dass sich die Relevanz der Regionen nicht linear gliederungstiefe erhöht. Wie zu erwarten, verdichten sich regionale Kontexte auf der niedrigsten regionalen Ebene. Kreise scheinen in dieser Hinsicht eine geeignete regionale Bezugsgröße darzustellen. Die hier gewählten Raumgliederungen unterscheiden sich jedoch nicht nur in der Gliederungstiefe, sondern stellen zudem auch räumliche Zuschnitte dar, die unterschiedlichen funktionalen und administrativen Zwecken dienen.

Donnerstag
12. März

9:00 Uhr

GA 03/49

Chair(s)

Rösken-Winter, Bettina

3

Hochholdinger, Sabine; Keller, Inka

Beziehungen zwischen affektiven Reaktionen von Teilnehmenden, Instruktionsstil der Lehrkräfte und Inhalt von Weiterbildungen

Theoretischer Hintergrund: Es ist naheliegend, dass erwachsene Teilnehmende ein Weiterbildungsangebot unterschiedlich positiv oder negativ erleben, abhängig von der Art des Angebots sowie von weiteren Bedingungen. In der Forschung zur beruflichen und betrieblichen Weiterbildung spielen Studien zu solchen affektiven und bewertenden Reaktionen eine eher nachgeordnete Rolle (z.B. Sitzmann et al. 2008). Dagegen stellt die psychologische Lernforschung die Bedeutung von Emotionen und Bewertungen als Antezedens-, Begleit- und Ergebnisvariablen für Lernprozesse und Lernmotivation heraus (z.B. Pekrun, 2011). Für erwachsene Lernende zeigten Towler et al. (2014) experimentell Zusammenhänge zwischen Charisma des Trainers, affektiven Reaktionen und Lernerfolg.

Auf der Grundlage der pädagogisch-psychologischen Lern- und Unterrichtsforschung erscheint es deshalb auch für die betriebliche und berufliche Weiterbildung interessant, die pädagogischen Bedingungen für solche Reaktionen zu betrachten. So hängt emotionales Erleben von SchülerInnen mit dem Schulfach und dem Instruktionsstil der Lehrkräfte zusammen: MINT-Fächer und anspruchsvolle LehrerInnen werden negativer erlebt als (Fremd-)sprachen und unterstützende Lehrkräfte (z.B. Götz et al., 2013). Weiterhin unterscheiden sich Instruktionsstil nach Domäne: Lehrende in „weichen“ Fächern orientieren ihre Lehre stärker an Teilnehmenden als Lehrende in „harten“ Fächern (z.B. Götz et al., 2013). Außerdem konnte die empirische Schulforschung zeigen, dass es relevant für den Unterrichtserfolg ist, wie sehr Lehrkräfte ihre Angebote auf ihre Zielgruppe hin ausrichten und diese aktivieren (z.B. Baumert et al., 2010). Diesem Lehrhandeln liegt die so genannte Lehrorientierung zugrunde, eine Facette pädagogischer Überzeugungen, zurückgehend auf Kember (1997). Sie beinhaltet auch, ob die Lehrperson ihre Rolle eher in einer von ihr ausgehenden Wissensvermittlung oder in der Anregung eigenständiger Lernprozesse bei den Lernenden sieht.

Fragestellung: Vor diesem Hintergrund wurden in der durchgeführten Studie für die betriebliche und berufliche Weiterbildung mehrere Fragenkomplexe untersucht: (1) Wie hängt der Instruktionsstil in Form von Lehrmethoden und Lehrorientierung der Lehrkräfte mit affektiven und bewertenden Reaktionen der Lernenden zusammen? (2) Unterscheiden sich affektive und bewertende Reaktionen der Teilnehmenden sowie der Instruktionsstil der Lehrenden nach Weiterbildungsinhalt?

Methode: Methodisch ist es wünschenswert, wenn kontextuelle Prädiktoren unabhängig von Ergebnisvariablen der Teilnehmenden erhoben werden (z.B. Sitzmann et al., 2008), um einen Common-Source-Bias zu vermeiden. Deshalb wurden bei typischen beruflichen Weiterbildungsangeboten 69 Lehrkräfte (Durchschnittsalter 48,3 Jahre, 22 Frauen und 44 Männer) sowie insgesamt 614 zugehörige Teilnehmende (Durchschnittsalter 32,5 Jahre, 174 Frauen und 427 Männer) befragt. Der Zugang zur Stichprobe erfolgte über Trainingsanbieter, Betriebe und IHK im Bodenseeraum. Die Lehrkräfte beschrieben ihre Lehrorientierung (Lernenden- und Lehrendenfokussierung) mit einem Instrument für erwachsene Lernende (Approaches to Teaching Inventory; Trigwell & Prosser 2004). Weiterhin berichteten sie ihre eingesetzten didaktischen Methoden und skizzierten die Weiterbildungsinhalte. Gleichzeitig schätzten ihre Teilnehmendem ihr emotionales Erleben in der Weiterbildung ein (PANAS, Krohne et al. 1996) und bewerteten ihre Zufriedenheit damit (MEI, Kauffeld et al., 2009) sowie die Effektivität der Lehrkraft (Donovan & Darcy, 2011). Die ermittelte Reliabilität der eingesetzten Skalen ist mit $.82 < \alpha < .89$ als gut einzustufen, mit Ausnahme der Skala zur Lehrendenorientierung mit $\alpha = .55$. Die so erhobene Datenstruktur besteht aus zwei Ebenen (Ebene 1: Teilnehmende, Ebene 2: Lehrkräfte), so dass sie im Sinne der o.g. Fragestellung mit mehrbenenanalytischen Korrelations- und Regressionsanalysen mit Mplus ausgewertet wurde.

Ergebnisse: Zusammenhangsanalysen zeigten erwartungsgemäß, dass die Lernendenorientierung und teilnehmeraktivierende Methoden der Lehrenden signifikant positiv mit günstigen affektiven Reaktionen und Bewertungen der Teilnehmenden korreliert, während die Lehrendenorientierung und lehrendenzentrierten Methoden negative, aber schwächere Beziehungen damit aufweisen. Daneben sind bedeutsame Kontexteffekte des Weiterbildungsinhalts festzustellen, der Unterschiede im Instruktionsstil der Lehrenden und in den Reaktionen der Teilnehmenden aufklärt: Softskills-Lerninhalte werden von den Teilnehmenden positiver bewertet, hier findet sich auch ein stärker lernendenzentrierter Instruktionsstil als bei Hardskills-Lerninhalten. Selektionseffekte sind nicht auszuschließen, so dass für zukünftige Studien eine Kontrolle der Kontextvariablen sinnvoll wäre.

Donnerstag
12. März

9:00 Uhr

GA 03/49

Chair(s)

Rösken-Winter, Bettina

Besser, Michael; Harks, Birgit; Leiss, Dominik

Wirkung von Lehrerfortbildungen auf die professionelle Handlungskompetenz von Lehrkräften: Formatives Assessment im kompetenzorientierten Mathematikunterricht

Im Kontext einer breiten Diskussion um erfolgreiches Lehren und Lernen in der Schule zeigen vielfältige empirische Studien die besondere Bedeutung professioneller Handlungskompetenz von Lehrkräften für ein Gelingen von Unterricht auf (siehe u. a. Kunter et al., 2011; Tatto et al., 2012). Weitestgehend unklar ist jedoch, wie sich professionelle Handlungskompetenz von Lehrkräften entwickelt und wie Lehrkräfte im Schuldienst bei der Entwicklung professioneller Handlungskompetenz – insbesondere in Form von Lehrerfortbildungen – gezielt unterstützt werden können. Zwar lassen sich einige generell notwendige Bedingungsfaktoren zur erfolgreichen Umsetzung von Lehrerfortbildungen aufzeigen (Desimone, 2009; Garet et al., 2001; Lipowsky, 2004), inwieweit diese jedoch hinreichend für eine professionelle Entwicklung von Lehrkräften im Schuldienst sind, ist als offene Forschungsfrage anzusehen.

Hieran anknüpfend untersucht das interdisziplinäre DFG-Forschungsprojekt Co²CA (Conditions and Consequences of Classroom Assessment) die Wirksamkeit von theoretisch fundierten und wissenschaftlich begleiteten Lehrerfortbildungen auf die Entwicklung ausgewählter Facetten der Professionalität von Lehrkräften. Am spezifisch-fachdidaktischen Beispiel lernförderlicher Leistungsbeurteilung als zentrales Element formativen Assessments (Maier, 2010) im kompetenzorientierten Mathematikunterricht wird insbesondere den beiden folgenden Forschungsfragen nachgegangen: (1) Inwieweit gelingt es, im Rahmen mehrwöchiger Lehrerfortbildungen das fachdidaktische Wissen und Können von Lehrkräften zu formativem Assessment im kompetenzorientierten Mathematikunterricht gezielt zu fördern? (2) Inwieweit lassen sich durch mehrwöchige Fortbildungen die formative Leistungsbeurteilungspraxis und die formative Leistungsbeurteilungsüberzeugungen von Mathematiklehrkräften (weiter-) entwickeln.

Im Rahmen der Studie haben insgesamt N = 67 Lehrkräfte entweder an Fortbildungen zu formativem Assessment im kompetenzorientierten Mathematikunterricht (Untersuchungsbedingung A; UB A; N = 30) oder aber an Fortbildungen zu allgemein-didaktischen Fragen eines kompetenzorientierten Mathematikunterrichts (Untersuchungsbedingung B; UB B; N = 37) teilgenommen. In beiden Untersuchungsbedingungen wurden die folgenden Evaluationsinstrumente administriert: (1) Allgemein-fachdidaktischer, fortbildungsunabhängiger Wissenstests des Forschungsprojekts COACTIV (siehe Krauss et al., 2008) zu Beginn der Fortbildungen (PCK-Test). (2) Spezifisch-fachdidaktischer, bzgl. der Inhalte von UB A fortbildungssensitiver Wissenstest zu formativem Assessment im kompetenzorientierten Mathematikunterricht am Ende der Fortbildungen (PCK-FA-Test). (3) Lehrerfragebögen zur Erfassung von formativer Leistungsbeurteilungspraxis und formativer Leistungsbeurteilungsüberzeugung mehrfach während der Fortbildungen.

Ergebnisse der Auswertungen der Testinstrumente zeigen auf: Lehrkräfte aus UB A schneiden am Ende der Fortbildungen im PCK-FA-Test signifikant besser ab als Lehrkräfte aus UB B ($t(65) = 6.66$; $p = .000$; $d = 1.63$). Sowohl innerhalb eines einfachen Regressionsmodells mit der Leistung im PCK-FA-Test als abhängiger Variable und der Fortbildungsbedingung als unabhängiger Variable (ANOVA) als auch in einem erweiterten Modell mit zusätzlicher Berücksichtigung der Leistung im PCK-Test als Kovariate (ANCOVA) lassen sich diese Unterschiede durch die Fortbildungen erklären. Etwa 40 % effektspezifischer Varianz im PCK-FA-Test werden durch die Fortbildungen erklärt (Forschungsfrage 1). Interessanterweise geht dieser Unterschied in der Expertise der Lehrkräfte jedoch nicht mit einem signifikanten Unterschied in der formativen Leistungsbeurteilungspraxis und der formativen Leistungsbeurteilungsüberzeugung der Lehrkräfte beider Untersuchungsbedingungen einher. Bezüglich dieser Facetten professioneller Handlungskompetenz lassen sich keinerlei Effekte der Fortbildungen auf diese Variablen aufzeigen (Forschungsfrage 2).

Im Kontext einer Auseinandersetzung mit Bedingungsfaktoren für erfolgreiches Unterrichten zeigt die Studie am spezifischen Beispiel formativen Assessments im kompetenzorientierten Mathematikunterricht somit Möglichkeiten zum Aufbau fachdidaktischer Expertise von Lehrkräften auf. Kritisch zu diskutieren ist jedoch die Tatsache, dass sich keine Effekte der Fortbildungen auf selbstberichtete Unterrichtspraxis und Überzeugungen zeigen.

Donnerstag
12. März

9:00 Uhr

GA 03/49

Chair(s)

Rösken-Winter, Bettina

Schüler, Sven; Stahnke, Rebekka; Rösken-Winter, Bettina

Professionalisierung von MathematikmultiplikatorInnen: Gestaltungsprinzipien und Kompetenzentwicklung

Motivation: Im Kontext zunehmender Standardisierung von Lehr-Lernerwartungen und einhergehender Professionsdebatte haben empirische Erhebungen zur Gestaltung von Fortbildungen und deren Wirkung auf professionelle Kompetenzen teilnehmender Lehrkräfte an Bedeutung gewonnen (Goldsmith, et al., 2014). Während die aktuelle Forschung die Ebene des Lernens von Lehrkräften in Fortbildungen schwerpunktmäßig thematisiert (Timperley, et al., 2007), fehlt ein systematischer Ansatz zur Erforschung der Wirkung von Fortbildungen auf die Kompetenzen von MultiplikatorInnen. Daher fokussiert dieser Beitrag zwei Aspekte, welche die Doppelrolle von MultiplikatorInnen als fach- und fortbildungsdidaktische ExpertenInnen widerspiegeln. Erstens untersuchen wir die kognitive Kompetenzentwicklung von MultiplikatorInnen nach Teilnahme an einer Qualifikationsmaßnahme des Deutschen Zentrums für Lehrerbildung Mathematik (DZLM). Zweitens betrachten wir, welche theoretischen Aspekte zur Gestaltung effektiver Fortbildungen die Teilnehmenden als relevant für die eigene MultiplikatorInnen-tätigkeit erachten.

Theoretische Einbindung: Ausgehend von Erkenntnissen der empirischen Bildungs- und Lehr-/Lernforschung postuliert das DZLM Gestaltungs- und Bedingungsfaktoren, welche als entscheidend für Wirkungen von Fortbildungsmaßnahmen gelten: Die inhaltliche Dimension wird über ein Kompetenzmodell (Blömeke, Suhl & Kaiser; Baumert et al., 2010), welches Professionswissen, Überzeugungen, technische Fähigkeiten und hinsichtlich der besonderen Rolle von MultiplikatorInnen auch fortbildungsdidaktische Elemente integriert, beschrieben. Die methodische Dimension umfasst strukturelle Aspekte von Fortbildungsangeboten, welche in der Forschungsliteratur als zielführend für nachhaltige Professionalisierungsprozesse erachtet werden: content focus, active learning, coherence, duration, und collective participation (Desimone, 2011; Garet, et al., 2001). Daraus abgeleitet diese folgenden DZLM-Gestaltungsprinzipien: Kompetenzorientierung, Teilnehmerorientierung, Fallbezogenheit, Methodenvielfalt, Kooperationsanregung und Reflexionsförderung. Die strukturelle Dimension des Theoriekonzepts der DZLM-Fortbildungsforschung adressiert Zusammenhänge entlang einer Wirkungskette, ausgehend von Kompetenzen für MultiplikatorInnen über Lehrkräfte bis zur Ebene der Lernenden.

Forschungsfragen: In Kontext der theoretischen Basierung der Gestaltungsprinzipien als Charakteristika effektiver Fortbildungen sind für unsere Studie drei Forschungsfragen richtungsweisend: (1) Wie entwickeln sich die selbsteingeschätzten Kompetenzen der MultiplikatorInnen entlang unterschiedlicher Kompetenzfacetten? (2) Welche Bedeutung messen die MultiplikatorInnen den DZLM- Gestaltungsprinzipien bezüglich der Wirkung von Fortbildungen bei? (3) Inwieweit sind die DZLM-Gestaltungsprinzipien in den Vorstellungen der MultiplikatorInnen zentral für die eigene Fortbildungstätigkeit?

Methode: Die Stichprobe besteht aus MultiplikatorInnen (n=12; neun weiblich, drei männlich mit durchschnittlich 20 Jahren Fortbildungs- und Lehrerfahrung), welche an einer DZLM-Qualifizierungsmaßnahme im Schuljahr 2012-2013 teilnahmen. Die Qualifizierungsmaßnahme folgte einem Sandwich-Prinzip, bestehend aus alternierenden Input- und Distanzphasen. Für die Datenerhebung wurden quantitative und qualitative Methoden verbunden, um Heterogenität und Komplexität von Professionalisierung multiperspektivisch zu erfassen. Zwei Fragebögen, bestehend aus offenen und geschlossenen Skalen zu Kompetenzfacetten und Gestaltungsprinzipien, wurden retrospektiv im Hinblick auf vor, direkt nach und sechs Monate nach der Fortbildungsmaßnahme eingesetzt. Retrospektive Selbstauskünfte können problematisch sein, weisen jedoch vergleichsweise hohe Validität auf (Lam & Bengo, 2003). Zusätzlich verwendeten wir qualitative Interviews (n=4), um Fragebogenaspekte vertiefend zu erfragen. Die zeitliche Veränderung der Kompetenzfacetten wurde durch nichtparametrische Friedmantests sowie anschließende paarweise Vergleich (unter Adjustierung des Signifikanzniveaus) überprüft. Die Analysen der Fragebögen und Interviews wurden über ein aus dem DZLM-Theorierahmen deduktiv antizipiertes und den gegebenen Antworten induktiv ausdifferenziertes Kategoriensystem basierend auf qualitativer Inhaltsanalyse strukturiert. Die Qualität des Kategoriensystems zeigt gute Interraterreliabilitäten (Cohen's kappa zwischen 0.67 und 0.71).

Ergebnisse: Für alle Kompetenzfacetten zeigte sich eine signifikante positive Entwicklung (Friedmantests mit p-Werten ≤ 0.023). Die anschließenden paarweisen Vergleiche zeigten für die Veränderung von vor zu direkt nach der Fortbildung für fünf von sieben Kompetenzfacetten einen mittleren bis großen Effekt, jedoch zeigte sich kein Kompetenzzuwachs bezüglich der Zeitspanne sechs Monate nach der Fortbildungsmaßnahme. Bei näherer Betrachtung der Interviews scheint dies an der Einführung neuer Inhalte im ersten Modul zu liegen, welche zwar im Rahmen der Fortbildung erprobt, danach jedoch nicht konsolidiert werden konnten. Hinsichtlich der DZLM-Gestaltungsprinzipien messen die MultiplikatorInnen Teilnehmerorientierung, Kompetenzorientierung und Kooperationsanregung die größte Bedeutung bei. Unsere qualitativen Ergebnisse zeigen, dass, obwohl die Gestaltungsprinzipien als zentral für die eigene Fortbildungstätigkeit bemessen werden, unterliegen sie in der Praxis verschiedenen Modifikationen, die Anlass zu vertiefenden Analysen bieten.

Donnerstag
12. März

9:00 Uhr

GA 03/49

Chair(s)
Rösken-Winter, Bettina

Lorke, Julia; Strippel, Christian Georg; Sommer, Katrin

„Für mich ist das 'ne Blume. 'Ne schöne, bunte Blume.“ – Naturwissenschaftliches Lernen durch „Science Busking“

Museen und Science Center sind Orte, die erwachsenen Besuchern vielfältige Lerngelegenheiten – auch zu naturwissenschaftlichen Themen – bieten (Waidacher, 2005). Ein Beispiel hierfür sind „Science Busking“-Aktivitäten. Dabei handelt es sich um kurze, interaktive Aktivitäten zu Inhalten, welche die Auseinandersetzung mit Naturwissenschaften bei den Teilnehmern fördern sollen (Sayer, Featherstone, & Gosling, 2014). Durch empirische Forschung aus dem Bereich Public Understanding of Science ist belegt, dass ein Defizit an naturwissenschaftlicher Grundbildung in der Bevölkerung besteht (Miller, 1998). Zudem ist als Bedingungen für die Lernwirksamkeit von Experimenten bekannt, dass das individuelle Lernverhalten instruktional unterstützt werden muss (Wirth et al., 2008). Unklar ist jedoch, ob derartiges Experimentieren auch an informellen Lernorten lernwirksam ist. Das führte zu folgender Forschungsfrage:

Wie verändert sich das chemische Fachwissen von Teilnehmern durch Lernen im Rahmen einer Science Busking-Aktivität quantitativ und qualitativ?

Es wurde eine Interventionsstudie mit Pre-Post-Testdesign zur Erfassung des Fachwissens zur Chromatographie konzipiert. Als Pre- und Post-Test dienen halbstrukturierte Interviews. Während der Intervention, einem Experiment zur Papierchromatographie, werden die Probanden instruktional bei der Durchführung und Auswertung durch fachkundige Personen, die auch die Datenerhebung durchführen, unterstützt. Der Post-Test schloss direkt an die Intervention an. Die Datenerhebung erfolgte am Deutschen Museum in München (DM), Techniqest Science Center in Cardiff/UK (TQ) und Alfred Krupp-Schülerlabor in Bochum (AKS). Es wurden Pre-Tests von N = 403 Personen erhoben (DM: n = 262; TQ: n = 106; AKS: n = 35). Von 144 Personen konnten sowohl Pre- als auch Post-Test erfasst werden.

Die Interviews wurden mittels qualitativer Inhaltsanalyse ausgewertet (Mayring, 2010). Durch Indexbildung wurden die Ergebnisse als Chromatographie-Index-Werte (CIS; Werte zwischen 0-100 möglich, je höher der CIS desto höher das Wissen über Chromatographie) quantifizierbar gemacht und anschließend statistisch analysiert (Lorke, 2010).

Die Auswertung des Pre-Tests zeigt, dass 73,2% der Besucher kein Vorwissen zur Chromatographie besitzen (CIS(pre)=0). Lediglich 26,8% der Besucher verfügen mindestens über geringes Vorwissen (CIS(pre)>0). Die Veränderungen der Chromatographie-Index-Werte von Pre- zu Post-Test wurden mittels Wilcoxon-Tests analysiert. Beim Wissenszuwachs zeigen sich Unterschiede in Abhängigkeit vom Vorwissen. Besucher ohne Vorwissen (N = 104; 72,2%) zeigen im Post-Test einen signifikanten Wissenszuwachs – ihr Lernzuwachs lässt sich quantitativ zeigen (Mdn(pre) = 0, Mdn(post) = 18, z = 7,38, p < .001, r = 0.51). Dass besonders Personen mit keinem bis geringem Vorwissen von Besuchern in informellen Lernumgebungen profitieren, wurde bereits in anderen Studien beobachtet (Beiers & McRobbie, 1992; J. Falk & Storksdiack, 2005; John H. Falk & Adelman, 2003). Auch in unserer Studie ließ sich bei Besuchern mit Vorwissen (N = 40; 27,8%) insgesamt kein signifikanter Lernzuwachs quantifizieren (Mdn(pre) = 19,53, Mdn(post) = 29,61, z = -.67, ns, r = -.08).

Betrachtet man jedoch das Wissen der Probanden mit Vorwissen über einzelne Teilbereiche der Chromatographie, so findet sich ein signifikanter Anstieg in den Kategorien, in denen diese Probanden im Pre-Test nur geringe Vorkenntnisse aufwiesen (Interaktion mit der mobilen Phase (Mdn(pre) = 0, Mdn(post) = 0, M(pre) = 16.46, M(post) = 39.25, z = -2.44, p < .05, r = -0.27), Interaktion mit der stationären Phase (Mdn(pre) = 0, Mdn(post) = 0, M(pre) = 3.66, M(post) = 19.51, z = -2.18, p < .05, r = -0.24) und Trennprinzip (Mdn(pre) = 0, Mdn(post) = 15, z = -.67, p < .01, r = -0.36)). Die Intervention hat somit einen Effekt auf die Qualität ihres Wissens über Chromatographie.

Diese Untersuchung zeigt für eine konkrete Science Busking-Aktivität, dass Lernzuwachs unter bestimmten Bedingungen wahrscheinlich ist und dieser unterschiedlichen Personengruppen ermöglicht wird. Die Untersuchung leistet auch einen grundsätzlichen Beitrag zur Evaluation von Science Busking-Aktivitäten. Es wird exemplarisch gezeigt, wie in einer von Freiwilligkeit geprägten Umgebung Lernzuwachs quantifizierbar gemacht werden kann, ohne dabei qualitative Veränderungen zu vernachlässigen.

Chair(s)
Lorke, Julia

Mathematisch-naturwissenschaftler Unterricht in Vorschule, Grundschule
und außerschulischen Lernorten

Dalehefte, Inger Marie; Rieck, Karen; Wendt, Heike; Kasper, Daniel

Neun Jahre SINUS in deutschen Grundschulen: Evaluation im Rahmen der TIMSS 2011-Erhebung – Befunde zum Zusammenhang zwischen Disparitäten und Kompetenzen

Theoretischer Hintergrund: SINUS an Grundschulen war ein Lehrerprofessionalisierungs- und Unterrichtsentwicklungsprogramm mit dem Ziel, Lehrkräfte darin zu unterstützen, den mathematisch-naturwissenschaftlichen Unterricht in der Grundschule weiterzuentwickeln und Schülerleistungen zu steigern. Der Erfolg des Programms wird mit Hilfe der Begleitforschung auf der Ebene der Schulen, der Lehrkräfte, des Unterrichts und der Leistungen der Schülerinnen und Schüler geprüft (Fischer et al., 2012). In diesem Beitrag wird das SINUS-Programm in Bezug zu einem programmunabhängigen Außenkriterium (Kontrollgruppe) mit Hilfe von Daten aus "Trends in International Mathematics and Science Study" (TIMSS 2011) (Bos et al., 2012) gesetzt. Ziel dieses Vergleichs ist es, Hinweise auf die Wirksamkeit des Programms zu erhalten. An dieser Stelle soll auf die Kompetenzen der Schülerinnen und Schüler im Bereich Naturwissenschaften und Mathematik unter Berücksichtigung der Disparitäten der Lernenden nach sozialer Herkunft, Migrationshintergrund und Geschlecht eingegangen werden.

Eine gelingende Professionalisierungsmaßnahme weist bestimmte Strukturen auf, die Veränderungen des Wissens, der Fähigkeiten und der Einstellungen der Lehrkräfte ermöglichen. Schulen und Lehrkräfte brauchen in diesem Prozess Unterstützung und Anregung durch inner- und außerschulische Expertise (Timperley, 2008). Im Programm SINUS an Grundschulen wurde eine solche Unterstützungsstruktur vom Programmträger (IPN, Kiel) durch die bundesweite Koordination des Programms gewährleistet. Verschiedene Programmaßnahmen, wie ein umfangreiches Angebot von Fortbildungsveranstaltungen und Handreichungen zu theoretisch und empirisch fundierten Entwicklungsbereichen des mathematischen und naturwissenschaftlichen Unterrichts in der Grundschule, unterstützten die Lehrkräfte bei der Weiterentwicklung ihres Unterrichts (Demuth et al., 2011). Studien haben gezeigt, dass in Deutschland nach wie vor ein substantieller Zusammenhang zwischen Kompetenz und sozialer Herkunft, Migrationshintergrund sowie Geschlecht besteht (Bos et al., 2012, Prenzel et al., 2013).

In diesem Beitrag wird dargestellt, ob und in welchem Ausmaß die Teilnahme am Programm SINUS an Grundschulen zu positiven Veränderungen der naturwissenschaftlichen bzw. mathematischen Kompetenz unter Betrachtung der sozialen Herkunft, des Migrationshintergrunds und des Geschlechts geführt hat.

Fragestellung: Folgende Fragestellung liegt der Untersuchung, die in diesem Beitrag vorgestellt wird, zugrunde: Unterscheiden sich die Testleistungen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in Deutschland (TIMSS-Stichprobe 2011) im Vergleich zu Testleistungen von Lernenden an SINUS-Grundschulen für die TIMSS-Gesamtskala Naturwissenschaften und Mathematik unter Betrachtung der sozialen Herkunft, des Migrationshintergrunds und des Geschlechts?

Stichprobe und Methode: Für die hier dargestellte Untersuchung wurden Daten von 1.581 Schülerinnen und Schülern aus 78 SINUS-Schulen einbezogen, die seit mindestens fünf Jahren am SINUS-Programm beteiligt waren. Die als Vergleichsmaßstab herangezogene Stichprobe aus TIMSS 2011 umfasst Daten von 3.995 deutschen Grundschulkindern. Beide Stichproben wurden unter identischen Erhebungsbedingungen und mit den gleichen Leistungstests untersucht. Die Leistungsdaten (Plausible Values) der Schülerinnen und Schüler wurden unter Berücksichtigung bestimmter Kontrollvariablen mit Tests auf Mittelwertsunterschiede analysiert (Wendt et al., 2012). Zunächst wurden die Gesamtwerte für die naturwissenschaftlichen sowie die mathematischen Kompetenzen der SINUS-Stichprobe auf der internationalen Metrik eingeordnet. Anschließend wurden die Schülerleistungen unter Betrachtung der Disparitäten nach sozialer Herkunft, Migrationshintergrund und Geschlecht mit der TIMSS-Stichprobe verglichen.

Ergebnisse: Die bisherigen Ergebnisse zeigen, dass im Hinblick auf die naturwissenschaftlichen und mathematischen Kompetenzen SINUS-Schülerinnen und -Schüler im Mittel deutlich höhere Testwerte erzielten als die Schülerinnen und Schüler der TIMSS-Stichprobe. Dieser relative Leistungsvorsprung zeigte sich für alle in TIMSS erhobenen Inhaltsbereiche und kognitiven Anforderungsniveaus. Auch ist festzustellen, dass in der SINUS-Stichprobe mehr Kinder Leistungen auf den oberen Kompetenzniveaus erzielten und sich weniger Schülerinnen und Schüler auf der niedrigsten Kompetenzstufe befanden. Ziel dieser Untersuchung ist es, darzustellen, ob SINUS-Lernende unter Betrachtung der sozialen Herkunft, des Migrationshintergrunds und des Geschlechts ihre Leistungen verbessern konnten und ob bzw. inwieweit die vorgestellten Befunde auf Programmaßnahmen zurückzuführen sind.

Donnerstag
12. März

9:00 Uhr

GA 03/142

Chair(s)
Lorke, Julia

Mathematisch-naturwissenschaftlicher Unterricht in Vorschule, Grundschule
und außerschulischen Lernorten

Flottmann, Julia; Ape, Marina; Leuchter, Miriam

Zusammenhänge von naturwissenschaftlichem Inhalts- und Prozesswissen im Vorschulalter

Theoretischer Hintergrund: Um positive Einstellungen gegenüber Naturwissenschaften möglichst früh vorzubereiten, sollte schon bei Kindern im Vorschulalter ein Verständnis für den naturwissenschaftlichen Erkenntnisprozess angebahnt werden (Eshach & Fried, 2005). Ziel naturwissenschaftlicher Bildung ist neben der Vermittlung von inhaltsbezogenem Wissen die Förderung von Prozesswissen (Dunbar & Klahr, 2012). Das Prozesswissen beinhaltet unter anderem das Planen von Experimenten (Zimmerman, 2007). In Bezug auf die Versuchsplanung ist die Variablenkontrollstrategie zentral, um aussagekräftige Ergebnisse zu erhalten (Chen & Klahr, 1999).

Die Anwendung der Variablenkontrollstrategie fällt je nach Kontext unterschiedlich aus (Bullock & Ziegler, 1993). Angemessenes inhaltliches Wissen kann einerseits die Auswahl geeigneter Variablen und Strategien fördern (Schauble, 1996). Andererseits kann inhaltliches Wissen dem Einsatz der Variablenkontrollstrategie im Wege stehen, da bei der Planung eines Experiments häufig auf inhaltliche Vorannahmen (z. B. ethische Überzeugungen in Bezug auf Lebewesen) anstatt auf die Variablenkontrolle zurückgegriffen wird (Croker & Buchanan, 2011). Demnach ist zu vermuten, dass die Unterscheidung zwischen lebendig und nicht lebendig (Jipson & Gelman, 2007) auch beim Einsatz der Variablenkontrollstrategie eine Rolle spielt (z. B. Kontext „Pflanzenwachstum“ vs. „Kugelbahn“).

Die vorliegende Studie untersucht: (a) ob Kinder im Vorschulalter die Variablenkontrolle bei der Versuchsplanung berücksichtigen können; (b) ob und wie sich der inhaltliche/ethische Kontext der Aufgabe auswirkt

Methodik: 59 Kinder (32 Jungen, 27 Mädchen) aus Kindertagesstätten im Münsterland in einem Alter von $M=6.67$ Jahren ($SD=.32$) haben an der Studie teilgenommen. Die bildgestützte Befragung bot jeweils drei Antwortoptionen an und wurde in den Kindergärten von geschulten Versuchsleitern durchgeführt. Um den Einfluss des Kontexts zu untersuchen, wurden vier Inhaltsbereiche ausgewählt, von denen angenommen wurde, dass Vorwissen und ethische Aspekte in unterschiedlichem Ausmaß relevant sind: schiefe Ebene, Schwimmen und Sinken, Reibung sowie Pflanzenwachstum. Jedes Kind bearbeitete Aufgaben zur Variablenkontrollstrategie zu jedem Kontext (Gesamtskala, 13 Items, Cronbachs $\alpha=.65$; schiefe Ebene, 3 Items, Cronbachs $\alpha=.64$; Schwimmen und Sinken, 4 Items, Cronbachs $\alpha=.61$; Reibung, 2 Items, Cronbachs $\alpha=.65$; Pflanzenwachstum, 4 Items, Cronbachs $\alpha=.79$). Auf Basis der Binomialverteilung (Ratewahrscheinlichkeit $< 10\%$) wurden die Kinder dahingehend kategorisiert, ob sie die Variablenkontrollstrategie anwenden, die fokussierte Variable konstant halten oder raten. Darüber hinaus wurden 4 Aufgaben zu Aspekten des inhaltlichen Wissens bearbeitet. Da bisherige Studien kontroverse Befunde hinsichtlich eines Zusammenhangs der Versuchsplanung mit der Fähigkeit zur Inhibition aufgabenirrelevanter Informationen erbrachten (Kwon & Lawson, 2000; Mayer, Sodian, Koerber, & Schwippert, 2014), absolvierte jedes Kind den Fruit-Stroop-Test (vgl. Archibald & Kerns, 1999).

Ergebnisse: Je nach Kontext wird die Variablenkontrollstrategie von 20% (schiefe Ebene), 27% (Schwimmen und Sinken), 27% (Reibung) und 10% (Pflanzenwachstum) der Kinder konsistent angewandt. 7% (schiefe Ebene), 12% (Schwimmen und Sinken), 19% (Reibung) und 56% (Pflanzenwachstum) der Kinder halten die fokussierte Variable konstant. 68% (schiefe Ebene), 51% (Schwimmen und Sinken), 44% (Reibung) und 24% (Pflanzenwachstum) der Kinder raten.

In Bezug auf das inhaltliche Wissen raten die Kinder in den Kontexten „schiefe Ebene“ und „Schwimmen und Sinken“, in den Kontexten „Reibung“ und „Pflanzenwachstum“ zeigen sich Deckeneffekte (95% richtige Antworten).

Chi-Quadrat-Tests zeigen keine Zusammenhänge zwischen den Aufgaben zur Variablenkontrollstrategie, dem Inhaltswissen und der Inhibitionsfähigkeit.

Diskussion: Die Anwendung der Variablenkontrollstrategie ist in den Kontexten „schiefe Ebene“, „Schwimmen und Sinken“ und „Reibung“ vergleichbar, obwohl sich in Bezug auf das inhaltliche Wissen große Unterschiede zeigen. Dies deutet darauf hin, dass die Anwendung der Variablenkontrollstrategie unabhängig von inhaltlichem Wissen ist. Im Themenbereich „Pflanzenwachstum“ hält der Großteil der Kinder die fokussierte Variable „Wasser“ jedoch konstant. Sind Antwortoptionen so konzipiert, dass die Variation der fokussierten Variable mit ethischen Gesichtspunkten konfiguriert, wird die Variablenkontrollstrategie kaum angewendet. Die Ergebnisse lassen vermuten, dass die Anwendung der Variablenkontrollstrategie unabhängig von inhaltlichem Vorwissen ist, insofern ethische Gesichtspunkte keine Rolle spielen.

Donnerstag
12. März

9:00 Uhr

GA 03/142

Chair(s)
Lorke, JuliaMathematisch-naturwissenschaftler Unterricht in Vorschule, Grundschule
und außerschulischen Lernorten

Lange, Kim; Ohle, Annika; Kleickmann, Thilo; Kauertz, Alexander; Möller, Kornelia; Fischer, Hans

Zur Bedeutung von Fachwissen und fachdidaktischem Wissen für Lernfortschritte von Grundschülerinnen und Grundschülern im naturwissenschaftlichen Sachunterricht

Bei der Suche nach Qualitätsmerkmalen im Lehr-Lernprozess, die die Erreichung von Lernfortschritten unterstützen, sind Lehrkräfte verstärkt in den Mittelpunkt pädagogisch-psychologischer und auch fachdidaktischer Forschung gerückt. Für die Initiierung erfolgreicher Lernprozesse gilt das professionelle Wissen von Lehrkräften als zentrale Komponente (Baumert & Kunter 2006; Bromme, 1997). Aus dem angloamerikanischen Raum existiert eine inzwischen weithin akzeptierte Unterteilung des professionellen Lehrerwissens in a) psychologisch-pädagogische, b) fachliche und c) fachdidaktische Anteile (Shulman, 1986, 1987). Für diese Unterscheidung gibt es erste empirische Belege (u.a. Blömeke et al. 2008). Während das generisch-pädagogische Wissen (PK) weitgehend vom Fachinhalt losgelöstes Wissen über die Organisation und Optimierung von Lehr- und Lernumgebungen sowie Lerntheorien und Lehrmethoden umfasst, sind die übrigen beiden Bereiche fachspezifisch. Das fachliche Wissen (CK) umfasst Wissen über die Inhalte und deren Struktur des zu unterrichtenden Unterrichtsfaches entsprechend der akademischen Bezugsdisziplin. Fachdidaktisches Wissen (PCK) ist Wissen darüber, wie Unterrichtsinhalte für Lernende zugänglich gemacht werden können (Shulman, 1986). Obwohl das PCK als eine Kombination und Integration von fachlichem und pädagogischem Wissen gilt, wird es in der Literatur als eigenständige und professionsprägende Komponente des professionellen Lehrerwissens beschrieben (Gess-Newsome, 1999), welches Lehrkräfte dazu befähigt, die Anforderungen eines kognitiv aktivierenden und gleichzeitig individuell unterstützenden Unterrichts zu erfüllen (Baumert et al., 2010).

Bestehende Studien geben Hinweise auf die Bedeutung des CK und des PCK für die Unterrichtsqualität und das Lernen von Schülern. Studien im Bereich Mathematik deuten darauf hin, dass das professionelle Wissen von Lehrkräften – vermittelt über die Unterrichtsqualität – in Beziehung zu Lernfortschritten von Schülern steht (Baumert et al. 2010; Hill, Rowan, & Ball, 2005). Dabei erwies sich das PCK in der Sekundarstufe als besonders relevant (Baumert et al. 2010). Studien im Bereich des naturwissenschaftlichen Sachunterrichts beschreiben die Bedeutung der fachspezifischen Bereiche des Professionswissens für die Unterrichtsqualität und das Erreichen von multiplen Zielen im Sachunterricht bislang nur in separaten Analysen (Ohle, Kauertz & Fischer 2011; Lange, Kleickmann, Tröbst & Möller, 2012). Offen bleibt hier die Frage, ob ein Wissensbereich von größerer Bedeutung für den Lernerfolg der Schüler ist als der andere.

Die vorliegende Studie greift zur Beantwortung dieser Frage auf Daten des DFG geförderten Projektes „Professionswissen von Lehrkräften, naturwissenschaftlicher Unterricht und Zielerreichung im Übergang von der Primar- zur Sekundarstufe“ (PLUS) zurück. Im Rahmen dieser Studie unterrichteten $N = 60$ Lehrkräfte $n = 1326$ Lernende der 3. und 4. Klasse zum Thema „Aggregatzustände“. Die leistungsbezogenen Zielkriterien aufseiten der Lernenden wurden in einem Prä-/Post-Design mit Fragebögen erfasst. CK und PCK der Lehrpersonen wurden in separaten Tests erhoben (Ohle et al., 2011; Lange et al., 2012). Zur Überprüfung der theoretisch angenommenen Struktur des Professionswissens wurde die manifeste, bivariate Korrelation zwischen den Scores des CK- und PCK-Tests berechnet; für die Berechnung der Zusammenhänge zu Lernfortschritten Mehrebenenmodelle spezifiziert und die prädiktive Validität der Wissenskomponenten verglichen.

Die Korrelation zwischen den beiden Wissensbereichen weist mit $r = .36$ ($p < .05$) einen mittleren Zusammenhang auf. Die Mehrebenenanalysen zeigen, dass unter Kontrolle individueller Lernvoraussetzungen und bedeutsamer Kontextmerkmale das PCK ein positiver Prädiktor für den Lernerfolg ist, nicht aber das CK. Das Modell, in das beide Prädiktoren simultan aufgenommen wurden, zeigt ebenfalls, dass das PCK, nicht das CK, leistungsrelevant ist.

Insgesamt belegen die Ergebnisse, analog zu den Befunden aus der Mathematik in der Sekundarstufe, dass auch im Sachunterricht das PCK von Grundschullehrkräften von großer Bedeutung ist, da sich trotz substanzieller Korrelationen zwischen den Wissensbereichen differenzielle Ergebnisse im Hinblick auf die Vorhersage von Schülerlernfortschritten zeigen. Mit Blick auf die Lehrerbildung ergeben sich aus diesen Befunden besondere Ansprüche an die fachbezogene Ausbildung: Sie werfen die Fragen auf, wie es gelingen kann, Grundschullehrkräfte beim Aufbau des PCK im Sachunterricht zu unterstützen und welche Rolle die fachlichen Anteile in der Ausbildung dafür spielen müssen.

Donnerstag
12. März

9:00 Uhr

GA 03/142

Chair(s)
Lorke, Julia

Mathematisch-naturwissenschaftler Unterricht in Vorschule, Grundschule
und außerschulischen Lernorten

Bohrmann, Mareike; Möller, Kornelia

Effekte der Expertise von (Sachunterrichts-)Lehrkräften auf die experimentellen Kompetenzen von Grundschulkindern in der dritten und vierten Klasse

Das Experiment ist eine wesentliche naturwissenschaftliche Methode (Bullock & Ziegler, 1999) und ein komplexer Problemlöseprozess, der (1) das Suchen und Aufstellen von Hypothesen, (2) die Planung und Durchführung von Experimenten und (3) die Analyse experimenteller Ergebnisse beinhaltet (Klahr, 2000).

Grundschulspezifische Untersuchungen zeigen, dass die Bewertung experimenteller Designs den Kindern leichter fällt als das eigene Entwickeln kontrollierter Experimente. Nur wenige Grundschüler/innen produzieren selbstständig kontrollierte Experimente (Bullock, Sodian & Koerber, 2009). Diese Fähigkeiten können durch gezielten Unterricht verbessert werden (Grygier, 2008), was notwendiges Wissen seitens der Lehrperson voraussetzt (Sodian, Thoermer, Kircher, Grygier & Günther, 2002). Hier setzt diese Untersuchung an.

Drei Forschungsfragen sind zentral: (1) Welche experimentellen Kompetenzen zeigen Kinder in dritten und vierten Klassen bezüglich der Bewertung und Planung kontrollierter Experimente sowie der Unterscheidung von Hypothesentestung und Effektproduktion? (2) Unterscheiden sich Kinder der dritten von denen der vierten Klasse bezüglich ihrer experimentellen Kompetenzen? (3) Hat die Expertise der unterrichtenden Lehrpersonen in Bezug auf den naturwissenschaftlichen Sachunterricht einen Effekt auf die experimentellen Kompetenzen der Kinder?

Ein gemeinsam mit der Arbeitsgruppe um Elsbeth Stern entwickelter Test berücksichtigt zwei a priori definierte Subkompetenzen: die Unterscheidung zwischen Hypothesentestung und Effektproduktion sowie den Einsatz der Variablenkontrollstrategie. Zu beiden Bereichen gibt es Items zu Bewertung und Entwicklung. Der Test besteht aus insgesamt 12 offenen, geschlossenen und Kombinationen offener und geschlossener Aufgaben (Cronbach's α für standardisierte Items = 0,775, N=292, MAlter=9,83). Zur Aufbereitung der Daten wurden die Antworten der Kinder mit Hilfe eines Codiermanuals drei möglichen Level zugeordnet (Level 0=keine oder naive Antworten (0 Punkte), Level 1=Zwischenvorstellungen (1 Punkt) und Level 2=wissenschaftlich adäquate Antworten (2 Punkte)). Zwei geschulte Codiererinnen ordneten die Antworten den drei Kompetenzniveaus zu (Kendalls T= 0.904). Die Auswertung der Daten erfolgte durch eine konfirmatorische Faktorenanalyse mit Hilfe des Programms MPlus sowie Korrelationsanalysen und Mittelwertvergleichen in SPSS.

Neben den experimentellen Kompetenzen wurden die unterrichtenden Lehrkräfte nach ihrem Aus- und Fortbildungshintergrund zu naturwissenschaftlichen Sachunterrichtsthemen befragt. Inhalte waren Aus- und Fortbildungen im Bereich des naturwissenschaftlichen Unterrichts, die Einstellung zum Experimentieren in der Grundschule sowie die Kooperation mit Universitäten.

Zur Bestätigung des gebildeten theoretischen Modells wurden ein 1-Faktor- (experimentelle Kompetenzen) und ein 2-Faktor-Modell (Variablenkontrollstrategie, Unterscheidung zwischen Hypothesentestung und Effektproduktion) an die Daten angepasst. Die Fitindices weisen auf einen guten Modellfit hin (1-Faktor-Modell: RMSEA=0,041, CFI=0,956, TLI=0,949; 2-Faktor-Modell: RMSEA=0,038, CFI=0,962, TLI=0,956).

Mit Hilfe der Angaben der Lehrpersonen wurden zwei Gruppen identifiziert: (1) Lehrkräfte mit hoher Expertise (Zusammenarbeit mit universitären Einrichtungen, Gefühl, im Bereich des Experimentierens sehr gut aus- und fortgebildet zu sein) und (2) Lehrkräfte mit niedriger Expertise.

Die Maximalpunktzahl des Tests beträgt 35 (fünf Aufgaben mit maximal 2 Punkten, drei Aufgaben mit maximal 3 Punkten, vier Aufgaben mit maximal 4 Punkte). Im Mittel erreichen die Kinder 15,97 Punkte (Max=32, Min=4, SD=5,9). Im Bereich der Bewertung erzielen die Kinder im Durchschnitt 51% und im Bereich der Entwicklung 44% der möglichen Punkte. Gleichzeitig korrelieren die beiden Bereiche Bewertung und Entwicklung miteinander mit 0,497. Folglich erreichen Kinder, die im Bereich der Evaluation höhere Punktzahlen erreichen, ebenfalls mehr Punkte im Bereich des Designs (Sig. 0,000). Kinder der 4. Klasse erreichen zudem im Durchschnitt 1,6 Punkte mehr als Kinder der 3. Klasse (Sig. 0,019, Cohen's $d=0,28$). Die von Lehrkräften mit hoher Expertise unterrichteten Kinder erreichen im Durchschnitt 5,3 Punkte mehr (Sig. 0,000). Hier liegt ein großer Effekt vor (Cohen's $d=1$). Der Einfluss der Lehrkraft ist somit wesentlich höher als der Einfluss des Alters bzw. der Klassenstufe.

Im Vortrag werden der Test präsentiert sowie Ergebnisse der Befragung berichtet und diskutiert.

Donnerstag
12. März

9:00 Uhr

GA 03/142

Chair(s)
Lorke, Julia

Mathematisch-naturwissenschaftlicher Unterricht in Vorschule, Grundschule
und außerschulischen Lernorten

Richter, Dirk

Die Bedeutung der Lehrbefähigung für den Kompetenzerwerb von Schülerinnen und Schülern: Ergebnisse aus nationalen Large-Scale-Assessments

Aktuell besteht in vielen Bundesländern eine Unterversorgung mit Fachlehrkräften insbesondere in den Fächern Mathematik und den Naturwissenschaften (KMK 2003). Infolge dieser Unterversorgung werden zum Teil Lehrkräfte eingesetzt, die selbst kein Studium in dem jeweiligen Fach abgeschlossen haben und somit keine Lehrbefähigung für dieses Fach besitzen. In Befragungen im Rahmen des IQB-Ländervergleichs zeigte sich, dass im Durchschnitt 27 Prozent der Mathematiklehrkräfte in Grundschulen und 14 Prozent der Mathematiklehrkräfte in weiterführenden Schulen fachfremd unterrichten (Richter et al. 2012, 2013). Bei diesen Lehrkräften kann davon ausgegangen werden, dass sie über weniger fachliche und fachdidaktische Kompetenzen verfügen als Lehrkräfte mit entsprechendem Fachstudium. Da jedoch vor allem das fachdidaktische Wissen bedeutsam für die Gestaltung kognitiv aktivierender Lerngelegenheiten und die konstruktive Unterstützung von Schülerinnen und Schülern ist (Baumert et al. 2010), kann angenommen werden, dass auch die Kompetenzen in fachfremd unterrichteten Klassen niedriger ausfallen als in Klassen mit einer studierten Fachlehrkraft.

Fragestellungen: (1) Wie stark unterscheiden sich Kompetenzen im Lesen und in Mathematik zwischen fachfremd und fachadäquat unterrichteten Klassen in der Primarstufe? (2) Wie stark unterscheiden sich Kompetenzen im Lesen, in Mathematik und in den Naturwissenschaften zwischen fachfremd und fachadäquat unterrichteten Klassen in den weiterführenden Schulen?

Die Analysen basieren auf den Stichproben der IQB-Ländervergleichsstudien aus den Jahren 2011 (Primarstufe) und 2012 (Sekundarstufe I). Im Ländergleich 2011 wurden Schülerinnen und Schüler aus 1295 Grundschulen in den Fächern Deutsch und Mathematik getestet und die jeweiligen Lehrkräfte befragt. Die zugrunde liegende Stichprobe umfasst 1227 Deutschlehrkräfte und 1200 Mathematiklehrkräfte. In diesen beiden Gruppen unterrichteten 17 Prozent das Fach Deutsch und 27 Prozent das Fach Mathematik ohne fachspezifische Lehrbefähigung. Die Gesamtstichprobe der Grundschullehrkräfte war im Durchschnitt 47.6 Jahre alt und 88.3 Prozent waren weiblich.

Im Ländervergleich 2012 nahmen Schülerinnen und Schüler aus 1266 weiterführenden Regelschulen teil und bearbeiteten einen Kompetenztest in den Fächern Mathematik, Biologie, Chemie und Physik. Auch in dieser Studie beantworteten die jeweils unterrichteten Fachlehrkräfte einen Fragebogen, in dem sie u.a. Angaben zu den erworbenen Lehrbefähigungen machten. Aus dem Fach Mathematik beteiligten sich 1660 Lehrkräfte (48.0 Jahre alt und 55 % weiblich) und aus den drei Naturwissenschaften 2927 Lehrkräfte (47.7 Jahre alt und 58 % weiblich). Der Anteil fachfremd unterrichtender Lehrkräfte lag in Mathematik bei 14 Prozent und in den Naturwissenschaften zwischen 10 Prozent (Chemie) und 16 Prozent (Physik). An nicht-gymnasialen Schularten fielen die Werte jedoch deutlich höher als an am Gymnasium.

Zur Untersuchung der Fragestellungen werden Mehrebenenmodelle in Mplus geschätzt. Im Fokus der Analysen steht der Zusammenhang zwischen der Lehrbefähigung der Lehrkraft und den Kompetenzen der Schülerinnen und Schüler in den jeweils untersuchten Domänen. Zusätzlich berücksichtigen die Modelle zentrale Hintergrundmerkmale von Schülerinnen und Schülern auf Individualebene (Zuwanderungshintergrund und sozialer Status) und die Schulart als Kovariate auf Klassenebene (für Daten der Sekundarstufe I). Eine Imputation fehlender Werte findet nicht statt.

Im IQB-Ländervergleich 2011 in der Primarstufe unterscheiden sich die mathematischen Leistungen von fachfremd unterrichteten Klassen statistisch bedeutsam von denen, die von Fachlehrkräften unterrichtet wurden. Im Lesen lassen sich hingegen keine systematischen Zusammenhänge nachweisen. Für den IQB-Ländervergleich 2012 zeigt sich ebenfalls, dass fachfremd unterrichtete Klassen in Mathematik im Durchschnitt schwächere Leistungen erreichen als diejenigen, die von einer Person mit Lehrbefähigung im Fach unterrichtet wurden. Für die Fächer Biologie und Physik ergeben sich signifikante Interaktionen zwischen der Lehrbefähigung und der Schulart; diese weisen darauf hin, dass die Kompetenzunterschiede zwischen fachfremd und fachadäquat unterrichteten Klassen an gymnasialen und nicht-gymnasialen Schularten unterschiedlich ausgeprägt sind. Ausschließlich an den nicht-gymnasialen Schularten zeigt sich ein signifikanter Leistungsvorsprung zugunsten der Klassen, die von einer Lehrkraft mit Lehrbefähigung unterrichtet werden.

Donnerstag
12. März

9:00 Uhr

GA 03/149

Chair(s)
Richter, Dirk

Porsch, Raphaela

Unterscheiden sich Mathematiklehrkräfte an Grundschulen mit und ohne Lehrbefähigung hinsichtlich ihrer berufsbezogenen Überzeugungen?

In den letzten Jahren haben sich zahlreiche Forschungsarbeiten mit der Beschreibung und Erfassung von Professionswissen bzw. -kompetenzen (angehender) Lehrkräfte beschäftigt (z.B. TEDS-M 2008: z.B. Blömeke, Kaiser & Lehmann 2010a, 2010b; COACTIV: z.B. Kunter, Baumert, Blum, Klusmann, Krauss & Neubrand 2011). Untersucht wurde u.a. die Wirkung des Wissens bzw. der Kompetenzen auf die Schülerinnen und Schüler (z.B. Krauss, Neubrand, Blum, Baumert, Brunner, Kunter & Jordan 2008). Bisher liegen jedoch wenige Arbeiten im deutschsprachigen Raum zum Thema „fachfremdes Unterrichten“ vor. Empirisch untersucht wurde der Einfluss der Qualifikation von Lehrkräften auf die Schülerleistungen, die über keine oder eingeschränkte Qualifikation in einem von ihnen unterrichteten Fach besitzen (Tiedemann & Billmann-Mahecha 2007; Richter, Kuhl, Reimers & Pant 2012; Richter, Kuhl, Haag & Pant 2013; Porsch & Wendt in Überarbeitung). Es zeigen sich in der Mehrzahl der Studien bedeutsame Unterschiede in den Schülerleistungen zugunsten derjenigen Schülerinnen und Schüler, die von Fachlehrerinnen und -lehrern unterrichtet wurden. Der Fokus der Arbeiten liegt jedoch nicht auf der Frage, ob sich die Lehrkräfte, denen die fachliche Ausbildung fehlt oder teilweise fehlt, im Hinblick auf andere Merkmale wie ihre unterrichts- oder selbstbezogenen Überzeugungen unterscheiden (zum Begriff vgl. Reusser, Pauli & Elmer 2011; Schlichter 2012). Bisher vorliegende – mehrheitlich qualitative – Arbeiten zu Überzeugungen von Lehrkräften sowie deren Einschätzung der Gelingensfaktoren von Unterricht und den wahrgenommenen Belastungen und Defiziten unter Berücksichtigung der Qualifikation verweisen auf die besondere Bedeutung dieses Untersuchungsgegenstands hin (Hammel 2011; Bosse & Törner, 2013; Mohd Salleh & Darmawan 2013).

Folgende Forschungsfrage soll beantwortet werden: Lassen sich Unterschiede zwischen Mathematiklehrkräften mit und ohne eine Lehrbefähigung für das Fach Mathematik in Bezug auf unterrichtsbezogene und selbstbezogene Überzeugungen feststellen? Folgende Konstrukte werden unterschieden: (a) Überzeugungen über grundlegende mathematischen Unterrichtsprinzipien, insbesondere einer eher konstruktivistischen Sichtweise auf das Lehren und Lernen in Mathematik, sowie die Einschätzung zur Realisierung im eigenen Unterricht, (b) Überzeugungen zu qualitätsfördernden Unterrichtsbedingungen im Fach Mathematik, (c) Einschätzung der Beeinträchtigung des eigenen Mathematikunterrichts durch verschiedene Schülergruppen, (d) Einschätzung der eigenen (ausreichenden) Vorbereitung zum Unterrichten in Mathematik differenziert nach Inhaltsbereichen

Stichprobe sind die Mathematiklehrkräfte der nationalen Erhebung von TIMSS 2007, die in einer 4. Klasse an Grundschulen unterrichteten. Insgesamt liegen als Grundlage der quantitativen Auswertung Daten von 237 Lehrkräften aus einem standardisierten Fragebogen mit der Angabe vor, ob sie für Mathematik eine Lehrbefähigung (LB) erworben haben oder nicht.

Die fachfremd unterrichtenden Lehrkräfte stimmen im Mittel einer konstruktivistischen Sichtweise seltener zu und realisieren diese weniger häufig im Unterricht. Die Fachlehrkräfte fühlen sich zudem – unter Berücksichtigung drei verschiedener Inhaltsdomänen – generell im Hinblick auf ihren Unterricht besser vorbereitet als die fachfremd unterrichteten Lehrkräfte. Zur Frage nach qualitätsfördernden Unterrichtsbedingungen zeigt sich, dass die Ausbildungszeit durch die Fachlehrkräfte als bedeutsamer eingeschätzt wird. Relevant erscheint zudem der Befund, dass die Fachlehrerinnen und -lehrer das eigene Interesse an Mathematik als wichtiger für die Unterrichtsqualität einschätzen als ihre fachfremd unterrichtenden Kollegen.

Mehrheitlich zeigen sich auf Grundlage dieser Vergleiche Unterschiede der Lehrkräfte in den Überzeugungen zum Lehren und Lernen sowie den selbstbezogenen Überzeugungen in Abhängigkeit von der Qualifikation (mit oder ohne eine Lehrbefähigung) im Fach Mathematik. Die Befunde können erste Hinweise für die Erstellung von Kompetenzprofilen bzw. Merkmalsbeschreibungen fachfremd unterrichtender Mathematiklehrkräfte liefern, so dass beispielsweise Fortbildungen zielgruppenspezifischer entwickelt werden können.

Chair(s)

Richter, Dirk

Oettinghaus, Lars; Korneck, Friederike; Kunter, Mareike

Überzeugungen und Unterrichtshandeln

Theoretischer Hintergrund: Die Bestimmung unterrichtlichen Handelns ist eine theoretisch oft hervorgehobene Eigenschaft der professionellen Überzeugungen von Lehrkräften. International existiert allerdings eine nicht ausreichende empirische Befundlage mit geringen Fallzahlen, schwer vergleichbaren theoretischen Konzeptionen und meist qualitativen Ergebnissen (Fives & Buehl, 2012).

Im deutschsprachigen Raum werden zwar vergleichbare Messmethoden eingesetzt (Seidel et al., 2008; Dubberke et al., 2008; Lamprecht, 2011), die auf die „pedagogical content beliefs“ von Fennema, Carpenter und Loef (1990) zurückgehen. Dennoch zeigen die Ergebnisse für den Zusammenhang von Überzeugungen und Unterrichtshandeln Varianzaufklärungen von null (Kunter, et al., 2013) bis 18 Prozent (Lipowsky et al., 2003). Eine mögliche Ursache für diese heterogene Befundlage liegt im variationsreichen methodischen Zugang zum Unterrichtshandeln (Aufgabenanalyse, Selbstbeurteilung, Schülerbeurteilung oder Unterrichtsbeobachtungen).

Die vorliegende Studie untersucht daher für die Physik den Zusammenhang von Überzeugungen und Unterrichtshandeln mehrperspektivisch in standardisierten Unterrichtsminiaturen. Dabei werden bezüglich der Überzeugungen die Skalen „Überzeugung zum selbständigen Lernen“ und „Überzeugung zur Vermeidung von Transmission“ (ebenfalls bezugnehmend auf die „pedagogical content beliefs“) und für die Unterrichtsqualität die beiden Basisdimensionen „konstruktive Unterstützung“ und „kognitive Aktivierung“ (Kunter & Trautwein, 2013) zugrunde gelegt.

Forschungsfrage: Hängen die „Überzeugungen zum selbständigen Lernen“ und die „Überzeugungen zur Vermeidung von Transmission“ positiv mit der „konstruktiven Unterstützung“ bzw. der „kognitiven Aktivierung“ zusammen?

Die Fragestellung wird in diesem Beitrag zunächst aus der Perspektive der Unterrichtsbewertungen hospitierender Studierender und Referendare beantwortet.

Methode: Der Unterricht wird in Rahmen einer phasenübergreifenden Lehrveranstaltung von Studierenden im Hauptstudium und Referendaren gehalten und ist in Bezug auf das Themengebiet (Freihandexperiment aus der Mechanik), die Unterrichtszeit (12-15 Minuten) und die Klassenstärke (halbe Klassen) komplexitätsreduziert und standardisiert, die Schulformen und Jahrgangsstufen variieren (Gymnasien oder Gesamtschulen, 7.-11. Klasse). Die angehenden Physiklehrkräfte unterrichten unabhängig von ihrem Studiengang in beiden Schulformen. Jede Lehrkraft unterrichtet zwei Mal jeweils eine Klassenhälfte. Diese beiden Sequenzen werden videografiert und von durchschnittlich 10 hospitierenden Peers sowie den durchschnittlich 12 Schülern hinsichtlich ihrer Unterrichtsqualität bewertet.

Die Stichprobe für diese Analyse setzt sich aus 76 Studierenden zusammen und umfasst eine Vollerhebung der Physiklehramtsstudierenden in Frankfurt seit dem Wintersemester 2012/13.

Die Überzeugungsmerkmale „Überzeugung zum selbständigen Lernen“ und „Überzeugung zur Vermeidung von Transmission“ konnten erfolgreich im Rahmen der klassischen Testtheorie konstruiert werden ($\alpha = .75$ bis $.79$). Die Skalen basieren auf dem Überzeugungsinstrument der proΦ-Studie und beinhalten 12 bis 16 Items pro Skala (Oettinghaus, in Arbeit). Ausgangspunkt des Messinstruments für die Unterrichtsqualität ist der Fragebogen der COACTIV-Studie (Baumert et al., 2009). Er besitzt in der hier eingesetzten und entsprechend adaptierten Form 7 bis 8 Items pro Skala.

Für die Skalenkonstruktion der Unterrichtsqualität wurden für jedes Unterrichtsqualitätsmerkmal drei latente Parcel-Scores (WLE-Schätzer aus einem dreidimensionalen Partial-Credit-Modell) als Datengrundlage für die Mehrebenen-Strukturgleichungsmodelle gebildet. Die Skalen „konstruktive Unterstützung“ und „kognitive Aktivierung“ konnten so reliabel modelliert werden ($ICC(2) = .82$ -.83).

Ergebnisse: Erste Analysen zeigen für die Zusammenhänge der „konstruktiven Unterstützung“ und der „kognitiven Aktivierung“ mit den Überzeugungen Varianzaufklärungen von vier bis 16 Prozent und übersteigen damit die berichteten Zusammenhänge anderer, ebenfalls auf Unterrichtsbeobachtungen basierender, Studien (Seidel et al., 2008). Im Vortrag wird der Einfluss der Überzeugungsskalen auf die beiden Unterrichtsqualitätsmerkmale, unter Kontrolle des Studiengangs der Lehrkraft sowie der Schulform der Schüler, in einem Mehrebenen-Regressionsmodell vertieft diskutiert.

Diskussion und Ausblick: Die Ergebnisse dieser Studie zeigen zum einen, dass Unterrichtshandeln auch mit Unterrichtsbeobachtungen effizient und reliabel erfasst werden kann. Zum anderen können für die fachbezogenen Überzeugungen stärkere Zusammenhänge zum Unterrichtshandeln nachgewiesen werden als für das fachbezogene Professionswissen (Baumert & Kunter, 2011; Vogelsang & Reinhold, 2012). Eine höhere Auflösung der Dimensionen der Unterrichtsqualität soll durch Videoanalyse im weiteren Projektverlauf erreicht werden, um den Zusammenhang der Überzeugungen auch zu den verschiedenen Subdimensionen der Unterrichtsqualitätsmerkmale diskutieren zu können.

Donnerstag
12. März

9:00 Uhr

GA 03/149

Chair(s)
Richter, Dirk

Jurik, Verena; Häusler, Janina; Stubben, Sina; Seidel, Tina

Das Zusammenspiel von Schülervoraussetzungen, Lehrereinschätzungen und Schülerbeteiligung: Ergebnisse einer Videostudie

Theoretischer Hintergrund: Eine aktive Beteiligung an Interaktionen im Unterricht wird als positiv für den Erfolg und die Lernentwicklung von SchülerInnen eingestuft (Mercer, 1996; Webb, 2009). Dabei ist die Beteiligung vom Zusammenspiel der individuellen Voraussetzungen abhängig, aber auch vom Verhalten, den Einstellungen und Erwartungen der Lehrkräfte (Hofer, 1997; Seidel & Reiss, 2014). Untersuchungen in Unterrichtsfächern zeigten, dass Mädchen und Jungen solche Fachdomänen oder schulrelevante Lernaktivitäten bevorzugen, die in Übereinstimmung mit Geschlechtsstereotypen sind (Hannover & Kessels, 2008). Weitere Studien machten sichtbar, dass generell der Anteil, zu welchem die SchülerInnen im Unterricht sprechen, gering ist und der Redeanteil der Lehrkräfte überwiegt (z.B. Seidel, 2003). Bezüglich des Aufrufverhaltens konnte gezeigt werden, dass Lehrkräfte häufiger mit leistungsstarken als mit schwächeren SchülerInnen interagieren (Brophy & Good, 1976; Gage & Berliner, 1986; Lipowsky et al., 2007) und nach dem Stellen einer Frage bei guten SchülerInnen länger warten, bis sie eine Antwort bekommen und häufiger Feedback geben (Hofer, 1981). Sacher konnte zudem herausstellen, dass im Allgemeinen ein hoher Anteil an SchülerInnen, der sich meldet, nicht aufgerufen wird und nicht dran kommt (Sacher, 1995).

Fragestellungen: Im vorliegenden Beitrag wurden die Beteiligung der SchülerInnen im Deutsch- und Mathematikunterricht durch die Anzahl und Dauer der Meldungen und die Anzahl der Aufrufe sowie die Anzahl und Dauer der Schülerbeiträge gemessen. Des Weiteren wurde die Schülerbeteiligung mit den gemessenen Schülervoraussetzungen (allgemeine kognitive Fähigkeiten, Vorwissen, Interesse und Selbstkonzept) und den Einschätzungen der Lehrkräfte von diesen Voraussetzungen in Verbindung gebracht.

Folgende Forschungsfragen stehen im Fokus dieses Beitrages: (1) Unterscheidet sich die Beteiligung der SchülerInnen in den Fächern Deutsch und Mathematik? (2) Welche der untersuchten Schülervoraussetzungen stehen in Zusammenhang mit einer aktiven Beteiligung am Deutsch- bzw. Mathematikunterricht? (3) Welche der Einschätzungen der Lehrkräfte von Schülervoraussetzungen stehen in Zusammenhang mit einer aktiven Beteiligung am Deutsch- bzw. Mathematikunterricht?

Methode.

Stichprobe: Der vorliegende Beitrag basiert auf Daten der Studie „Interaction“. Die Stichprobe umfasst N = 20 Gymnasialklassen der 8. Jahrgangsstufe mit insgesamt N = 472 Schülerinnen und Schülern (3 Mädchenklassen, 53% Mädchen in koedukativen Klassen) aus Bayern.

Design: Die SchülerInnen wurden bezüglich ihrer kognitiven Fähigkeiten (KFT) zu Beginn des Schuljahres, zu ihrem Vorwissen (16 TIMSS-Items) und ihrem Interesse am Fach sowie zu ihrem Selbstkonzept zu Beginn und zum Ende des Schuljahres in einer Unterrichtsstunde (45 Minuten) befragt. Während des Schuljahrs wurde je eine Stunde im Deutsch- und Mathematikunterricht videografiert.

Statistische Analysen: Die Videos wurden mit der Software Mangold Interact kodiert und über die Einheiten der Sprecher ausgewertet. Des Weiteren wurden Korrelations- und Regressionsanalysen mit SPSS (Version 22) durchgeführt.

Ergebnisse: Erste Ergebnisse zeigen, dass sich im Durchschnitt mehr Mädchen am Deutsch- und mehr Jungen am Mathematikunterricht beteiligen. Für die individuellen SchülerInnen zeigen sich jedoch keine klar gegenläufigen Ergebnisse in den beiden unterschiedlichen Fächern. Im Gegenteil: Wer sich in Deutsch oft beteiligt, tut dies auch in Mathematik – über die Geschlechter hinweg. Dabei zeigen Jungen in beiden Fächern eine höhere Beteiligung. Bezüglich des Zusammenspiels von Beteiligung und Schülermerkmalen, scheinen in Mathematik Vorwissen, Interesse und Selbstkonzept mit der Beteiligung zusammenzuhängen, vor allem bei Mädchen. Für Deutsch zeigt sich hier jedoch kein deutlicher Zusammenhang. Bezogen auf die Einschätzungen der Lehrkräfte zeigt sich, dass die Schülerbeteiligung mit der Schülereinschätzung der Lehrkräfte zu Beginn des Schuljahres enger zusammenhängt als mit der zum Ende des Schuljahres. Vor allem in Deutsch hängt die Einschätzung des Schülerelbstkonzepts mit einer vermehrten Beteiligung zusammen. Für Jungen scheint der Zusammenhang zwischen Schülerbeteiligung und Lehrereinschätzungen insgesamt enger. Die Ergebnisse zeigen, dass es interessante Unterschiede zwischen den Geschlechtern und unterschiedlichen Schulfächern in der näheren Betrachtung von Unterrichtsbeteiligung gibt. Sie deuten auch darauf hin, dass je nach Unterrichtsfach unterschiedliche Schülervoraussetzungen Beteiligungsentscheidungen beeinflussen und diese Beteiligung von den Lehrkräften unterschiedlich gedeutet wird.

Donnerstag
12. März

9:00 Uhr

GA 03/149

Chair(s)
Richter, Dirk

Herppich, Stephanie; Wittwer, Jörg; Nückles, Matthias; Renkl, Alexander

Negative Effekte von Lehrerfahrung? Interaktivität im Tutoring fördert Lernen. Aber erfahrene Tutoren sind weniger interaktiv als unerfahrene.

Theoretischer Hintergrund: Strategien, die man zur Instruktion anwendet, bestimmen maßgeblich deren Lernwirksamkeit (Hattie, 2009). Die Effektivität von Eins-zu-Eins-Tutoring wird häufig auf die besonderen Instruktionsstrategien im Zusammenhang mit den Kommunikationsmöglichkeiten dieser Instruktionsform zurückgeführt (Graesser, Person & Magliano, 1995). Wenn der Tutor durch seine Instruktionsstrategien ermöglicht, dass Tutor und Tutee sich in einem interaktiven Austausch befinden, gilt dies als besonders lernförderlich (Chi, 2009). Forschungsergebnisse zeigten, dass Tutoren sich kaum optimal interaktiv verhalten (Graesser et al., 1995). Zusammenhänge zwischen Interaktionsverhalten und Lernen wurden jedoch selten untersucht. Weiterhin wird diskutiert, welche Rolle Lehrerfahrung für Instruktionseffektivität spielt (VanLehn, 2007). Studien deuten an, dass erfahrene Tutoren häufiger interaktive Instruktionsstrategien verwenden als unerfahrene (Chae, Kim & Glass, 2005). Direkte Vergleiche, bei denen Instruktionseffektivität berücksichtigt wird, fehlen jedoch.

Fragestellung: In dieser empirischen Studie untersuchten wir, in welcher Beziehung interaktivitätsförderliches Verhalten des Tutors im Tutoringdialog zum Lernen eines Tutees steht und wie Lehrerfahrung die Interaktivität beeinflusst. Wir testeten folgende Hypothesen: (1) Interaktive Instruktionsstrategien des Tutors fördern Lernen. (2) Erfahrene Tutoren verwenden häufiger interaktive Instruktionsstrategien als unerfahrene. (3) (a) Erfahrene Tutoren fördern das Lernen des Tutees stärker als unerfahrene. (b) Dies ist so, weil erfahrene Tutoren häufiger interaktive Instruktionsstrategien verwenden als unerfahrene.

Methode: N = 46 zufällig zusammengesetzte Tutor-Tutee Dyaden diskutierten einen Text zum menschlichen Herz-Kreislauf-System. Lehrerfahrene Tutoren waren 21 Biologielehrkräfte. Lehruferfahrene Tutoren waren 25 Biologiestudierende. Tutees waren Siebtklässlerinnen und Siebtklässler. Die abhängigen Variablen waren (1) die Interaktivität der Tutor-Interaktionsstrategien, kodiert aus den transkribierten Tutoringdialogen unter Adaption des Statement Coding von Chi, Siler, Jeong, Yamauchi und Hausmann (2001). Nach Chi, Roy und Hausmann (2008) bestimmten wir die Interaktivität der Instruktionsstrategien als Quotient zwischen Scaffoldingstatements und Erklärungsstatements des Tutors. Weitere abhängige Variable war (2) der Lernzuwachs des Tutees, gemessen als relativierte Differenz (a) zwischen Prä- und Posttest zu Konzepten des Herz-Kreislauf-Systems und (b) zwischen Prä- und Posttest zum mentalen Modell des Herz-Kreislauf-Systems.

Ergebnisse: Lineare Regressionsanalysen zeigten, dass Interaktivität das Lernen von Konzepten nicht, $R\text{-Quadrat} = .01$, $F(1, 44) = 0.29$, $b = .03$, $p > .05$, $f = 0.08$ (kleiner Effekt), das Lernen des mentaler Modelle aber durchaus förderte, $R\text{-Quadrat} = .07$, $F(1, 44) = 3.40$, $b = 0.13$, $p < .05$, $f = 0.28$ (mittlerer Effekt). Erfahrene Tutoren verwendeten seltener interaktive Instruktionsstrategien als unerfahrene, $R\text{-Quadrat} = .16$, $F(1, 44) = 8.24$, $b = -0.50$, $p < .05$, $f = 0.43$ (großer Effekt). Eine Mediationsanalyse (Hayes, 2008) für das Konzeptlernen zeigte, dass erfahrene Tutoren, ohne Kontrolle weiterer Variablen, Lernen stärker fördern als unerfahrene, $R\text{-Quadrat} = .09$, $F(1, 44) = 4.26$, $b = 0.15$, $p < .05$, $f = 0.31$ (mittlerer Effekt; vgl. Herppich, Wittwer, Nückles & Renkl, 2014). Vermittelt über die Interaktivität förderten unerfahrene Tutoren das Konzeptlernen jedoch mehr als erfahrene, wie der Test des indirekten Effekts von -0.05 ergab ($\kappa\text{-Quadrat} = .10$ (mittlerer Effekt, Preacher & Kelley, 2011), 95% Bootstrappingkonfidenzintervall mit 10000 Stichprobenziehungen: KIB [-0.10, -0.005]). Eine analoge Mediationsanalyse für das Lernen mentaler Modelle zeigte keinen absoluten Unterschied für Lernen bei unerfahrenen und erfahrenen Tutoren $R\text{-Quadrat} < .01$, $F(1, 44) = 4.26$, $b = 0.01$, $p > .05$, $f = 0.01$ (kleiner Effekt). Als Folge von Interaktivität förderten unerfahrene Tutoren jedoch wiederum das Lernen mehr als erfahrene (indirekter Effekt = -0.08, $\kappa\text{-Quadrat} = .12$ (mittlerer Effekt), KIB [-0.14, -0.03]).

Interaktive Instruktionsstrategien förderten somit das Lernen. Dies galt eindeutig für den Erwerb komplexer mentaler Modelle als für das Erlernen einzelner Konzepte. Früheren Befunden widersprechend, waren lehrunerfahrene Tutoren interaktiver und förderten Lernen vermittelt über ihre Interaktivität stärker als lehrerfahrene. Möglicherweise neigen gerade erfahrene Lehrkräfte, stärker als Studierende, bei einer konzeptuellen Domäne wie dem Herz-Kreislauf-System zum ‚Dozieren‘. Interpretationen zum Instruktionsverhalten erfahrener und unerfahrener Tutoren, welche die Domäne nicht berücksichtigen, sollten überprüft werden.

Prediger, Susanne; Poehler, Birte; Weinert, Henrike

Sind Textaufgaben für sprachlich Schwache wirklich schwieriger? Analysen zu Prozent-Aufgaben

Sprachlich bedingte Leistungsdisparitäten in Mathematiktests sind inzwischen oft beschrieben, aber bislang nur begrenzt gut erklärt (Abedi, 2006; OECD, 2007; Prediger et al., 2013). Oft wurde festgestellt, dass sprachlich schwache Lernende gerade in Textaufgaben die größten Schwierigkeiten haben, und zwar auch signifikant mehr als ihre sprachlich stärkeren Peers (Duarte et al. 2011). Dagegen sprechen andere Studien, die zeigen, dass Kontextanbindung die Zugänglichkeit von Aufgaben auch erhöhen kann (van den Heuvel-Panhuizen 2005). Dies wird auch für eine graphische Stützung von Aufgaben, z.B. durch einen Prozent-streifen angenommen.

In einem schriftlichen Test zur Prozentrechnung (mit einer Stichprobengröße von $n = 250$) wurden da-her Prozentaufgaben in drei Formaten verglichen: Textaufgaben, graphisch gestützte und entkleidete Auf-gaben, jeweils zu drei verschiedenen Problemtypen: „Prozentwert gesucht“, „Grundwert gesucht“, „Grundwert gesucht bei angegebener Reduktion“ (Pöhler, Prediger, Weinert 2015). Die Sprachkompetenz der getesteten Siebtklässlerinnen und Siebtklässler wurde mit einem C-Test erfasst. Die schriftlichen Tests wurden durch klinische Interviews zur Erfassung der Lösungsprozesse („cognitive lab“) trianguliert, so dass erste Erklärungsansätze für die statistisch gewonnenen Befunde identifiziert werden können.

Überprüft im schriftlichen Test wurden folgende Hypothesen: (H1) Textaufgaben sind schwieriger als entkleidete Aufgaben. (H2) Graphisch gestützte Aufgaben sind leichter als Textaufgaben, aber schwerer als entkleidete Aufgaben. (H3) Sprachlich schwache Lernende haben mit anderen Aufgabenformaten Probleme als sprachlich starke, insbesondere haben sie relativ mehr Schwierigkeiten mit Textaufgaben.

Die quantitative Auswertung des Tests ergab, dass die Haltbarkeit der ersten beiden Hypothesen vom Be-kanntheitsgrad der Problemtypen abhängt. Für die vertrauten Problemtypen „PW gesucht“ und „GW ge-sucht“ wird die entkleidete Aufgabe signifikant am häufigsten gelöst, die Textaufgabe am seltensten, die Hypothesen bestätigen sich also. Für den weniger vertrauten und komplexeren Problemtyp „Grundwert gesucht bei angegebener Reduktion“ dagegen drehen sich Schwierigkeitsrangfolgen um.

Die Analysen der Lösungsprozesse in den klinischen Interviews stützen die Interpretation, dass für we-niger vertraute und komplexere Problemtypen der Kontext der Textaufgaben die Zugänglichkeit erhöhen kann, so das der Abstand geringer wird.

Die dritte Hypothese muss verworfen werden. Zwar lösen die sprachlich Schwachen die Textaufgaben signifikant seltener als die sprachlich Starken, doch ist das Verhältnis der Aufgabenformate zueinander in beiden Subsamples genau gleich. Eine Aufgabe zur Erhebung des konzeptuellen Verständnisses dagegen wird von den sprachlich Schwachen mehr als statistisch erwartbar schlechter gelöst.

Dies führt zur neuen Hypothese, dass nicht allein das Textverständnis, sondern vor allem das konzeptu-elle Verständnis für Konzepte der Prozentrechnung für die Leistungsdisparitäten zwischen sprachlich star-ken und schwachen Lernenden entscheidend sind.

Donnerstag
12. März

9:00 Uhr

GA 03/149

Chair(s)
Richter, Dirk

Schüle, Christoph; Besa, Kris-Stephen; Denger, Corinna; Feßler, Felix; Arnold, Karl-Heinz
Die Entwicklung allgemeiner Selbstwirksamkeitserwartungen im Rahmen schulpraktischer Studienelemente

Neben pädagogischem, fachlichem und fachdidaktischem Wissen brauchen Lehrkräfte für die psychische Regulationsfähigkeit innerhalb ihres professionellen Handlungskontextes ein Vertrauen in die eigenen Fähigkeiten mit Hilfe derer Sie schwierige Anforderungen ihrer beruflichen Tätigkeit bewerkstelligen können (Baumert & Kunter, 2006). Ein solches Vertrauen wird entsprechend der sozial-kognitiven Theorie Banduras (1986) als individuelle Selbstwirksamkeitserwartung bezeichnet.

Bandura (1986) folgend sind insbesondere eigene sowie stellvertretene Erfolgserfahrungen aber auch soziale Überzeugungen sowie positive affektive und physiologische Zustände potentielle Quellen wahrgenommener Selbstwirksamkeit. Die Forschung zeigt hierzu, dass Studierende solchen Quellen vor allem im Schulpraktikum begegnen (Bach, 2013; Fives, Hamman & Olivarez, 2007; Klassen & Durksen, 2014; Moulding, Stewart & Dunmeyer, 2014). Die Mehrzahl der Untersuchungen konzentriert sich hierbei allerdings vorrangig auf das bereichsspezifische Konstrukt der individuellen Lehrerselbstwirksamkeitserwartung. Unbeachtet bleibt in diesem Zusammenhang jedoch die Frage, inwiefern Schulpraktika auch einen Beitrag zur allgemeinen Wahrnehmung eigener Ressourcen des Umganges mit alltäglichen Schwierigkeiten und Barrieren leisten und so die intentionale sowie regulative Selbst- und Persönlichkeitsentwicklung der Studierenden beeinflussen. Der vorliegende Beitrag prüft daher die Frage, wie sich die bereichsübergreifende allgemeine Selbstwirksamkeitserwartung im Schulpraktikum entwickelt.

Hierzu wurden insgesamt 460 Lehramtsstudierende zu sechs Messzeitpunkten mittels der Skala zur Allgemeinen Selbstwirksamkeitserwartung von Schwarzer & Jerusalem (1999) befragt. Die erste Befragung (t1) erfolgte hierbei in der Einführungswoche, vor dem eigentlichen Beginn der Lehrveranstaltungen. Der zweite Messzeitpunkt (t2) schloss sich an ein Hospitationspraktikum des ersten Semesters an. Die dritte Befragung (t3) wurde nach einem Tagespraktikum des zweiten Semesters durchgeführt, in dem die Studierenden ihre erste Unterrichtsstunde hielten. Der vierte Messzeitpunkt (t4) wurde vor einem vierwöchigen allgemeinen Blockpraktikum, welches in den Semesterferien zwischen dem dritten und vierten Semester, stattfand, realisiert. Daran schloss sich ein fünfter Messzeitpunkt (t5) an. Darüber hinaus wurde im Rahmen einer Pflichtvorlesung des vierten Semesters eine Follow-up-Befragung durchgeführt (t6).

Die Ergebnisse eines Baseline-und Neighbour-Change Modells (Chi-Quadrat = 54.37; df = 33, $p > .1$; CFI = .997; RMSEA = .017; SRMR = .048) zeigen über die sechs Messzeitpunkte hinweg einen kontinuierlichen linearen Anstieg der allgemeinen Selbstwirksamkeitserwartung, welcher sich insbesondere im Verlauf des vierwöchigen Blockpraktikums als signifikant erweist.

Vor dem Hintergrund dieser Befunde bestätigt sich die Annahme, dass schulpraktische Erfahrungen einen Beitrag zur Wahrnehmung allgemeiner Ressourcen des Umganges mit Schwierigkeiten und Anforderungen leisten.

Donnerstag
12. März

9:00 Uhr

GB 03/49

Chair(s)
Schüle, Christoph

Biermann, Antje; Karbach, Julia; Spinath, Frank; Brünken, Roland

Effekte von Merkmalen des Praktikums und der Persönlichkeit auf die selbsteingeschätzte Kompetenz von Lehramtsstudierenden

Schulpraktische Studien während der universitären Phase der Lehramtsausbildung werden als bedeutende Faktoren für die Entwicklung der Handlungskompetenz von Lehramtsstudierenden betrachtet. Praxisphasen werden dann als effektiv angesehen, wenn eine strukturierte Betreuung durch Universität und Praktikumschule gewährleistet ist (z. B. Gröschner & Seidel, 2012; Hascher & Kittinger, 2014), eine effektive Theorie-Praxis-Verknüpfung hergestellt werden kann (z. B. Darling-Hammond, Hammerness, Grossman, Rust & Shulman, 2005) und Studierende sich in einem Team austauschen können (Hudson, Miller, Salzberg & Morgan, 1994). Darüber hinaus werden personenbezogene Merkmale wie kognitive Leistungen oder die Persönlichkeit als Determinanten professioneller Kompetenz diskutiert (Hanfstingl & Mayr, 2007; Kunter, Klusmann, Baumert, Richter, Voss & Hachfeld, 2013). Studien, die die Wirksamkeit einzelner Faktoren in Hinblick auf die Handlungskompetenz untersuchen, liegen sowohl im deutsch- als auch im englischsprachigen Raum vor; eine gemeinsame Betrachtung von Merkmalen von Schulpraktika und Merkmalen der Lehramtsstudierenden ist eher selten.

Im Rahmen der Studie SioS-L (Studie zu individuellen und organisationalen Einflüssen auf den Studien- und Ausbildungserfolg in der Lehrerausbildung) wurden 443 Studierende vor dem ersten Schulpraktikum hinsichtlich personenbezogener Merkmale (Abiturnote, Persönlichkeit) sowie nach dem Praktikum hinsichtlich ihrer Einschätzung der Bedingungen des Praktikums (Betreuung, Theorie-Praxis-Verknüpfung, Zusammenarbeit im Team) und ihrer Selbsteinschätzung der eigenen Kompetenz (in Anlehnung an Oser & Oelkers, 2001) befragt.

Die Ergebnisse der hierarchischen Regressionsanalysen zeigen, dass durchgehend die wahrgenommene Theorie-Praxis-Verknüpfung als Merkmal des Praktikums den höchsten Anteil der Varianz in Bezug auf die vier selbst eingeschätzten Kompetenzmerkmale aufklärt. Als Persönlichkeitsmerkmal können Gewissenhaftigkeit und Neurotizismus für unterschiedliche Kompetenzbereiche jeweils einen zusätzlichen Varianzanteil aufklären.

Die Ergebnisse liefern wichtige empirische Evidenz für die Wirksamkeit einzelner Merkmale von Schulpraktika, insbesondere der Relevanz einer effektiven Anleitung zur Theorie-Praxis-Verknüpfung für die Entwicklung von Handlungskompetenz schon in einer frühen Phase der Ausbildung. Vor dem Hintergrund der querschnittlichen Erhebung an einer Universität müssen die Ergebnisse jedoch hinsichtlich einer eingeschränkten Generalisierbarkeit betrachtet werden.

Donnerstag
12. März

9:00 Uhr

GB 03/49

Chair(s)

Schüle, Christoph

Lohse, Karoline; Kreutzmann, Madeleine; Lasson, Andrea; Boekhoff, Jannes; Tietze, Wolfgang

Lernort Praxis – Evaluation eines Pilotprogramms zur Professionalisierung frühpädagogischer Fachkräfte

Bei der Beurteilung der Ausbildungsqualität frühpädagogischer Fachkräfte wird zunehmend auf eine nur wenig systematische Verzahnung theoretischer und praktischer Ausbildungsinhalte hingewiesen (Behr & Walter, 2012, Flämig, 2011). Mit dem Bundesprogramm Lernort Praxis (gefördert durch das BMFSFJ) soll die Kooperation zwischen den Lernorten Fach-/ Hochschule und Kindertagesstätte gestärkt und durch eine verbesserte Anleitungskonzeption am Lernort Praxis ein wesentlicher Beitrag zur Professionalisierung zukünftiger frühpädagogischer Fachkräfte geleistet werden. An 76 Pilotstandorten in sieben Bundesländern werden zur Umsetzung dieses Vorhabens in den teilnehmenden Kindertagesstätten Praxismenator/inn/en eingesetzt, die als zusätzliche Ressource in den Einrichtungen eine Anschubfunktion für die Qualitätsentwicklung des Lernorts Praxis innehaben. Ihre Tätigkeit soll dazu beitragen, dass in den Einrichtungen Organisations- und Managementstrukturen für die Praxisausbildung etabliert, Ausbildungs- und Anleitungskonzepte entwickelt und Kooperationsstrukturen mit den Fachschulen bzw. (Fach-)Hochschulen aufgebaut und institutionalisiert werden. Für diese Aufgaben werden Praxismenator/inn/en gemeinsam mit den Führungskräften der Einrichtungen über ein eigens entwickeltes Curriculum geschult. Das Weiterbildungscurriculum ist darauf ausgerichtet den Lernort Praxis als System zu entwickeln. Dabei werden die unterschiedlichen Akteure, die an der Ausbildung am Lernort Praxis beteiligt sind, ihrer Rolle und Qualifikation entsprechend berücksichtigt.

Die formative Evaluation des Programms begleitet die Entwicklung und Implementierung des Curriculums in der Weiterbildung. Die summativ Evaluation soll Aussagen zu Wirkungen des Pilotprogramms unter der Berücksichtigung individueller und kontextueller Bedingungen am Lernort Praxis ermöglichen und der Identifikation spezifischer Wirkmechanismen dienen. Im Rahmen einer Vorher-Nachher-Messung sollen programmbedingte Veränderungen in Struktur- und Kontextmerkmalen (z.B. Rahmenbedingungen der Praxisanleitung, Qualität der Lernortkooperation) als auch in Aspekten des professionellen Handelns (z.B. pädagogische Orientierungen und Einstellungen, Selbst- und Rollenverständnis) der verschiedenen Akteure am Lernort Praxis ermittelt werden. Zu diesem Zweck werden im Rahmen von Prä- und Posterhebung die unterschiedlichen Perspektiven der Akteure am Lernort Praxis erfasst.

Die Präerhebung zur Beschreibung der Ausgangslage in den teilnehmenden Einrichtungen erfolgte im Zeitraum 07.-09.2014 und umfasst Daten von 76 Praxismenator/inn/en, 170 Einrichtungsleitungen, 350 Praxisanleiter/inne/n sowie 41 Praxisdozent/inn/en von Fach-/Hochschulen und 200 Auszubildenden und Studierenden, welche zum Zeitpunkt der Erhebung in den Einrichtungen tätig waren.

In diesem Beitrag sollen zunächst das Design sowie zentrale Untersuchungsvariablen der summativen Evaluation erläutert und die Daten der Präerhebung im Rahmen einer Gegenstandsbeschreibung des Lernorts Praxis vorgestellt werden. Ein vorläufiger Überblick bestätigt zunächst die in früheren Veröffentlichungen bemängelte unsystematische Zusammenarbeit der Lernorte Schule und Praxis. So geben 36% der Praxismenator/innen (N=75) an, dass in ihrer Einrichtung für keine der Fachschulen/(Fach-)Hochschulen, von denen Auszubildende/Studierende betreut werden, schriftliche Kooperationsvereinbarungen vorliegen. Nur 25% der befragten Dozenten (N=41) vom Lernort Schule haben Kenntnis von einem Kooperationskonzept ihrer Institution für die Zusammenarbeit mit dem Lernort Praxis. Immerhin 80% der Praxisanleiter/innen (N=329) geben an innerhalb des vergangenen Kitajahres Kontakt zu Praxisdozent/innen gehabt zu haben. 25% der Praxisanleiter/innen (und 38% der Leitungen) kennen jedoch keinen konkreten Ansprechpartner am Lernort Schule, welchen sie bei Fragen zum Praktikum kontaktieren könnten.

Ein ähnliches Bild ergibt sich für den Entwicklungsstand des Ausbildungsortes Lernort Praxis. 59% der Einrichtungsleitungen (N=145) und ebenso 60% der Praxisanleiter/innen (N=321) geben an, dass es in ihrer Einrichtung ein schriftliches Konzept für die Praxisanleitung gibt, welches sich unter anderem an die anzuleitenden Personen richtet. Von 184 befragten Auszubildenden und Studierenden geben jedoch nur 56% an dieses Konzept zu kennen.

Weitere Analysen werden detaillierte Einblicke für die Beschreibung der Ausgangslage in den am Programm beteiligten Praxislernorten ermöglichen. Bei ihrer Darstellung werden insbesondere Aspekte hervorgehoben, die durch das Weiterbildungscurriculum angesprochen und in ihrer Entwicklung gefördert werden sollen. Darüber hinaus sollen Fragestellungen hinsichtlich der Mehrebenenstruktur der Daten und Möglichkeiten ihrer Interpretation diskutiert werden.

Donnerstag
12. März

9:00 Uhr

GB 03/49

Chair(s)
Schüle, Christoph

Spaude, Elena; Starker, Ulrike

Erwerben Lehramtsstudierende Handlungskompetenzen in einem Planspiel „Schulalltag“?

Im Schulalltag sind oft schnelle Reaktionen auch in komplexen und emotional zugespitzten Situationen erforderlich. Lehrer und Lehrerinnen wägen ihr Verhalten nicht ab, sondern zeigen Überreaktionen (Humpert & Dann, 2001). So kommt es zu Verhaltensweisen wie Sarkasmus, aus Ärger schreien oder das Blamieren und Bestrafen, die unweigerlich negative Folgen für den weiteren Unterrichtsverlauf haben (Lewis, Romi, Katz & Qui, 2008; Mainhard, Brekelmans & Wubbels, 2011; Romi, Lewis, Roache & Riley, 2011).

Um Lehramtsstudierende für diese Situationen zu sensibilisieren und die Entwicklung und Reflexion von Handlungsalternativen zu ermöglichen, bietet die Johannes Gutenberg-Universität (JGU) Mainz seit 2013 das Planspiel „Schulalltag“ an.

Fragestellung: Die vorliegende Studie beschäftigt sich mit der Frage, ob sich das Planspiel „Schulalltag“ sich zum Erwerb von Handlungskompetenzen eignet. Dazu werden folgende Hypothesen untersucht: (1) Nach der Teilnahme am Planspiel ändern die Teilnehmer ihre Strategien für die Lösung schwieriger Unterrichtssituationen häufiger als die Kontrollgruppe. (2) Nach der Teilnahme am Planspiel geben die Teilnehmer im Vergleich zur Kontrollgruppe mehr aufgreifende Reaktionen an als vor dem Planspiel. (3) Nach der Teilnahme am Planspiel geben die Teilnehmer im Vergleich zur Kontrollgruppe weniger punitive Reaktionen an als vor dem Planspiel.

Methode.

Design und Durchführung: Das Planspieltraining besteht aus drei Terminen: Ein Vorbereitungstag zur Einführung in die Thematik, ein Planspiel-Tag, und ein Nachbereitungstag zur Reflexion sowie die Einbettung der Erfahrungen in einen theoretischen Kontext. Die Untersuchung fand im Prä-Post-Design mit zwei Kontrollgruppen statt.

Stichprobe: Am Planspiel „Schulalltag“ im Sommersemester 2014 nahmen N = 34 Lehramtsstudierende (19 davon weiblich) teil, die Kontrollgruppe bestand aus N = 59 Studierenden (47 davon weiblich) im gleichen Studienabschnitt.

Versuchsmaterial und Kodierung: Der Fragebogen enthält vier Fallvignetten zu kritischen Unterrichtssituationen. Die Befragten geben an, wie sie mit der Situation umgehen würden. Die offenen Antworten auf die Fallvignetten wurden mithilfe des Kategoriensystems BAVIS von Humpert und Dann (2001) kategorisiert. BAVIS enthält 10 Kategorien für Lehrerreaktionen, zuzüglich einer Kategorie für „Sonstige Lehrermaßnahmen“; die Kategorien „Kontern“, „Thema integrieren“ und „Aufgreifen nach der Stunde“ wurden ergänzt. Die Kategorien werden zusammengefasst als Neutrale Maßnahmen: Beobachten/Ignorieren, Abbrechen, Mahnen; Punitiv Maßnahmen: Drohen, Bestrafen, Herabsetzen; Sozial-integrative Maßnahmen: Kompromiss vorschlagen, Integrieren, Ermutigen, Einfühlen, Kontern, Thema integrieren, nach der Stunde aufgreifen. Registriert wurden jeweils die Kategorie der ersten Reaktion und die Anzahl der unterschiedlichen Reaktionen. Die Übereinstimmungen von zwei unabhängigen Ratern fallen mit $\kappa = .37$ bis $.62$ bei der Kategorisierung der ersten Reaktion sowie mit $r = .58$ bis $.79$ bei der Anzahl der vergebenen Kategorien annehmbar bis gut aus.

Ergebnisse: In der einfaktoriellen Varianzanalyse ergab sich für die Änderung der ersten Reaktion kein signifikanter Unterschied zwischen den Planspiel-Teilnehmern und der Kontrollgruppe. Entgegen Hypothese 1 änderten die Teilnehmer ihre Strategien für die Lösung schwieriger Unterrichtssituationen also nicht häufiger als die Kontrollgruppe. Entgegen Hypothese 2 ergab sich weder für die Anzahl aufgreifender Verhaltensweisen als erste Reaktion noch für den Anteil aufgreifender Verhaltensweisen ein signifikanter Unterschied zwischen den beiden Gruppen.

Für die Anzahl punitiver Verhaltensweisen als erste Reaktion ($F(1, 59) = .08, p > .05$), ergab sich zwar kein signifikanter Unterschied zwischen den beiden Gruppen, wohl aber für den Anteil punitiver Verhaltensweisen mit $F(1,57) = 5.80, p < .05$. Entsprechend Hypothese 3 geben die Teilnehmer nach ihrer Teilnahme am Planspiel im Vergleich zur Kontrollgruppe also anteilig signifikant weniger punitive Reaktionen an als vor dem Planspiel. Das Ergebnis zeigt, dass das Verhaltensspektrum der Lehramtsstudierenden in kritischen Situationen durch das Planspiel „Schulalltag“ etwas erweitert wird und das punitive Element etwas zurückgenommen wird. Offen bleibt jedoch, ob sich dies auch in eine reale Situation übersetzen lässt.

Chair(s)

Schüle, Christoph

Egbert, Björn

Die Unternehmensgründung im Schülerverständnis – Eine phänomenographische Untersuchung in der Sekundarstufe I

Angesichts einer sich wandelnden Arbeits- und Lebenswelt und den damit verbundenen Herausforderungen des täglichen Lebens, welche zunehmend komplex und durch Unsicherheit geprägt sind, verwundert es nicht, dass persönliche und unternehmerische Selbstständigkeit und die sie bestimmenden personengebundenen Eigenschaften – unabhängig von Bildungsweg, Schulform und Abschluss – in den Fokus bildungspolitischer Zielsetzungen und pädagogischen Bemühungen rücken (siehe Europäische Kommission, 2006, S. 17).

Pädagogische Maßnahmen, die der Förderung von Selbstständigkeit und unternehmerischen Kompetenzen dienen, werden als Entrepreneurship-Education bezeichnet (vgl. Jung, 2012, S. 18). Die Adressaten sind dabei keineswegs auf künftige Unternehmer zu beschränken, sondern Individuen aus allen Bereichen der Lebens- und Arbeitswelt (vgl. Aff, 2008, S. 306).

Aus didaktischer Sicht geht mit dem Ziel der unternehmerischen Kompetenzentwicklung die Herausforderung einer legitimierten Umsetzung in Form von Lehrveranstaltungen einher. Bislang basiert die Lehre jedoch weitgehend auf theoretischen Unterrichtsmodellen. Dabei ist unklar, wie sich der Lernweg für die Lernenden real gestaltet und welche gegenstandsbezogenen Verständnisschwierigkeiten sie aufweisen (vgl. Egbert, 2014, S. 171).

Vor diesem Hintergrund besteht das Ziel des Beitrags darin, empirisch eruierte Ergebnisse eines Forschungsprojektes darüber vorzustellen, wie der Lerngegenstand der Unternehmensgründung von Lernenden der Sekundarstufe I wahrgenommen wird, um die darauf aufbauende, didaktische Rekonstruktion von Ökonomieunterricht zu ermöglichen. Die folgenden Fragen leiteten die Untersuchung: (1) Welche qualitativ unterscheidbaren Verständnisweisen lassen sich zum Phänomen der Planungsaufgaben von Unternehmensgründern feststellen? (2) Welche Aspekte des Phänomens werden von den Lernenden berücksichtigt und welche nicht? (3) Welche gegenstandsbezogenen Verständnisschwierigkeiten lassen sich feststellen?

Zur Beantwortung wurde der Forschungsansatz der Phänomenographie verwendet, bei der die „Art der Wahrnehmung von Phänomenen der uns umgebenden Welt durch das Individuum“ (Birke & Seeber, 2012, S. 223f) im Fokus, also im Speziellen die Wahrnehmung durch Schülerinnen und Schüler. Phänomene werden in diesem Kontext als abgrenzbare Einheiten des Erlebens verstanden. Es wird angenommen, dass sich Schülerinnen und Schüler relevante Phänomene in einer bestimmten Art erklären, die sich von Expertenerklärungen unterscheiden (können). Während Lehrkräfte zur Gestaltung des Unterrichts im Idealfall auf ein Fachkonzept zurückgreifen, erklären sich Schülerinnen und Schüler Phänomene auf ihre eigene Weise und unter Verwendung ihrer bis dato existenten kognitiven wie affektiven Fähigkeiten und gesammelten Erfahrungen. Ausgangspunkt für die Anwendung der Phänomenographie im Rahmen didaktischer Forschungen ist die Annahme, dass die Beschaffenheit des Verständnisses von Aufgaben und Lerninhalten für die Lernenden zentral für ihren weiteren Lernerfolg ist (vgl. Murmann, 2008, S. 187).

Um zu erforschen, wie Schülerinnen und Schüler ihre subjektive Realität konzeptualisieren, wurden 44 Interviews durchgeführt, um das vorhandene Antwortspektrum zu erheben und den Ergebnisraum abzubilden. Die Interviews basierten auf zwei Fragensets aus je zwei problemzentrierten Fragen, wobei das Antwortformat offen war.

Um einen vollständigen Ergebnisraum für die Sekundarstufe I abzubilden, wurde ein variations-maximierendes Sample (Klassenstufen 8, 9 & 10, Gender, Leistung, Motivation, Integration) gezogen (siehe Morse, 1994, S. 229). Die Analyse und Auswertung der Interviews basierte auf den angefertigten Transkripten der Untersuchung und erfolgte weitgehend qualitativ. Dabei wurden die Aussagen der Schülerinnen und Schüler nach der Methode der qualitativen Inhaltsanalyse mittels Zusammenfassung (nach Mayring, 2010) zunächst reduziert und anschließend durch eine mehrschrittige induktive Kategorienbildung in die Hauptkategorien des Ergebnisraums überführt und die erste Forschungsfrage beantwortet (siehe Deppermann, 1999, S. 19 und Mayring, 2010, S. 84). Die Ergebnisse offenbaren, dass Fachvorstellungen und Schülerkonzepte dabei – unabhängig von Klassenstufe und Schulform – deutlich differieren.

Um zu eruieren, welche Aspekte des Phänomens von den Lernenden angemessen berücksichtigt wurden und welche nicht (Fragestellung 2), erforderte hingegen eine deduktive Kategorienanwendung, die sich der induktive Kategorienbildung anschloss. Dazu war die theoriebasierte Erarbeitung des Fachkonzeptes unabdingbar, die bereits vor der Interviewphase vollzogen wurde und die entsprechenden Sub- und Hauptkategorien des Fachkonzeptes lieferte.

Abschließend wurden die phänomenbezogenen Verständnisschwierigkeiten der 44 Interviewten auf Grundlage der Originaltranskripte ermittelt (Fragestellung 3).

Donnerstag
12. März

9:00 Uhr

GB 03/49

Chair(s)
Schüle, Christoph

Wirthwein, Linda

Determinanten und Auswirkungen von Zielorientierungen aus unterschiedlichen methodischen Blickwinkeln

Seit mehr als drei Jahrzehnten stehen Zielorientierungen als zentrale motivationale Determinanten des Lern- und Leistungsverhaltens im Fokus der Forschung. Dabei erweisen sich jeweils unterschiedliche Zielorientierungen als prädiktiv für verschiedene Leistungs- und Motivationsvariablen. Vor allem Lernziele (Ziel, eigene Fähigkeiten zu erweitern), aber auch Annäherungs-Leistungsziele (Ziel, besser als andere zu sein) weisen dabei in der Regel die höchsten (positiven) Zusammenhänge zu unterschiedlichen motivationalen Variablen und auch Leistungsindikatoren auf. Ziel des Symposiums ist es, die Relevanz von unterschiedlichen Zielorientierungen im Lern- und Leistungskontext aufzuzeigen. Insgesamt werden in diesem Symposium sowohl Determinanten als auch Auswirkungen von Zielorientierungen untersucht. Darüber hinaus sollen auch etwaige Schwierigkeiten der aktuellen Zielorientierungsforschung (z.B. Erfassung von Zielorientierungen) thematisiert werden. Das Symposium umfasst vier Beiträge, die jeweils einen unterschiedlichen methodischen Zugang wählen: Dabei wird sowohl auf quer- und längsschnittliche Daten, als auch auf ein quasi-experimentelles Design sowie meta-analytische Techniken zurückgegriffen.

Schöne und Stiensmeier-Pelster untersuchen die Selbstwertkontingenz als Determinante von Zielorientierungen. Dazu wurde ein quasi-experimentelles Design mit zwei Messzeitpunkten gewählt. Die Ergebnisse stützen die Annahmen der Autoren, dass sich das Ausmaß der Selbstwertkontingenz auf die Art der motivationalen Zielorientierung auswirkt: Personen mit kompetenz-kontingentem Selbstwert präferieren eher Leistungszielorientierungen.

Steinmayr befasst sich in ihrer längsschnittlich angelegten Studie mit den (reziproken) Zusammenhängen zwischen Need for Cognition, Zielorientierungen und schulischer Leistung. Des Weiteren geht es um die Frage, ob Zielorientierungen die Zusammenhänge zwischen Need for Cognition und Schulleistung vermitteln. Die Autorin findet Hinweise darauf, dass sich Need for Cognition auf die Zielorientierungen (mit Ausnahme der Annäherungs-Leistungsziele) auswirkt. Darüber hinaus kann gezeigt werden, dass der Einfluss von Need for Cognition auf die Schulleistung teilweise über Lernzielorientierungen mediiert wird.

Schwinger, Pütz, Lemmer und Wirthwein betrachten in ihrer Meta-Analyse die Zielorientierungen als Determinante einer maladaptiven Selbstregulationsstrategie, dem sogenannten Self-Handicapping. Die höchsten Zusammenhänge können hier für Vermeidungs-Leistungsziele sowie Lernziele gefunden werden. Des Weiteren ermitteln die Autoren wichtige Moderatoren dieses Zusammenhangs: Vor allem die Operationalisierung der Konstrukte scheint für die Höhe der Korrelation relevant zu sein.

Die Operationalisierung von Zielorientierungen wird auch im letzten Beitrag des Symposiums behandelt. Wirthwein und Steinmayr finden in ihrer Studie Belege für die differentielle Validität von mehrdimensionalen Annäherungs- und Vermeidungs-Leistungszielen und schlussfolgern, dass in zukünftigen Studien möglichst mehrere Komponenten zur Erfassung von Leistungszielen vorgegeben werden sollten.

Susanne Buch stellt im Anschluss an die Vorträge eine Gesamtdiskussion der Beiträge vor.

Donnerstag
12. März

9:00 Uhr

HGB 20

Chair(s)

Wirthwein, Linda
Schwinger, Malte**Determinanten und Auswirkungen von Zielorientierungen aus unterschiedlichen methodischen Blickwinkeln**

Schöne, Claudia; Stiensmeier-Pelster, Joachim

Selbstwert und Motivation - Einfluss der Selbstwertkontingenz auf die situative Zielorientierung von Lernenden bei der Bearbeitung von Aufgaben

Theoretischer Hintergrund: Die Bedeutung des Selbstwerts (Höhe des Selbstwerts), für die Motivation von Lernenden (Höhe der Motivation) ist in vielen Studien gezeigt worden – die Effekte sind üblicherweise relativ stabil, jedoch vergleichsweise gering. In den letzten Jahren rücken zunehmend andere Facetten des Selbstwerts als dessen Höhe stärker in den Fokus der Forschung. Eine dieser Facetten – die Abhängigkeit des Selbstwerts (Kontingenz) – wird in der vorliegenden Studie als Einflussfaktor für die Qualität der Motivation, konkreter gesagt für Art der situativen Lern- und Leistungsmotivation, in den Blick genommen. Lernende unterscheiden sich darin, ob, wie sehr, und welchen Bereichen sie ihren Wert vom Erreichen von Standards abhängig machen (Selbstwertkontingenz). Ein insbesondere für Lernende relevanter Bereich (neben anderen Bereichen wie z. B. Attraktivität oder Anerkennung Anderer) ist der Bereich kognitiver Kompetenz. Machen Personen ihren Selbstwert von Erfolgen und Misserfolgen in diesem Bereich abhängig, spricht man von Kompetenzkontingenz des Selbstwerts.

Fragestellung: In der Studie wird die Bedeutung der Selbstwertkontingenz im Bereich Kompetenz für die Art der Motivation (Zielorientierung) bei der Aufgabebearbeitung betrachtet. Personen mit abhängigem Selbstwert sollten solche Situationen stärker als Chance zur Steigerung ihres Selbstwertgefühls bzw. als potenzielle Selbstwertbedrohung wahrnehmen. Als Reaktion darauf verfolgen sie bei der Aufgabebearbeitung in höherem Maße so genannte Selbstvalidierungsziele. Dies sollte in Lern-Leistungssituationen zu einem erhöhten Bestreben, Kompetenz zu demonstrieren bzw. mangelnde Kompetenzen zu verbergen führen und dadurch zugleich Aufmerksamkeit von der Aufgabe abziehen. Es wird die Annahme überprüft, dass ein kompetenz-kontingenter Selbstwert die Leistungszielorientierungen fördert und zugleich die Lernzielorientierung möglicherweise behindert.

Methode: In der quasi-experimentellen Online-Studie wurde zur Überprüfung der Annahmen bei N = 380 Studierenden zu einem 1. MZP zunächst die Selbstwertkontingenz (UV) und die dispositionelle Zielorientierung (trait) per Fragebogen erhoben (SELLMO, Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2012; bzw. Skala „Kompetenzkontingenz“ der SESKON; Schöne, Herrmann & Stiensmeier-Pelster, in Vorb). Zum 2. MZP bearbeiteten die Versuchspersonen zunächst Aufgaben aus verschiedenen Bereichen (u. a. Matrizen, Wortschatz, Wissen, Anagramme). Während der Bearbeitung wurde dabei entweder leicht positives („weiter so“) oder eher negatives („konzentriere dich“) Feedback gegeben (Moderator). Anschließend gaben die Probanden unter eine Coverstory an, unter welchen „emotionalen und motivationalen Bedingungen“ sie den folgenden Aufgabenblock bearbeiten möchten. Hierfür wurden adaptierte Zielorientierungssitems (SELLMO) genutzt (situative Zielorientierung: Wunsch; AV1). Anschließend wählten sie nochmals zwischen drei Bearbeitungsmodi aus, die jeweils einer Zielorientierung entsprechen (situative Zielorientierung: Entscheidung; AV2). Die situative Zielorientierung wurde somit auf zwei unterschiedliche Weisen operationalisiert.

Ergebnisse: Zur Prüfung wurden lineare (AV1) bzw. logistische (AV2) Regressionsanalysen durchgeführt. Die Befunde zeigen, dass die zum ersten Messzeitpunkt erhobene Selbstwertkontingenz - über die zum selben Messzeitpunkt erhobene, dispositionelle Zielorientierung hinaus - zum Wunsch nach (AV1) und Entscheidung für (AV2) leistungszielorientierter Bearbeitungsmodi führt. Die Analysen ergeben darüber hinaus einen negativen, jedoch nur für die Entscheidung (AV2) auch statistisch signifikanten Effekt der Selbstwertkontingenz auf die situative Lernzielorientierung. Die Wahrscheinlichkeit für die Wahl einer vermeidungsleistungszielorientierten Aufgabebearbeitung (im Vergleich zur lernzielorientierten Bearbeitung) steigt zudem nach negativem Feedback mit ansteigender Kompetenz-Kontingenz um den Faktor 3.6 mit jeder Einheit der Kontingenz.

Interpretation: Zusammenfassend betrachtet unterstützen die Befunde die Annahmen. Der Selbstwert beeinflusst die aktuelle Motivation der Lernenden insofern, als sich das Ausmaß der Selbstwertkontingenz auf die Art der motivationalen Zielorientierung bei der Aufgabebearbeitung auswirkt. Im konkreten Fall wünschten die Probanden nicht nur in stärkerem Ausmaß einen typisch leistungszielorientierten Modus für die Aufgabebearbeitung, sondern Sie entschieden sich dann auch mit höherer Wahrscheinlichkeit für diesen Modus. Die Befunde werden bzgl. der Bedeutung der Selbstwertkontingenz als Determinante der ZO sowie in ihrer Bedeutung für schulisches Lernen und für Bildungsprozesse diskutiert.

Chair(s)

Wirthwein, Linda
Schwinger, Malte

Determinanten und Auswirkungen von Zielorientierungen aus unterschiedlichen methodischen Blickwinkeln

Steinmayr, Ricarda

Zur Rolle von Need for Cognition für Zielorientierungen in der Schule

Unter „Need for Cognition“ verstehen Cacioppo und Petty (1982) die Tendenz eines Individuums, sich gerne mit kognitiv anregenden Herausforderungen auseinander zu setzen. Ursprünglich aus der Persönlichkeitspsychologie stammend, wird dieses Konstrukt zunehmend auch im schulischen Kontext untersucht. So weist Need for cognition substantielle Zusammenhänge zu unterschiedlichen motivationalen Variablen, beispielsweise Zielorientierungen, auf (z.B. Day et al., 2007; Meier, Vogl & Preckel, 2014). Bisherige Studien verweisen außerdem auf die Relevanz von Need for cognition als Prädiktor für den Wissenserwerb und damit für die Schulleistung (z.B. Bertrams & Dickhäuser, 2009). Bislang existieren jedoch noch keine empirischen Studien, die sich dem Zusammenspiel von Need for cognition, Zielorientierungen und Schulleistung längsschnittlich widmen, so dass die Wirkrichtung des Zusammenhangs zwischen diesen Variablen empirisch bislang nicht überprüft wurde. Es ist ein Ziel des vorliegenden Beitrags, dieser Fragestellung nachzugehen. Darüber hinaus ist wenig darüber bekannt, über welche Prozesse der Zusammenhang zwischen Need for Cognition und Schulleistung vermittelt wird. Ausgehend von theoretischen Überlegungen von McCrae und Costa (1999) lässt sich vermuten, dass die Zusammenhänge zwischen Need for cognition und Schulleistung über Zielorientierungen vermittelt werden. Die positiven Korrelationen zwischen Need for Cognition und Schulleistung geben erst Hinweise, dass diese Annahme tatsächlich korrekt ist. Hierbei zeigten sich stärkere Zusammenhänge zwischen Lernzielen (Ziel, seine Kompetenzen zu erweitern) und Need for Cognition als zwischen Leistungszielen (Ziel, seine Kompetenzen zu demonstrieren) und Need for Cognition (Meier et al., 2014), so dass eher bei Lernzielen als bei Leistungszielen ein Mediationseffekt erwartet wird. Diese Mediationseffekte zu überprüfen ist ein weiteres Ziel des vorliegenden Beitrages. Hierbei wird in Anlehnung an Cole und Maxwell (2003) der postulierte Mediationseffekt von Zielorientierungen aufgrund der längsschnittlichen Datenbasis überprüft.

Dazu wurden N = 483 Schüler/innen der 11. Klasse von 5 Gymnasien (Alter: M = 16,43; SD = .55) im Abstand von sechs Monaten zweimal mit Fragebögen zur Erfassung von Need for Cognition (Bless et al., 1994), Zielorientierungen (Spinath et al., 2002) und Schulleistung (Zeugnisnoten) befragt. Zur Überprüfung der oben formulierten Fragestellungen wurden cross-lagged Modelle berechnet, wobei zunächst die reziproken Beziehungen zwischen Need for Cognition und Zielorientierungen sowie zwischen Need for Cognition und Schulleistung untersucht wurden. Im Anschluss wurden in Anlehnung an Cole und Maxwell (2003) Mediationsanalysen zur Rolle von Zielorientierungen hinsichtlich des Zusammenhangs von Need for Cognition und Schulnoten durchgeführt.

Die Analysen mittels SEM cross-lagged Modellen zeigten, dass sich Need for Cognition positiv auf die Veränderung von Lernzielen auswirkt, während der reziproke Effekt nicht gegeben ist. In Bezug auf Need for Cognition und Annäherungsleistungsziele konnten keine wechselseitigen Zusammenhänge nachgewiesen werden. Need for Cognition ging jedoch mit einer Abnahme der Vermeidungsleistungsziele und der Arbeitsvermeidung einher. Reziproke Zusammenhänge konnten hier ebenfalls nicht gezeigt werden. Need for Cognition wirkte sich positiv auf die Veränderung in der Schulleistung aus, während das gleiche reziprok der Fall war. Die Mediationsanalysen waren in Einklang mit den Erwartungen. Es konnte gezeigt werden, dass der Einfluss von Need for Cognition auf die Schulleistung teilweise über Lernzielorientierungen mediiert wurde. Die Ergebnisse werden vor dem Hintergrund ihrer Bedeutung von Need for Cognition in akademischen Leistungskontexten diskutiert.

Donnerstag
12. März

9:00 Uhr

HGB 20

Chair(s)

Wirthwein, Linda
Schwinger, Malte**Determinanten und Auswirkungen von Zielorientierungen aus unterschiedlichen methodischen Blickwinkeln**

Schwinger, Malte; Pütz, Nadine; Lemmer, Gunnar; Wirthwein, Linda

Zielorientierungen als Determinanten von Self-Handicapping: Eine Metaanalyse

Theoretischer Hintergrund: In der aktuellen Bildungsforschung befassen sich zahlreiche Studien mit der Frage, welche Ziele Schüler und Studierende beim Lernen verfolgen und welche Arten von Zielen als besonders adaptiv in Bezug auf pädagogisch bedeutsame Outcome-Variablen wie z.B. Leistung, adaptive Selbstregulation während des Lernens oder Wohlbefinden anzusehen sind (zsf. Senko et al., 2011). Besondere Aufmerksamkeit hat diesbzgl. die Forschung zu den sogenannten Zielorientierungen erfahren, bei denen im Wesentlichen drei Arten unterschieden werden. Lernzielorientierte Personen sind bestrebt, ihre Kompetenzen zu erweitern sowie ihre individuellen Fähigkeiten zu steigern. Annäherungs-Leistungszielorientierte wollen besser sein als andere und ihre Fähigkeiten demonstrieren, wohingegen Vermeidungs-Leistungszielorientierte darum bemüht sind, nicht schlechter zu sein als andere und ihre mangelnden Fähigkeiten zu verbergen (Elliot & Murayama, 2008). Bezüglich der Zusammenhänge mit pädagogisch bedeutsamen Ergebnisvariablen zeigten sich für die Lernzielorientierung in zahlreichen Studien positive Effekte, während die Vermeidungs-Leistungszielorientierung eindeutig negative Folgen nach sich zog. Die Befundlage für Annäherungs-Leistungsziele kann als gemischt bezeichnet werden. In der vorliegenden Studie soll der Zusammenhang dieser drei Zielorientierungen mit einer spezifischen Selbstregulationsstrategie, dem sogenannten Self-Handicapping, metaanalytisch untersucht werden. Self-Handicapping bezeichnet eine Strategie zum Schutz des eigenen Selbstwerts, bei der eine Person sich vor einer Bewertungssituation (z.B. Prüfung, Vorstellungsgespräch) ein Handicap verschafft, welches ihr bei einem eventuellen Misserfolg als Ausrede dient (Schwinger & Stiensmeier-Pelster, 2012). Ein typisches Beispiel stellen Schüler oder Studierende dar, die vor einer wichtigen Prüfung zu spät mit dem Lernen beginnen. Im Fall einer schlechten Prüfungsnote begründen sie diese mit ihrer zu späten Vorbereitung. In einer Reihe von Studien zeigten sich bedeutsame Zusammenhänge zwischen den verschiedenen Zielorientierungen und dem Ausmaß an Self-Handicapping. Insgesamt ist die diesbzgl. Befundlage jedoch recht heterogen und reicht von nicht signifikanten bis zu hoch negativen bzw. positiven Korrelationen.

Fragestellung: Da habituelles Self-Handicapping nachweislich maladaptive Konsequenzen für die Leistung und das Wohlbefinden der jeweiligen Personen nach sich zieht, sollen anhand der hier vorgestellten Metaanalyse die wichtigsten Risiko- und Resilienzfaktoren für die Nutzung von Self-Handicapping als Regulationsstrategie identifiziert werden. Neben den oben beschriebenen Zielorientierungen werden noch der allgemeine Selbstwert und das fachspezifische Fähigkeitsselbstkonzept als wichtige Determinanten von Self-Handicapping angenommen (vgl. Schwinger & Stiensmeier-Pelster, 2012). Aufgrund der in der Literatur berichteten Zusammenhänge werden positive Korrelationen mit der Annäherungs- und Vermeidungs-Leistungszielorientierung sowie negative Korrelationen mit der Lernzielorientierung, dem allgemeinen Selbstwert und dem Fähigkeitsselbstkonzept erwartet. Durch die hier vorgestellte Metaanalyse sollen a) erstmalig der mittlere Zusammenhang zwischen den verschiedenen Prädiktoren und Self-Handicapping spezifiziert und b) potentielle Moderatoren dieses Zusammenhangs ermittelt werden.

Methode: In die Metaanalyse wurden insgesamt 118 Studien mit $N = 73,330$ Schülern und Studierenden aufgenommen, in denen Korrelationen zwischen Fragebogenmaßen für Self-Handicapping und Fragebogenmaßen für die drei Zielorientierungen, den allgemeinen Selbstwert und/ oder das Fähigkeitsselbstkonzept berichtet wurden. Als Moderatoren wurden die Operationalisierung der jeweiligen Konstrukte, Art der Stichprobe (Schüler vs. Studierende), Herkunft der Stichprobe, Schultyp, Geschlecht, Alter und Ethnizität in den Analysen berücksichtigt.

Ergebnisse: Mit Hilfe des R-Pakets metafor (Viechtbaucher, 2010) wurden univariate Fixed- und Random-Effects-Modelle berechnet. Die höchsten mittleren Zusammenhänge mit Self-Handicapping zeigten sich für den allgemeinen Selbstwert ($r = -.38$), das Fähigkeitsselbstkonzept ($r = -.37$) und die Vermeidungs-Leistungsziele ($r = .33$, alle $p < .01$). Lernziele ($r = -.18$, $p < .01$) und Annäherungs-Leistungsziele ($r = .05$, $p < .05$) wiesen im Mittel deutlich niedrigere Korrelationen mit Self-Handicapping auf. Mit Ausnahme des Fähigkeitsselbstkonzepts erwies sich in Bezug auf alle untersuchten Determinanten die Operationalisierung des Self-Handicapping-Konstrukts (d.h. der verwendete Fragebogen) als signifikanter Moderator der jeweiligen Zusammenhänge. Insgesamt zeigen die Befunde, dass bei Schlussfolgerungen über Risiko- und Resilienzfaktoren von Self-Handicapping zwingend das eingesetzte Messinstrument beachtet werden sollte.

Donnerstag
12. März

9:00 Uhr

HGB 20

Chair(s)

Wirthwein, Linda
Schwinger, Malte

Determinanten und Auswirkungen von Zielorientierungen aus unterschiedlichen methodischen Blickwinkeln

Wirthwein, Linda; Steinmayr, Ricarda

Zur Relevanz differentieller Leistungsziele

Zielorientierungen gelten als relevante Prädiktoren schulischen Lern- und Leistungsverhaltens (vgl. Elliot, 2005). Vor allem die Zusammenhänge zwischen Annäherungs-Leistungszielen, Vermeidungs-Leistungszielen und Indikatoren schulischer Leistungen fallen dabei laut mehrerer Meta-Analysen (z.B. Huang, 2012; Hulleman, Schragger, Bodmann & Harackiewicz, 2010; Wirthwein, Sparfeldt, Pinquart, Wegerer & Steinmayr, 2013) heterogen aus. Als relevanter Moderator dieses Beziehungsmusters hat sich u.a. die Operationalisierung der Leistungsziele herausgestellt. In diesem Zusammenhang wird z.B. von Senko, Hulleman und Harackiewicz (2011) geschlossen, dass vor allem normativorientierte Annäherungs-Leistungsziele (Ziel, besser als andere sein zu wollen) höhere positive Korrelationen zur Leistung aufweisen als kompetenzorientierte Annäherungs-Leistungsziele (Ziel, hohe Kompetenzen zu zeigen). Vermeidungs-Leistungsziele weisen negative Zusammenhänge zur Schulleistung auf – eine entsprechende Unterscheidung verschiedener Vermeidungs-Leistungsziele wurde in der einschlägigen Literatur bislang jedoch noch nicht getroffen. Hulleman et al. (2010) konnten darüber hinaus aufzeigen, dass die Operationalisierung der Leistungsziele auch ein relevanter Moderator für die Zusammenhänge zu anderen motivationalen Variablen (z.B. Interesse) darstellt.

Um die Relevanz differentieller Leistungsziele zu untersuchen, wurden einer schulischen Stichprobe (N = 310; Alter: M = 14.91, SD = 1.72) Items mehrerer Instrumente zur Erfassung von Zielorientierungen vorgegeben. Neben den Zielorientierungen wurden zusätzlich Schulnoten, weitere motivationale Variablen (schulisches Selbstkonzept, schulische Werte, Leistungsängstlichkeit) sowie das schulische Wohlbefinden erfasst. Sämtliche Konstrukte wurden für die Schulfächer Mathematik und Deutsch erfragt. Konfirmatorische Faktorenanalysen wiesen auf die Mehrdimensionalität von Annäherungs- und Vermeidungs-Leistungszielen hin. Neben einer kompetenzorientierten und einer normativorientierten Annäherungs-Leistungszielkomponente resultierte eine weitere evaluative Annäherungs-Leistungsziel-Komponente, die das Ziel fokussiert, anderen gegenüber hohe Kompetenzen zu demonstrieren. Der Modellfit für die jeweiligen Modelle war durchweg zufriedenstellend. Bei den Vermeidungs-Leistungszielen ergaben sich ebenfalls drei Faktoren, obwohl der Modellfit verglichen mit den differentiellen Annäherungs-Leistungszielen etwas schlechter ausfiel: Auch hier resultierten neben kompetenzorientierten Vermeidungs-Leistungszielen (Verbergen von Inkompetenz), normativorientierte (Vermeiden eines schlechteren Abschneidens verglichen mit anderen) sowie evaluative Vermeidungs-Leistungsziele (Vermeidung, falsche Antworten zu geben).

Es ergaben sich differentielle Zusammenhänge zwischen den unterschiedlichen Annäherungs-Leistungszielen und der Schulleistung – entgegen der Vermutung von Senko et al. (2011) zeigten sich jedoch auch substantielle Zusammenhänge mit den kompetenzorientierten Annäherungs-Leistungszielen. Die Zusammenhänge zwischen den differentiellen Vermeidungs-Leistungszielen und den Schulnoten waren vernachlässigbar. Auch hinsichtlich der weiteren motivationalen Variablen zeigten sich in Abhängigkeit von der jeweiligen Leistungszielkomponente überwiegend differentielle Zusammenhänge. Die höchsten positiven Zusammenhänge resultierten zwischen den kompetenzorientierten Annäherungs-Leistungszielen und schulischem Selbstkonzept, schulischen Werten und schulischem Wohlbefinden. Bei den Vermeidungs-Leistungszielen ergaben sich beispielsweise nur bezüglich der kompetenz- und der evaluativorientierten Ziele positive Zusammenhänge zur Leistungsängstlichkeit; die Zusammenhänge zu den übrigen motivationalen Variablen fielen – mit Ausnahme der normorientierten Vermeidungs-Leistungsziele – überwiegend gering aus. Insgesamt zeigen sich vor allem für die mehrdimensionalen Annäherungs-Leistungsziele, aber teilweise auch für die mehrdimensionalen Vermeidungs-Leistungsziele differentielle Beziehungsmuster zu Schulnoten sowie zu ausgewählten motivationalen Variablen. Zukünftige Studien sollten daher möglichst mehrere Komponenten zur Erfassung von Leistungszielen vorgeben.

Donnerstag
12. März

9:00 Uhr

HGB 20

Chair(s)

Wirthwein, Linda
Schwinger, Malte

Determinanten und Auswirkungen von Zielorientierungen aus unterschiedlichen methodischen Blickwinkeln

Oberle, Monika

Politisches Selbstkonzept von Jugendlichen

Es wird davon ausgegangen, dass das Selbstkonzept der eigenen Fähigkeiten ein wichtiger Faktor für erfolgreiche Lernprozesse ist. Das Fähigkeitsselbstkonzept gilt als multidimensionales Konstrukt (Shavelson et al., 1976), das sich u.a. in akademische und nicht-akademische Facetten aufgliedert. Für das akademische Selbstkonzept konnte nachgewiesen werden, dass dieses aus weitgehend unabhängigen verbalen und mathematischen Facetten besteht (Möller et al., 2009). Es gibt Forschung zu Determinanten des akademischen Selbstkonzepts wie bspw. der Leistungsstärke des Umfelds (Marsh & Craven, 2006), wie auch zu dessen Auswirkungen auf Leistung, Motivation und der Wahl von „Leistungskursen“ und Studienfächern (Köller et al., 2000). Trotz dieser umfangreichen Forschungsaktivitäten fand das Konstrukt eines politischen Selbstkonzepts bislang nur wenig Beachtung in der Selbstkonzeptforschung. Dennoch ist anzunehmen, dass sich ein solches im Jugendalter herausbildet, da das Jugendalter als Zeit der politischen Sozialisation und einer zunehmenden politischen Involviertheit gilt. Tatsächlich ist das politische Selbstkonzept, wenn auch zumeist mit anderen Fragestellungen als denen der Bildungsforschung, Gegenstand der politischen Sozialisationsforschung und politischen Kulturforschung (bereits Almond & Verba, 1963). Fokussiert wurden bspw. das subjektive Wissen über Politik sowie das interne politische Effektivitätsgefühl (internal political efficacy, vgl. Vetter 1997), wobei die Selbsteinschätzung der eigenen Kenntnisse manchmal auch als Teilfacette der politischen Selbstwirksamkeitsüberzeugungen (efficacy) modelliert wird.

Bislang nicht befriedigend geklärt ist zunächst die Frage, inwiefern das politische Fähigkeitsselbstkonzept eine eigenständige, vom generellen akademischen Selbstkonzept sowie anderen nicht-akademischen Selbstkonzepten zu unterscheidende Disposition darstellt. Zum anderen stellt sich die Frage nach schulischen und außerschulischen Einflussfaktoren des politischen Selbstkonzepts. Eine bislang unzureichend adressierte Frage betrifft schließlich die Kausalstruktur angenommener Zusammenhänge. Diese Forschungsdefizite geben Anlass für ein Symposium zum Politischen Selbstkonzept von Jugendlichen.

Das Symposium vereint mit der pädagogischen Psychologie, der politischen Soziologie, der Politikdidaktik und der Erziehungswissenschaft (Diskutant) unterschiedliche fachliche Perspektiven auf das politische Selbstkonzept. Die vier Beiträge

- (1) Arens & Watermann: Validierung des Konstrukts des politischen Selbstkonzepts
- (2) van Deth & Quintelier: Henne oder Ei? Eine Exploration kausaler Zusammenhänge zwischen politischem Selbstbewusstsein und Partizipation
- (3) Landwehr & Weißeno: Selbstkonzept, Interesse und Wissen im Politikunterricht
- (4) Forstmann & Oberle: Politisches Selbstkonzept – Facetten und geschlechtsspezifische Einflussfaktoren

bieten zum einen spannende Überschneidungen hinsichtlich der zentralen Fragestellungen nach (a) Validierung des Konstrukts politisches Selbstkonzept sowie (b) dessen Einflussfaktoren, mit unterschiedlichen methodischen Zugängen (u.a. Strukturgleichungen in MPlus, cross-lagged-panel-designs) und Datensätzen. Zum anderen werfen sie jeweils spezifische Schlaglichter auf relevante Teilfragen, wie die nach schulischen und außerschulischen Sozialisationsfaktoren (u.a. Leistungsstärke der Schulklasse, Medienkonsum und familiärem Austausch), nach (geschlechtsspezifischen) Einflussfaktoren wie Geschlechterrollenvorstellungen sowie nach dem kausalen Zusammenhang von Selbstkonzept und politischem Verhalten.

Donnerstag
12. März

9:00 Uhr

HGB 30

Chair(s)

Oberle, Monika

Arens, A. Katrin; Watermann, Rainer

Validierung des Konstrukts des politischen Selbstkonzepts

In der Selbstkonzeptforschung geht man von einem multidimensionalen Selbstkonzept aus, das aus verschiedenen Inhaltsfacetten besteht. Dabei wird grob zwischen akademischen und nicht-akademischen Selbstkonzepten unterschieden, die in sich wiederum in spezifischere Selbstkonzepte (z.B. verbales Selbstkonzept, mathematisches Selbstkonzept) ausdifferenzieren. Das Konstrukt eines politischen Selbstkonzepts fand bislang nur wenig Beachtung in der Selbstkonzeptforschung. Dennoch ist anzunehmen, dass sich ein solches im Jugendalter herausbildet, da das Jugendalter als Zeit der politischen Sozialisation und eines zunehmenden politischen Interesses gilt. In dieser Studie sollten daher verschiedene Ansätze zur Validierung des politischen Selbstkonzepts vorgestellt werden. Ebenso wurde die Entwicklung des politischen Selbstkonzepts über das Jugendalter hinweg untersucht, wobei sich die normative Stabilität und die mittlere Veränderung betrachtet wurden. Die Daten stammten aus der Studie zu Bildungsverläufen und psychosozialer Entwicklung im Jugendalter (BIJU), in denen Schüler (N = 2504) in den Klassenstufen 7 und 10 zu ihrem politischen Selbstkonzept befragt wurden. Dazu dienten vier Items, bei denen die Schüler nach ihrer Selbstwahrnehmung zu ihrem politischen Wissen und ihrem politischen Verständnis befragt wurden und diese auf einer vierstufigen Likert-Skala einschätzen sollten. Als Validierungsvariable sollten die Schüler zusätzlich Auskunft über ihr politisches Informationsverhalten geben, wobei es um die Häufigkeit ging, zu der sie sich durch Medien oder Gespräche mit der Familie oder Freunden über das politische Geschehen informierten. Ebenso wurde das politische Wissen der Schüler anhand eines standardisierten Leistungstest erfasst. Zudem wurden das verbale Selbstkonzept, das mathematische Selbstkonzept und das Selbstwertgefühl der Schüler erhoben: Beim verbalen und mathematischen Selbstkonzept ging es um die Selbsteinschätzung eigener mathematischer und verbaler Fähigkeiten (jeweils fünf Items); beim Selbstwertgefühl wurde anhand von vier Items nach der globalen Selbstakzeptanz gefragt, wobei das Selbstwertgefühl vorherigen Befunden zufolge einen höheren Zusammenhang zu nicht-akademischen als zu mit akademischen Selbstkonzeptfacetten aufweist. In allen mit Mplus Version 7.1 durchgeführten Analysen wurde für die Klassenzugehörigkeit der Schüler kontrolliert. Da die Fragestellungen teilweise auch längsschnittliche Analysen beinhalteten wurde zudem vorab die Invarianz der Konstrukte über die Zeit geprüft und bestätigt. Konfirmatorische Faktorenanalysen unterstützten die Existenz eines politischen Selbstkonzepts als separate Selbstkonzeptfacette, das zu anderen Selbstkonzepten abgrenzbar ist. Die geringen Korrelationen zum verbalen Selbstkonzept ($r = .115$), mathematischen Selbstkonzept ($r = .146$) und Selbstwertgefühl ($r = .155$) verdeutlichten, dass das politische Selbstkonzept weder dem akademischen noch dem nicht-akademischen Bereich des Selbstkonzepts zugeordnet werden kann. Die Korrelation zwischen den Faktoren des politischen Selbstkonzepts zu beiden Messzeitpunkten ($r = .490$) verweist auf eine moderate normative Stabilität ($r = .490$) über die Klassenstufen 7 bis 10. Ein latent change model zeigte, dass das politische Selbstkonzept über diesen Zeitraum im Mittel ansteigt. Es ergaben sich substanzielle querschnittliche Zusammenhänge zum politischen Informationsverhalten und politischen Wissen. Im Längsschnitt wurden diese Zusammenhänge durch cross-lagged effects model analysiert, in dem bei Kontrolle der Stabilität die gegenseitige Beeinflussung der Variablen der Konstrukte abgebildet wird. Dabei zeigten sich reziproke Effekte, so dass das politische Selbstkonzept gleichzeitig als Outcome und Prädiktor für politisches Wissen und politisches Informationsverhalten fungiert. Entsprechend liefert die Studie zahlreiche empirische Befunde für die Existenz eines politischen Selbstkonzepts, die im Hinblick auf ihre Implikationen für Forschung und Praxis diskutiert und in die vorhandene Selbstkonzeptforschung eingebettet werden.

Donnerstag
12. März

9:00 Uhr

HGB 30

Chair(s)

Oberle, Monika

van Deth, Jan W.; Quintelier, Ellen

Henne oder Ei? Eine Exploration kausaler Zusammenhänge zwischen politischem Selbstbewusstsein und Partizipation

Das Konzept ‚politisches Selbstbewusstsein‘ wird auf unterschiedliche Weisen eingesetzt. Viele Forschungsansätze untersuchen Konzeptualisierungen und Dimensionen des Begriffes sowie Zusammenhänge des politischen Selbstbewusstseins mit anderen Persönlichkeitsmerkmalen. Eine explizite Verknüpfung des Konzeptes mit Verhalten wird jedoch entweder überhaupt nicht thematisiert oder aber politisches Selbstbewusstsein einseitig als Prädisposition für zukünftige Handlungen verstanden (attitudes-behaviour model). Weit weniger verbreitet sind Ansätze, welche das politische Selbstbewusstsein nicht als Prädisposition, sondern vielmehr als Konsequenz bestimmter Verhaltensweisen verstehen (z.B. self-perception und cognitive dissonance theories). Obwohl ein rekursiver Zusammenhang am plausibelsten erscheint, fehlt es bis heute an systematischen Vergleichen dieser beiden Ansätze. Die wesentliche Forschungsfrage ist also, ob politisches Selbstbewusstsein zu politischem Verhalten führt oder politisches Verhalten umgekehrt als Determinante politischen Selbstbewusstseins zu betrachten ist.

Zur Beantwortung dieser Frage ist ein Forschungsentwurf erforderlich, welcher (i) longitudinal angelegt ist und intra-individuelle Entwicklungen verfolgt (Paneldaten), (ii) zumindest Messungen zu drei Zeitpunkten umfasst, damit rekursive und zyklische Effekte berücksichtigt werden können, und (iii) sich nicht auf politisches Selbstbewusstsein im engeren Sinne beschränkt, sondern einen Vergleich mit anderen Orientierungen ermöglicht. Der Belgian Political Panel Survey (BPPS) erfüllt diese Bedingungen optimal: Die Umfrage umfasst eine umfangreiche Stichprobe der belgischen Jugend, welche im Zeitraum 2006-2011 drei Mal ausführlich befragt wurde. Mehr als 3000 Jugendliche beteiligten sich an allen drei Wellen (jeweils im Alter von 16, 18 und 21 Jahren). Somit ist eine einzigartige Datengrundlage geschaffen worden, mit welcher die Persistenz von Einstellungen und Verhaltensweisen in der sehr wichtigen Übergangsphase vom Jugendalter bis in die Adoleszenz untersucht werden kann. Zudem ermöglicht diese Datengrundlage, Entwicklungsprozesse und wechselseitige Abhängigkeiten zwischen Einstellungen und Verhalten zu analysieren.

Die möglichen kausalen Zusammenhänge zwischen politischem Selbstbewusstsein und politischem Verhalten schätzen wir mit einem „Drei-Wellen-Zwei-Variablen Pfadmodell“ mit verschiedenen sozio-demographischen Kontrollvariablen. Dabei werden für politisches Verhalten verschiedene Varianten politischer Partizipation untersucht (Protestformen und politischer Konsum) und für politische Einstellungen Messungen von politischem Selbstbewusstsein (internal political efficacy), aber auch von politischem Interesse, politischem Vertrauen und bürgerschaftlichen Normen berücksichtigt. Auf der Basis eines Vergleichs der cross-lagged Effekte in diesen Modellen (z. B.: Haben 16-Jährige mit relativ viel politischem Selbstbewusstsein eine größere Chance, sich als 18- und 21-Jährige zu beteiligen als andere? Oder haben 16-Jährige, die politisch aktiv sind, nach zwei bzw. fünf Jahren ein relativ höheres politisches Selbstbewusstsein?) kann die relative Stärke der angenommenen Wirkungsrichtungen beurteilt werden.

Die Ergebnisse der Studie zeigen, dass politische Einstellungen nicht unbedingt dem Verhalten vorausgehen und demnach also auch nicht ausschließlich als Prädispositionen betrachtet werden können. Im Gegenteil: Es ist viel wahrscheinlicher, dass politische Partizipation politische Einstellungen stärkt, als dass umgekehrt politische Einstellungen im späteren Alter politische Aktivitäten auslösen. Dies bedeutet nicht, dass Einstellungen wie politisches Selbstbewusstsein irrelevant für politisches Verhalten sind. Jedoch kann die umgekehrte Wirkungsrichtung empirisch viel deutlicher nachgewiesen werden. Diese Ergebnisse können als Grundlage für eine (erneute) Diskussion über die möglichen Wirkungen verschiedener didaktischer Konzepte in der politischen Bildung genutzt werden: für späteres politisches Verhalten ist gegenwärtige und tatsächliche Beteiligung offensichtlich wichtiger als die Entwicklung verheißungsvoller Einstellungen.

Donnerstag
12. März

9:00 Uhr

HGB 30

Chair(s)
Oberle, Monika

Landwehr, Barbara; Weißeno, Georg

Selbstkonzept, Interesse und Wissen im Politikunterricht

Theoretischer Hintergrund: Im Modell der Politikkompetenz von Detjen et al. (2012) sind schulische Selbstkonzepte und das Fachinteresse theoretisch verortet. Aus anderen Domänen ist bekannt, dass das allgemeine und das fachspezifische Selbstkonzept sowie das Fachinteresse prädiktive Effekte für die Lernergebnisse haben. Aus theoretischer Perspektive fundieren sie die Erreichbarkeit von Leistungen. In der politikdidaktischen Theorie wird angenommen, dass diese motivationalen Orientierungen für die Befähigung zur Bewältigung von Aufgaben im Unterricht bedeutsam sind (Weißeno 2015). Das theoretische Modell von Eccles & Wigfield (2002) postuliert gleichfalls Effekte des Selbstkonzepts (Erwartungskomponente) und des Interesses (Wertkomponente) auf die Leistung. Leistung ist danach zu beschreiben durch die subjektive Bewertung des Leistungserlebens und die subjektive Erwartung darüber, dass einer intendierten Handlung ein Ereignis folgt oder nicht. Gefragt wird in der Forschung darüber hinaus, ob es Klasseneffekte durch den Vergleich der eigenen Leistung mit der der Mitschüler/-innen gibt. Die Höhe schulischer Selbstkonzepte fällt bei Existenz einer starken Vergleichsgruppe niedriger aus (Marsh 1991; Köller, Schnabel & Baumert 2000). Deshalb wird teilweise auch angenommen, dass die Wirkrichtung vielmehr als wechselseitig anzusehen ist.

Fragestellungen: In diesem Beitrag gehen wir am Beispiel von zwei Studien der Frage der Wirkrichtung nach und prüfen, inwieweit sich bezogen auf das Fachwissen unterschiedliche Formen des Selbstkonzepts (fachbezogenes und allgemeines) und des Interesses an Politik (am Schulfach und an politischen Themen) ausdifferenzieren lassen. Gefragt wird weiter, ob soziale Vergleichsprozesse im Politikunterricht wirksam sind.

Methode: In zwei Befragungen werden das Selbstkonzept und das Interesse variiert. Befragt wurden einmal 502 Schüler/-innen der 9. und 10. Klassen an Realschulen und Gymnasien. Wir gehen dabei theoriegeleitet primär zwei Hypothesen nach. Das fachspezifische Selbstkonzept wird erstens mit einer Skala (5 Items) von Möller & Köller (2000) auf das Wissen gerichtet. Das Fachinteresse am Politikunterricht wurde mit fünf Items von Köller, Schnabel & Baumert (2000) erfasst. Der Wissenstest umfasst 34 Items zur Europäischen Union. Der Wissensschnitt der Klasse wird zweitens zentriert am Gesamtwissensschnitt und mit dem individuellen Wissen über die Note auf das fachspezifische Selbstkonzept und Interesse am Fach gerichtet. In einer weiteren Stichprobe mit 1071 Schüler/-innen der 9. und 10. Klassen an Realschulen und Gymnasien werden das allgemeine schulische Selbstkonzept mit der Skala von Schulz et al. (2011) (8 Items), das Interesse an politischen Themen mit einer Skala aus ICCS (2009) (4 Items) sowie das Wissen zum politischen System (35 Items) dargestellt. Es werden jeweils Strukturgleichungen gerechnet.

Ergebnisse: Erwartungskonform wirken das Fachinteresse wie auch das fachspezifische Selbstkonzept positiv auf die Leistung. Das fachspezifische Selbstkonzept ist zusätzlich stark mit dem Fachinteresse assoziiert. Das allgemeine schulische Selbstkonzept wirkt etwas schwächer auf das Wissen und hängt weniger mit dem Interesse an politischen Themen zusammen. Das Interesse an politischen Themen hat einen etwas stärkeren Effekt auf das Wissen. Die Ergebnisse stehen in Einklang mit den theoretischen Annahmen der Politikdidaktik und mit der Theorie von Eccles und Wigfield. In Klassen mit variierenden Leistungsniveaus zeigt sich nur der erwartete Effekt der individuellen Leistung auf das fachspezifische Selbstkonzept. Dieser Effekt ist auch nach Kontrolle der Note gegeben. Ein Strukturgleichungsmodell zeigt, dass der um den Gesamtmittelwert der Klasse zentrierte Wissensschnitt der Klasse als sozialer Vergleich keine Effekte hat. Der soziale Vergleichsaspekt ist nicht darstellbar. Die Frage der Wirkrichtung lässt sich theoretisch, aber nicht mit den Ergebnissen entscheiden.

Donnerstag
12. März

9:00 Uhr

HGB 30

Chair(s)

Oberle, Monika

Forstmann, Johanna; Oberle, Monika

Politisches Selbstkonzept: Facetten und geschlechtsspezifische Einflussfaktoren

Theoretischer Hintergrund: Das Selbstkonzept der eigenen Fähigkeiten gilt als wichtiger Faktor erfolgreicher Lernprozesse. Nicht befriedigend geklärt ist bislang die Frage, inwiefern das politische Selbstkonzept eine vom generellen akademischen Selbstkonzept sowie anderen nicht-akademischen Facetten des Selbstkonzepts abgrenzbare Disposition darstellt. Erhebungsinstrumente des fachspezifischen Selbstkonzepts der Begabung können sich auf das Fach Politik beziehen, jedoch auch auf unterrichtsunabhängige Fähigkeiten. Mit Blick auf das politische Selbstkonzept haben zahlreiche Studien – insbesondere der politischen Sozialisations-, Jugend- und Kulturforschung – das subjektive politische Wissen sowie das interne politische Effektivitätsgefühl (bspw. die Fähigkeit, andere von der eigenen politischen Meinung zu überzeugen) von Jugendlichen erhoben. Hierbei wurden Geschlechterdifferenzen zu Lasten von Mädchen bzw. Frauen gezeigt, auch unter Kontrolle des Fachwissens (Westle, 2006; Oberle, 2012). Dass für einen solchen gender gap im fachspezifischen Selbstkonzept ein stereotype threat (Steele, 1997) verantwortlich sein kann, konnten Studien zu den MINT-Fächern zeigen (Spencer et al., 1999). Dass ein solcher Erklärungsansatz für den gender gap im politischen Selbstkonzept vielversprechend ist, legen Studien der politischen Wissensforschung nahe (McGlone et al., 2006).

Fragestellungen: Bilden die drei Facetten (unterrichts-)fachspezifisches Selbstkonzept der Begabung, subjektives politisches Wissen bzw. Verständnis sowie internes politisches Effektivitätsgefühl (internal political efficacy) unterscheidbare Faktoren eines gemeinsamen Messmodells? Lässt sich das politische Selbstkonzept von Schüler/-innen von ihrem allgemeinen akademischen Selbstkonzept abgrenzen? Sind die stereotypen Geschlechterrollenvorstellungen der Schülerinnen und Schüler ein relevanter Prädiktor ihres politischen Selbstkonzepts? Zeigen sich geschlechtsspezifische Unterschiede hinsichtlich der Einflüsse politischer Sozialisationsinstanzen (Eltern, peers, Medien) auf das politische Selbstkonzept der Jugendlichen?

Methode: Die untersuchte Stichprobe besteht aus N = 714 Schülerinnen und Schülern der Jahrgänge 9 und 10 aus insgesamt 29 Klassen (Gymnasien und Gesamtschulen, überwiegend Niedersachsen; 67 % Gymnasium; 52,5 % weiblich; Altersdurchschnitt 15,4 Jahre). Die Datenerhebung erfolgte 2013 und 2014 mittels standardisiertem Fragebogen durch geschulte Versuchsleiter/-innen im Klassenverband. Messmodelle der interessierenden Konstrukte wurden ebenso wie latente Regressionen und Strukturgleichungen in MPlus 7.0 gerechnet, unter Berücksichtigung der hierarchischen Struktur der Daten. Dabei kamen überwiegend 4-stufig likert-skalierte Items zum Einsatz. Die latent modellierten Konstrukte weisen eine akzeptable Reliabilität und Datenpassung auf: Das akademische Selbstkonzept wurde mit 4 Items erhoben (nach Schilling et al. 2005 sowie Arens et al. 2011; Fitwerte: Cronbach's α = .79; CFI= 1.00; RMSEA= .03; WRMR= .30; χ^2 = 3.68(2)ns). Das politische Selbstkonzept wurde als 3-faktorielles Konstrukt modelliert: (i) subjektives Wissen (6 Items; u.a. nach Vetter 2006 und Kerr et al. 2010), (ii) internes Effektivitätsgefühl (9 Items; nach Kerr et al. 2010 und Westle 2006), (iii) fachspezifisches Selbstkonzept der Begabung (4 Items; nach Köller et al. 2006, Kessels 2002); Fitwerte: Cronbach's α : I= .84, II= .86, III= .89; CFI= .98; RMSEA= .055; WRMR= 1.58; χ^2 = 476.30(149)***. Geschlechterrollenstereotype wurde als 2-faktorielles Konstrukt modelliert: I = privat bzw. Haushalt (5 Items), II= politisch (4 Items); Fitwerte: Cronbach's α : I= .77, II=.80; CFI= .98; RMSEA= .06; WRMR= 1.00; χ^2 = 87.75(26)***.

Ergebnisse: Ein dreidimensionales Konstrukt des politischen Selbstkonzepts (auf das Schulfach bezogenes Selbstkonzept der Begabung; subjektives politisches Wissen bzw. Verständnis; internes politisches Effektivitätsgefühl) lässt sich erfolgreich modellieren und deutlich von einem allgemeinen akademischen Selbstkonzept abgrenzen. Es zeigen sich geschlechtsspezifische Unterschiede sowohl in der Ausprägung des politischen Selbstkonzepts der Schüler/-innen, als auch bezüglich dessen Prädiktoren. So spielen bei Mädchen ihre Geschlechterrollenvorstellungen (politische Dimension) eine bedeutsame Rolle. Weitere relevante Einflussfaktoren sind die politischen Sozialisationsinstanzen Mutter, Vater, peers und Massenmedien, wobei sich wiederum geschlechtsspezifische Muster zeigen.

Donnerstag
12. März

9:00 Uhr

HGB 30

Chair(s)

Oberle, Monika

Gantner, Stephan

Operationalisierung und Messung berufsfachlicher Handlungskompetenz in dualen und schulischen Ausbildungsberufen

Von der beruflichen Ausbildung wird aus ökonomischer und gesellschaftlicher Perspektive ein wichtiger Beitrag erwartet (Neumann, 2006). Deshalb ist es notwendig sicherzustellen, dass die in den Vorgaben der verschiedenen Ausbildungsberufe festgelegten Bildungsziele auch erreicht werden. Um dies zu gewährleisten, werden Messinstrumente benötigt, die eine valide Erfassung der in der Ausbildung de facto entwickelten Kompetenzen in den verschiedenen Ausbildungsberufen ermöglichen. In diesem Symposium werden an diese Problematik anschließend vier Projekte aus unterschiedlichen disziplinären Kontexten vorgestellt, die sich mit der Operationalisierung und Messung berufsfachlicher Handlungskompetenz in dualen und schulischen Ausbildungsberufen auseinandersetzen. Adressiert werden die beruflichen Ausbildungsgänge „Laborassistent(inn)en in Biologie und Medizin“, „Medizinische Fachangestellte“, „Fachinformatiker(inne)n“ und „Elektroniker(inne)n der Automatisierungstechnik“. Trotz der hohen inhaltlichen Heterogenität der verschiedenen Ausbildungsgänge, finden sich Übereinstimmungen in den strukturellen Annahmen zur Kompetenzstruktur. Der Erwerb der beruflichen Handlungskompetenzen, stellt einen wesentlichen Bestandteil der beruflichen Ausbildung dar (KMK, 1996). Sie umfasst neben dem Fachwissen verschiedene prozedurale Kompetenzen (Fertigkeiten und Fähigkeiten), die sich wiederum jeweils aus berufsspezifischen Facetten zusammensetzen. Die Erfassung und empirische Überprüfung von Kompetenzmodellen sind seit etwa einer Dekade Gegenstand der Bildungsforschung (Seeber et. al 2010).

Auf der Basis berufsspezifischer theoretischer Kompetenzmodelle zu den oben genannten Ausbildungsberufen, die die jeweiligen berufsfachlichen Kompetenzen in ihren Subdimensionen und deren Zusammenhänge beschreiben, wurden in verschiedenen Projekten Testinstrumente entwickelt und die unterstellten Kompetenzmodelle empirisch überprüft. Die Kompetenzmodelle und die entwickelten Instrumente werden in diesem Symposium vorgestellt und vergleichend diskutiert. Im Einzelnen handelt es sich dabei um die folgenden Projekte, Instrumente und Kompetenzdimensionen: (1) Im Projekt ManKobE (bei Laborant(inn)en und Technischen Assistent(inn)en der Biologie und Medizin) wurde ein ‚Paper & Pencil‘-Test entwickelt, der das Handlungswissen im Bereich der betrieblichen Abläufe in den Facetten Arbeitsorganisation, verantwortliches Handeln und Labortechniken erfasst. (2) Im Projekt CosMed (bei Medizinischen Fachangestellten) wurde ein adaptiver ‚Paper & Pencil‘-Test zur Erfassung des Fachwissens entwickelt und unter Verwendung von Video Clips simulationsgestützt prozedurales Wissen erfasst. (3) Im Projekt Aqua.Kom (bei Fachinformatiker(inne)n) wurde ein ‚Paper & Pencil‘ - Test zur Erfassung des Fachwissens entwickelt sowie ein PC basierter Simulationstest zur Erfassung der fachlichen Problemlösefähigkeit. (4) Im Projekt KOKO-EA (bei Elektroniker(inne)n der Automatisierungstechnik) wurde eine Computersimulation als auch ‚Paper & Pencil‘-Tests entwickelt, um verschiedene Subdimensionen beruflicher Kompetenzen valide zu erfassen.

Donnerstag
12. März

9:00 Uhr

HGB 40

Chair(s)

Gantner, Stephan
Harms, Ute**Operationalisierung und Messung berufsfachlicher Handlungskompetenz in dualen und schulischen Ausbildungsberufen**

Gantner, Stephan; Harms, Ute

Messung des Handlungswissens in der beruflichen Ausbildung von Laborant(inn)en und Technischen Assistent(inn)en in der Biologie und Medizin

Theoretischer Hintergrund: Wie die schulische Ausbildung (z. B. KMK, 2004) folgt inzwischen auch die berufliche Ausbildung kompetenzorientierten Beurteilungskriterien zur Bewertung des Ausbildungserfolgs (Winther & Klotz, 2014). Testinstrumente zur Erfassung der zu erwerbenden Handlungskompetenz liegen zum jetzigen Zeitpunkt allerdings nur in Ansätzen vor (Rosendahl & Straka, 2010). Das am IPN angesiedelte Projekt ManKobE (Mathematisch-naturwissenschaftlichen Kompetenz in der beruflichen Erstausbildung) widmet sich daher der Erfassung eines Kerns der professionellen Handlungskompetenz, dem Professionswissen. Dieses beschreibt all jene Wissensbestände, die es den späteren Absolventen erlauben, den Anforderungen des beruflichen Alltags gerecht zu werden. Bezüglich des Professionswissens wird an die in der Kognitionspsychologie etablierte Differenzierung in theoretisch-formales, verbal beschreibbares Professionswissen (deklaratives Wissen) angeknüpft und dieses dem praktisch-handlungsorientierten Wissen gegenüber gestellt (prozedurales Wissen; de Jong & Ferguson-Hessler, 1996; König, 2010). Letzteres wird hier als berufsspezifisches Handlungswissen bezeichnet.

Ziel dieses Beitrags ist die Entwicklung eines Testinstruments zur Erfassung berufsspezifischen Handlungswissens von Laborant(inn)en und Technischen Assistent(inn)en in der Biologie (BL, BTA) und Medizin (MTLA). Basierend auf den bundes- und landesgesetzlichen Vorgaben und Rahmenlehrplänen sowie den Vorgaben der IHK wurde dieses in drei Wissensfacetten operationalisiert: ‚verantwortliches Handeln‘ (VH), ‚Arbeitsorganisation‘ (AO) und ‚Labortechniken‘ (LT). VH beschreibt das Handlungswissen in den Bereichen Laborsicherheit und Umweltschutz, AO das Handlungswissen in den Bereichen Informationsbeschaffung und Labordatenerfassung/–auswertung und LT das Handlungswissen, das sich auf den Einsatz von Geräten und Methoden bei Laboranalysen bezieht.

Fragestellungen: (1) Sind die Wissensfacetten des berufsspezifischen Handlungswissens (VH, AO, LT) empirisch trennbar? (2) In welchem Zusammenhang stehen die Wissensfacetten VH, AO und LT zueinander?

Methode: Auf der Basis einer Analyse aktuell gültiger Curricula wurde ein curricular valider Itempool für die drei Wissensfacetten VH (33 Multiple Choice Items), AO (48 Multiple Choice Item) und LT (54 Multiple Choice Items) entwickelt. Zur systematischen Gestaltung der Itementwicklung wurde ein Aufgabenraaster aufgespannt, in dem zwischen vier biologischen Inhaltsbereichen (Biochemie, Mikrobiologie, Molekularbiologie und Zellbiologie) und zwei kognitiven Anforderungen (‚Erinnern & Abrufen‘ bzw. ‚Verstehen & Anwenden‘) unterschieden wird. Der Itempool wurde einer Stichprobe aus 250 Auszubildenden vorgelegt, die an vier unterschiedlichen Ausbildungsorten ihrer Ausbildung in den Bereichen BL, BTA und MTLA nachgingen. Die Auszubildenden waren im Durchschnitt 20 Jahre (SD=4,8) alt (67% weiblich); 70% hatten Fachhochschulreife, 29% Realschul-, und 1% Hauptschulabschluss. Zur Überprüfung der diskriminanten Validität wurden multidimensionale Rasch-Analysen durchgeführt (einfaches logistisches Modell).

Ergebnisse: Gemäß unserer ersten Forschungsfrage wurde überprüft, ob sich die empirischen Daten durch die Annahme dreier latenter Fähigkeiten (VH, AO, LT) besser erklären lassen als durch die Annahme einer einzigen Fähigkeit. Zur Beantwortung der Forschungsfrage wurde ein dreidimensionales Rasch-Modell an die Daten angepasst, in dem jedes Item als Indikator für eine bestimmte Wissensfacette operiert. Diesem Modell wurde ein eindimensionales Modell gegenübergestellt, in dem alle Items als Indikatoren einer einzigen latenten Fähigkeit fungieren. Die Rasch-Analysen zeigen, dass die Informationskriterien für das dreidimensionale Modell (AIC: 12833,6; BIC: 13165,7) niedriger ausfallen – und damit auf eine bessere Modellpassung hindeuten – als für das eindimensionale Modell (AIC: 12934,6; BIC: 13253,8). Ein χ^2 -Differenztest ($\chi^2[5, N=250]=37,56; p<,001$) zeigt zudem, dass dieser Unterschied signifikant ausfällt. Als Maß für die Reliabilität der Subskalen wurde die EAP/PV-Reliabilität bestimmt (VH=,80; AO=,77; LT=,82), die auf eine gute Messgenauigkeit des entwickelten Instruments hindeuten.

Gemäß unserer zweiten Forschungsfrage wurden latente Korrelationsanalysen zwischen den Facetten des berufsspezifischen Handlungswissens berechnet. Dabei ergeben sich durchwegs hohe Zusammenhänge zwischen den Facetten des Handlungswissens: $r=,85$ ($p<,001$) zwischen VH und LT; $r=,82$ ($p<,001$) zwischen AO und LT; $r=,83$ ($p<,001$) zwischen AO und VH.

Unsere Studie zeigt, dass die Wissensfacetten VH, AO und LT zwar deutlich korrelierte, aber dennoch empirisch trennbare Fähigkeiten darstellen. Das Messinstrument soll im weiteren Projektfortschritt dazu genutzt werden, um Unterschiede in der dualen industriebezogenen und staatlich- oder privat-schulischen Ausbildung zu identifizieren.

Donnerstag
12. März

9:00 Uhr

HGB 40

Chair(s)

Gantner, Stephan
Harms, Ute

Operationalisierung und Messung berufsfachlicher Handlungskompetenz in dualen und schulischen Ausbildungsberufen

Seeber, Susan; Ketschau, Thilo; Rueter, Theresa

Kognition und Prozedur: Strukturanalysen auf der Grundlage adaptiver und simulationsgestützter Testformate bei Medizinischen Fachangestellten

Theoretischer Hintergrund: In Anlehnung an das Evidence-Centered Assessment-Design (Mislevy & Haertel 2006) wurde für die Messung beruflicher Kompetenzen Medizinischer Fachangestellter ein Domänenmodell und darauf aufbauend ein mehrdimensionales Kompetenzmodell spezifiziert, das (1) Kompetenzen in der Betreuung von Patienten, (2) medizinische Fachkompetenzen im Hygiene- und Präventionsmanagement sowie in der Labordiagnostik und (3) kaufmännisch-verwaltende Fachkompetenzen als Dimensionen unterscheidet. Neben den Wissensarten wurden bei der Konstruktion der Testaufgaben weitere kognitionspsychologische sowie fachdidaktische Anforderungen zugrunde gelegt: (a) Unterschiedliche institutionelle Lernkontexte (Betrieb, Schule), (b) Bezug zum curricularen Kontext und zu realen Arbeitsprozessen, (c) systemische Perspektive auf Prozesse, (d) Einbindung der Assessmentaufgaben in Sequenzen mit fachlichen und fachübergreifenden Anforderungen, (e) verschiedene Komplexitätsgrade der Aufgaben, (f) nicht eindeutig definierte berufliche Handlungsanforderungen und (g) unterschiedliche kognitive Anforderungen (vgl. Schavelson, Ruiz-Primo & Wiley 2005; Achtenhagen & Winther 2009).

Bei der Unterscheidung der Wissensarten wurde der Klassifikation nach Anderson und Krathwohl (2001) gefolgt. Unter Faktenwissen wird explizites, verbalisierbares und für eine bestimmte Domäne relevantes Wissen verstanden. Konzeptuelles Wissen wird als vernetztes Begriffswissen in Form von Kategorien, Modellen, Schemata betrachtet. Prozedurales Wissen umfasst hingegen einfache bis komplexe Routinen und Handlungsmuster und liegt überwiegend implizit vor.

Fragestellung: In dem Beitrag wird der Frage nachgegangen, in welchem Verhältnis prozedurales und konzeptuelles Wissen bei Medizinischen Fachangestellten steht.

Methode: Für die Erfassung von Faktenwissen und konzeptuellem Wissen wurde ein adaptiver Test entwickelt. Um die beruflichen Handlungskompetenzen und das darin eingebettete prozedurale Wissen zu messen, wurden authentische berufliche Handlungssituationen mit Videoclips dargestellt und neben Arbeitsdokumenten in eine webbasierte Testumgebung eingebunden. Die faktorielle Validität der mehrdimensionalen Modelle wurde mittels der Rasch-Analyse geprüft.

Ergebnis: Skaliert wurde der adaptive Test zur Messung deklarativen Fachwissens über das einparametrische Rasch-Modell. In der Pilotierung (N=598) wurde eine EAP/PV-Reliabilität von 0,765 erreicht. Dimensionsanalysen lieferten keine Evidenzen für das angenommene dreidimensionale Modell. Eine zweidimensionale Modellierung nach medizinischen und kaufmännischen Handlungsbereichen lieferte ebenso keinen besseren Modellfit im Vergleich zum eindimensionalen Modell. Auch untermauerte die Korrelation zwischen den beiden Dimensionen von 0,96 ein eindimensionales Modell.

Die 64 Items aus dem Videotest (N=128), der auf die Messung prozeduralen Wissens zielte, erreichten in der Pilotierung eine EAP/PV-Reliabilität von 0,821. Modellprüfungen auf der Grundlage mehrdimensionaler Rasch-Modelle legten beim Videotest eine zweidimensionale Struktur mit der Differenzierung in den medizinischen und kaufmännischen Handlungskontext nahe. Die bessere Modellpassung und die messfehlerbereinigte Korrelation von .476 stützten ein solches zweidimensionales Modell.

Die Haupterhebung wurde im Juli 2014 mit einer umfangreichen Stichprobe abgeschlossen (N=1.155, 66 Klassen, 37 Schulen). Es soll nun geprüft werden, ob die empirischen Belege zum Verhältnis von konzeptuellem und prozeduralem Wissen, in denen eine relative Stärke von sog. „Handlungskompetenzen“ (prozedural) im Vergleich zu „Verstehensbasierten Kompetenzen“ (konzeptuell) bei Auszubildenden im kaufmännischen Bereich nachgewiesen werden konnte (vgl. Winther & Achtenhagen 2009), auch für die Medizinischen Fachangestellten gelten. Die Analysen werden auf der Grundlage von Strukturgleichungsanalysen mit MPLUS vorgenommen. Dabei wird von einem Within-Modell ausgegangen, da sowohl eine Reihe von Items disjunkt in Bezug auf die Wissensarten ist, als auch verschiedene Wissensarten in einer substanziellen Zahl an Items jeweils angesprochen werden.

Donnerstag
12. März

9:00 Uhr

HGB 40

Chair(s)

Gantner, Stephan
Harms, Ute

Operationalisierung und Messung berufsfachlicher Handlungskompetenz in dualen und schulischen Ausbildungsberufen

Velten, Stefanie; Dietzen, Agnes; Nitzschke, Alexander; Maier, Annette

Erklärungsmodelle der Fachkompetenz von Fachinformatikern und Fachinformatikerinnen

Theoretischer Hintergrund: Im beruflichen Bildungsbereich lassen sich domänenübergreifend empirisch zumindest das Fachwissen und die Fähigkeit dieses Wissen in problemhaltigen Kontexten anzuwenden (fachspezifische Problemlösefähigkeit) als eigene Subdimensionen unterscheiden (Nickolaus/Seeber 2013). Erklärungsmodelle zur Entwicklung berufsfachlicher Kompetenzen liegen im kaufmännischen und gewerblich-technischen Bereich vor (z.B. Abele u.a. 2012; Nickolaus u.a. 2010). Das Fachwissen wird darin meist als Prädiktor der fachspezifischen Problemlösefähigkeit modelliert und besitzt dafür domänenübergreifend die größte prädiktive Kraft. Die gemeinsame Varianz des Fachwissens und der fachspezifischen Problemlösefähigkeit liegt häufig bei ca. 50% (Nickolaus u.a. 2011). Anderen kognitiven Voraussetzungen (fluide Intelligenz, dynamische Problemlösefähigkeit sensu Greiff/Funke) werden lediglich indirekte Einflüsse auf die fachspezifische Problemlösefähigkeit attestiert (Abele u.a. 2012). Für das Fachwissen erweisen sich das fachspezifische Vorwissen, mathematische Fähigkeiten und die Lesekompetenz als erklärungsrelevant. Motivation und Qualitätsmerkmale der Ausbildung gehen mit geringem Gewicht ein. Von zentralem Interesse scheint, inwieweit sich diese Erklärungsmuster im berufsfeldübergreifend geschnittenen Ausbildungsberuf Fachinformatiker/-in bestätigen lassen. Laut BERUFENET (2014) planen und konfigurieren Fachinformatiker/-innen IT- Systeme und analysieren und beheben auftretende Störungen. Zugleich werden auch kaufmännische Kompetenzen relevant, was in den Rahmenlehrplänen in Form hoher kaufmännischer Anteile Ausdruck findet. Vor diesem Hintergrund erwarten wir, dass bezogen auf technische Problemstellungen das eindimensional skalierte Fachwissen deutlich geringere prädiktive Kraft für die Problemlöseleistung besitzt, als in rein technisch geschnittenen Berufen. Dafür sprechen auch erste Analyseergebnisse im Ausbildungsberuf Elektroniker/-in, für die Walker (2014) zeigen konnte, dass lediglich jenen Fachwissensdimensionen prädiktive Kraft für die Problemlösekompetenzen zukommt, die inhaltlich eng assoziiert sind.

Fragestellung: Lässt sich für Fachinformatiker/-innen die hohe prädiktive Kraft des fachspezifischen Vorwissens und der Basiskompetenzen für das Fachwissen bestätigen und erweisen sich nur inhaltlich affine Subdimensionen des Fachwissens als prädiktiv für die fachliche Problemlöseleistung?

Im Vortrag werden zunächst Ergebnisse zur Modellierung der berufsfachlichen Kompetenz berichtet und anschließend darauf bezogene Erklärungsmodelle, in welchen kognitive Merkmale (Intelligenz, Vorwissen, Mathematikkenntnisse, Leseverständnis) und Merkmale der Lernumgebungen sowie die Motivation Berücksichtigung finden.

Methode: Die Kooperationsstudie des BIBB und der Universität Stuttgart ist längsschnittlich angelegt und überspannt in zwei geschachtelten Längsschnitten mit rund 1500 Auszubildenden die gesamte Ausbildungsspanne. Leistungsdaten wurden zum Ausbildungsbeginn, Mitte des zweiten Ausbildungsjahres und am Ausbildungsende in vier Bundesländern (2010 bis 2012) erhoben (Dietzen u.a. 2014). Motivationsdaten und die Ausbildungsqualitäten wurden zweimal erfasst.

Theoriegeleitet wurden vier Modelle des Fachwissens geschätzt und anschließend mittels Chi²-Test gegeneinander getestet, für die Erklärungsmodelle wurden Strukturgleichungsmodelle berechnet.

Ergebnisse: Die Reliabilität des Fachwissenstests (PPT) liegt eindimensional bei EAP/PV=.77. Der Problemlösetest (realisiert in einer Client-Server-Umgebung) erreicht eine Reliabilität von EAP/PV=.63. Die relativ geringe Reliabilität resultiert aus der begrenzten Testzeit und den langen Bearbeitungszeiten der authentischen Problemstellungen.

Für das Fachwissen zum Zwischen- und Abschlusstest erweisen sich fünfdimensionale Modelle als am besten passend, wobei sich auch in dieser Domäne die Ausdifferenzierung entlang curricularer Inhaltsbereiche vollzieht. Die latenten Zusammenhänge zwischen den Subdimensionen des Fachwissens variieren zwischen .49 und .83, wobei die Zusammenhänge für jene Subdimensionen mit engeren inhaltliche Bezüge starke Ausprägungen zeigen (zwischen den kaufmännischen Subdimensionen ($r=.75$) bzw. den Subdimensionen Computer- und Netzwerktechnik ($r=.83$)).

Zusammenhangsanalysen bis zum Zwischentest ergaben, dass die Lese- und Mathematikleistungen 46% der Varianz des Eingangstest erklären und der IQ nur indirekt eingeht. Das Vorwissen hat mit einem Pfadkoeffizienten von .77 den stärksten Einfluss auf das Fachwissen zum Zwischentest (erklärte Varianz unter Einbezug der Motivation 62%). Ähnliche Befundmuster werden auch für den Abschlusstest erwartet.

Zwischen der berufsfachlichen Problemlöseleistung und dem Fachwissen zum Ende der Ausbildung ergeben sich deutlich geringere Zusammenhänge als für andere Domänen berichtet wurde. Diese liegen manifest in der Größenordnung von .2-.3. Erwartungskonform zeigt sich, dass die Korrelationen ausschließlich zwischen den Problemlöseleistungen und den technischen Subdimensionen des Fachwissens signifikant ausfallen. Eine latente Modellierung wird gegenwärtig erstellt.

Donnerstag
12. März

9:00 Uhr

HGB 40

Chair(s)

Gantner, Stephan
Harms, Ute

3

Walker, Felix; Nickolaus, Reinhold

Facetten fachspezifischer Problemlösekompetenz von Elektronikern der Automatisierungstechnik: Diagnostische Zugänge und Zusammenhänge

Theoretischer Hintergrund: Berufsfachliche Problemlösekompetenzen stellen relativ heterogene Konstrukte dar, deren Operationalisierung im gewerblich-technischen Bereich überwiegend mit Bezug zu analytischen Anforderungskontexten als Fehleranalysefähigkeit in technischen Systemen erfolgte. Für die Fehleranalysefähigkeit, als analytische Facette fachspezifischer Problemlösekompetenz, konnte domänenübergreifend die empirische Trennbarkeit von Fachwissen und fachspezifischer Problemlösekompetenz bestätigt werden (s. NICKOLAUS 2011, S. 334). Des Weiteren weisen Arbeiten aus dem gewerblich-technischen (s. NICKOLAUS u.a. 2012, S. 265ff.) und dem allgemeinbildenden Bereich (s. SCHERER 2012, S. 40ff. und S. 16) die hohe prädiktive Kraft des Fachwissens für die fachspezifische Problemlösekompetenz nach.

Bezogen auf die Facetten der fachspezifischen Problemlösekompetenz wurde für elektrotechnische Berufe neben einer analytischen mit einer konstruktiven Facette der fachspezifischen Problemlösekompetenz bereits in älteren Beiträgen ein mehrdimensionales Konstrukt unterstellt (s. NICKOLAUS 2011, S. 334). Diese Vermutung konnte durch Analysen betrieblicher Tätigkeiten in elektrotechnischen Berufen (s. KALLIES/HÄGELE/ZINKE 2014; ZINKE/SCHENK/KRÖLL 2014) erhärtet werden, in der sich neben analytischen auch konstruktive Anforderungssituationen als hoch bedeutsam und als typische berufliche Tätigkeiten erwiesen. Zusätzlich werden die Vermutungen durch Ergebnisse der PISA-Studien zur Problemlösekompetenz gestützt. Analysen von LEUTNER u.a. (2012) ergaben, dass analytische Anforderungssituationen (z.B. Fehlersuche) und konstruktive Anforderungssituationen (z.B. Systeme entwerfen) zwar hoch korrelierte jedoch empirisch trennbare Facetten der Problemlösekompetenz sind (s. LEUTNER u.a. 2012, S. 37).

Fragestellung: Es wird untersucht, ob analytische und konstruktive Anforderungssituationen im Beruf des Elektroniklers für Automatisierungstechnik empirisch trennbare Facetten der fachspezifischen Problemlösekompetenz darstellen und wie sich der Einbezug des Fachwissens und der kognitiven Grundfähigkeit in die Modellierung auf den Zusammenhang beider Facetten auswirkt.

Methode: Im Beitrag werden ausgehend von den theoretischen Konzeptionen der beiden Facetten (analytisch/konstruktiv) der fachspezifischen Problemlösekompetenz zwei diagnostische Zugänge für deren Erfassung vorgestellt und bzgl. ihrer Validität diskutiert. Anschließend wird die Struktur der Facetten unter Einbezug des Fachwissens und der kognitiven Grundfähigkeit analysiert. Gerechnet werden hierfür Strukturgleichungsmodelle mit (a) einer frei geschätzten Korrelation zwischen den Facetten (analytisch/konstruktiv) und (b) einem eindimensionalen Modell, was einer Fixierung der Korrelation zwischen den Facetten auf eins entspricht. Basierend auf Fitwerten (z.B. ΔCFI bzw. LR-Test) wird die Passung der Modelle verglichen.

Ergebnisse: Bisherige, auf einem Teildatensatz beruhende, Analysen ($n=124-150$) stützen die Annahme, dass es sich bei den Dispositionen zur Bewältigung analytischer und konstruktiver Anforderungssituationen um empirisch trennbare Facetten fachspezifischer Problemlösekompetenz handelt (latente Korrelation $r=.6$). Unter Einbezug des Fachwissens und der kognitiven Grundfähigkeiten sinkt der Zusammenhang zwischen den Facetten weiter ab (latente Korrelation $r=.4$).

Diskussion: Die Befunde zeigen einen großen Einfluss des Fachwissens auf die beiden Facetten fachspezifischer Problemlösekompetenz. Weiter sprechen sie dafür, dass es sich um eigenständige Konstrukte handelt. Inwiefern die inhaltliche Nähe beider Facetten eine mögliche Erklärung für die verbleibende gemeinsame Varianz (ca. 16%) darstellt, soll abschließend diskutiert werden.

Donnerstag
12. März

9:00 Uhr

HGB 40

Chair(s)

Gantner, Stephan
Harms, Ute

Operationalisierung und Messung berufsfachlicher Handlungskompetenz in dualen und schulischen Ausbildungsberufen

Daniel, Annabell

Soziale Ungleichheit beim Hochschulzugang und Möglichkeiten der Intervention

Im internationalen Vergleich ist die Beteiligung an Hochschulbildung in Deutschland niedrig und zudem in hohem Maße sozial selektiv. Verschiedene Untersuchungen konnten nachweisen, dass die Chance des Hochschulzugangs systematisch mit dem sozioökonomischen Status der Familie variiert und die gestiegene Zahl an Veröffentlichungen hierzu belegt eindrucksvoll die Aktualität des Themas (Lörz 2012, Becker & Hecken 2009). So konnten neben Unterschieden in der Studienaufnahme auch herkunftsspezifische Unterschiede bei der Studienfachwahl sowie beim Übergang in ein Masterstudium festgestellt werden. Auf welche Mechanismen soziale Ungleichheiten beim Hochschulzugang zurückzuführen sind ist in der bildungssoziologischen Forschung gut dokumentiert. Anknüpfend an die Arbeit von Boudon (1974) wurden verschiedene entscheidungstheoretische Modelle vorgelegt, die sozial selektive Beteiligungsquoten an Hochschulbildung als eine Folge von individuellen, zwischen Sozialschichten variierenden Leistungen (primäre Herkunftseffekte) und schichtspezifischen Kosten-Nutzen-Abwägungen bei der Bildungsentscheidung (sekundären Herkunftseffekten) betrachten (Erikson & Jonsson 1996, Breen & Goldthorpe 1997). Ebenso als nützlich für die Analyse ungleicher Bildungschancen haben sich theoretische Ansätze kultureller Reproduktion erwiesen (Bourdieu & Passeron 1971).

Die Erkenntnisse vorliegender Studien liefern Hinweise darauf, wo Interventionen zur Verringerung sozialer Ungleichheit beim Hochschulzugang ansetzen müssten. Es fehlte bislang jedoch an belastbaren Daten, um die Effektivität entsprechender Maßnahmen untersuchen zu können. Das Symposium widmet sich daher der Frage, inwieweit es gelingen kann, den Zugang zur Hochschule weniger sozial selektiv zu gestalten. Aus der Perspektive unterschiedlicher Fachdisziplinen und unter Anwendung verschiedener methodischer Verfahren werden sowohl individuelle als auch institutionelle Maßnahmen betrachtet.

Der erste Beitrag von Daniel, Watermann und Maaz untersucht im Kontrollgruppendesign den Einfluss einer schulischen Informationsveranstaltung auf zentrale Determinanten der Studienentscheidung. Die Ergebnisse zeigen, dass eine gezielt an Studienberechtigte nichtakademischer Herkunft gerichtete Intervention mit unterschiedlichen Konsequenzen für die Einschätzung von studienbezogenen Erträgen, Kosten und Erfolgswahrscheinlichkeiten einhergeht. Aus bildungsökonomischer Sicht beschäftigt sich der zweite Beitrag von Peter, Zambre, Spieß & Storck mit der Frage, inwieweit sich Informationsdefizite von Studienberechtigten nichtakademischer Herkunft verringern lassen und in der Folge zu höheren Studienaspirationen führen. Sie nutzen dazu Daten eines randomisierten Feldexperiments und finden differenzielle Interventionseffekte für SchülerInnen unterschiedlicher Herkunft. Basierend auf dergleichen Datengrundlage untersuchen Finger, Rusconi, Solga und Ehlert im dritten Beitrag, inwiefern herkunftsspezifische Studienfachaspirationen durch Informationen zu Arbeitsmarkterträgen beeinflusst werden und Ungleichheiten in der Fächerwahl ausgleichen können.

Schließlich geben Quast & Lörz im vierten Beitrag einen Ausblick wie sich vor dem Hintergrund der Studienstrukturreform Entscheidungsprozesse am Bachelor-/Masterübergang gestalten. Die Ergebnisse liefern Hinweise, welche Interventionsansätze an der zweiten Übergangsschwelle erfolgversprechend sein könnten, um auch den Zugang zum weiterführenden Studium von der sozialen Herkunft zu entkoppeln.

Donnerstag
12. März

9:00 Uhr

HGB 50

Chair(s)

Daniel, Annabell
Watermann, Rainer**Soziale Ungleichheit beim Hochschulzugang und Möglichkeiten der Intervention**

Daniel, Annabell; Watermann, Rainer; Maaz, Kai

Effekte einer Schulintervention auf zentrale Determinanten der Studienentscheidung

Zahlreiche Studien der empirischen Bildungs- und Ungleichheitsforschung belegen eine hohe soziale Selektivität des Hochschulzugangs. Als ein Erklärungsansatz hierfür hat sich insbesondere die von Boudon (1974) eingeführte Unterscheidung von primären und sekundären Herkunftseffekten bewährt. Analysen zu deren relativen Bedeutsamkeit beim Hochschulzugang zeigen, dass sich die sozialen Ungleichheiten weniger auf Leistungsunterschiede, sondern vielmehr auf sekundäre Effekte, also Unterschiede im Entscheidungsverhalten, zurückführen lassen (Lörz & Schindler 2011). Basierend auf den Annahmen der Rational-Choice-Theorie wird die Studienentscheidung demnach als das Ergebnis eines herkunftsspezifischen Abwägungsprozesses verstanden, bei dem subjektiv erwartete Erträge, Kosten und Erfolgswahrscheinlichkeiten berücksichtigt werden (Erikson & Jonsson 1996, Breen & Goldthorpe 1997). Dabei sind es vor allem die wahrgenommenen Kosten eines Studiums sowie die geringe Erfolgserwartung, die Schülerinnen und Schüler nichtakademischer Herkunft von einem Studium ablenken (Becker & Hecken 2007). Individuelle Maßnahmen mit dem Ziel der Verringerung sozialer Ungleichheiten beim Hochschulzugang sollten daher an diesen sekundären Effekten ansetzen. Bislang liegen jedoch keine Befunde dazu vor, inwieweit es entsprechenden Interventionen gelingen kann, die Determinanten der Studienentscheidung zu beeinflussen.

Die vorliegende Studie geht dieser Frage nach und untersucht Effekte einer Schulveranstaltung, die sich insbesondere an Studienberechtigte nichtakademischer Herkunft richtet und studienrelevante Informationen bereitstellt. Mit Hilfe eines Kontrollgruppendesigns wird geprüft, ob die Teilnahme an der Intervention einen Einfluss auf die subjektive Einschätzung der von einem Studium erwarteten Erträge, Kosten und Erfolgswahrscheinlichkeiten hat. Insgesamt wurden hierzu N = 2353 Studienberechtigte aus 21 Schulen zu drei Messzeitpunkten über einen Zeitraum von einem Jahr befragt, wobei die erste Befragung im Mittel fünf Wochen nach der Schulveranstaltung stattfand.

Mittels eines Propensity Score Matching unter Berücksichtigung demografischer Variablen, des familiären Bildungshintergrunds sowie der Schulleistung wird untersucht, inwieweit die Effekte auf einen kausalen Einfluss der Schulveranstaltung zurückzuführen sind. Für die längsschnittliche Analyse der Effekte wurden zweifaktorielle ANOVAs mit Messwiederholung berechnet.

Die Ergebnisse zeigen, dass sich die Teilnahme an der Intervention positiv auf die subjektive Einschätzung des Studierertrags auswirkt, dieser Effekt jedoch nicht von zeitlicher Dauer ist. Für die wahrgenommenen Kosten lässt sich ebenfalls ein Interventionseffekt nachweisen: Während in der Kontrollgruppe mit Herannahen des Hochschulübergangs zunehmend höhere Studienkosten wahrgenommen werden, erweist sich die Kosteneinschätzung in der Interventionsgruppe über den gesamten Untersuchungszeitraum hinweg als stabil. Allerdings zeigt sich darüber hinaus, dass eine durch die Intervention angestoßene Auseinandersetzung mit den Studienanforderungen insbesondere bei Schülerinnen und Schülern nichtakademischer Herkunft mit negativen Folgen für die Einschätzung der subjektiven Erfolgswahrscheinlichkeit verbunden ist. Ein signifikanter Interventionseffekt auf die Studienintention der Teilnehmerinnen und Teilnehmer konnte nicht beobachtet werden.

Der vorliegende Beitrag liefert im Rahmen einer Feldstudie damit erste Hinweise auf die Effektivität von Maßnahmen zur Verringerung sekundärer Herkunftseffekte beim Hochschulzugang.

Donnerstag
12. März

9:00 Uhr

HGB 50

Chair(s)

Daniel, Annabell
Watermann, Rainer

Soziale Ungleichheit beim Hochschulzugang und Möglichkeiten der Intervention

Peter, Frauke; Zambre, Vaishali; Spieß, C. Katharina; Storck, Johanna

Preference for College and Educational Inequality: Do Students Lack Information? Evidence from a Randomized Field Experiment

Students from non-academic family backgrounds are still highly underrepresented in German universities. Even when neglecting the early educational inequalities that lower the probability to obtain a university entrance certificate (Abitur), inequality persists. Focusing only on students obtaining the Abitur, students from a non-academic family background are still significantly less likely to enroll in university. The economic literature mainly suggests two reasons for this underrepresentation: (1) information asymmetry and (2) financial constraints. The sociological literature, in addition, emphasizes that (3) cultural distance and (4) status maintenance prevent students from deciding in favor of university education. Whereas the latter three explanations have been widely studied, studies focusing on the first mechanism only began to emerge in recent years and constitute a new but growing literature.

This paper sheds light on whether information deficits prevent students from enrolling in university by evaluating the effect of an information intervention as a potential policy. We investigate how students' intention to go to university and their actual application change as a result of altering their information set. In particular, we assume that students from a non-academic family background lack information on labor market benefits of university education because they cannot draw on their parents' experience as an information source. Thus, students from a non-academic family background might be more responsive to information than their class-mates with an academic background.

In May 2013 we conducted a randomized field experiment and treated students one year prior their graduation from high school. We randomly selected eight of 27 upper secondary schools in Berlin. These schools were mainly sampled from disadvantaged neighborhoods to reach a sufficient number of high school students with a non-academic family background. The information treatment consisted of a 20min in-class presentation on the costs and benefits of university education as well as on possible financing strategies. At the end of the session a 3min video summarized the information for standardization purposes. In August 2013 and 2014 we surveyed the sample again, and thus, are able to compare differences in students' preferences and their application behavior after graduation in the treatment and the control group. To estimate the effect of information on the study intention, we apply a difference-in-differences approach.

We find a positive change in students' preference for university education if they come from a non-academic family background, while we find no effect for students from an academic background. This result supports the hypothesis of asymmetric information. Further, we examine students' pre and post treatment information set to identify whether the provided information led to an update of beliefs. Using the differences in treatment-related beliefs, we also analyze possible channels by which the information intervention affects students' evaluation of university education. We thereby utilize answers on beliefs about unemployment risk, income premium and income over the life cycle. Besides students' preference for university education we also evaluate the effect of our information intervention on students' actual application to university. (Data on students' application behavior will only become available next month and thus will be included in the presentation in March 2015.)

Donnerstag
12. März

9:00 Uhr

HGB 50

Chair(s)

Daniel, Annabell
Watermann, Rainer

Soziale Ungleichheit beim Hochschulzugang und Möglichkeiten der Intervention

Finger, Claudia; Rusconi, Alessandra; Solga, Heike; Ehlert, Martin

Stabilität der Studienfachwahlpläne von Abiturient/innen und die Rolle der sozialen Herkunft

Nicht nur der Hochschulübergang ist von Herkunftsunterschieden gekennzeichnet, Gleiches gilt für die Studienfachwahl (Lörz 2012). Akademikerkinder wählen häufiger prestigereiche Fächer wie Medizin und Jura, um ihren sozialen Status zu reproduzieren und sich von der wachsenden Zahl an Studierenden zu distinguieren (Reimer/Pollak 2010). Nicht-Akademikerkinder wählen hingegen häufiger wie Fächer etwa aus den Ingenieur- oder Sozialwissenschaften — d.h. Fächer mit möglichst geringen direkten Kosten (Dauer des Studiums, benötigte Materialien etc.) und sozialen Kosten (kultureller Distanz zum Elternhaus) (Becker et al. 2010).

Bisherige Forschungsergebnisse beruhen meist auf Querschnittsdaten; Veränderungsprozesse bleiben verborgen. Die Forschung zur Studienaspiration zeigt allerdings, dass mit dem Näherrücken des Entscheidungszeitpunkts bei Studienberechtigten aus Nicht-Akademikerhaushalten überdurchschnittlich häufig ein „Abkühlungsprozess“ stattfindet (Hanson 1994). Desiderate der vorliegenden Forschung sind die Bedeutung von Informationen über Arbeitsmarkterträge unterschiedlicher Fächer für Studienfachwahlen, deren Einfluss auf die Dynamiken der Studienfachwahl generell sowie für unterschiedliche soziale Herkunftsgruppen. Damit befasst sich unser Beitrag.

Verwendet werden Paneldaten des DIW-WZB-Projekts „Berliner-Studienberechtigten-Panel“ (Best Up) für etwa 1000 Berliner Studienberechtigten an 27 Berliner Schulen mit gymnasialer Oberstufe. Die erste Befragung fand im Frühjahr 2013 statt – etwa ein Jahr vor den Abiturprüfungen. Zu den Befragungsinhalten gehörten die (idealistischen und realistischen) Studienfachaspirationen. Anschließend an die Klassenraumbefragung nahmen Schüler/innen von acht zufällig ausgewählten Schulen an einem Informationsworkshop über Finanzierungsmöglichkeiten, Kosten und Erträge eines Studiums, darunter auch Informationen zum Durchschnittseinkommen verschiedener Studienfächer, teil. In zwei weiteren Wellen – etwa neun Monate vor den Abiturprüfungen (Spätsommer 2013) und im Spätsommer 2014 (d.h. während der Bewerbungsphase für Studiengänge) – wurden sie erneut befragt. Diese (derzeit) einmaligen Längsschnittdaten zu (idealistischen und realistischen) Studienfachaspirationen in Kombination mit weiteren erhobenen wichtigen Faktoren (wie Noten, kognitive Grundfähigkeiten, soziale Netzwerke etc.) werden mit Durchschnittseinkommen verschiedener Studienfächer (als Indikatoren studienfachspezifischer Arbeitsmarkterträge) – berechnet auf Basis von Mikrozensusdaten (Glocker/Storck 2014) – erweitert.

Untersucht werden die Veränderungen der „realistischen“ Studienfachaspirationen von der 1. bis zur 3. Welle – ein Zeitraum, in dem es vermehrt zu Reflexions- und Austauschprozessen kommt. Die abhängige Variable ist kategorial (Verbleib auf ertragreicher bzw. ertragsniedriger Studienfachwahl, Wechsel in ertragsreichere bzw. ertragsniedrigere Fächer oder Ausbildung). Wir verwenden entsprechende Regressionsverfahren sowie die KHB-Dekompositionsmethode zur Abschätzung der Einflussstärke der jeweiligen Faktoren. Da die Daten der 3. Welle noch nicht vorliegen, können hier noch keine Ergebnisse berichtet werden.

(1) Aus einer Rational-Choice-Perspektive kann erwartet werden, dass sich – unter Kontrolle von wahrgenommener Erfolgswahrscheinlichkeit und Kosten – das Wissen über höhere Arbeitsmarkterträge bestimmter Fächer (Nutzen) positiv auf deren Wahl auswirkt. Aufgrund geringerer Netzwerkressourcen haben Nicht-Akademikerkinder eingeschränkten Zugang zu diesbezüglichen Informationen. Daher sollten sie überdurchschnittlich stark von externen Informationen über studienfachspezifische Arbeitsmarkterträge profitieren. Die Frage ist somit, ob aufgrund der zusätzlichen Informationen zu Erträgen verschiedener Studienfächer (auch wenn sie nicht durch die Eltern erfolgt) Studienberechtigte aus nicht-akademischem Elternhaus ihre Fachwahlpläne eher in Richtung „ertragreicherer“ Fächer anpassen?

(2) Schulleistungen könnten hierbei eine vermittelnde Rolle spielen, da – unabhängig von den Ertragsinformationen – ein Wechsel zu ertragsreicheren Studienfächern, wie etwa Humanmedizin oder Rechtswissenschaften, aufgrund von Zulassungsbeschränkungen häufig nur für leistungsstarke Studienberechtigte erfolgsversprechend ist. Ergebnisse früherer Studien zeigen, dass es (a) einen Herkunftseffekt in den Abiturleistungen gibt und (b) der Herkunftseffekt in der Fachwahl am stärksten bei Studienberechtigten mit den höchsten Leistungen ist (Reimer/Pollak 2010). Informationen über zu erwartende studienfachspezifische Arbeitsmarkterträge könnten insofern zu einer Verbesserung des Leistungsprofils beim Endspurt zum Abitur bei „leistungsschwächeren“ Studienberechtigten führen und/oder Angleichung bei „Leistungsstarken“.

(3) Idealistische Fachwünsche könnten für Stabilität/Veränderung von Fachwahlplänen – in Kombination mit dem Informationstreatment – von Bedeutung sein. Findet bei einer Veränderung der Fachwahlpläne ein Rückgriff auf diese Fachwünsche statt? Und geschieht dies insbesondere, wenn Studienberechtigten (insbesondere aus Nicht-Akademikerfamilien) vermittelt wird, dass ihr gewünschtes Studienfach zu den ertragreicheren gehört?

Donnerstag
12. März

9:00 Uhr

HGB 50

Chair(s)

Daniel, Annabell
Watermann, Rainer

Soziale Ungleichheit beim Hochschulzugang und Möglichkeiten der Intervention

Quast, Heiko; Lörz, Markus

Warum bestehen am Übergang vom Bachelor- ins Masterstudium herkunftsspezifische Ungleichheiten?

Im Zuge des Bologna-Prozesses wurden in Deutschland die traditionellen Diplom- und Magisterabschlüsse auf ein dreijähriges Bachelor- und ein darauf aufbauendes ein- bis zweijähriges Masterprogramm umgestellt. Mit dieser Studienstrukturreform wurde eine zusätzliche Selektionsschwelle eingeführt und zwei Hochschulabschlüsse etabliert, die mit unterschiedlichen Berufs- und Karrierechancen auf dem Arbeitsmarkt einhergehen. Obwohl mit der neuen Studienstruktur soziale Ungleichheiten abgebaut werden sollten, zeigen aktuelle Studien, dass insbesondere Bachelorabsolventen aus weniger privilegierten Elternhäusern oftmals auf ein Masterstudium verzichten. Aus theoretischer Perspektive lassen sich verschiedene Erklärungsansätze für dieses junge Phänomen skizzieren, jedoch finden sich bislang nur wenig empirisch abgesicherte Erkenntnisse für die in dieser Übergangsphase stattfindenden Entscheidungsprozesse. Der vorliegende Beitrag knüpft an diese Ausgangssituation an und beschäftigt sich sowohl in theoretischer als auch empirischer Hinsicht mit der Frage, welche Mechanismen dazu führen, dass Bachelorabsolventen aus weniger privilegierten Familien seltener ein Masterstudium aufnehmen.

Aus Perspektive einer rationalen Entscheidungstheorie (Breen & Goldthorpe 1997), vor dem Hintergrund kultureller Reproduktionsprozesse (Bourdieu & Passeron 1977) und einer Lebensverlaufperspektive (Hillmert & Jacob 2010) werden zunächst vier Mechanismen skizziert, die möglicherweise den sozialen Unterschieden am Bachelor-/Masterübergang zugrunde liegen. Diese theoretischen Überlegungen werden anhand eines repräsentativen Studienberechtigtenpanels hinsichtlich ihrer empirischen Evidenz betrachtet und erste Zusammenhänge aufgezeigt. Auf Basis logistischer Regressionsanalysen und der KHB-Dekompositionsmethode lässt sich zeigen, dass die sozialen Unterschiede am Bachelor-/Masterübergang vorwiegend auf vorgelagerte Bildungsentscheidungen sowie leistungsbezogene Unterschiede und eine höhere Kostensensibilität der weniger privilegierten Gruppen zurückzuführen sind.

Donnerstag
12. März

9:00 Uhr

HGB 50

Chair(s)

Daniel, Annabell
Watermann, Rainer**Soziale Ungleichheit beim Hochschulzugang und Möglichkeiten der Intervention**

Schwab, Susanne

Soziale Netzwerke in heterogenen (Klassen-)Settings

Im Zuge aktueller Schulentwicklung gewinnt der Umgang mit Heterogenität immer mehr an Bedeutung, da die individuellen Potentiale aller Lernenden genutzt werden sollen. Ziel ist es, mit Heterogenität wertschätzend umzugehen. In diesem Sinne werden bspw. bewusst Klassenzusammensetzungen mit Kindern verschiedener Muttersprachen initiiert, oder SchülerInnen mit sonderpädagogischem Förderbedarf (SPF) werden – immer häufiger – in Regelschulklassen integriert. Zielsetzung ist u.a., dass auch SchülerInnen lernen, die Heterogenität ihrer Peers wertzuschätzen. Demzufolge gewinnen kooperative Lernformen und soziale Netzwerke immer stärker an Bedeutung. Im Rahmen des Symposiums ist das Ziel, soziale und kooperative Netzwerke aus interdisziplinärer Sichtweise zu analysieren. Dies geschieht zudem mittels unterschiedlicher methodischer Ansätze. Die Frage der Wertschätzung der Heterogenität wird auf SchülerInnen-Ebene beantwortet, indem untersucht wird, inwieweit soziale Netzwerke in heterogenen SchülerInnenschaften funktionieren.

Im Beitrag von Rauch und Ziener werden regionale bildungsbezogene Netzwerke auf Bundesebene in Österreich in Bezug auf ihre Ziel- und Schwerpunktsetzungen evaluiert. Im Fokus steht dabei die Frage, inwieweit sich Netzwerke, die zur innovativen Weiterentwicklung des Unterrichts geschaffen wurden, entwickeln und inwiefern diese eine Informationsfunktion, Lernfunktion, politische Funktion und psychologische Funktion erfüllen. Methodisch werden sowohl Ergebnisse einer Cross-Case-Analyse der Jahresberichte regionaler Netzwerke als auch eine Interviewerhebung präsentiert.

Nach diesem einleitenden Beitrag auf Systemebene werden Helm und Grabmer auf Klassenebene Lern-Netzwerke von SekundarstufenschülerInnen berufsbildender Schulen aus Österreich analysieren. Zentrale Fragestellung ist, ob die Lern-Netzwerke der SchülerInnen von Schulklassen, welche laut Schulprofil verstärkt kooperatives, offenes Lernen durchführen, tatsächlich stärker ausgeprägt sind, als jene von SchülerInnen aus traditionell unterrichteten Kontrollklassen. Auch wird analysiert, inwiefern schülerbezogene Netzwerk-Charakteristika mit den fachlichen Leistungen und den sozialen Kompetenzen der SchülerInnen einhergehen.

Jörg Nicht berichtet im Anschluss über Heterogenität aus Sicht von SchülerInnen bilingual-binationaler Schulklassen. Dabei werden zwei Schulprojekte im Hinblick auf den Zusammenhang zwischen Sprachlernen und sozialen Beziehungen untersucht. Methodisch wird quantitativ auf die Rekonstruktion von Beziehungsstrukturen und Netzwerken und qualitativ auf Gruppendiskussionen zurückgegriffen.

Im Rahmen des Beitrags von Susanne Schwab werden abschließend soziale Beziehungen in Integrationsklassen untersucht, wobei der Fokus auf die Frage der sozialen Integration von SchülerInnen mit SPF gelegt wird. Dazu wurden die SchülerInnen mittels Fragebogen zu ihren Netzwerken befragt und ihre Selbstsicht der Sicht der Peers gegenübergestellt.

Tobias C. Stubbe bündelt abschließend alle vier Beiträge und wird sowohl methoden- als auch inhaltlich auf die Ergebnisse zu den sozialen Netzwerken in den heterogenen (Klassen-)Settings eingehen.

Donnerstag
12. März

9:00 Uhr

HGC 20

Chair(s)

Schwab, Susanne

3

Rauch, Franz; Ziener, Karen

Regionale Netzwerke zur Förderung von MINDT-Unterricht in Österreich: Evaluation bundesländerspezifischer Ziel- und Schwerpunktsetzungen

Ausgangslage: Die Anfänge des Projektes IMST – Innovationen Machen Schulen Top! – reichen bis Ende der 1990er Jahre zurück. Ziel ist eine systematische und nachhaltige Stärkung der Innovationskultur im MINDT-Unterricht (Mathematik, Informatik, Naturwissenschaften, Deutsch, Technik). Die Regionalen Netzwerke (Bundesländernetzwerke) sind eine von mehreren IMST-Maßnahmen (www.imst.ac.at) und haben in allen österreichischen Bundesländern sukzessive zu neuen Kooperationen zwischen LehrerInnen verschiedener Fächer und Schultypen, Schulbehörde, Aus- und Fortbildung und Forschung geführt, die in einem ständigen Prozess qualitativ und quantitativ weiterentwickelt werden. Im Rahmen von IMST wurde außerdem die Einrichtung von 18 Regionalen Fachdidaktikzentren (RFDZ) in 7 Bundesländern gefördert und für deren Weiterentwicklung das Qualitätslabel „Regional Educational Competence Centre (RECC)“ geschaffen. Zu den Aufgaben der Regionalen Netzwerke gehören die Koordination durch eine Steuergruppe, die Einrichtung einer Ansprechstelle für Schulen und LehrerInnen, die Organisation von Erfahrungsaustausch und Weiterbildung und eine regelmäßige Evaluation. Das genaue Aufgabenprofil eines Regionalen Netzwerks orientiert sich an den Bedürfnissen der Schulen im Bundesland und den bestehenden Ressourcen.

Theoretischer Hintergrund: Im Bildungsbereich sind Netzwerke in den letzten Jahren attraktiv geworden und fungieren als intermediäre Strukturen (Czerwanski, Hameyer, & Rolff 2002) die Kompetenzen bündeln (OECD 2003). Zu den theoretischen Ansätzen bildungsbezogener Netzwerke gehören unter anderem: gemeinsame Intention und Ziele, die Orientierung an einem von allen getragenen Rahmenthema und Zielhorizonten (Liebermann & Wood 2003), gegenseitiges Vertrauen als Voraussetzung, um Wissen auszutauschen und zu teilen und innovative Wege zu gehen (McLaughlin et al. 2008). Es geht in Bildungsnetzwerken nicht um gelegentliche Interaktionsbeziehungen sondern auch um institutionalisierte Kooperationen. Nach Dalin (1999) haben Netzwerke im Bildungsbereich vier Funktionen: Informationsfunktion, Lernfunktion, Politische Funktion und Psychologische Funktion.

Fragestellung: Wie wurden die Ziele von IMST zur Weiterentwicklung des Unterrichts in den Bundesländern angesichts der spezifischen Bedingungen und Ressourcen konkretisiert und umgesetzt? Wie bilden sich die Funktionen nach Dalin (1990) ab? Wie dynamisch entwickeln sich die Netzwerke?

Methode: Aufbauend auf der Analyse früherer Evaluationen zu den Regionalen Netzwerken (Erlacher 2006) und den Regionalen Netzwerken in Salzburg (Erlacher 2009) und der Steiermark (Strametz 2009) erfolgen eine Cross-Case-Analyse der Jahresberichte der Regionalen Netzwerke 2011-2014 und eine Interviewerhebung 2014-2015 unter Mitgliedern der Steuergruppen und LehrerInnen in den neun Bundesländernetzwerken. Es werden ca. 25 Leitfadeninterviews geführt und inhaltsanalytisch ausgewertet (Mayring 2003).

Ergebnisse: Netzwerke können als strukturelle Versuche verstanden werden, in den Widerspruchs- und Spannungsfeldern Autonomie und Vernetzung, Struktur und Prozess, Vorgabe und Freiraum, Freiwilligkeit und Verbindlichkeit, Differenz und Partizipation, Verwaltungshierarchie und Basisbewegung, Praxis und Wissenschaft neue Wege in der Gestaltung des Lernens und der Zusammenarbeit von Personen und Institutionen zu betreten. Langfristig besteht die Herausforderung darin, eine dynamische, „flüssige“ Balance von Innovation und Kontinuität aufrechtzuerhalten (Rauch 2013).

Vor dem Hintergrund dieses allgemeinen Befundes werden die Ergebnisse bezogen auf die Fragestellung unter Einarbeitung der aktuellen Interviewdaten in Bundesländerfallstudien (Yin 2003) dargestellt und in Thesen verdichtet zusammengefasst. Die Fallstudien und Thesen bieten sowohl Hinweise für praktische Interventionen in den Netzwerken und führen darüber hinaus zu Hypothesen für weiterführende Forschung.

Donnerstag
12. März

9:00 Uhr

HGC 20

Chair(s)

Schwab, Susanne

Helm, Christoph; Grabmer, Susanne

Cooperatives Offenes Lernen – Eine Analyse von Lern-Netzwerken in traditionellen und reformpädagogischen Klassen sowie ihrer Zusammenhänge mit fachlichen und sozialen Schülerkompetenzen.

Die Förderung sozialer und kooperativer Kompetenzen sind unbestrittene Unterrichtsziele. Als didaktisch-methodisches Mittel zur Erreichung dieser eignen sich in besonderer Weise kooperative Lernformen, da sie Schüler/innen Lerngelegenheiten (bspw. Aufeinandertreffen verschiedener Sichtweisen) bieten, die zur Entwicklung kooperativer, sozialer, moralischer und intellektueller Skills beitragen (Dewey 1900; Piaget 1932/1965; Vygotsky 1978; Cohen 1994).

Um Schüler/innen verstärkt solche Lerngelegenheiten zu ermöglichen, wurde 1996 das Modell des COoperativen Offenen Lernens (COOL; Neuhauser & Wittwer 2008) in den berufsbildenden Vollzeitschulen Österreichs eingeführt. Dieses Modell beruht auf den Ideen von John Dewey und steht damit im Einklang reformpädagogischer, schülerzentrierter Konzepte. Inwiefern es diesem Konzept gelingt soziale Kompetenzen zu fördern, ist Gegenstand dieses Beitrags.

Theoretischer Hintergrund: In der Literatur herrschen zwei zentrale theoretische Ansätze zum kooperativen Lernen vor, die die Wirkungen dieser Lernform zu erklären versuchen: die soziokonstruktivistische Perspektive (Piaget 1985) und die soziokulturelle Perspektive (Vygotsky 1986). Sie gehen von der Annahme aus, dass soziale Interaktionen zwischen Schüler/innen beim gemeinsamen Lernen zu sogenannten soziokognitiven Konflikten (z.B. durch das Aufeinandertreffen verschiedener Sichtweisen) führen. Die erfolgreiche Auflösung dieser Konflikte durch bspw. wechselseitiges Geben von Erklärungen (Webb & Farvar 1999) führt zur (Ko-)Konstruktion neuen Wissens bzw. Modifikation bestehender Schemata. Die soziokulturelle Perspektive nimmt zudem an, dass kooperatives Lernen vor allem unter Anleitung eines/r kompetenteren Schülers/in (Zone der proximalen Entwicklung) passiert.

Auf den ersten Blick verweisen diese Ansätze vor allem auf die Potentiale kooperativen Lernens in Hinblick auf die Erreichung kognitiver Ziele. Allerdings sind diese allgemeinen Lerntheorien auch auf kooperative, soziale Lernziele anwendbar (Euler 2007).

Fragestellung: Im Zentrum der Analysen steht die Frage, inwiefern das COOL-Konzept zur Förderung kooperativer und sozialer Kompetenzen beiträgt. Konkret werden folgenden Hypothesen geprüft: (1) COOL-Klassen verfügen im Vergleich zu den traditionell unterrichteten Kontrollklassen (an derselben Schule) über ein Lern-Netzwerk, das signifikant stärker ausgeprägt ist (z.B. Density, Mean-Degree). Zusätzlich werden Dyadic und Triadic Effects explorativ analysiert. (2) Das Ausmaß an wahrgenommenen (Rating-Skala) und „tatsächlichem“ kooperativem Lernen (z.B. In- and Outdegree, Reciprocity) hängt signifikant positiv mit den Schülerleistungen im Fach Rechnungswesen sowie den allgemeinen Sozialkompetenzen als auch dem un-terrichtsbezogenen pro-sozialen Verhalten der Schüler/innen zusammen.

Methode: Stichprobe. 720 Schüler/innen aus 24 Klassen der 10. Schulstufe in 7 berufsbildenden mittleren und höheren Schulen in Österreich wurden mit einem Onlinefragebogen befragt sowie mit einem Papier-Bleistifttest getestet.

Messinstrumente: Zur Erfassung des Lern-Netzwerks wurden Klassenlisten im Onlinefragebogen vorgelegt und folgende „Indikatorfrage“ gestellt: „Kreuze an, wie oft du mit deinen Klassenkollegen gemeinsam Arbeitsaufträge/(Lern-)Aufgaben im RW-Unterricht bearbeitest!“ Facetten der selbsteingeschätzten sozialen Kompetenz wurden mit dem Verfahren Interpersonal Competence Questionnaire (Riemann & Allgöwer 1993) sowie Skalen zum pro-sozialen Verhalten des DIPF (DaQS o.d.) erfasst. Der wahrgenommene Einsatz kooperativer Lernformen wurde mit selbsterstellten Items erfasst. Die fachliche Schülerleistung in Rechnungswesen wurde mit dem IRT-basierten Test „Wissensüberprüfung von Basiskonntnissen der Buchhaltung“ gemessen (Helm 2014).

Analyseverfahren: Zur Prüfung der Hypothese 1 werden Netzwerkanalysen mit R-Programmen (Skvoretz 2014) für jedes Netzwerk einer Schulklasse berechnet. Diese Programme erlauben Signifikanztest vor dem Hintergrund der Biased Net Theory. Die Ergebnisse dieser 24 Netzwerkanalysen werden sodann mittels meta-analytischen Ansatzes (Snijders & Baerveldt, 2003) zusammengefasst, um die Unterschiede zwischen COOL- und traditionellen Klassen darzustellen. Zur Prüfung von Hypothese 2 werden Ergebnisse der Netzwerkanalysen mit den Ergebnissen der Onlinebefragung und des Rechnungswesen-Kompetenztests korreliert, wobei für die hierarchische Struktur der Daten kontrolliert wird (Type Complex in Mplus).

Ergebnisse: Bisher wurden zwei Klassen analysiert, wobei die COOL-Klasse zwar eine signifikant höhere Density aber eine gleiche Mean-Degree und Reciprocity aufweist. In den beiden Klassen stehen die Degree-Werte der Schüler/innen im moderaten, signifikanten Zusammenhang mit den fachlichen Rechnungswesen-Leistungen und dem sozialen Engagement. Die Ergebnisse werden vor dem Hintergrund qualitativer Studien zum kooperativen, offenen Lernen im COOL-Konzept diskutiert. Diese geben Hinweise auf die Qualität der Umsetzung des kooperativen Lernens vor dem Hintergrund des Forschungsstandes.

Chair(s)
Schwab, Susanne

Soziale Netzwerke in heterogenen (Klassen-)Settings

Nicht, Jörg

Heterogenität aus Sicht von Schüler_innen bilingual-binationaler Schulklassen. Ergebnisse einer Netzwerkanalyse

Binational-bilinguale Schulprojekte stellen einen schulischen Sonderfall dar: Sprachliche Heterogenität ist nicht nur erwünscht, sondern wird absichtlich erzeugt, indem die Zusammensetzung der Schulklassen bewusst heterogen gestaltet wird. Zwei solcher Projekte stehen im Zentrum des Beitrags: Am Augustum-Annen-Gymnasium in Görlitz lernen deutsche und polnische Schüler_innen gemeinsam, am Friedrich-Schiller-Gymnasium in Pirna deutsche und tschechische Schüler_innen. Beide Projekte achten darauf, dass die Schüler der jeweils zwei Nationen zu etwa gleichen Teilen in den Schulklassen vertreten sind. Zentrales Strukturelement sind die heterogenen Klassen, die entsprechend Zwei-Wege-Modellen zu effektivem Sprachlernen führen sollen. Damit versuchen solche Schulprojekte, kognitive und soziale Lernprozesse unter der Bedingung von Heterogenität auf spezifische Weise zu verknüpfen. Ob und, wenn ja, wie es den Projekten gelingt, Sprachlernen und soziale Beziehungen der Schüler_innen aufeinander zu beziehen, ist eine in der Schulforschung bislang nicht geklärte Frage.

Theoretischer Hintergrund: Die Frage, was unter schulischen Peer-Beziehungen zu verstehen ist, wird mithilfe eines relationalen, beziehungs- und netzwerktheoretischen Ansatzes präzisiert und operationalisiert. Eine relationale Forschungsperspektive (vgl. White 2008; Nicht 2013) erweist sich in zweifacher Hinsicht als fruchtbar: Sie ermöglicht es einerseits, die Eigenheit und Emergenz sozialer Beziehungen zu berücksichtigen und ihre Relevanz für die Strukturierung des schulischen Alltags zu untersuchen. Sie ermöglicht es andererseits, das Innovative und Neuartige der beiden binational-bilingualen Schulprojekte in den Blick zu nehmen, wobei individuelle und netzwerktheoretische Perspektive miteinander verknüpft werden können.

Untersuchungsfrage: Im Mittelpunkt des Beitrags stehen zwei Fragen: (1) Welche sozialen Beziehungen können zwischen den Schüler_innen in den beiden binational-bilingualen Schulprojekten rekonstruiert werden – und zwar in formaler und in inhaltlicher Hinsicht? (2) Wie bewerten die Schüler_innen diese Beziehungen?

Untersuchungsdesign und methodisches Vorgehen: Aus den theoretischen Überlegungen ergeben sich zwei Konsequenzen für das Design der Untersuchung: Erstens werden die Beziehungen zwischen den Schüler_innen mit netzwerkanalytischen Methoden untersucht. Die Grundlage für die Rekonstruktion von Beziehungsstrukturen und Netzwerken bildet eine schriftliche Befragung, d. h. es werden Fragebögen eingesetzt, in denen die Schüler_innen ihre schulischen Peer-Beziehungen auf verschiedenen Dimensionen einschätzen. Aus den Ergebnissen der quantitativen Netzwerkanalyse leitet sich zweitens die Fallauswahl für den qualitativen Untersuchungsteil – die Gruppendiskussionen – ab. Die Kombination von schriftlicher Befragung und Gruppendiskussionen ermöglicht eine vertiefte Rekonstruktion von Beziehungsstrukturen und Netzwerken in ausgewählten Klassen der beiden untersuchten Schulprojekte.

Ergebnisse: In Bezug auf die verschiedenen Formen schulischer Peer-Beziehungen sind fünf Befunde festzuhalten. Erstens hat die sprachlich-nationale Herkunft allein keine Bedeutung dafür, wie viele bzw. in welcher Intensität die Schüler_innen Beziehungen eingehen. Zweitens lässt sich für alle hier untersuchten Formen sozialer Beziehungen zeigen, dass sie eher zu Schüler_innen der eigenen sprachlich-nationalen Herkunft eingegangen werden als zu den Schüler_innen aus der jeweiligen Partnation. Drittens unterscheiden sich die Schüler_innen deutscher und nicht-deutscher Herkunft hinsichtlich der national-monolingualen Beziehungen bzw. der binational-bilingualen Beziehungen, wenn man die Größe der eigenen Gruppe in der Schulklasse (bezogen auf das Geschlecht und die sprachlich-nationale Herkunft) berücksichtigt. Damit erweist sich die Komposition der Schulklasse als bedeutsam für die schulischen Peer-Beziehungen: Je nach eigener Gruppengröße sinken oder steigen die Chancen für binational-bilinguale Beziehungen. Allerdings lassen sich auch abweichende Muster finden. Viertens lässt sich die Bedeutung institutioneller Faktoren (hier: gemeinsame Unterrichtszeit) für die Aufnahme von Beziehungen nur bedingt nachweisen. Fünftens muss festgehalten werden, dass die relationalen Daten zeigen, dass sich für einen Teil der Schüler_innen (bei Freundschaften: 56,5 Prozent) keine direkten bilingual-binationale Beziehungen rekonstruieren lassen.

In zwei Fallstudien werden diese Ergebnisse vertieft und ergänzt. Zeigen lässt sich, dass die Diskutierenden aus dem deutsch-polnischen Projekt tendenziell ein universalistisch-identifikatorisches Beziehungsmuster artikulieren; von Schüler_innen deutscher Herkunft in Pirna wird ein individualistisch-leistungsorientiertes Beziehungsmuster zum Ausdruck gebracht; das Beziehungsmuster, das für die Schüler_innen tschechischer Herkunft rekonstruiert wird, kann als gemeinschaftlich-kulturalistisch bezeichnet werden. Diese Beziehungsmuster stehen in einem engen Zusammenhang mit der Frage, wie die Schüler_innen Heterogenität bewerten.

Donnerstag
12. März

9:00 Uhr

HGC 20

Chair(s)

Schwab, Susanne

Schwab, Susanne

Soziale Netzwerke in Integrationsklassen – sind SchülerInnen mit SPF integriert?

Hochselektive Schulsysteme beginnen, im Zuge der Inklusionsdebatte, SchülerInnen mit sonderpädagogischem Förderbedarf (SPF) in Regelklassen zu integrieren. Eine wichtige, dahinterstehende Zielsetzung ist, dass SchülerInnen mit SPF eine größtmögliche Teilhabe in der Gesellschaft gewährt wird, da bei einer Sonderbeschulung auch die soziale Ausgrenzung als problematisch gesehen wird. Soziale Integration, soziale Inklusion und soziale Partizipation sind dabei immer wieder Schlagwörter, wobei diese im Hinblick auf die Forschung auf vier Subthemen fokussieren: Interaktionen zwischen Peers, wechselseitige Freundschaften, soziale Akzeptanz und die Selbstsicht der eigenen sozialen Teilhabe (siehe dazu Koster et al., 2009). Aktuelle Befunde aus der Literatur lassen derzeit darauf schließen, dass sich die Situation der sozialen Partizipation in Integrationsklassen für SchülerInnen mit SPF eher schwierig gestaltet und dass Letzere in Bezug auf das soziale Peer-Netzwerk einem deutlichen Ausgrenzungsrisiko unterliegen. Frostad und Pijl (2007) berichten darüber, dass SchülerInnen mit SPF weniger Freunde haben als ihre Peers. Pijl et al. (2011) zeigen ein ähnliches Bild im Bezug auf Interaktionen. Die Ergebnisse von Huber und Wilbert (2012) sprechen dafür, dass SchülerInnen mit SPF im Vergleich zu den MitschülerInnen ohne SPF im Hinblick auf die Akzeptanz in der Peer-Group schlechter abschneiden und Schwab (in press) berichtet über ein stärkeres Einsamkeitsempfinden. Die vorliegende Studie hat das Ziel, die Peer-Netzwerke von SchülerInnen mit SPF näher zu analysieren und dabei deren eigene Perspektive und jene der Peers näher zu analysieren. Dazu werden 2 Hypothesen aufgestellt: (1) SchülerInnen mit SPF fühlen sich im Vergleich zu ihren Peers weniger stark in die sozialen Netzwerke der Klasse integriert. (2) SchülerInnen mit SPF werden im Vergleich zu ihren Peers weniger stark in die sozialen Netzwerke der Klasse integriert.

Da grundsätzlich davon ausgegangen wird, dass die soziale Integration stark von klassenspezifischen Aspekten abhängt, wird zusätzlich untersucht, inwieweit die soziale Integration zwischen den einzelnen Schulklassen variiert.

Methode: Insgesamt werden die Angaben von 601 SchülerInnen (35% SchülerInnen der 4. Schulstufe, 65% SchülerInnen der 7. Schulstufe) analysiert. Von diesen hatten 126 SchülerInnen einen SPF-Bescheid vom Bezirksschulrat zugewiesen bekommen. Mit der Frage „Wie oft verbringst du deine Pause mit folgendem/r MitschülerIn?“ wurde ein soziometrisches Rating durchgeführt. SchülerInnen bewerteten somit die Interaktionen mit allen MitschülerInnen jeweils auf einer fünfstufigen Ratingskala. Berechnet werden im Bezug auf die Interaktionen sowohl die Selbstsicht als auch die Fremdsicht. Zudem wurden die SchülerInnen gebeten, ihre (maximal) fünf besten FreundInnen in der Klasse zu nennen, wobei die Anzahl der wechselseitigen Freundschaftsnennungen als Indikator für die soziale Teilhabe verwendet wird.

Ergebnisse: Während lediglich 2.1% jener SchülerInnen ohne SPF nicht eine einzige wechselseitige Freundschaftsbeziehung in ihrer Klasse haben, ist dieser Anteil bei SchülerInnen mit SPF mit 14.7% deutlich höher. Anhand einer Multi-level-Analyse zeigte sich, dass die Varianzaufklärung für wechselseitige Freundschaften keine signifikanten Varianzanteile auf Klassenebene aufweist. Der SPF-Status ist auf der Ebene der SchülerInnen hingegen ein starker Prädiktor ($\beta = -.30$, $p < .001$). Bei der Selbsteinschätzung der Interaktionen mit den Peers wird 17% der Varianz auf Klassenebene und 83% auf der Individualebene aufgeklärt, wobei SPF keinen signifikanten Prädiktor darstellt. In Bezug auf die Peerangaben zu den Interaktionen liegt der Anteil der Varianzaufklärung auf der individuellen Ebene bei 71%, wobei SPF signifikant zur Varianzaufklärung beiträgt ($\beta = -.32$, $p < .001$). Auf der Klassenebene werden 29% Varianz aufgeklärt. Zusammenfassend kann gesagt werden, dass SchülerInnen mit SPF weniger stark in die sozialen Netzwerke in der Klasse integriert sind. Dies zeigt sich jedoch nicht für die Selbsteinschätzung der SchülerInnen mit SPF. Zudem liefert die Studie Hinweise darauf, dass in den einzelnen Klassen unterschiedliche soziale Integration vorherrscht. Für die Präsentation werden zudem die sozialen Netzwerke der Klassen grafisch gegenübergestellt.

Donnerstag
12. März

9:00 Uhr

HGC 20

Chair(s)

Schwab, Susanne

Neumann, Marko

Abitur und Matura im Wandel: Schulleistungen und Studienerfolg von Abiturienten im Kontext veränderter schulorganisatorischer und curricularer Rahmenbedingungen

Die schulorganisatorischen und curricularen Rahmenbedingungen auf dem Weg zum Abitur haben in den letzten Jahren wesentliche Veränderungen erfahren. In Politik und Öffentlichkeit wird dabei gegenwärtig vor allem die „richtige“ Dauer der Schulzeit zum Abitur diskutiert. Die nahezu flächendeckend erfolgte Umstellung auf das achtjährige Gymnasium hat infolge vielfach zu vernehmender Kritik inzwischen in mehreren Bundesländern verschiedene Modifikationen erfahren, die zum Teil auf eine Rückkehr zum neunjährigen Gymnasium hinauslaufen. Aus Sicht der empirischen Bildungsforschung ist dabei zu konstatieren, dass sowohl die Umstellung auf das achtjährige Gymnasium als auch die gegenwärtig teilweise zu beobachtende Rück-Reform nur auf einer schmalen empirischen Befundlage basierten bzw. basieren. Auch mit Blick auf die inhaltlich-organisatorische Ausgestaltung der gymnasialen Oberstufe kam es in den letzten Jahren zu deutlichen Veränderungen, was sich unter anderem in der in vielen Bundesländern erfolgten Aufhebung der Differenzierung in Grund- und Leistungskurse in den Kernfächern Deutsch, Mathematik und Fremdsprache (zum Teil umgesetzt über so genannte „Profiloberstufen“), der Einführung nationaler Bildungsstandards für das Abitur sowie zentraler Abiturprüfungen manifestiert (vgl. im Überblick Neumann, 2014). Die Tendenz der Ausweitung des verpflichtenden Fächerkanons zulasten vorhandener Wahlmöglichkeiten lässt sich in ähnlicher Weise auch in der Schweiz beobachten.

Die Beiträge des Symposiums greifen unterschiedliche Aspekte dieser vielfältigen Veränderungsprozesse auf. Nele Kampa, Michael Leucht und Olaf Köller gehen in ihrem Beitrag auf der Datengrundlage der LISA-Studie der Frage nach, inwieweit nach Einführung der Profileroberstufe in Schleswig-Holstein in den verschiedenen Oberstufenprofilen von einem hinreichenden voruniversitären Mathematikleistungsniveau der Abiturienten ausgegangen werden kann. Stanislav Ivanov, Roumiana Nikolova und Ulrich Vieluf untersuchen in ihrem Beitrag anhand der LAU- und KESS-Daten die Auswirkungen der Umstellung auf das achtjährige Gymnasium in Hamburg auf die Fachleistungen der Gymnasiasten am Ende der gymnasialen Oberstufe. Die G8-Reform steht auch im Zentrum des Beitrags von Wolfgang Wagner, Jochen Kramer, Nicolas Hübner und Ulrich Trautwein, die die Auswirkungen der Schulzeitverkürzung auf die Fachleistungen der Abiturienten in Baden-Württemberg auf der Datengrundlage einer NEPS-Zusatzstudie untersuchen und dabei auf die Leistungsdaten von drei aufeinanderfolgenden Abiturjahrgängen zurückgreifen können. Der Beitrag von Maren Oepke und Franz Eberle untersucht die Frage, inwieweit sich durch das voruniversitäre Fachleistungsniveau in den Fächern Deutsch und Mathematik als Indikator von Studierfähigkeit der spätere Studienerfolg Schweizer Maturanden vorhersagen lässt. Die Analysen basieren auf der EVAMAR II-Studie (vgl. Eberle et al., 2008) und einer späteren Follow-Up-Erhebung. Die vier Beiträge werden von Svenja Mareike Kühn diskutiert.

Donnerstag
12. März

9:00 Uhr

HGC 40

Kampa, Nele; Leucht, Michael; Köller, Olaf

Mathematische Kompetenzen in unterschiedlichen Profilen der gymnasialen Oberstufe

Das Fach Mathematik leistet einen grundlegenden Beitrag zu den Bildungszielen der gymnasialen Oberstufe und der Kompetenzentwicklung bis zur Allgemeinen Hochschulreife. Durch die Veröffentlichung der Oberstufenbefunde der Third International Mathematics and Science Study (TIMSS; Baumert, Bos & Lehmann, 2000) wurde für Mathematik auf Kompetenzdefizite hingewiesen. Weitere Studien untermauerten diese Befunde für die Sekundarstufe II (z. B. Transformation des Sekundarschulsystems und akademische Karrieren [TOSCA]; Köller, Watermann, Trautwein & Lüdtke, 2004). Zeitgleich zu TIMSS wurden Reformen in der gymnasialen Oberstufe zur Sicherung fachlicher Standards in den Kernfächern Deutsch, Mathematik und einer Fremdsprache angestoßen. Durch die folgenden Mainzer und Husumer Beschlüsse ([Kultusministerkonferenz] KMK, 1995; KMK, 1999) wurde festgehalten, dass die Fächer Deutsch, Mathematik und eine Fremdsprache in den letzten beiden Schuljahren der Sekundarstufe II bindend zu belegen sind. Nachfolgend haben viele Bundesländer das Kurssystem zugunsten von Profiloberstufen aufgegeben. Schülerinnen und Schüler wählen inhaltliche Schwerpunkte (Profile), müssen allerdings die Kernfächer Deutsch, Mathematik und eine Fremdsprache auf erhöhtem Anforderungsniveau belegen. In Schleswig-Holstein müssen z. B. die Schülerinnen und Schüler aller Profile mit vier Stunden Mathematik unterrichtet werden. Jedoch zeigt die Literatur, dass Schwerpunktsetzungen in der gymnasialen Oberstufe eng an die am Ende der Sekundarstufe I erreichten Leistungsstände gekoppelt sind (z. B. Köller, Daniels, Schnabel & Baumert, 2000).

Wir nutzen diese strukturellen Veränderungen, um zu klären, ob durch die verpflichtende Belegung des Faches Mathematik auf erhöhtem Anforderungsniveau substanzielle Anteile der Jugendlichen profilunabhängig (naturwissenschaftlich, sprachlich, gesellschaftswissenschaftlich, ästhetisch oder sportlich) die Ziele voruniversitärer mathematischer Bildung erreichen. Wir erwarten (a) signifikant höhere kognitive Grundfähigkeiten und Mathematikleistungen von Schülerinnen und Schülern im naturwissenschaftlichen Profil. Dementsprechend erwarten wir, dass (b) in dieser Gruppe der Anteil der Jugendlichen höher ist, der die Ziele des voruniversitären Mathematikunterrichts erreicht. Auch gehen wir davon aus der Frage nach, dass (c) Leistungsunterschiede zwischen Profilen durch die Eingangsmerkmale Geschlecht und kognitive Grundfähigkeiten vermittelt werden, also nach Kontrolle beider Merkmale diese Unterschiede abnehmen.

Um diesen Fragen nachzugehen, analysieren wir die Mathematikleistungen von N = 1.360 (17 Schulen) Abiturientinnen und Abiturienten an allgemeinbildenden Gymnasien in Schleswig-Holstein aus dem Jahr 2013. Die Schülerinnen und Schüler waren zum Zeitpunkt der Testung 19;6 Jahre alt. Der eingesetzte mathematical literacy Test für Studierende des Nationalen Bildungspanels (NEPS) besteht aus 20 Items und erlaubt die Definition von Kompetenzstufen, in der Erwartungen im Sinne von Regelstandards für die Allgemeine Hochschulreife abgebildet werden können. Die kognitiven Grundfähigkeiten wurden mit zwei Untertests des Kognitiven Fähigkeitstest (KFT 4-12+ - Revision; Heller & Perleth, 2000) erfasst. Die mit Mplus durchgeführten Analysen mit gewichteten Daten wurden mit fünf Plausible Values durchgeführt.

Die Analysen zeigen, dass die Verteilung auf die Profile bezüglich des Geschlechts und der kognitiven Grundfähigkeiten unterschiedlich ausfällt und somit gleiche Eingangsvoraussetzungen nicht gegeben sind. Insbesondere im naturwissenschaftlichen Profil handelt es sich um eine mit Blick auf die kognitiven Grundfähigkeiten positiv selektierte Stichprobe. Die Unterschiede in den mathematischen Fähigkeiten fallen im Vergleich zu den kognitiven Grundfähigkeiten noch prägnanter aus. Die Abstände des naturwissenschaftlichen Profils zu den übrigen Profilen entsprechen einem Leistungszuwachs von mehreren Schuljahren. Der größte Kompetenzunterschied liegt zwischen dem naturwissenschaftlichen (566 Kompetenzpunkte) und ästhetischen Profil (464 Kompetenzpunkte; $d = 1.13$).

Bezogen auf die Standarderreichung verfehlen 69 % die standardbasierten Ziele voruniversitärer mathematischer Bildung. Die Standarderreichung ist wiederum profilabhängig. So liegen 52 % der Schülerinnen und Schüler auf den Stufen III und IV und genügen damit standardbezogenen Erwartungen der gymnasialen Oberstufe; im ästhetischen Profil sind es nur 12 %. In multiplen Regressionsanalysen konnten wir zudem zeigen, dass die Profilverunterschiede durch die Eingangsvoraussetzungen vermittelt werden. Die Abstände zwischen den Profilen werden geringer. Der Abstand zwischen dem naturwissenschaftlichen (559 Kompetenzpunkte) und ästhetischen Profil (495 Kompetenzpunkte) sinkt, bleibt jedoch der größte. Die Ergebnisse werden vor dem Hintergrund der Bildungsstandards für die Abiturprüfung diskutiert.

Donnerstag
12. März

9:00 Uhr

HGC 40

Ivanov, Stanislav; Nikolova, Roumiana; Vieluf, Ulrich

G 8 und G 9 im LAU-KESS-Kohortenvergleich – eine Vergleichsanalyse zur Frage der Schulzeitverkürzung in Hamburg

Auch wenn vor allem demographische und gesellschaftspolitische Gründe für die Einführung des achtjährigen Gymnasiums entscheidend waren (vgl. hierzu Klemm 2013 und Kühn 2013), so steht die Umsetzung dieser Entscheidung in Zusammenhang mit verschiedenen weiteren qualitätssichernden schulischen Reformmaßnahmen. Die konkrete Umsetzung der Schulzeitverkürzung ist gleichwohl abhängig von den regionalen demografischen Strukturen und schulpädagogischen Konzepten, die innerhalb und zwischen den einzelnen Bundesländern zum Teil zu deutlich unterschiedlichen G-8-Modellen führten. In unserem Beitrag fokussieren wir die Situation in Hamburg und stellen die Besonderheiten sowie die spezifischen schulischen Kontextbedingungen dar, die vor und nach der Einführung vom G-8 in die Hamburger Schulstruktur kennzeichnend waren. Im Mittelpunkt steht dabei die vergleichende Analyse der Lernausgangslage zu Beginn der Mittelstufe (Jahrgangsstufe 7) und der Lernstände am Ende der gymnasialen Oberstufe zwischen den sogenannten LAU- und KESS-Jahrgängen. Die Grundlage des Vergleichs sind also zum einen die Leistungs- und Hintergrunddaten des Hamburger LAU-Jahrgangs, ein kompletter Schülerjahrgang, der unter den Bedingungen des G-9 im Schuljahr 2005 das Abitur erworben hat und dessen Lernstände u.a. in den Kompetenzdomänen Englisch, Mathematik und Naturwissenschaften im Rahmen der Schulleistungsuntersuchung „Aspekte der Lernausgangslage und der Lernentwicklung“ (LAU) von Jahrgangsstufe 7 bis zur Jahrgangsstufe 13 erfasst wurden. Dem LAU-Jahrgang werden die Leistungs- und Hintergrunddaten des sieben Jahre später untersuchten KESS-Jahrgangs gegenübergestellt, der vom Schuljahr 2004 bis Schuljahr 2012 unter den Bedingungen von G-8 zum Abitur geführt wurde. Im Rahmen der KESS-Studie („Kompetenzen und Einstellungen von Schülerinnen und Schülern“) wurden wie in LAU zuvor mit den gleichen bzw. leicht modifizierten Leistungstests die Lernstände in den Kompetenzdomänen Englisch, Mathematik und Naturwissenschaften von Beginn der Jahrgangsstufe 7 bis Ende der gymnasialen Oberstufe (Jahrgangsstufe 12) erfasst. Damit ist Hamburg das einzige Bundesland, das über längsschnittliche miteinander verknüpfbare Schülerleistungsdaten zweier vollständiger Schülerjahrgänge, eines G-9 und eines G-8 Jahrgangs, von Beginn der Mittelstufe bis zum Ende der gymnasialen Oberstufe verfügt, wodurch sich Entwicklungstrends im Kontext des achtjährigen Gymnasiums im Vergleich zu den Lernkontexten des früheren neunjährigen gymnasialen Bildungsgangs beschreiben und analysieren lassen.

Die zentrale Frage, die im Mittelpunkt der Analysen steht, ist, inwieweit unter Berücksichtigung der Lernausgangslagen und ausgewählter leistungsrelevanter soziodemografischer und -kultureller Hintergrundmerkmale generelle sowie differenzielle Effekte des G 8 am Ende der Oberstufe feststellbar sind. Damit greift der Beitrag ein wichtiges Forschungsdesiderat auf, zumal es zu den Auswirkungen der mittlerweile in fast allen Bundesländern eingeführten Schulzeitverkürzung bislang kaum empirische Evidenz gibt. Zur Beantwortung der Fragestellung wurden in einem ersten Schritt die Ergebnisse der Abiturientinnen und Abiturienten des KESS-Jahrgangs in den Tests Deutsch-Leseverständnis, Englisch-Sprachverständnis und Mathematische Grundbildung zu Beginn der Mittelstufe (KESS 7) in einem Verankerungsverfahren auf die Skalenmetrik der jeweiligen LAU-Skalen transformiert. Die drei für die beiden Jahrgänge gemeinsamen Skalen wurden zu einem latenten Faktor Allgemeiner Fachleistungsindex zu Beginn der Jahrgangsstufe 7 (AFI 7) zusammengefasst mit dem Ziel, die individuellen leistungsbezogenen Lernvoraussetzungen der Gymnasiastinnen und Gymnasiasten beider Schülerjahrgänge zu Beginn der Mittelstufe abzubilden. Mithilfe von Regressionsanalysen mit dem Programmpaket MPLUS wurde anschließend unter Berücksichtigung der AFI 7 sowie weiterer leistungsrelevanter Hintergrundmerkmale wie Geschlecht, Familiensprache, Geburtsland der Eltern, höchster Schulabschluss der Eltern und Buchbestand im Elternhaus der Effekt der Kohortenzugehörigkeit (LAU/KESS) für die am Ende der gymnasialen Oberstufe erreichten Lernstände untersucht. Die Befunde deuten darauf hin, dass die Einführung des achtjährigen Bildungsgangs zum Abitur in Hamburg nicht zu einer generellen Absenkung des Leistungsniveaus in den untersuchten Kompetenzbereichen geführt hat. Dies ist umso bemerkenswerter, als dass sowohl der Anteil von Schülerinnen und Schülern am Gymnasium als auch der Anteil an Schülerinnen und Schülern aus eher bildungsfernen Elternhäusern im Vergleich zum LAU-Jahrgang deutlich gestiegen ist. Allerdings zeigen sich sowohl domänenspezifische als auch differenzielle Effekte in Abhängigkeit der Lernausgangslage, auf die bereits Baumert und Watermann (2000) hingewiesen haben.

Donnerstag
12. März

9:00 Uhr

HGC 40

Chair(s)

Neumann, Marko
Kramer, Jochen
Trautwein, Ulrich

Abitur und Matura im Wandel: Schulleistungen und Studienerfolg von Abiturienten im Kontext veränderter schulorganisatorischer und curricularer Rahmenbedingungen

Wagner, Wolfgang; Kramer, Jochen; Hübner, Nicolas; Trautwein, Ulrich

Schulzeitverkürzung und Fachleistungen: Die G8-Reform in Baden-Württemberg

Theoretischer Hintergrund: Im letzten Jahrzehnt fand in vielen Bundesländern eine Schulzeitverkürzung am Gymnasium um ein Jahr bei gleichzeitiger Erhöhung der Schülerwochenstunden statt (Trautwein & Neumann, 2008), die als G8-Reform breite Aufmerksamkeit auf sich zog und nicht unumstritten war und ist. Dies zeigt sich beispielsweise daran, dass Niedersachsen mit dem Schuljahr 2015/2016 flächendeckend zum G9 zurückkehrt (KMK, 2014). Die Diskussion der G8-Reform wird allerdings auf einer überschaubaren empirischen Basis geführt. Der wissenschaftliche Forschungsstand zu den intendierten und nicht-intendierten Reformeffekten der Schulzeitverkürzung kann derzeit als unbefriedigend bezeichnet werden (vgl. Kühn, van Ackeren, Bellenberg, Reintjes & im Brahm, 2013).

Eine häufig geäußerte Kritik am achtjährigen Gymnasium bezieht sich auf erwartete negative Effekte im Hinblick auf die Leistungsentwicklung aufgrund der erhöhten Lernintensität (vgl. Kühn et al., 2013). Erste empirische Studien scheinen zumindest teilweise diese Befürchtungen zu belegen, wenngleich sich eher geringe Leistungsunterschiede zwischen G8- und G9-Schülern ergeben (Büttner & Thomsen, im Druck; Vieluf, Ivanov, & Nikolova, 2012).

Fragestellung: Der Beitrag behandelt die Frage, ob sich Leistungsunterschiede für Schülerinnen und Schüler aus G8- und G9-Jahrgängen in Baden-Württemberg im Hinblick auf die Domänen Mathematik, Englisch, Biologie und Physik ergeben. Diesbezüglich sollen auf Basis von Daten des Nationalen Bildungspanels (NEPS Zusatzstudie Baden-Württemberg; Blossfeld, Rossbach & Maurice, 2011; Wagner et al., 2011) drei Abschlussjahrgänge eingehend miteinander verglichen werden (Die Daten der ersten reinen G8-Kohorte aus dem Abschlussjahr 2013 sollen hier ebenfalls berücksichtigt werden [voraussichtliche Veröffentlichung: Dezember 2014]).

Methode: Bei den Kohorten handelt es sich um den letzten ausschließlichen G9-Abschlussjahrgang aus dem Jahr 2011, den aus G8- und G9-Schülerinnen und Schülern bestehenden Abschlussjahrgang („Doppeljahrgang“) aus dem Jahr 2012 (jeweils etwa 1200 Schülerinnen und Schüler aus insgesamt 48 Schulen), sowie den ersten ausschließlichen G8-Abschlussjahrgang aus dem Abschlussjahrgang 2013.

Ergebnisse: Erste Analysen auf Basis der Daten der ersten beiden Wellen verweisen auf eine akzeptable psychometrische Qualität der Leistungstests. Eine gemeinsame Skalierung über alle drei Wellen und entsprechende Differential Item Functioning-(DIF-)Analysen wird erfolgen, sobald die Daten aus Welle 3 verfügbar sind. Zuverlässige Aussagen über Leistungsunterschiede zwischen G8- und G9-Schülern sind nur auf der Datenbasis sämtlicher drei Erhebungswellen möglich, da die G8-Kohorte im gemeinsamen Abiturjahrgang einige Besonderheiten aufweist (u.a. deutlich niedrigere Klassenwiederholungsrate bei einer erhöhten Schulwechselrate für G9-Schüler, die offensichtlich den Wechsel in den G8-Zug vermeiden wollten; www.bildungsmonitoring.de, Indikator D13.1).

Die hierarchische Struktur der Daten sowie fehlende Werte werden bei den Analysen durch Verwendung von robusten Schätzverfahren sowie Full-Information-Maximum-Likelihood-(FIML-)Verfahren berücksichtigt. In den Analysen soll auch der Einfluss relevanter Kovariaten (z.B. Klassenwiederholung, Kurswahl, soziale Herkunft) auf potenzielle Leistungsunterschiede untersucht werden.

Anmerkung: Diese Arbeit nutzt Daten des Nationalen Bildungspanels (NEPS) Zusatzstudie Baden-Württemberg, doi:10.5157/NEPS:BW:2.0.0. Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde. Seit 2014 wird NEPS vom Leibniz-Institut für Bildungsverläufe e.V. (IfBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

Donnerstag
12. März

9:00 Uhr

HGC 40

Chair(s)

Neumann, Marko
Kramer, Jochen
Trautwein, Ulrich

Abitur und Matura im Wandel: Schulleistungen und Studienerfolg von
Abiturienten im Kontext veränderter schulorganisatorischer und curricularer
Rahmenbedingungen

Oepke, Maren; Eberle, Franz

Die Bedeutung von Deutsch- und Mathematikkenntnissen für Studierfähigkeit und Studienerfolg

Theoretischer Hintergrund: Als eines der Hauptziele des (Schweizer) Gymnasiums gilt die Erlangung einer allgemeinen Studierfähigkeit (MAR 95, 1995). Inhaltliche Definitionen der Studierfähigkeit nehmen häufig Bezug auf den Studienerfolg. Danach soll Studierfähigkeit „Studienerfolg ermöglichen und diesen erwartbar werden lassen“ (Konegen-Grenier, 2002, S. 29) oder Studierfähigkeit wird als Gesamtheit aller Kompetenzen (Kenntnisse, Fähigkeiten, Fertigkeiten und Bereitschaften) verstanden, die dazu befähigen, „ein Studium erfolgreich zu beginnen, durchzuführen und abzuschließen“ (Huber, 2009, S. 108). Allgemeine Studierfähigkeit meint dann, dass sich diese Kompetenzen in Übereinstimmung mit der allgemeinen Hochschulzugangsberechtigung auf jedes Hochschulstudium in jedem Fach erstrecken. Dem gegenüber ist zu berücksichtigen, dass sich fachliche Anforderungen verschiedener Studienrichtungen sehr stark unterscheiden können. Zwar kristallisierten sich in Theorie und einigen empirischen Untersuchungen zur allgemeinen Studierfähigkeit sogenannte "basale" fachliche Kompetenzen heraus, die von vielen Studienfächern vorausgesetzt werden (Eberle et al., 2008). Zu diesen Kompetenzen gehören Wissen und Können aus der Erstsprache, der Wissenschaftssprache Englisch und aus der Mathematik sowie Informatikanwendungskompetenzen (siehe auch Huber, 2009, S. 118 ff.). Die Frage, welche Bedeutung solchen basalen Kompetenzen faktisch für Studienerfolg und damit auch für die allgemeine Studierfähigkeit zukommt, ist bisher jedoch wenig untersucht (vgl. SKBF, 2010; vgl. Wolter, Diem & Messer, 2013). Der vorliegende Beitrag konzentriert sich auf Kompetenzen aus dem Bereich Deutsch als Erstsprache und Mathematik.

Fragestellung: Inwiefern eignen sich Kompetenztests aus den Bereichen Erstsprache Deutsch und Mathematik für die Vorhersage von Studienerfolg als ein Indikator für Studierfähigkeit für verschiedene Studienrichtungen mit unterschiedlichen Anforderungen an das Wissen und Können ihrer Studierenden?

Methode: Die vorliegenden Daten wurden einerseits im Rahmen der nationalen Studie EVAMAR II (Eberle et al., 2008) erhoben, in der ausgewählte Subkomponenten allgemeiner Studierfähigkeit mittels Tests und Selbsteinschätzung erfasst wurden. Dazu wurden Abiturientinnen und Abiturienten kurz vor ihrer Abschlussprüfung 2007 auch mittels Rasch-skalierten Leistungstests aus den Bereichen Erstsprache und Mathematik getestet (N = 3773). Zum anderen liegen Daten aus einer vom Schweizer Nationalfonds geförderten Nachfolgestudie vor, in der eine Deutschschweizer Teilstichprobe von EVAMAR II im Winter 2011/12 (N ca. 1200) und 2012/13 (N ca. 1000) zu Studienverlauf und Studienerfolg befragt wurde. Studienerfolg wurde dabei anhand unterschiedlicher objektiver (Studiennoten, Studiendauer, Studienabbruch, Studienleistungen) und subjektiver Kriterien (Studienzufriedenheit, subjektiv wahrgenommener Erfolg) operationalisiert. Weiterhin wurden die Studierenden nach ihrer Einschätzung gefragt, für wie wichtig sie verschiedene Facetten schulisch vermittelter Kompetenzen für ihr jeweiliges Hauptfach des Studiums halten.

Ergebnisse: Die varianzanalytischen Ergebnisse zeigen bezüglich der subjektiv wahrgenommenen fachlichen Anforderungen erhebliche Unterschiede zwischen den verschiedenen Studienrichtungen. Die Leistungstests eignen sich für die Vorhersage von Studienerfolg unterschiedlich gut: Die Regressionsanalysen ergeben, dass über die Durchschnittsabiturnote hinaus zum Zeitpunkt des Abiturs gute Deutschleistungen keine zusätzliche Varianz des Studienerfolgs in solchen Studienrichtungen erklären können, in denen Deutschkenntnisse von den Studierenden als besonders wichtig angesehen werden, wohl aber in eher mathematisch orientierten Studiengängen. Analog zeigen sich hohe Mathematikkompetenzen eher nicht als gute Vorhersagefaktoren für Studienerfolg in Studienrichtungen, in denen Mathematikkenntnisse von den Studierenden als besonders wichtig bezeichnet werden, dagegen aber in eher sprachlastigen Studienrichtungen mit gering eingeschätzter Bedeutung der Mathematik. Diese Befunde lassen sich mit Deckeneffekten erklären, denn zum Zeitpunkt der Erhebungen 3-4 Jahre nach dem Abitur werden Studienrichtungen mit hohen mathematischen Ansprüchen im Durchschnitt von Personen der Stichprobe mit guten bis sehr guten Mathematikkompetenzen, „sprachlastige“ Fächer von solchen mit guten bis sehr guten Deutschkompetenzen studiert; diese guten bis sehr guten Kompetenzen vermögen dann innerhalb dieser Studienrichtungen nicht mehr zu differenzieren.

Donnerstag
12. März

9:00 Uhr

HGC 40

Wüstenberg, Sascha

Nutzung und Nutzen der Bearbeitungszeit als Verhaltensmaß in computerbasierten Large Scale Assessments

Computerbasiertes Testen findet in der nationalen wie auch internationalen empirischen Bildungsforschung immer häufigere Verwendung, wie etwa im Rahmen weltweiter Large Scale Assessments (LSA). Beispiele hierfür sind die Erfassung der digitalen Lesekompetenz in PISA 2009, der domänenübergreifenden Problemlösefähigkeit in PISA 2012 und der Kompetenz im Umgang mit Informations- und Kommunikationstechnologie in PIAAC 2012. Der Einsatz computerbasierter Testungen erlaubt dabei nicht nur ein automatisiertes Scoring der Ergebnisse der Aufgabenbearbeitung, sondern im Vergleich zu papierbasierten Testungen auch eine exakte Erfassung von Verhaltensdaten wie zum Beispiel der Bearbeitungszeit. Das Symposium bringt Beiträge aus vier Projekten zusammen, die sich alle mit der Nutzung und dem Nutzen der Bearbeitungszeit als Verhaltensmaß im Rahmen von computerbasierten Testungen in LSAs beschäftigen. Die Beiträge repräsentieren dabei heterogene Konstrukte (Problemlösen, digitales Lesen, Umgang mit Informations- und Kommunikationstechnologie), Stichproben (Schüler, Erwachsene) und LSAs (PIAAC Feldtest, nationales Bildungsmonitoring in Schulen, eigene Studien).

In Beitrag 1 (Scherer et al.) wird basierend auf Daten des finnischen Bildungsmonitorings gezeigt, dass der Zusammenhang zwischen schulischem Selbstkonzept und Problemlöseleistung partiell durch die Bearbeitungszeit mediiert wird. Ferner wird dargelegt, dass bei Kontrolle von Alter und Reasoning die Bearbeitungszeit einen indirekten Effekt über die Problemlösefähigkeit auf schulische Leistungen aufweist. Ergänzend zum ersten Beitrag wird in Beitrag 2 (Rudolph et al.) anhand von Schülerdaten aus Deutschland gezeigt, dass die Beziehung zwischen Reasoning und Need for Cognition (d.h., die intrinsische Motivation zu denken und Probleme zu lösen) zum Kriterium Problemlöseleistung durch die Bearbeitungszeit mediiert wird. Während die ersten beiden Beiträge die Bedeutsamkeit von Bearbeitungszeit als relevantes Verhaltensmaß für domänenübergreifendes Problemlösen herausstellen, steht in Beitrag 3 (Stelter et al.) der Umgang mit Informations- und Kommunikationstechnologie im Vordergrund. Basierend auf PIAAC-Daten wird in diesem Beitrag gezeigt, dass Unterschiede zwischen Alterskohorten in der Lösung von technologiebasierten Aufgabenstellungen auf längere Bearbeitungszeiten für automatisierte basale Teilhandlungen (wie beispielsweise dem Setzen eines Lesezeichens oder dem Verschieben einer Mail) zurückgehen. In Beitrag 4 (Naumann et al.) wird anhand von PISA 2009-Daten zu digitalem Lesen untersucht, wie Aufgabenschwierigkeit, Lesekompetenz und deren Interaktion mit der Bearbeitungszeit zusammenhängen. Ein Hauptbefund der Autoren ist dabei, dass vor allem Personen mit niedriger Lesekompetenz ihre Bearbeitungszeit nicht an die Aufgabenschwierigkeit anpassen.

Zusammengenommen diskutieren die Beiträge das Potential und die Beschränkungen der Bearbeitungszeit als Verhaltensmaß anhand unterschiedlicher Konstrukte. Das übergeordnete Ziel des Symposiums ist es, Rückschlüsse über beteiligte motivationale und kognitive Prozesse zu liefern und eine Grundlage für pädagogische Interventionen und Trainings zu bilden.

Donnerstag
12. März

9:00 Uhr

HGC 50

Chair(s)

Wüstenberg, Sascha
Greiff, Samuel**Nutzung und Nutzen der Bearbeitungszeit als Verhaltensmaß in computerbasierten Large Scale Assessments**

Scherer, Ronny; Greiff, Samuel; Hautamäki, Jarkko

Über das Potenzial von Aufgabenbearbeitungszeiten beim komplexen Problemlösen: Vorhersage von schulischen Leistungen und Zusammenhang mit Selbstkonzept

Im Zuge der computerbasierten Erfassung von Konstrukten in der empirischen Bildungsforschung stehen neben reinen leistungsbasierten Variablen auch Prozessdaten zur Verfügung. Prozessdaten enthalten sowohl Informationen über das strategische Vorgehen der Schülerinnen und Schüler während der Testadministration, als auch die zur Aufgabenbearbeitung benötigte Zeit. Insbesondere bei der Messung kognitiver Grundfähigkeiten (Reasoning) kommt Aufgabenbearbeitungszeiten eine besondere Bedeutung zu. So indizieren lange Bearbeitungszeiten die Anstrengung der Schülerinnen und Schüler in komplexen Problemstellungen (Kupiainen et al., 2014). Somit lassen sich Bearbeitungszeiten unter anderem zur Vorhersage schulischer Leistungen heranziehen (Demetriou et al., 2014). Ein wesentlicher Vorteil der Verwendung dieser Zeiten besteht darin, dass sie aufgrund ihrer objektiven Messung nicht durch soziale Erwünschtheit, die etwa in selbstberichteten Anstrengungsvariablen auftreten kann, determiniert sind.

Es stellt sich die Frage, inwieweit die Befunde zur Vorhersage schulischen Erfolgs auf Konstrukte übertragbar sind, die über kognitive Grundfähigkeiten hinausgehen. Im Kontext cross-curricularer Kompetenzen spielt beispielsweise die Fähigkeit, komplexe Probleme durch Modellbildung und Wissensanwendung zu lösen eine entscheidende Rolle. Greiff und Fischer (2013) konnten zeigen, dass die komplexe Problemlösefähigkeit einen substanziellen Effekt auf schulische Ergebnisse selbst nach Kontrolle kognitiver Grundfähigkeiten aufweist. Das Zusammenspiel zwischen schulischem Erfolg, Aufgabenbearbeitungszeiten und Leistungen beim komplexen Problemlösen ist jedoch weitestgehend ungeklärt.

Das Hauptziel der vorliegenden Studie besteht somit darin, die Zusammenhänge zwischen Aufgabenbearbeitungszeiten und Leistungen beim komplexen Problemlösen mit schulischen Leistungen und Selbstkonzept zu beschreiben. Es wird erwartet, dass der Zusammenhang zwischen Aufgabenbearbeitungszeiten und schulischen Leistungen durch die Performanz beim komplexen Problemlösen mediiert wird. Zusätzlich wird erwartet, dass der Zusammenhang zwischen dem Selbstkonzept und der Problemlöseleistung durch die Aufgabenbearbeitungszeiten mediiert wird. Grundlage für diese Hypothesen bildet die Annahme, dass Aufgabenbearbeitungszeiten mit der investierten Anstrengung der Schülerinnen und Schüler bei der Problembearbeitung korrespondieren und somit die Problemlöseleistung, die schulischen Erfolg vorherzusagen vermag, direkt beeinflussen (z.B. Dodonova & Dodonov, 2012; Vock, Preckel, & Holling, 2011).

Basierend auf den Daten einer finnischen „Large-Scale-Studie (N=4,029) wurden neun komplexe Problemlöseszenarien hinsichtlich der Aufgabenbearbeitungszeiten sowie Performanz von Sechst- und Neuntklässlern untersucht. In den komplexen Problemlöseszenarien wurden die Komplexitäten der Szenarien systematisch variiert sowie verschiedene kognitive Prozesse erfasst (MicroDYN; Greiff & Fischer, 2013). Zu diesen Prozessen gehören Modellbildung und Wissensanwendung (Greiff & Fischer, 2013). In einem ersten Schritt des komplexen Problemlösens wurden die Schülerinnen und Schüler aufgefordert, Informationen über ein System durch systematische Interaktion zu generieren (Modellbildung), um diese anschließend in einem zweiten Schritt zur Erreichung eines Zielzustandes anzuwenden (Wissensanwendung). Neben den computerbasierten Problemlösetests wurden schulische Leistungen in Form von Noten sowie verschiedene Dimensionen des schulischen Selbstkonzepts erhoben. Außerdem wurde ein Reasoning-Test zur figuralen Analogienbildung eingesetzt. Die Zusammenhänge zwischen den Konstrukten wurden mit Hilfe latenter Strukturgleichungsmodelle geprüft.

Die Befunde zeigen, dass die Aufgabenbearbeitungszeiten für die verschiedenen Dimensionen komplexen Problemlösens schulische Leistungen erklären ($\beta=.18-.28$). Diese Zusammenhänge wurden selbst nach Kontrolle von Alter und figuralem Reasoning durch die Leistungen im komplexen Problemlösen vollständig mediiert. Insofern ist von einem indirekten Einfluss der Aufgabenbearbeitungszeiten auf schulische Leistungen auszugehen. Schulisches Selbstkonzept konnte die Bearbeitungszeiten positiv vorhersagen ($\beta=.16-.18$), wenngleich der Zusammenhang zwischen Selbstkonzept und Problemlöseleistung nur partiell durch die Aufgabenbearbeitungszeiten mediiert war.

Die vorliegende Studie zeigt den Nutzen der Aufgabenbearbeitungszeiten beim Verständnis des komplexen Problemlösens zur Vorhersage schulischer Ergebnisse auf und erweitert bisherige Befunde zur Mediationsfunktion von intelligenzbasierten Konstrukten um die komplexe Problemlösefähigkeit (Demetriou et al., 2014). Aufgabenbearbeitungszeiten stellen hiernach objektiv und schnell zugängliche Prozessdaten dar, die schulischen Erfolg erklären können. Auf Basis der dargestellten Resultate wird argumentiert, dass Aufgabenbearbeitungszeiten in Verbindung mit Leistungen zur Vorhersage schulischer Leistungen herangezogen werden können. Selbstkonzept stellt zudem eine Determinante des Zusammenhangs zwischen Bearbeitungszeiten und Leistungen beim komplexen Problemlösen dar.

Donnerstag
12. März

9:00 Uhr

HGC 50

Chair(s)

Wüstenberg, Sascha
Greiff, Samuel

Nutzung und Nutzen der Bearbeitungszeit als Verhaltensmaß in
computerbasierten Large Scale Assessments

Rudolph, Julia; Niepel, Christoph; Wüstenberg, Sascha; Greiff, Samuel

Prädiktoren komplexen Problemlösens besser verstehen: Bearbeitungszeit als partieller Mediator für den Effekt von Reasoning und Need for Cognition auf komplexes Problemlösen

Die vorliegende Studie untersucht, inwieweit (1) Reasoning und (2) Need for Cognition (NFC; die intrinsische Motivation zu denken und Probleme zu lösen; Cacioppo & Petty, 1982) Leistung bei einem computerbasierten, komplexen Problemlösetest vorhersagen. Hierbei steht die Frage im Fokus, inwieweit der Einfluss von NFC und Reasoning durch die Bearbeitungszeit mediiert wird.

Komplexes Problemlösen (KPL) bezeichnet die Fähigkeit, sich selbständig Wissen durch Exploration anzueignen und dieses Wissen zielführend anzuwenden (Funke, 2001). KPL beinhaltet, ebenso wie Reasoning, das Erkennen und Anwenden von Regeln. Diese Gemeinsamkeit zwischen Reasoning und KPL zeigt sich empirisch in bedeutsamen Zusammenhängen beider Konstrukte ($r \geq .56$; Wüstenberg, Greiff & Funke, 2012). Der Zusammenhang von NFC und KPL wurde dagegen bisher noch nicht untersucht. Allerdings konnte bisherige Forschung zeigen, dass Schüler_innen mit höherem NFC motivierter sind, komplexe Aufgaben zu bearbeiten (Cacioppo & Petty, 1982) und besser im Adaptive Decision Making sind als Schüler_innen mit einem niedrigeren NFC (Levin, Huneke & Jasper, 2000). Basierend auf obigen Befunden erwarten wir einen signifikanten Zusammenhang von Reasoning und NFC mit KPL.

In der vorliegenden Studie wird zudem untersucht, inwiefern die Bearbeitungszeit den Einfluss von Reasoning und NFC mediiert. Schüler mit höheren kognitiven Fähigkeiten arbeiten häufiger selbstreguliert (Veenman, Wilhelm & Beishuizen, 2004) und können somit besser einschätzen, wie viel Zeit sie für eine (komplexe) Aufgabe verwenden sollten. Über NFC ist bekannt, dass es mit erhöhter Persistenz bei der Aufgabenbearbeitung einhergeht (Hill, Foster, Elliott, Shelton, McCain & Gouvier, 2013). Jüngste Ergebnisse zeigen weiterhin, dass eine längere Bearbeitungszeit mit einer besseren Problemlöseleistung einhergeht (Goldhammer, Naumann, Steller, Tóth, Rölke, & Klieme, 2014). Zusammengefasst erwarten wir, dass die Regression von Reasoning und NFC auf KPL partiell durch die Bearbeitungszeit mediiert wird.

Es liegen Daten von 539 deutschen Schüler_innen aller Schulformen im 7. Schuljahr vor (Alter: $M=13.9$; $SD=.69$; 52.2% Mädchen; 8.2% ohne Geschlechtsangabe). NFC wurde mittels einer deutschen Adaption der Cacioppo-Skala gemessen (Preckel, 2013). Reasoning wurde mittels des Culture Fair Tests (Weiß, 2008) und KPL mittels des computerbasierten MicroDYN Ansatzes (Greiff, 2012) erfasst. Die Konstrukte werden latent modelliert, lediglich die Bearbeitungszeit von KPL geht als manifester Mittelwert aller MicroDYN-Aufgaben in die Analyse ein. Alle Faktoren weisen ausreichende bis zufriedenstellende Modellfits auf ($RMSEA \leq .048$; $CFI \geq .905$; $TLI \geq .902$). Die Hypothesen werden anhand eines latenten Strukturgleichungsmodells in Mplus (Muthén, & Muthén, 1998–2010) geprüft, in dem sowohl direkte Pfade von NFC und Reasoning auf das Kriterium KPL modelliert wurden, als auch indirekte Pfade über den Mediator Bearbeitungszeit.

Die standardisierten Ergebnisse zeigen, dass die Wirkungen von NFC und Reasoning auf KPL über die Bearbeitungszeit partiell mediiert werden (Modellfit: $RMSEA = .009$; $CFI = .911$; $TLI = .908$). Reasoning ($\beta=.42$) und NFC ($\beta=.11$) sind direkt mit KPL assoziiert. Ferner zeigt sich, dass Reasoning ($\beta=.30$) und NFC ($\beta=.16$) mit der Bearbeitungszeit verbunden sind, welche als Mediator wiederum die Leistung in KPL vorhersagt ($\beta=.30$). Der indirekte Pfad von Reasoning über die Bearbeitungszeit auf KPL beträgt .04, während der indirekte Pfad von NFC über die Bearbeitungszeit .07 beträgt. Alle direkten und indirekten Pfade sind signifikant ($p < .01$).

Die Ergebnisse zeigen, dass ein signifikanter Mediatoreffekt der Bearbeitungszeit vorliegt, welcher allerdings keinen großen Effekt darstellt. Dennoch zeigen die Analysen, dass die Beziehung von Reasoning und NFC zu KPL besser erklärt werden kann, wenn Prozessdaten berücksichtigt werden. Um ein vollständiges Bild zur Wirkung von NFC und Reasoning auf das KPL-Verhalten zu gewinnen, müssten jedoch noch weitere Prozessdaten Berücksichtigung finden, wie das strategische Vorgehen oder die Zeit zwischen den Bearbeitungsschritten. Ein tiefgreifendes Verständnis von KPL-Verhalten und seinen Prädiktoren entsteht erst, wenn diese Informationen im Zusammenspiel mit der Bearbeitungszeit betrachtet werden.

Donnerstag
12. März

9:00 Uhr

HGC 50

Chair(s)

Wüstenberg, Sascha
Greiff, Samuel

Nutzung und Nutzen der Bearbeitungszeit als Verhaltensmaß in
computerbasierten Large Scale Assessments

Stelter, Annette; Naumann, Johannes; Goldhammer, Frank; Rölke, Heiko

Technologischer Wandel und die Kompetenz Probleme technologiebasiert zu lösen – Die Rolle von Routinen im Kohortenvergleich

Theoretischer Hintergrund: Der Umgang mit Informations- und Kommunikationstechnologie (IKT) ist im Alltag inzwischen von grundlegender Bedeutung und zum Beispiel obligatorisch für die Arbeits- oder auch Wohnungssuche. Die in den letzten Jahrzehnten erfolgten technologischen Veränderungen bergen insbesondere für Erwachsene neue Herausforderungen, und speziell ältere Erwachsene haben mehr Probleme im Umgang mit IKT (Norris, 2001; OECD, 2013). Diese Herausforderungen können durch lebenslanges und informelles Lernen gemeistert werden. Hierbei muss sowohl deklaratives Wissen über IKT als auch prozedurales Wissen (Fitts & Posner, 1967) erworben werden. Vorteilhaft ist es, wenn der Erwerb prozeduralen Wissens eine möglichst weitgehende Automatisierung basaler Teilhandlungen einschließt (Schneider & Stern, 2010). Beim Problemlösen werden Routinen zunächst keine zentrale Rolle zugeschrieben, da Probleme sich definitionsgemäß nicht allein durch den Einsatz von Routineoperationen lösen lassen (Mayer 1994). Eine fortgeschrittene Automatisierung von für das Lösen eines Problems relevanten Teilhandlungen kann jedoch den Problemlöseprozess unterstützen, weil in diesem Fall kognitive Kapazitäten für kontrollierte Informationsverarbeitung frei werden (Stelter, Goldhammer, Naumann, & Rölke, 2014).

Hypothesen: Stelter et al. (2014) konnten zeigen, dass die Automatisierung von relevanten Teilhandlungen prädiktiv für das erfolgreiche technologiebasierte Lösen von Problemen ist. Dies legt die Vermutung nahe, dass sich beobachtete Unterschiede zwischen Jahrgangskohorten in der Kompetenz beim technologiebasierten Problemlösen auch aus Unterschieden in der Automatisierung relevanter Teilhandlungen erklären lassen.

Methode: Dem vorliegenden Beitrag liegen Daten des deutschen Feldtestes (2011) der PIAAC-Studie zugrunde. Analysiert wurden Daten von 297 Personen (Altersdurchschnitt=38.2 Jahre; SD=13.5; Min = 16; Max = 65) und drei Kohorten gebildet (jüngste 16-29 Jahre, mittlere 30-44 Jahre und älteste 45-65 Jahre). Die Wahrscheinlichkeit der richtigen Lösung von insgesamt sechs technologiebasierten Problemlöseaufgaben wurde zunächst aus der Kohortenzugehörigkeit (zwei Dummyvariablen; Referenzkategorie: jüngste Kohorte), und aus der Automatisierung basaler Teilhandlungen vorhergesagt. Als Automatisierungsindikator wurde die mittlere Zeit angesetzt, die die Personen für drei automatisierbare Teilhandlungen (Setzen eines Lesezeichens, Verschieben einer Mail und Schließen eines Popups) benötigten. Schließlich wurden beide Prädiktoren simultan verwendet. In allen Modellen wurden Geschlecht und Migrationshintergrund kontrolliert. Analysiert wurden generalisierte lineare Mischmodelle (GLMM) mit Residuen für Aufgaben und Personen.

Ergebnisse: Die Lösungswahrscheinlichkeiten der jüngsten und mittleren Kohorte unterschieden sich nicht signifikant (siehe Modell 1, Tabelle 1). Ein signifikanter negativer Effekt konnte für die älteste Kohorte gefunden werden. Angehöriger dieser Kohorte lösen demnach technologiebasierte Probleme mit geringerer Wahrscheinlichkeit als Angehörige der jüngsten Kohorte. In Modell 2 zeigte sich der erwartete negative Effekt des Automatisierungsindikators. Längere Bearbeitungszeiten bei automatisierbaren Teilschritten hängen negativ mit dem Lösungserfolg zusammen. Auch im dritten Modell findet sich ein signifikanter negativer Effekt des Automatisierungsindikators, wohingegen der Effekt der Kohortenzugehörigkeit nicht mehr signifikant ist. Dies legt nahe, dass Personen – ungeachtet ihres Alters – mit ähnlich automatisierten Handlungen gleiche Lösungswahrscheinlichkeiten erreichen.

Prädiktoren	Modell 1		Modell 2		Modell 3	
	Feste Effekte	p	Feste Effekte	p	Feste Effekte	p
Interzept	0,57 (0,62)	n.s.	0,45 (0,59)	n.s.	0,18 (0,48)	n.s.
Kohorte 30-44 Jahre	0,01 (0,24)	n.s.			0,12 (0,24)	n.s.
Kohorte 45-65 Jahre	0,54 (0,25)	0,28			0,34 (0,24)	n.s.
Automatisierungsindikator			-0,50 (0,12)	<0,01	-0,46 (0,13)	<0,01
Geschlecht Männlich	-0,18 (0,20)	n.s.	-0,25 (0,20)	n.s.	-0,20 (0,20)	n.s.
Kein Migrationshintergrund	-0,37 (0,42)	n.s.	-0,43 (0,42)	n.s.	-0,32 (0,41)	n.s.

Diskussion: Die hier vorgestellten Ergebnisse sind vereinbar mit der Annahme, dass Unterschiede zwischen Alterskohorten im technologiebasierten Problemlösen unter anderem auch auf Unterschiede in der Automatisierung basaler Teilhandlungen zurückgehen. Denkbar ist, dass ältere Personen (45-65 Jahre) in Schule und Ausbildung relativ wenig Kontakt mit IKT hatten, was sich im späteren Erwachsenenalter in entsprechend eingeschränkte Möglichkeiten zur Übung einschlägiger Handlungen niedergeschlagen haben könnten. Übung wiederum ist zentral für Automatisierung (Fitts & Posner, 1967). Sollten diese Überlegungen durch künftige Analysen Stützung erfahren, wäre dies ein weiteres Argument für die Forcierung „lebenslangen Lernens“ im Bereich von IKT.

Donnerstag
12. März

9:00 Uhr

HGC 50

Chair(s)

Wüstenberg, Sascha
Greiff, Samuel



Naumann, Johannes

Gute Leser passen beim Lesen digitaler Texte die Bearbeitungszeit der Aufgabenschwierigkeit an

Theoretischer Hintergrund und Fragestellung: Bearbeitungszeiten bei Leseaufgaben sind ein zentraler Kompetenzindikator (z.B. Richter et al., 2012), und prädiktiv für das Bearbeitungsergebnis. Der Zusammenhang zwischen Bearbeitungszeit und Bearbeitungsergebnis wird allerdings durch Aufgabenschwierigkeit und Personenfähigkeit moderiert. Während sich bei einfachen Aufgaben und fähigen Personen keine oder negative Zusammenhänge zeigen, finden sich bei schweren Aufgaben und wenig fähigen Personen positive Zusammenhänge (Goldhammer et al., 2014; Naumann & Goldhammer, 2014). Im Prinzip könnten also weniger fähige Personen in der Lage sein, mangelnde Fähigkeiten durch längere Bearbeitungszeiten zu kompensieren (vgl. Walczyk, 2000). Fraglich ist allerdings, ob es tatsächlich eher weniger fähige Personen sind, die bei schwierigen digitalen Leseaufgaben mehr Zeit investieren. In Laborexperimenten sind es oft die kompetenteren Leser, die mehr Zeit in relevante Teile des Aufgabenbearbeitungsprozesses wie Planung von Navigation (Rouet & Bigot, 2007) oder Verarbeitung relevanter Textteile (Rezende & DeSousa, 2008) investieren.

In der vorliegenden Studie wurde vor diesem Hintergrund untersucht, wie Bearbeitungszeiten bei digitalen Leseaufgaben mit Lesekompetenz zusammenhängen, und in welcher Weise dieser Zusammenhang durch die Aufgabenschwierigkeiten moderiert wird.

Methode: Die Datengrundlage bildete das PISA 2009 Digital Reading Assessment (N=34.506, 19 Länder 29 computerbasierte Leseaufgaben, OECD, 2011). Die Textmaterialien waren typische digitale Texte (z.B. Websites, Blogs). Alle Textstimuli waren Hypertexte. Die Aufgabebearbeitungszeiten wurden aus Logfiles extrahiert, logarithmisch transformiert und standardisiert. Die Aufgabenschwierigkeitsparameter wurden aus OECD (2012) entnommen. Lesekompetenz wurde über die PISA Print-Leseskala operationalisiert. Die Aufgabenbearbeitungszeit wurde in einem LMM vorhergesagt, das als feste Effekte Aufgabenschwierigkeit und Lesekompetenz, sowie deren Interaktion enthielt. Aufgaben, Schüler und Schulen wurden als Zufallseffekte modelliert. Alle Analysen wurden länderweise vorgenommen und die Ergebnisse metaanalytisch integriert (random effects model, s. Cheung, 2008).

Ergebnisse und Diskussion: Wir berichten die hypothesenrelevanten festen Effekte. Sämtliche Zufallseffekte waren signifikant, ebenso die Varianz der festen Effekte zwischen Ländern ($\tau^2 \leq 0.004$).

Aufgabenschwierigkeit: In allen Ländern fand sich ein deutlicher positiver Haupteffekt der Aufgabenschwierigkeit, $0.36 \leq b \leq 0.58$, metaanalytisch: $b = 0.46$ (SE = 0.02), $p < .001$.

Lesekompetenz: Der Haupteffekt von Lesekompetenz war in allen Ländern positiv, variierte allerdings deutlich in seiner Ausprägung, $0.004 \leq b \leq 0.19$, metaanalytisch: $b = 0.10$ (SE = 0.01), $p < .001$.

Interaktion Aufgabenschwierigkeit x Lesekompetenz: Die Haupteffekte von Aufgabenschwierigkeit wurden in allen Ländern durch eine positive Interaktion moderiert, $0.05 \leq b \leq 0.11$, metaanalytisch: $b = 0.09$ (SE = 0.004), $p < .001$. Um diese Interaktion besser interpretieren zu können wurden simple slopes bestimmt. Der Effekt von Lesekompetenz wurde bedingt auf leichte ($\beta = 2$ logits) und schwere ($\beta = +2$ logits) Aufgaben, der Effekt der Aufgabenschwierigkeit wurde bedingt auf Personen mit niedriger Lesefähigkeit (-2SD) und hoher Lesefähigkeit (+2SD). Diese Analysen zeigten für leichte Aufgaben konsistent einen negativen Effekt der Lesekompetenz: Leichte Aufgaben wurden von guten Lesern schneller gelöst als von schlechten Lesern, $0.15 \leq b \leq 0.01$, metaanalytisch: $b = 0.08$ (SE = 0.01), $p < .001$. Für schwere Aufgaben zeigte sich dagegen ein deutlicher positiver Effekt der Lesekompetenz, $0.10 \leq b \leq 0.38$, metaanalytisch: $b = 0.27$ (SE = 0.02), $p < .001$. Dieses Ergebnis lässt sich so deuten, dass gute Leser eher bereit waren, ausreichend Zeit in die Bearbeitung schwerer Aufgaben zu investieren. Entsprechend zeigte sich bei guten Lesern ein deutlicher Effekt der Aufgabenschwierigkeit, $0.46 \leq b \leq 0.73$, metaanalytisch: $b = 0.63$ (SE = 0.02), $p < .001$. Bei schwachen Lesern dagegen war dieser Effekt ebenfalls positiv, aber deutlich abgeschwächt, $0.18 \leq b \leq 0.49$, metaanalytisch: $b = 0.28$ (SE = 0.02), $p < .001$.

Insgesamt also passen vor allem diejenigen Personen, die von mehr Zeiteinsatz besonders profitieren könnten, ihre Aufgabenbearbeitungsgeschwindigkeit besonders wenig an die Aufgabenschwierigkeit an. Warum dies so ist (metakognitive Regulation? Motivation?) muss künftige Forschung klären.

Donnerstag
12. März

9:00 Uhr

HGC 50

Chair(s)

Wüstenberg, Sascha
Greiff, SamuelNutzung und Nutzen der Bearbeitungszeit als Verhaltensmaß in
computerbasierten Large Scale Assessments

Hoffmann-Erz, Ruth

Entwicklung generalisierender Operationen im Zusammenhang mit sprachsystematischen Gegebenheiten. Empirische Untersuchung am Beispiel der Länge- und Kürzemarkierung

In der Forschung wird der Orthographieerwerb seit einigen Jahrzehnten als individueller eigenaktiver Lernprozess verstanden, der weitgehend durch unbewusste Regelprozesse bestimmt ist. Bei der Frage nach den mentalen Prozessen zeigen sich besonders für den fortgeschrittenen Rechtschreiberwerb zwei grundlegend unterschiedliche Zugriffsweisen, die als lexikalische und generalisierende Lernertypen oder auch Rechtsschreib-Operationen bezeichnet werden (Thomé 1999: 244). Während bei der lexikalischen Operation die Speicherung auf wort- oder morphemspezifischen Informationen in einem inneren orthographischen Lexikon beruht, werden bei der generalisierenden Operation innere orthographische Regeln meist implizit wirksam.

Neuere neurowissenschaftliche Untersuchungen zum impliziten Lernen zeigen, dass im Gehirn anhand von Beispielen eine allgemeine innere Repräsentation aufgebaut wird, die auf neue Situationen übertragen werden kann. Da unsere Welt regelhaft ist, müssen wir uns nicht jede Einzelheit merken. Unser Gehirn ist vielmehr auf das Lernen von Allgemeinem angelegt, indem es aus Beispielen selbst Regeln produziert, die uns nicht bewusst sind (Spitzer 2006: 76 f.). Regelhafte Strukturen werden dabei implizit aus statistischen Abhängigkeiten von Umweltreizen generiert, was einen Einfluss sprachsystematischer Gegebenheiten auf die Lernprozesse vermuten lässt (vgl. Treimann, Kessler 2006).

Den Erkenntnissen der Gehirnforschung zum impliziten Lernen folgend, ergibt sich für die Orthographie-Erwerbsforschung die Frage, wie die Generalisierungskompetenz im Entwicklungsverlauf aufgebaut wird und inwieweit regelhafte Strukturen der Orthographie implizit aus Beispielen erlernt werden.

Der Vortrag stellt eine empirische Studie vor, die Lernprozesse des impliziten Lernens im Orthographieerwerb am Beispiel der Länge und Kürzemarkierung untersucht (vgl. Hoffmann-Erz 2014). Ziel der Studie ist es, implizite Lernprozesse zu erkunden, die normgerechte Schreibungen bei noch nicht erlernten Wörtern generieren. Die Untersuchung beruht auf einem Mixed-Methods-Design. Im ersten Untersuchungsteil wurden 106 erwachsenen kompetenten Schreibern (Lehramtsstudierenden) Pseudowörter diktiert (Quasiexperiment). Postaktional fand mit der Methode des lauten Denkens eine schriftliche, zunächst offene, Befragung statt, um introspektiv die kognitiven Prozesse zu beleuchten. Eine geschlossene, schriftliche Befragung schloss sich an, bei der Echtwörter zu den Pseudowörtern assoziiert werden sollten. Der letzte Untersuchungsteil bildet eine Korrelationsstudie mit 307 Grundschulern der Jahrgangsstufen 2 bis 4 (Pseudolängsschnitt). Durch ein selbstentwickeltes Testdiktat werden Generalisierungen methodisch durch einen Vergleich der Echt- und Pseudowörtergraphien bestimmt.

Die Ergebnisse zeigen, dass die Generalisierungskompetenz mit der Regelhaftigkeit orthographischer Phänomene zusammenhängt. Während Phänomene, die regelhaft beschrieben werden können, in großem Maße übereinstimmend generiert werden, zeigen unsystematische Phänomene keine eindeutigen Generalisierungen. Als häufigste Begründung für die gewählten Pseudowörtergraphien wurden in dem offenen Format Analogien zu Echtwörtern genannt. Durch einen Vergleich der Pseudowörter mit den assoziierten Echtwörtern konnten sechs verschiedene Strategien der Analogiebildung operationalisiert werden. Die assoziierten Echtwörter zeigen dabei eine weitgehend ähnliche phonologische Struktur wie die Pseudowörter. Auf der Grundlage des Schülertests ermittelt eine Regressionsanalyse einen hohen Zusammenhang ($\beta = .88$, $R^2 = .77$, $p < .01$, $N = 307$) zwischen der Kompetenz bei den Echtwörtern und der Generalisierungskompetenz bei den Pseudowörtern.

Die Ergebnisse der Studie lassen die Interpretation zu, dass regelhafte Strukturen der deutschen Orthographie implizit durch Generalisierungen erschlossen werden. Darüber hinaus konnte gezeigt werden, dass die Generalisierungskompetenz im Entwicklungsverlauf kontinuierlich durch den Zuwachs wortspezifischer Kenntnisse aufgebaut wird bzw. Systematiken der Schrift mit zunehmendem orthographischem Wortschatz erschlossen werden.

Donnerstag
12. März

9:00 Uhr

GC 03/142

Chair(s)

Hoffmann-Erz, Ruth

Pfost, Maximilian; Fischer, Melanie Y.

Die Rolle der phonologischen Bewusstheit für den Schriftspracherwerb im Deutschen

Theoretischer Hintergrund: Das Konzept der phonologischen Bewusstheit, das heißt die Fähigkeit der bewussten Reflektion über die Lautstruktur der gesprochenen Sprache einschließlich deren Manipulation (vgl. Melby-Lervåg, Lyster & Hulme, 2012; Schnitzler, 2008), hat im Verlauf der letzten Jahrzehnte besondere Beachtung bei der Erforschung der Entwicklung schriftsprachlicher Kompetenzen erfahren. Im Rahmen querschnittlich, längsschnittlich sowie experimentiell angelegter Studien, konnten vielfach Zusammenhänge zwischen Maßen der phonologischen Bewusstheit und der Lese- und Rechtschreibleistung nachgewiesen werden (vgl. Ehri et al., 2001; Goswami & Bryant, 1990; Melby-Lervåg et al., 2012; Wagner & Torgesen, 1987). Viele Arbeiten stammen dabei aus dem englischen Sprachraum. Die englische Sprache ist jedoch, im Vergleich zum Deutschen, durch eine relativ hohe Inkonsistenz der Graphem-Phonem-Korrespondenz gekennzeichnet (Borgwaldt, Hellwig & Groot, 2005; Seymour, Aro & Erskine, 2003). Sprach- bzw. Orthographievergleichende Arbeiten zeigen, dass in Sprachen mit höherer Inkonsistenz der Graphem-Phonem-Zuordnung auch höhere Korrelationen zwischen der phonologischen Bewusstheit und dem Lesen von Wörtern zu erwarten sind (Mann & Wimmer, 2002; Ziegler et al., 2010).

Fragestellung: Der vorliegende Beitrag beschäftigt sich mit dem Stellenwert der phonologischen Bewusstheit für den Schriftspracherwerb im Deutschen. Es stellt sich die Frage, ob verglichen mit internationalen Meta-Analysen, auch für das Deutsche vergleichbar hohe längsschnittliche Zusammenhänge von phonologischer Bewusstheit und schriftsprachlichen Kompetenzen bestehen. Darüber hinaus sollen Effekte einer Förderung der phonologischen Bewusstheit für das Lesen und Rechtschreiben im Deutschen bestimmt und mit den in internationalen Arbeiten berichteten Effekten verglichen werden.

Methode: Im vorliegenden Beitrag werden die Ergebnisse von zwei Meta-Analysen von Studien aus dem deutschen Sprachraum zusammengefasst. Die erste Meta-Analyse fokussiert auf die Rolle der phonologischen Bewusstheit in der längsschnittlichen Vorhersage der Lese- und Rechtschreibleistung. Die zweite Meta-Analyse fasst Befunde zur Effektivität von Fördermaßnahmen der phonologischen Bewusstheit für die Entwicklung schriftsprachlicher Kompetenzen zusammen.

Ergebnisse: In der ersten Meta-Analyse wurden die empirischen Befunde von 21 unabhängigen längsschnittlichen Beobachtungsstudien, die den Inklusionskriterien entsprachen, zusammengefasst. Die durchschnittliche mittlere Korrelation von Maßen der phonologischen Bewusstheit mit dem Lesen und Rechtschreiben beträgt $r = 0.31$ ($Zr = 0.32$).

Im Rahmen der zweiten Meta-Analyse konnten die Ergebnisse von 22 Interventions-Kontrollgruppen-Vergleiche von Förderstudien der phonologischen Bewusstheit integriert werden. Die Ergebnisse zeigen substantielle Fördereffekte für die phonologische Bewusstheit in mittlerer Höhe sowie positive Transfereffekte auf das Lesen und Rechtschreiben in kleiner Höhe.

Diskussion: Insgesamt belegen die Befunde, dass Maße der phonologischen Bewusstheit das Potential aufweisen, rund ein Zehntel der Varianz der späteren Lese- und Rechtschreibkompetenzen vorherzusagen. Im Vergleich zu internationalen Arbeiten mit einem großen Anteil von Studien aus englischsprachigen Ländern (Melby-Lervåg et al., 2012; Scarborough, 1998), sind die im deutschen Sprachraum berichteten Korrelationen phonologischer Bewusstheit mit den schriftsprachlichen Kompetenzen allerdings geringer. Auch eine Förderung der phonologischen Bewusstheit mit dem Ziel einer (langfristigen) Förderung schriftsprachlicher Kompetenzen scheint, insbesondere im Verhältnis zu den vielversprechenden Befunden internationaler Meta-Analysen (Bus & van Ijzendoorn, 1999; Ehri et al., 2001), nur vergleichsweise gering effektiv. Die Ergebnisse deuten zusammenfassend darauf hin, dass eine Übertragung empirischer Befunde von Studien aus anderen Orthographien, insbesondere aus dem Englischen, nur eingeschränkt möglich ist. Die höhere Konsistenz der deutschen Orthographie, insbesondere mit Hinblick auf die Graphem-Phonem-Korrespondenz (Leserichtung), im Vergleich zur englischen Orthographie scheint mit einer geringeren Bedeutsamkeit der phonologischen Bewusstheit vor allem für die Leseentwicklung im Deutschen einherzugehen.

Donnerstag
12. März

9:00 Uhr

GC 03/142

Chair(s)

Hoffmann-Erz, Ruth

Flöter, Manja; Weigel, Sabine; Schmerse, Daniel

Effekte in der sprachlichen Bildung von Kindern durch den Besuch von Kindertageseinrichtungen

Theoretischer Hintergrund: Internationale Studien belegen, dass sich Bildungsbenachteiligung für Kinder aus bildungsfernen Elternhäusern bereits im vorschulischen Alter nachweisen lässt (z.B. Geoffroy et al., 2010). Zudem zeigt sich, dass eine geringe familiäre Anregungsqualität in nachteiligem Zusammenhang mit der kindlichen Entwicklung zu sehen ist (Melhuish, 2010; NICHD ECCRN;1998).

In diesem Kontext wird auch in Deutschland seit langem die Frage gestellt, ob solche Disparitäten durch den Besuch von Kindertagesstätten verringert werden können (vgl. Hasselhorn & Kuger, 2014). Insbesondere im Bereich der sprachlichen Entwicklung zeigten einige Studien positive Effekte für Kinder aus sozial benachteiligten Familien, die eine vorschulische Einrichtung besuchten (vgl. Siraj-Blatchford et al., 2011; Taggart et al., 2006). Weiterführend hierzu steht die Überlegung, ob eine hohe Qualität in der Kindertagesbetreuung zur Verringerung der Unterschiede in besonderem Maße beiträgt und dementsprechend im Sinne kompensatorischer Wirkung zu betrachten ist. Hasselhorn und Kuger (2014) weisen darauf hin, dass internationale Studien hierzu bisher widersprüchliche Ergebnisse zeigen und für Deutschland die systematische Reflexion der Möglichkeiten und Grenzen der vorschulischen Förderung derzeit noch ausstehen.

Ziel der vorliegenden Untersuchung ist es, einen Beitrag zur Beantwortung dieser Fragestellung zu leisten.

Fragestellung: Ausgehend vom Modell pädagogischer Qualität in Kindertageseinrichtungen (Orientierungs-, Struktur- und Prozessqualität, Tietze et al., 1998) wird in der vorliegenden Studie untersucht, ob und in welcher Weise die Qualität von Kindertagesstätten mit dem Verlauf der kindlichen Entwicklung in Zusammenhang steht. Dabei werden verschiedene Teilfragestellungen betrachtet: (a) Hat die Qualität der Kindertageseinrichtungen einen Effekt auf die kindliche Sprachentwicklung? (b) Kann mit guter Qualität von Kindertageseinrichtungen ein niedriger Sprachanregungsgehalt in Familien ausgeglichen werden (kompensatorische Effekte)?

Methode: Die vorliegenden Daten stammen aus der Evaluation des Bundesprogramms „Schwerpunkt-Kitas Sprache & Integration“, welches sich insbesondere auf Kindertagesstätten konzentriert, die überdurchschnittlich häufig von Kindern mit Migrationshintergrund und/oder aus sozial benachteiligten Familien besucht werden.

Im Rahmen der Evaluation wurden zu Beginn des Jahres 2013 in 335 Kindertagesstätten verschiedene Indikatoren sprachförderlicher Qualität durch direkte Beobachtung und durch schriftliche Befragungen erhoben. Zudem wurden 1331 Kinder und deren Familien zu Hause besucht. Hierbei wurden u.a. der Entwicklungsstand der Kinder im Wort- und Satzverstehen getestet (z.B. SETK Subskala Satzverstehen; Grimm, 2001; Grimm, Aktas & Frevert, 2000) sowie der Sprachanregungsgehalt des häuslichen Umfeldes von geschulten Beobachtern bewertet. Der zweite Messzeitpunkt begann im März 2014 und erreicht eine Wiedererhebungsrate von rund 84% der Ausgangsstichprobe.

Die Fragestellung (a) soll im Rahmen von multiplen schrittweisen Regressionsmodellen zur Erklärung des Verlaufs der sprachlichen Entwicklung bewertet werden.

Für Fragestellung (b) werden die Indikatoren des häuslichen Anregungsgehaltes sowie auch die institutionellen Qualitätsindikatoren zu einem Compositescore zusammengefasst. Entsprechend diesen Maßen werden jeweils drei Typen der Anregungsqualität gebildet (niedrig, mittel und hoch). Mittels Gruppenvergleichsanalysen soll die Größe der Interaktionseffekte der unterschiedlichen Qualitätstypen auf die kindliche Entwicklung betrachtet werden.

Ergebnisse und Ausblick: Im Rahmen der regressionsanalytischen Voranalysen des ersten Messzeitpunktes für Fragestellung A konnten bei Berücksichtigung struktureller und prozessualer Merkmale bisher tendenzielle direkte Effekte der Einrichtungsqualität auf die Indikatoren der kindlichen Entwicklung gefunden werden. Bedeutend höheren Anteil an der Erklärung haben die erhobenen Maße der familiären Prozessqualität. Es gibt jedoch Hinweise, die für einen Zusammenhang zwischen Indikatoren der institutionellen Prozessqualität und der familiären Prozessqualität in Abhängigkeit von der bisherigen Betreuungsdauer sprechen, so dass davon auszugehen ist, dass zum 2. Messpunkt (1 Jahr später) die Effekte deutlicher zu Tage treten werden.

Die Untersuchung der Fragestellung B wird auf den Daten der Längsschnitterhebung von Messzeitpunkt 1 und 2 basieren. Die Datenerhebung des 2. Messzeitpunktes befindet sich im Moment in der Endphase, so dass die Ergebnisse des Prä-Post-Vergleichs am Tagungstermin präsentiert werden können.

Donnerstag
12. März

9:00 Uhr

GC 03/142

Chair(s)

Hoffmann-Erz, Ruth

Jurecka, Astrid; Bündgens-Kosten, Judith; Elsner, Daniela

Erfassung mündlicher fremdsprachlicher Interaktion in der Primarstufe: Ein Vergleich von Interview- und Peer-to-Peer-Testsettings

Um der Forderung der europäischen Kommission (1998/2002) nach Förderung frühen Fremdspracherwerbs nachzukommen, wurde in Deutschland flächendeckend obligatorischer fremdsprachlicher Unterricht ab der dritten Klasse eingeführt (vgl. Elsner, 2010). Infolgedessen wurden in deutschen Bundesländern auf dem Gemeinsamen Europäischen Referenzrahmen für Sprachen (GER; Europarat, 2001; kommunikativer Sprachansatz, z.B. Bachman/Palmer 1996; Savignon 2002) basierende Bildungsstandards für fremdsprachliche Kompetenzen in der Grundschule entwickelt (vgl. Dausend 2014). Dabei werden primär mündliche Kompetenzbereiche (Hören/Sprechen/Interaktionen) fokussiert (vgl. Engel 2009). Für mündliche Kommunikation ist dabei insbesondere auch die Interaktion relevant: So werden im GER (S. 88ff) kognitive und kooperative Strategien (z.B. Sprecherwechsel/Kooperation/Bitten um Klärung) sowie diesen jeweils zugehörige Sprachaktivitäten (z.B. Vorschlagen/Evaluieren/Fragenstellen/Zusammenfassen), die von einem Sprachverwender zur Steuerung von Interaktion angewendet werden können, anhand gesonderter Kompetenzskalen beschrieben.

Tests zur Erfassung mündlicher fremdsprachlicher Kompetenzen in der Primarstufe (z.B. EVENING-Studie (Keßler, 2009); Cambridge YLE (UCLES, 2013)) beinhalten zumeist jedoch keine explizite Erfassung dieser interaktiven Sprachaktivitäten; zudem werden häufig lediglich sehr kurze, stark gesteuerte, thematisch eingeschränkte Interviewverfahren eingesetzt. Beides führt mutmaßlich zu einer Einschränkung der Konstruktvalidität, da bereits erworbene Sprachkompetenzen der SchülerInnen nicht in ihrer kompletten Breite erfasst werden: So existieren aus der Erwachsenenforschung Hinweise, dass weniger gesteuerte Test-Settings (wie Peer-to-Peer-Settings; z.B. Taylor/Wigglesworth, 2009) es den Getesteten erlauben, eine größere Bandbreite interaktionaler Kompetenzen aufzuzeigen als Interview-Settings (z.B. Ducasse/Brown 2009).

Ziel dieser Studie ist es, für eine konstruktvalide Erfassung mündlicher fremdsprachlicher Interaktionskompetenzen in der Primarstufe Interview- sowie Peer-to-Peer-Testsettings (gesteuert/weniger gesteuert) am Beispiel von Englisch als Fremdsprache hinsichtlich Anzahl und Art induzierter interaktiver Sprachaktivitäten zu untersuchen und zu vergleichen: (1) Inwieweit unterscheiden sich die beiden Testsettings hinsichtlich der Menge produzierter Sprache (Quantitativer Sprachhandlungsaspekt)? Hypothese: aufgrund stärkerer thematischer Lenkbarkeit seitens der Getesteten werden eine größere Anzahl von Wörtern sowie Interaktionen in Peer-to-Peer-Settings produziert. (2) Inwieweit unterscheiden sich die beiden Testsettings hinsichtlich der Art produzierter Interaktionen (Qualitativer Sprachhandlungsaspekt)? Hypothese: In beiden Settings werden unterschiedliche Sprachaktivitäten produziert (größere Bandbreite unterschiedlicher Interaktionen aufgrund freierer thematischer Gestaltbarkeit in Peer-to-Peer-Settings; Interview-Setting: primär Beantworten von Fragen; vgl. Ducasse/Brown; Europarat, 2001), und somit jeweils auch unterschiedliche Bereiche des Konstrukts erfasst.

Die Datenerhebung erfolgte paarweise an n=38 PrimarschülerInnen (n=19 Paare; 4. Klasse; M(Alder)=9,4 Jahre; männlich=18; weiblich=20). Hintergrund der Test-Settings bildete eine computergestützte Geschichte (MuViT/Elsner 2011; Input: visuell (Bild/Schrift); auditiv), welche die Kinder gemeinsam möglichst auf Englisch bearbeiten sollten. Die Operationalisierung der Testsettings erfolgte anhand einer sequentiellen Bearbeitung der Geschichte in mehreren, abwechselnd aufeinanderfolgenden Phasen (Einganginterview & gesteuerte Interview-Sequenzen während Geschichte („Prompts“)=Interview-Setting; freie Bearbeitungs-/Kommunikationssequenzen=Peer-to-Peer-Setting). Die paarweise Bearbeitung wurde videografiert und transkribiert. Als Maß für die interaktive Sprachaktivierung (quantitativer Sprachhandlungsaspekt) in den beiden Test-Settings diente die Anzahl produzierter Wörter/Interaktionen. Die Codierung hinsichtlich der Art der Sprachaktivitäten (qualitativer Sprachhandlungsaspekt) erfolgte auf Basis der im GER beschriebenen interaktiven Sprachaktivitäten (2 Rater; ICC=.76).

Insgesamt verwendeten die SchülerInnen durchschnittlich deutlich mehr deutsche als englische Wörter zum Kommunizieren (t-Test; $\bar{x}_1=535,84$, $\bar{x}_2=39,49$; $p=.00$). Hinsichtlich des quantitativen Sprachhandlungsaspektes zeigt sich annahmegemäß, dass im Rahmen der Peer-to-Peer-Settings signifikant mehr Wörter insgesamt (Deutsch+Englisch) produziert wurden (t-Test; $\bar{x}_1=429,26$, $\bar{x}_2=146$; $p=.000$). Dahingegen wurden mehr englische Wörter in Interview-Settings als in Peer-to-Peer-Settings gesprochen ($\bar{x}_1=27,84$; $\bar{x}_2=11,58$; $p=.002$). Die (sprachübergreifende) Untersuchung des qualitativen Sprachhandlungsaspektes erfolgte anhand einer Varianzanalyse mit Messwiederholung. Entsprechend der Annahmen zeigen multivariate Tests einen signifikanten Haupteffekt des Messwiederholungs-Faktors „Test-Setting“ ($p=.000$; $\eta^2=.978$); univariate Varianzanalysen zeigen für 11 der 13 Sprachaktivitäten signifikante Mittelwertsunterschiede. Dabei wurde „Fragen beantworten“ im Interview-Setting signifikant häufiger durchgeführt, während zehn weitere interaktive Sprachaktivitäten signifikant häufiger im Peer-to-Peer-Setting vorkamen (=größere Bandbreite interaktiver Sprachhandlungen). Die Ergebnisse weisen darauf hin, dass die Testsettings unterschiedliche Bereiche des Konstrukts erfassen, und somit ein additiver Einsatz beider Settings die Konstruktvalidität von Testverfahren erhöhen kann. Implikationen der Ergebnisse für die Entwicklung von Testverfahren werden diskutiert.

Donnerstag
12. März

9:00 Uhr

GC 03/142

Chair(s)

Hoffmann-Erz, Ruth

Paetsch, Jennifer; Felbrich, Anja; Stanat, Petra

Longitudinale Zusammenhänge zwischen sprachlichen Kompetenzen und elementaren mathematischen Modellierungskompetenzen bei Kindern mit Deutsch als Zweitsprache

Lernende, deren Familiensprache nicht der Instruktionssprache entspricht, verfügen häufig über eingeschränkte Sprachkompetenzen (z. B. Schwippert, Wendt & Tarelli, 2012), sodass die sprachlichen Anforderungen des Mathematikunterrichts für sie eine Hürde beim Erlernen der Fachinhalte bilden können. Es wird deshalb davon ausgegangen, dass der Kompetenznachteil in Mathematik von Schülerinnen und Schülern mit Zuwanderungshintergrund unter anderem auf die unzureichende Beherrschung der Instruktionssprache dieser Lernenden zurückzuführen ist (z. B. Stanat, 2006). Insbesondere für die Bearbeitung von mathematischen Modellierungsaufgaben, die den Lernenden in Textform präsentiert werden (sog. Text- bzw. Sachaufgaben), sind ausreichende sprachliche Kompetenzen erforderlich. Eine solche Aufgabe kann nur gelöst werden, wenn die Lernenden einerseits das zur Lösung notwendige mathematische Wissen –in der überwiegend auf Sprache basierenden Unterrichtssituation– erlangt haben, und wenn sie andererseits in der Lage sind, den im Text beschriebenen Sachverhalt zu lesen und zu verstehen. In den letzten Jahren wurde in der mathematikdidaktischen Diskussion verstärkt die Notwendigkeit betont, realitätsbezogene Fragestellungen in den Mathematikunterricht zu integrieren (z. B. Blum, 2010). In diesem Zusammenhang ist mathematisches Modellieren (außermathematischer Situationen) zu einem zentralen Gegenstand des Mathematikunterrichtes geworden, der auch in den Bildungsstandards betont wird (Borromeo Ferri, Greefrath & Kaiser, 2013; KMK, 2004). Bereits in der Grundschule soll die Förderung der mathematischen Modellierungsfähigkeit durch den Einsatz von Sach- und Textaufgaben erfolgen (sog. elementares Modellieren, vgl. Schwarzkopf, 2006).

Zwar liegen zahlreiche empirische Studien vor, die den Zusammenhang zwischen Sprachkompetenz und mathematischer Modellierungskompetenz belegen (z. B. Fuchs et al., 2006), Zusammenhänge zwischen den Entwicklungsverläufen der Kompetenzen wurden bislang jedoch nicht überprüft. Um Zweitsprachenlernende bei der Bewältigung sprachlicher Anforderungen gezielt unterstützen zu können, stellt sich zudem die Frage, welche sprachlichen Teilkompetenzen für die mathematische Kompetenzentwicklung von besonderer Bedeutung sind. Der vorliegende Beitrag knüpft an diese Forschungslücke an und untersucht die längsschnittlichen Zusammenhänge zwischen Sprachkompetenzen in den Bereichen Lesen und Grammatik und der mathematischen Modellierungskompetenz bei Kindern mit Deutsch als Zweitsprache. Es wird erwartet, dass beide sprachlichen Kompetenzen einen signifikanten Beitrag zur Vorhersage der Entwicklung mathematischer Modellierungskompetenz leisten. Zudem wird der Frage nachgegangen, ob und inwieweit Veränderungen der sprachlichen Kompetenzen der Zweitsprachenlernenden (in den Bereichen Lesen und Grammatik) mit Veränderungen in ihrer mathematischen Modellierungskompetenz einhergehen.

Für die Analysen wurde auf Daten der Prä- und Posttests der Interventionsstudie BeFo (Bedeutung und Form: Fachbezogene und sprachsystematische Förderung in der Zweitsprache) zurückgegriffen, die darauf abzielte, zwei Sprachförderansätze vergleichend zu untersuchen (Stanat & Rösch, 2011). An der Studie nahmen 370 Kinder aus insgesamt 15 Berliner Grundschulen teil, die mindestens mit einem Elternteil eine andere Sprache als Deutsch sprechen. Es wurden die Daten der Erhebungszeitpunkte am Beginn und am Ende des dritten Schuljahres verwendet. Die Veränderung über die Zeit wurde mit Hilfe von Latent-Change-Modellen analysiert (vgl. Steyer, Eid & Schwenkmezger 1997). Zur Prüfung der Hypothesen wurden die mathematische Modellierungskompetenz, die Lesekompetenz und die Grammatikkompetenz simultan in ein Modell aufgenommen. Als Kontrollvariablen wurden die arithmetische Kompetenz und die allgemeine kognitiven Grundfähigkeiten einbezogen.

Die Schülerinnen und Schüler zeigten in allen Leistungsbereichen einen erheblichen durchschnittlichen Lernzuwachs vom ersten Messzeitpunkt zu Beginn der dritten Jahrgangsstufe bis zum zweiten Messzeitpunkt am Ende der dritten Jahrgangsstufe. Es ergab sich nicht nur für Leseverstehen, sondern darüber hinaus auch für die Grammatikkompetenz ein Zusammenhang mit der mathematischen Modellierungskompetenz. Während die sprachlichen Kompetenzen zu Beginn der Untersuchung keinen bedeutsamen Beitrag zur Erklärung der Entwicklung der mathematischen Modellierungskompetenz im Verlauf der dritten Jahrgangsstufe leisteten, waren die Veränderungen des Leseverständnisses und die Veränderungen in der Grammatikkompetenz erwartungskonform positiv mit der Entwicklung der mathematischen Modellierungskompetenz assoziiert. Insgesamt liefern die Befunde wichtige Hinweise darauf, dass Zweitsprachenlernende mit der Herausforderung konfrontiert sein dürften, gleichzeitig mathematische Kompetenzen und Fähigkeiten in der Instruktionssprache entwickeln zu müssen.

Donnerstag
12. März

9:00 Uhr

GC 03/142

Chair(s)
Hoffmann-Erz, Ruth

3

Buddeberg, Klaus; Riekman, Wibke

Studie zum mitwissendem Umfeld funktionaler Analphabetinnen und Analphabeten

Theoretischer Hintergrund: In der Erwachsenenbildungsforschung spielt Forschung zu Grundkompetenzen (Literalität, Alltagsmathematik) eine wachsende Rolle. Ungeachtet großer Zahlen betroffener Personen, wie sie PIAAC (Rammstedt 2013) oder die leo.-Studie (Grotlüschen und Riekman 2012) beziffert haben, finden nur wenige von ihnen den Weg zur institutionalisierten Erwachsenenbildung. Mehreren Millionen Personen mit geringer Literalität stehen jährlich etwa 20.000 Teilnehmende an Lese- und Schreibkursen der Volkshochschulen gegenüber. Die zentrale Frage, die sich daraus ergibt, ist die der Erreichbarkeit/Adressierbarkeit.

Bekanntermaßen sind Teilnahmequoten an Erwachsenenbildung von Geringqualifizierten (Bilger et al. 2013, S. 74–79) und in noch stärkerem Maße von Personen mit mangelnden Grundqualifikationen (Bilger 2012; Autorengruppe Bildungsberichterstattung, S. 149) niedrig. Im Förderschwerpunkt des BMBF „Arbeitsplatzorientierte Alphabetisierung und Grundbildung Erwachsener“ widmen sich zahlreiche Projekte der Erreichbarkeit von Betroffenen über Betriebe, bei denen sie beschäftigt sind. Die „Umfeldstudie – Studie zum mitwissendem Umfeld funktionaler Analphabetinnen und Analphabeten“ weitet diesen Fokus und betrachtet das gesamte Umfeld von Personen mit geringen Lese- und Schreibfähigkeiten. Dies basiert auf der Annahme, dass Betroffene in ihrem Umfeld Vertrauenspersonen haben, die sie bei der Bewältigung von Lese- und Schreibaufgaben unterstützen (Döbert und Hubertus 2000, S. 70) und als „doppelter Adressat“ für die Erreichbarkeit dienen können (Grotlüschen et al. 2014, S. 119–121).

Fragestellung: Dass Personen mit geringer Literalität in der Regel Unterstützung aus ihrem Umfeld erhalten, kann als sicher gelten. Weitgehend unbekannt sind das quantitative Ausmaß dieses „Mitwissens“ sowie die verschiedenen Ausformungen der Dyaden von Betroffenen und Mitwissenden im Sinne einer Typologie des Mitwissens (pragmatisch, paternalistisch, offen, verdeckt, tabubelastet etc.). Zudem stellt sich die Frage, in welchen Feldern derartiges Wissen vorliegt (Familie, Beruf, Freizeit). Und schließlich ist aus Sicht der Erwachsenenbildungsforschung zu fragen, ob die Unterstützung betroffener Personen zum informellen Lernen oder gar zum Lernen im Kurs ermutigt oder ob sie eher zu einem Verharren in der Situation führt.

Von großer Heterogenität ist auf zwei Ebenen auszugehen: Innerhalb der Gruppe der Betroffenen dürfte es sich sowohl um Personen handeln, die im Sinne der leo.-Diktion (funktionale) Analphabet/inn/en sind, als auch um Personen mit stark fehlerhafter Schreibung (Grotlüschen et al. 2012). Von Heterogenität kann auch auf Seiten der Typen des Mitwissens ausgegangen werden. So soll geprüft werden, inwieweit sich berufliches Mitwissens qualitativ von familiärem Mitwissen unterscheidet.

Methode: Die Studie besteht aus zwei gleichberechtigten Teilstudien. Im Rahmen der qualitativen Teilstudie wurden qualitative Leitfadeninterviews geführt und ausgewertet (orientiert an der Grounded Theory). Im Rahmen der quantitativen Teilstudie wurden 1.500 Erwachsene in Hamburg telefonisch befragt. Beide Teilstudien sind im Sinne eines Mixed-Methods-Ansatzes vielfach aufeinander bezogen. In der Vorbereitung und Durchführung sowie aktuell in der Auswertung stehen die jeweiligen methodischen Stärken beider Teilstudien im Vordergrund (Kelle und Erzberger 2000 / 2013; Denzin 1978), um die Teilfragen mit einem möglichst angemessenen methodischen Instrumentarium beantworten zu können.

Ergebnisse: Beide Teilstudien befinden sich zurzeit (Herbst 2014) in der Phase erster Auswertungen. Die Inzidenz des Mitwissens ist erwartet hoch, d.h. etwa jeder dritte Erwachsene kennt eine oder mehrere Personen, die ungewöhnliche Probleme beim Lesen und/oder Schreiben haben. Dieses Wissen führt häufig zu einem offenen Umgang mit dem Phänomen (die Lesart des funktionalen Analphabetismus als großes Tabu ist einer kritischen Prüfung zu unterziehen), es resultiert aber nur selten daraus, dass Betroffene aktiv zum Lernen ermutigt werden. Hinsichtlich der Typologie des Mitwissens ist für das Frühjahr 2015 mit soliden Ergebnissen zu rechnen. Bereits jetzt wird aber deutlich, dass sich das familiäre Mitwissen von dem im beruflichen oder sonstigen Umfeld strukturell unterscheidet.

Die Gesamtschau der Ergebnisse wird ermöglichen, die für die Adressierung von Lernenden relevanten gesellschaftlichen Felder zu identifizieren und das dort zu vermutende Mitwissen zu quantifizieren. Auf Basis der Kenntnis über verschiedene Formen des Mitwissens (Typologie) sollen praxisrelevante Empfehlungen über geeignete Ansprache- und Unterstützungsformen entwickelt werden.

Donnerstag
12. März

9:00 Uhr

GC 03/149

Chair(s)
Buddeberg, Klaus

Greiten, Silvia

Unterrichtsplanung im Kontext von Inklusion - Befunde zu Konzeptveränderungen von Sonderpädagogen und Regelschullehrkräften durch den Unterricht mit integrativen Lerngruppen

Theoretischer Hintergrund: Die Heterogenität der Lerngruppen nimmt aufgrund der Auswirkungen von Schulstrukturveränderungen und der Umsetzung von Vorgaben zur Inklusion zu. Mit diesen Entwicklungen stellt sich für Lehrpersonen die Frage nach Unterrichtskonzeptionen, die den Anforderungen an Lernsituationen in zielgleichen und zieldifferenten Unterrichtssettings gerecht werden. Für in Regelschulen der Sekundarstufen unterrichtende Sonderpädagogen ergeben sich weitere Fragen, da sie sich in der jeweiligen Schulform über die förderspezifischen Curricula hinaus auch an fachspezifischen Curricula orientieren müssen (Greiten, 2014a, 2014b, 2015).

Aus der Forschung ist bekannt, dass Unterricht an Regelschulen häufig gleichförmig nach dem Rhythmus Einstieg, Erarbeitung und Sicherung verläuft (Arnold & Koch-Priewe, 2010) und Lehrkräfte sich an dem Wahrnehmungsraster ‚Klasse‘ orientieren (Kunze, 2008) – mit wenig Raum für individuelles Lernen. Unterrichtskonzeptionen für inklusiven Unterricht stellen das ‚Gemeinsame Lernen‘ in den Mittelpunkt (Feyerer & Prammer, 2003). Forschungsdefizite zum Handlungswissen auf der Basis einer inklusiven Didaktik sind festzustellen (Klauß, 2014).

Fragestellung: In der seit Mai 2012 laufenden qualitativ empirischen Studie „Einstellungen und Erfahrungen von Lehrerinnen und Lehrern zur individuellen Förderung und Inklusion im Unterricht“ wird u. a. der Frage nachgegangen, welche Erfahrungen und Konzepte Lehrpersonen in die Unterrichtsplanung einfließen lassen, um ihren Unterricht auf individuelle Förderung und Inklusion auszurichten. Die Stichprobenwahl der Interviewten orientiert sich an folgender Clusterzuordnung: Regelschullehrkräfte ohne integrative Erfahrung, Sonderpädagogen ohne integrative Erfahrung, Regelschullehrkräfte mit integrativer Erfahrung und Sonderpädagogen mit integrativer Erfahrung.

Die hier vorgestellte Teilstudie „Konzeptveränderung von Sonderpädagogen und Regelschullehrkräften durch den Unterricht mit integrativen Lerngruppen“ konzentriert sich auf zwei Fragestellungen: (1) Inwiefern verändern sich Konzepte zur Unterrichtsplanung, wenn Sonderpädagogen und Regelschullehrkräfte der Sekundarstufen in integrativen Lerngruppen unterrichten und die Lehrkräfte damit eine Kontrastierung ihrer bisherigen und neuen Erfahrungen erleben? (2) Welche Ansätze für die Konzeptentwicklung zur Unterrichtsplanung für heterogene Lerngruppen lassen sich aus den Erfahrungen der Befragten erheben, um diese für empirischer Studien zu Konzepten inklusiven Unterrichts nutzen zu können?

Methode: Durch Leitfadenterviews mit Lehrpersonen aller Schulformen der Sekundarstufen und insbesondere auch mit Sonderpädagogen werden u. a. Beschreibungen und Erfahrungen zur Unterrichtsplanung für heterogene Lerngruppen erhoben. Dieser Teilstudie liegen jeweils fünf Transkripte aus den Clustern ‚Sonderpädagogen mit integrativer Erfahrung‘, Fachrichtungen Sprache, Lernen sowie soziale und emotionale Entwicklung, und ‚Regelschullehrer mit integrativer Erfahrung‘ zugrunde. Die Auswertung erfolgt mit der qualitativen Inhaltsanalyse. Es ist das Ziel, Veränderungsprozesse der Konzepte zur Unterrichtsplanung von Sonderpädagogen und Regelschullehrkräften im Kontrast des Unterrichts in integrativen und nicht integrativen Lerngruppen nachzuzeichnen und daraus Ansätze für Konzepte zum Unterricht mit heterogenen Lerngruppen zu entwickeln.

Ergebnisse: In der Studie kann zunächst dokumentiert werden, dass die jeweilige Rollenzuweisung in der Kooperation zwischen Sonderpädagogen und Regelschullehrkräften im integrativen Unterricht über die Weiterentwicklung der Kompetenzen und Konzepte zur Unterrichtsplanung mitbestimmt. Werden die Rollen im Sinne des Spezialisten für Förderung einerseits und für Fachunterricht andererseits stabilisiert, ändern sich die jeweiligen Planungskonzepte kaum. Planen aber beispielsweise Sonderpädagogen auch Unterricht für die gesamte Lerngruppe, sind es v. a. die Kompetenzen der inhalts- und niveaubezogenen Differenzierung sowie der gleichzeitigen förderbezogenen Perspektive, die auf die Unterrichtsplanung für die gesamte Klasse wirken. Für die Regelschullehrkräfte lässt sich beschreiben, dass sie ihre Konzepte trotz der Unterrichtssituation in der integrativen Lerngruppe häufig nur auf der strukturellen Ebene wie Gruppenzuweisungen oder Niveaudifferenzierung modifizieren, deutlich weniger sind förderdiagnostische Akzente zu erkennen. Hingegen ist für sie die Orientierung an den fachlichen Zielen und den Leistungen ein wichtiger Bezugspunkt ihrer Unterrichtsplanung. Wie die Interviews dokumentieren, können beide Professionen ihre Konzepte durch die Auseinandersetzung mit Lerngruppen im integrativen Unterricht und durch die kollegiale Kooperation im Hinblick auf kompetenz- und förderorientierten Unterricht verändern. Aus diesen Prozessbeschreibungen lassen sich Ansätze für Konzepte zur Unterrichtsplanung für heterogene Lerngruppen entwickeln.

Keywords: Unterrichtsplanung, heterogene Lerngruppen, Inklusion, Lehrerbildung

Donnerstag
12. März

9:00 Uhr

GC 03/149

Chair(s)
Buddeberg, Klaus

Kullmann, Harry; Geist, Sabine; Lütje-Klose, Birgit

Schulisches Wohlbefinden als Qualitätsmerkmal inklusiver Schulen – Ausgewählte Befunde einer jahrgangsübergreifenden Längsschnittstudie an der Laborschule Bielefeld

Theoretischer Hintergrund: Schulisches Wohlbefinden ist eine Grundbedingung gelingender Lernprozesse bei Kindern und Jugendlichen. Es bildet einen Teil der je individuellen, emotional-kognitiven Rahmung aller vom Schul- und Unterrichtsbesuch ausgehenden Bewältigungs- und Entwicklungsaufgaben. Die Ausprägung des Wohlbefindens beruht auf der Balance des Erlebens positiv konnotierter Emotionen und Kognitionen sowie des Fehlens von negativen Emotionen und Kognitionen.

Gerade Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SFB) haben in inklusiven Settings durch den sozialen Vergleich mit mehrheitlich leistungsstärkeren Peers aufgrund von Bezugsgruppeneffekten besondere Herausforderungen zu bewältigen (z.B. Seaton et al. 2010). Das Erreichen eines hohen Wertes subjektiv wahrgenommenen Wohlbefindens der gesamten Lerngruppe ist ein Gradmesser schulischer Inklusivität und eine Moderatorvariable zur Erklärung des schulischen Leistungsniveaus (z.B. Fend und Sandmeier 2004; Hascher und Hagenauer 2011).

Zur begleitenden Evaluation der schulischen Maßnahmen im Rahmen der Implementation der VN-Behindertenrechtskonvention ist es wünschenswert, das Wohlbefinden in inklusiven Lerngruppen sowohl zuverlässig als auch ökonomisch erfassen zu können. Eine besondere Schwierigkeit auf dem Weg zur Realisierung dieses Zieles liegt darin begründet, dass es für die Sekundarstufe I keine hinreichend geprüften Fragebogeninstrumente aus Erhebungen in inklusiven Lerngruppen gibt. Letzteres wiederum ist u.a. dem Umstand geschuldet, dass bislang nur sehr wenige Schulen dieser Schulstufe eine hinreichend lange Erfahrung als inklusive Schule aufweisen um als Basis für betreffende Evaluationsstudien in Frage zu kommen.

Die Laborschule Bielefeld verfügt als Versuchsschule des Landes NRW über jahrzehntelange Erfahrung in der gemeinsamen Unterrichtung von Schülerinnen und Schülern mit und ohne einen SFB, die grundsätzlich in allen und nicht nur in ausgewählten Lerngruppen umgesetzt wird (Thurn und Tillmann 2011). Die Laborschule arbeitet nach einem inklusiven Ansatz, wie er zukünftig verstärkt auch in Regelschulen vorzufinden sein wird und bietet insofern einen idealen Untersuchungsrahmen für die Entwicklung und Erprobung eines Instruments zur Erfassung des schulischen Wohlbefindens in inklusiven Lerngruppen der Sekundarstufe I.

Fragestellungen: Erstens: Lässt sich ein Fragebogeninstrument konstruieren, welches verschiedene Facetten schulischen Wohlbefindens in mehreren Jahrgängen der Sekundarstufe I und in allen relevanten Subgruppen der Schülerschaft zuverlässig misst?

Zweitens: Wie ist das schulische Wohlbefinden jeweils ausgeprägt – in Gruppen, zwischen Gruppen und jeweils längsschnittlich – und welche Zusammenhänge mit Bedingungsfaktoren lassen sich herausarbeiten?

Von besonderem Interesse sind jeweils – vor dem Hintergrund eines erweiterten Inklusionsbegriffs – die Befunde entlang inklusionsrelevanter Diversitätskriterien wie SFB, Geschlecht und Migrationshintergrund.

Methode: Dreimalig durchgeführt wurde eine standardisierte Fragebogenerhebung in den Jahrgangsstufen 6 bis 9 (im Jahr 2012) bzw. 6 bis 10 (2013 und 2014) an der Laborschule Bielefeld. Die Ausschöpfungsquote betrug jeweils mindestens 86% bei N = 212, 292 bzw. 300 und einem Anteil der Schülerinnen und Schülern mit SFB bei ca. 10%.

Die Datenauswertung erfolgte anhand von deskriptiver Statistik, Inferenzstatistik sowie mittels Strukturgleichungsmodellierungen.

Ergebnisse: Der Fragebogen misst alle sechs Facetten des schulischen Wohlbefindens (Einstellungen und Emotionen gegenüber der Schule, Affinität zur Stammgruppe/Klasse, Schulischer Selbstwert, Abwesenheit von Sorgen in der Schule, von sozialen Problemen sowie von körperlichen Beschwerden) auf der Basis von jeweils mindestens vier Items hinreichend zuverlässig (mit Cronbachs $\alpha \geq .779$).

Konfirmatorische Faktorenanalysen zeigen, dass die beste Modellanpassung ein Modell mit sechs korrelierten Faktoren liefert. Ein Vergleich der Schülerinnen und Schüler mit und ohne SFB ergibt in mehreren Erhebungswellen keine Unterschiede für die Wohlbefindensdimensionen über die Abwesenheit schulischer Sorgen, sozialer Probleme sowie körperlicher Beschwerden. Diese liegen jeweils nur in geringem Umfang vor. Die größten Unterschiede zeigen sich durchgehend für den schulischen Selbstwert, der am stärksten von Bezugsgruppeneffekten beeinflusst wird.

Die Gruppenunterschiede für die Hintergrundvariablen Geschlecht und Migrationshintergrund sind in Bezug auf viele Teildimensionen des Wohlbefindens gar nicht vorhanden oder liegen im Bereich einer geringen Effektstärke. Diverse Gruppenvergleiche, z.B. zum Verlauf über die Jahrgangsstufen der Sekundarstufe I, und Analysen zum Einfluss ausgewählter Bedingungsfaktoren auf das schulische Wohlbefinden, werden im Vortrag vorgestellt und diskutiert.

Donnerstag
12. März

9:00 Uhr

GC 03/149

Chair(s)
Buddeberg, Klaus

Nessler, Stefan

Inklusive Fachdidaktik in der Lehramtsausbildung für das Fach Biologie in der Sekundarstufe I - Entwicklung und Evaluation eines interdisziplinären Seminars

Theoretischer Hintergrund: Das Lehren und Lernen in inklusiven Schulsettings ist seit Ratifizierung der UN-Bundesrechtskonvention (2009) eine der größten Herausforderungen, der sich die Lehramtsausbildung in Deutschland stellen muss. Dies liegt nicht zuletzt daran, dass der Erfolg von inklusivem Unterricht vor allem von einer gelungenen praktischen Umsetzung abhängt und somit insbesondere in der Verantwortung von LehrerInnen liegt. Deswegen ist es für eine LehrerInnen-Professionalisierung unumgänglich, dass Inklusion neben der fachlichen und didaktischen Ausbildung ein fester Bestandteil der Lehrer(aus)bildung werden muss.

Allerdings ist die universitäre Lehramtsausbildung nur unzureichend auf eine inklusive Lehramtsausbildung vorbereitet (Schuppener 2014). Dies liegt unter anderem auch daran, dass die Ausbildung von Sonderpädagogen in ihren Spezialgebieten, wie z.B. Förderbedarf Lernen oder Sprache, und der Regelschullehrern im Sekundarbereich häufig getrennt voneinander stattfinden (Merz-Atalik 2014). Auf diesem Weg ist es schwer, eine Vernetzung zwischen Inklusion, Fachdidaktik und inklusiver Pädagogik beider Studiengänge zu erreichen, wodurch die Umsetzung von Inklusion in Regelschulen erschwert wird (Demmer-Dieckmann 2010; Reich 2012).

Vielversprechende Ansätze für eine inklusive Lehrerausbildung sind vor allem integrierte (integrated models) und verschmolzene Ausbildungsgänge (merged models; Plugach & Blanton 2009). Jedoch gibt es derzeit kaum empirische Untersuchungen, die den Erfolg dieser Modelle unterstützen (Pugach et al. 2014).

An der Universität zu Köln wurde deswegen ein Seminar „Inklusive Fachdidaktik Biologie für die Sekundarstufe I“ nach dem Konzept des „integrierten Ausbildungsmodell“ entwickelt, in welchem die Studiengänge der Sonderpädagogik aller Förderschwerpunkte mit dem Lehramtsstudium der Sekundarstufe I im Fach Biologie mit dem Ziel verzahnt wurden, Lehramtsstudierende für inklusiven Unterricht vorzubereiten. Dafür werden interdisziplinäre Teams aus Sonderpädagogen und Regelschullehrern gebildet, welche i) fachspezifisches Wissen der Förderbedarfe und der Fachdidaktik der Sekundarstufe I austauschen können und ii) eine selbst erarbeitete Unterrichtsstunde mit einer inklusiven Schulklasse im Team-Teaching durchführen. Für die Entwicklung des Seminars wurde sich an verschiedenen Merkmalen von Plugach & Blanton (2009) orientiert: (a) fakultätsübergreifende Kooperation: Das Seminar wurde durch ein multiprofessionelles Team aus Mitarbeitern der Humanwissenschaftlichen Fakultät, Department Heilpädagogik, und der Mathematische-Naturwissenschaftlichen Fakultät, Institut für Biologie und Ihre Didaktik konzipiert (Krämer et al. 2014), (b) Tiefe des Wissens: Inhaltlich wird der Fokus auf die drei Bereiche (1) Inklusion, (2) Forschendes Lernen und (3) Fachdidaktik gelegt. In den jeweiligen Bereichen bereiten sich Studierende intensiv mittels Expertenlisten auf die jeweiligen Fachinhalte vor, (c) Beurteilung von Lehr- und Lernergebnissen: Mittels eines Lerntagebuches und ConceptMaps werden die Lernprozesse der Studierenden festgehalten, (d) Zusammenarbeit zwischen Universität und Schule: Die durchzuführende Unterrichtsstunde als auch eine Hospitation finden in Kooperation mit inklusiven Schulen statt.

Fragestellung: Vor dem Hintergrund, eine möglichst vielschichtige inklusive Lehramtsausbildung zu gewährleisten, werden folgende Fragen gestellt: (1) Kann das Seminar dazu beitragen, ein umfangreiches und vernetztes Wissen in Bezug auf Inklusion, Fachdidaktik und Praxis herzustellen? (2) Kann das Seminarkonzept Einstellungen, Haltungen und Bedenken zu inklusiver Pädagogik sowie die Selbstwirksamkeit beeinflussen?

Methoden: Die Evaluation des Seminars erfolgt zum einen durch ein Pre-Post-Test-Design, in welchem Studierende zu Beginn jeweils eine ConceptMap (Novak & Gowin 1984) zu der Fokusfrage „Wie sind Biologie, Didaktik und Inklusion vernetzt?“ anfertigen. Des Weiteren werden für die Erhebung der Einstellungen, Haltungen, Bedenken zu inklusiver Pädagogik sowie der Selbstwirksamkeit die Fragebögen SACIE-R und TEIP (Feyerer et al. 2014) verwendet. Zum anderem wird der Lernprozess über Lerntagebücher festgehalten. Die Fokusaufgabe des Lerntagebuchs ist eine persönliche Stellungnahme der Studierenden zu dem Thema „Inklusiver Biologieunterricht in der Sekundarstufe I“ und wird zu Beginn des Seminars verfasst. Danach bekommen die Studierenden einmal während des Seminars und zum Abschluss die Gelegenheit ihre verfassten Texte jeweils zu überarbeiten. Das Lerntagebuch wird qualitativ nach Mayring (2010) ausgewertet.

Ergebnisse: Da sich bei der Evaluation um ein laufendes Seminar im Wintersemester 2014/2015 handelt, liegen die ersten Ergebnisse Anfang Februar 2015 vor und können somit für die Tagung präsentiert werden.

Donnerstag
12. März

9:00 Uhr

GC 03/149

Chair(s)
Buddeberg, Klaus

Glock, Sabine; Böhmer, Ines

Der Migrationshintergrund von Schülern als implizite und explizite Assoziationen: Ein Vergleich von Lehrerexperten und -novizen

Theoretischer Hintergrund: Im gesamten Bildungsverlauf von Schülern spielt die Leistungsbewertung durch die Lehrkraft eine zentrale Rolle. Studien zeigen auf, dass bei der Notenvergabe nicht nur die tatsächliche Leistung, sondern auch sachfremde Kriterien wie der Migrationshintergrund (MH) einfließen (Ditton & Krüsken, 2009). Türkischstämmige Schüler erhalten bei gleicher Leistung schlechtere Noten als ihre deutschstämmigen Mitschüler (Kristen, 2006). In diesem Zusammenhang kann die Wirkung von stereotypen Lehrererwartungen gegenüber Schülern mit MH einen Beitrag zur Erklärung der Notendisparitäten liefern (Flam, 2007). Stereotype Erwartungen beeinflussen sowohl die Lehrer-Schüler-Interaktion (van den Bergh, Denessen, Hornstra, Voeten & Holland, 2010) als auch das Urteil der Lehrpersonen über ihre Schüler (Glock & Krolak-Schwerdt, 2013). Stereotype Erwartungen werden als kognitive Assoziationen konzeptualisiert und in dualen Prozessmodellen in ihrem Einfluss auf das Verhalten spezifiziert. Ein duales Prozessmodell ist das Reflective-Impulsive-Model (Strack & Deutsch, 2004), welches implizite, automatische und explizite, kontrollierte Prozesse unterscheidet, die von impliziten und expliziten Assoziationen beeinflusst werden. In Anlehnung an das Reflective-Impulsive-Model (Strack & Deutsch, 2004) untersuchte diese Studie die expliziten und impliziten Assoziationen als Bestandteil stereotyper Erwartungen von Lehrpersonen und Lehramtsstudierenden über ihre Schüler mit und ohne MH.

Fragestellung: (1) Sind die impliziten Assoziationen zwischen Schülern mit MH und Eigenschaften negativ? (2) Sind auch die expliziten Assoziationen negativ? (3) Haben Experten, die über mehr Erfahrung mit Schülern mit MH verfügen, andere explizite und implizite Assoziationen?

Methode: Insgesamt nahmen 46 Gesamtschullehrkräfte (25 weiblich, Alter: 44.65, SD=12.64, Erfahrung: 15.88 Jahre, SD=12.64) und 40 Lehramtsstudierende (29 weiblich, Alter: 25.63, SD=3.13, Berufserfahrung: 13 Wochen, SD=12.58).

Zur Erfassung der impliziten kognitiven Assoziationen zwischen Schülern mit bzw. ohne MH und positiven bzw. negativen schulischen Eigenschaften wurde ein Implicit Association Test (IAT; Greenwald, McGhee & Schwartz, 1998) durchgeführt. Die Aufgabe der Probanden bestand darin, sowohl Bilder von Schülern mit bzw. ohne MH als auch positive bzw. negative schulische Eigenschaften per Tastendruck in Kategorien einzuordnen. In einem Durchgang wurden Schüler mit MH und negative Eigenschaften auf einer Taste und Schüler ohne MH und positive Eigenschaften auf der anderen Taste gepaart. In einem anderen Durchgang wurde die Paarung gedreht. Erfasst wurde die Reaktionszeit. Nach dem IAT, wurden die expliziten Assoziationen mit Hilfe des Fragebogens „Multikulturelle Überzeugungen“ erfasst (Hachfeld, et al., 2012). Den impliziten Assoziationen lag ein 2x2 mixed Design mit Expertise (Experten vs. Novizen) als Zwischensubjektfaktor und Kombination (Schüler mit MH und negative Eigenschaften vs. Schüler mit MH und positive Eigenschaften) zugrunde. Den expliziten Assoziationen lag ein einfaktorielles Between-Subjects-Design mit Expertise als Faktor zugrunde.

Ergebnisse: Die impliziten Assoziationen zeigten keinen Unterschied zwischen Experten und Novizen. Bei beiden Gruppen konnte eine starke implizite Assoziation zwischen Schülern mit MH und negativen Eigenschaften gefunden werden. Im Gegensatz dazu zeigten sich im Fragebogen deutlich positive Ausprägungen expliziter Assoziationen gegenüber Schülern mit MH.

Diskussion: Die Befunde stützen die Annahme, dass ein Bestandteil stereotyper Erwartungen die impliziten Assoziationen zwischen Schülern mit MH und negativen Eigenschaften sind. Dies steht in starkem Kontrast zu den positiven expliziten Assoziationen. Dabei liegt der Schluss nahe, dass Lehrkräfte, wenn sie explizit gefragt werden, ihre impliziten Assoziationen nicht bewusst und offen berichten (van den Bergh et al., 2010). Die theoretischen Annahmen des RIM-Modells spezifizieren den Einfluss automatischer und kontrollierter Prozesse auf das Verhalten, wobei automatische Prozesse und somit auch implizite Assoziationen als an allen kognitiven Prozessen beteiligt angesehen werden (Strack & Deutsch, 2004). Dies impliziert, dass Lehrkräfte bei jedem unterrichtlichem Handeln durch die negativen impliziten Assoziationen gegenüber Schülern mit MH beeinflusst werden. Daher sollte in weiteren Studien untersucht werden, ob sich implizite Assoziationen auf die Lehrer-Schüler-Interaktion und auf die Beurteilung von Schülern mit MH auswirken. Daraus könnten wertvolle Hinweise gewonnen werden, ob die negativen impliziten Assoziationen einen Beitrag zur Chancengleichheit im deutschen Bildungssystem liefern.

Donnerstag
12. März

9:00 Uhr

GC 03/149

Chair(s)
Buddeberg, Klaus

Schell, Annika; Kühl, Tim; Münzer, Stefan

Lernen mit zweidimensionalen im Vergleich zu dreidimensionalen Visualisierungen im Mathematikunterricht unter Berücksichtigung des räumlichen Vorstellungsvermögens

Theoretischer Hintergrund: In der Studie wurde das Lernen mit zweidimensionalen Visualisierungen (2D-Gruppe) im Vergleich zu dreidimensionalen Visualisierungen (3D-Gruppe) im Mathematikunterricht in Abhängigkeit der räumlichen Fähigkeiten von Schülerinnen und Schüler (SuS) untersucht.

Die für eine sechste Klassenstufe curricular validen Lerninhalte adressierten das Erkennen rechter Winkel innerhalb räumlicher Objekte (z.B. Quader, Schrägdach, Pyramide). Um einen optimalen dreidimensionalen Eindruck der Objekte zu vermitteln wurden die Visualisierungen in der 3D-Gruppe stereoskopisch dargeboten (hierfür mussten SuS entsprechende polarisierende 3D-Brillen aufsetzen). Der 2D-Gruppe wurden die gleichen Visualisierungen dargeboten – allerdings nicht stereoskopisch. Es wurde angenommen, dass für das Erlernen des Erkennens der rechten Winkel innerhalb räumlicher Objekte dreidimensionale Visualisierungen lernförderlich sind. Darüber hinaus wurde entsprechend der Ability-as-Compensator-Hypothese (vgl. Höffler, 2010) angenommen, dass die 3D- im Vergleich zu 2D-Visualisierungen besonders SuS mit niedrigen räumlichen Fähigkeiten zu besseren Lernerfolgen verhelfen sollte, während für SuS mit hohen räumlichen Fähigkeiten der Einfluss des Visualisierungsformats eine untergeordnete Rolle spielen sollte.

Methode: Für die Studie wurde ein 2x3-Versuchsplan mit den unabhängigen Variablen Visualisierungsart (3D- vs. 2D-Visualisierung) und räumliche Fähigkeit (gering, mittel, hoch) verwendet. Die Studie wurde mit SuS (n = 64) mehrerer sechsten Klassen an einer Gemeinschaftsschule durchgeführt. Die räumliche Fähigkeit wurde mithilfe des Paper Folding Tests (PFT), der die Komponente Visualization testet, und einem Figurenidentifikationstests (FIT), der die Komponente Closure Flexibility testet, erhoben. Anhand eines Wissenstests wurde der Lernerfolg überprüft. Ferner wurde mittels eines Fragebogens die Auswirkung auf die Motivation erfasst.

Ergebnisse: Die Auswertung des Motivationsfragebogens ergab, dass die unterschiedlichen Bedingungen keinen Einfluss auf die Motivation der SuS hatten. Die Ergebnisse des Wissenstest zeigten keinen Haupteffekt hinsichtlich der Visualisierungsart: Dass die 3D-Lerneinheit ein besseres Treatment darstellte, als die herkömmliche 2D-Lerneinheit, konnte für die Gesamtheit der 2D- bzw. 3D-Gruppe nicht nachgewiesen werden. Für den Faktor räumliches Vorstellungsvermögen zeigte sich ein Haupteffekt: Ein höheres räumliches Vorstellungsvermögen führte zu einem höheren Lernerfolg. Die Annahme, dass Lernende mit niedrigen räumlichen Fähigkeiten in der 3D-Gruppe bessere Performanz im Wissenstest zeigen als Lernende der 2D-Gruppe, wurde nicht bestätigt. Die 3D-Lerneinheit konnte die niedrigen räumlichen Fähigkeiten nicht kompensieren. Für SuS mit höheren räumlichen Fähigkeiten (im FIT) zeigte sich ein überraschendes Befundmuster: Hier schnitten SuS mit 2D-Visualisierungen besser ab als SuS mit 3D-Visualisierungen!

Diskussion: Dieser letztgenannte Befund steht nicht im Einklang mit der ursprünglich angenommenen Ability-as-Compensator-Hypothese. Er steht jedoch auch nicht im Einklang mit der in der Literatur gelegentlich vertretenen Ability-as-Enhancer-Hypothese (z.B. Huk, 2006). Diese besagt, dass der Vorteil von 3D-Visualisierungen vornehmlich nur bei Lernenden, die hohe räumliche Fähigkeiten besitzen, zum Tragen kommt und somit ein Lernvorteil von 3D- gegenüber 2D-Visualisierungen nur für Lernende mit hohen räumlichen Fähigkeiten sichtbar wird.

Der Befund, dass für SuS mit hohen räumlichen Fähigkeiten 2D-Visualisierungen zu einem besseren Ergebnis führen als 3D-Visualisierungen, lässt sich wohlmöglich mit einer Verständnisillusion und damit einhergehenden Hemmfunktion der 3D-Visualisierungen erklären (inhibitory function; Schnotz & Rasch, 2005): Die 3D-Visualisierungen wirken zu leicht um tiefere und lernförderliche mentale Verarbeitungsprozesse anzustoßen, wodurch das Potential von SuS mit hohen räumlichen Fähigkeiten nicht aktiviert wird und ein besseres Verständnis ausbleibt.

Zusammenfassend lässt sich festhalten, dass die stereoskopisch dargebotenen 3D-Visualisierungen für das Erlernen des Erkennens der rechten Winkel innerhalb räumlicher Objekte nicht lernförderlich waren, sondern für SuS mit höheren räumlichen Fähigkeiten sogar lernhinderlich. Um die Stabilität dieses überraschenden Effekts zu überprüfen ist eine Replikationsstudie in einem vergleichbaren Setting erforderlich und geplant.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Häusler, Janina

Leistungsstarke Schülerinnen und Schüler im Deutsch- und Mathematikunterricht

Theoretischer Hintergrund: SchülerInnen unterscheiden sich hinsichtlich ihrer kognitiven und motivational-affektiven Voraussetzungen, welche eine wichtige Rolle im Klassenzimmer einnehmen (Perry, Turner, & Patrick, 2004; Snow, Corno, & Jackson, 1996). Ergebnisse der IPN-Videostudie zeigten, dass sich leistungsstarke SchülerInnen im Fach Physik durchweg durch hohe individuelle Voraussetzungen auszeichnen und die Unterrichtsqualität höher einschätzen als ihre Mitschüler (Seidel et al., 2006). Allerdings muss beachtet werden, dass sich bedeutsame Geschlechterunterschiede in den Schülerleistungen zeigen. Im mathematisch-naturwissenschaftlichen Bereich sind Jungen den Mädchen überlegen, während Mädchen im Fach Deutsch im Allgemeinen bessere Leistungen zeigen als Jungen. Hinsichtlich der Unterrichtsbeteiligung wurden starke SchülerInnen schon früh in den 1980ern seitens ihrer Lehrkräfte in zwei verschiedene Gruppen eingeteilt: starke SchülerInnen, die sich rege am Unterrichtsgespräch beteiligen, und eine zweite Gruppe an starken SchülerInnen, die ruhiger und zurückhaltender sind (Hofer, 1981). Begehr, 2004 und Jones et al., 2006 stützen diese Hypothese, denn sie konnten herausstellen, dass sich die Mehrheit der SchülerInnen nicht am Unterrichtsgespräch beteiligt, während es einige „Stars“ gibt, die häufig von ihren Lehrkräften aufgerufen werden. Generell gilt es jedoch zu beachten, dass die Schülerbeteiligung im Unterricht ist nicht nur von den individuellen Schülercharakteristika, sondern auch vom Lehrerverhalten, den Einstellungen und Erwartungen der Lehrkräfte abhängig sind (Jurik et al., 2013; Seidel & Reiss, 2014).

Fragestellungen: Der vorliegende Beitrag beschäftigt sich mit starken SchülerInnen im Deutsch- und Mathematikunterricht. Die individuellen Schülercharakteristika (kognitive Fähigkeiten, Interesse, Selbstkonzept, Vorleistung) und die Einschätzung der Unterrichtsqualität wurden per Fragebogen erhoben und zusammen mit Daten zur Beteiligung am Unterricht, gemessen durch die Häufigkeit und Dauer der Schülerbeiträge, aus Videoaufzeichnungen weiter analysiert.

Folgende Forschungsfragen stehen im Fokus dieses Beitrages: (1) Gibt es Unterschiede in der Anzahl an starken SchülerInnen und dem Geschlecht der SchülerInnen für die Fächer Deutsch und Mathematik? (2) Unterscheidet sich die Beteiligung starker SchülerInnen am Unterricht in den beiden Fächern von anderen SchülerInnen? (3) Schätzen starke SchülerInnen die Unterrichtsqualität in beiden Fächern anders ein als ihre Mitschüler?

Methode.

Stichprobe: Die Stichprobe aus dem DFG geförderten Projekt Interaction umfasst insgesamt N = 472 SchülerInnen aus 20 8. Klassen an bayerischen Gymnasien.

Design: Die SchülerInnen wurden zu Beginn und am Ende des Schuljahres 2013/14 hinsichtlich ihrer kognitiven Fähigkeiten, ihrem Interesse, ihrem Selbstkonzept und ihrer Vorleistung sowie der eingeschätzten Unterrichtsqualität befragt. Zum Schulhalbjahr wurden jeweils eine Deutsch- und Mathematikunterrichtsstunde videografiert.

Statistische Analysen: Die SchülerInnen wurden mittels LCA (Mplus) unterschiedlichen Schülerprofilen zugeordnet. Die Videos wurden mit der Software Mangold Interact kodiert und über die Sprechereinheiten ausgewertet. Des Weiteren wurden weitere statistische Analysen mit SPSS (Version 22) durchgeführt.

Ergebnisse: Erste Ergebnisse zeigen, dass es insgesamt im Fach Deutsch mehr starke SchülerInnen (n=149) gibt als in Mathematik (n=102). 47 der starken SchülerInnen sind in beiden Fächern stark. Im Fach Deutsch sind etwas mehr Mädchen leistungsstark als Jungen, in Mathematik etwas mehr Jungen leistungsstark als Mädchen. Bezüglich der Beteiligung am Unterrichtsgespräch stützen die Ergebnisse die Hypothese, dass es sowohl aktive, als auch ruhigere und zurückhaltende starke SchülerInnen gibt. Die aktiven melden sich häufiger und haben insgesamt einen höheren Sprechanteil in einer Unterrichtsstunde. Bei der Analyse der Unterrichtseinschätzungen zeigten sich keine Unterschiede. Starke SchülerInnen schätzten den Unterricht nicht positiver oder negativer ein als ihre Mitschüler.

In den kommenden Wochen folgen noch weitere Analysen zum Unterrichtsgespräch z.B. hinsichtlich der Qualität an Fragen und Rückmeldungen an starke Schüler im Vergleich zu den Mitschülern.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Laging, Angela

Der Einfluss der Erhebungsinstrumente auf den Zusammenhang von Selbstwirksamkeit und Leistung im Bereich Mathematik – Ergebnisse einer Meta-Analyse

Die Erwartungen der eigenen Selbstwirksamkeit (SWK) beeinflussen motivationale Prozesse wie die Wahl von Aufgaben/Aktivitäten, die investierte Anstrengung, die Persistenz und emotionale Reaktionen sowie Prozesse des selbstregulierten Lernens (Zimmerman, 2000). Der Einfluss auf die Motivation, das Lernen und die Leistung wurde bereits für verschiedene Fächer in einigen Studien bestätigt (Pajares, 1996), so auch für Mathematik. Multon, Brown und Lent (1991) konnten den positiven Zusammenhang zwischen Selbstwirksamkeitserwartungen und akademischen Leistungen in ihrer Meta-Analyse mit einer durchschnittlichen Effektstärke (ES) von 0,38 nachweisen (n=38). Es wurden vier Moderatorvariablen identifiziert, die jedoch keine in sich homogenen Untergruppen gebildet haben: Leistungserfassung, Zeitpunkt der ES-Berechnung, Leistungsniveau und Klassenstufe bzw. Alter. In die Analysen wurden weder die Domänen noch die Art der Erhebung der Selbstwirksamkeit einbezogen. Dabei ist die Erfassung der Selbstwirksamkeit von besonderer Bedeutung. Aufgrund der Aufgaben- und Bereichsspezifität ist eine globale Erfassung der Selbstwirksamkeit problematisch und schwächt die Vorhersagekraft (u. a. Pajares, 1996). In Studien von Bong (u. a. 1999) erreichen aufgaben- und problemspezifische Instrumente zur Messung der SWK in Mathematik höhere Korrelationen zu verschiedenen Leistungsmaßen als eine reine domainspezifische Erfassung.

Die hier vorgestellte Meta-Analyse konzentriert sich auf den Bereich Mathematik mit besonderem Fokus auf die Charakteristika der Erhebungsinstrumente als mögliche Moderatorvariablen. Zur Auswahl der Primärstudien wurden bereits vorhandene Reviews und Meta-Analysen, mehrere Online-Datenbanken und Referenzen innerhalb der ausgewählten Studien verwendet. Selektionskriterien waren u. a. bereichsspezifische Erhebungsinstrumente (mathematische Selbstwirksamkeit und Mathematikleistung) und ausreichend Informationen zur Berechnung der ES. Anhand der Stichwörter und Abstracts wurden über 400 Studien ausgewählt, von denen 94 in die Meta-Analyse eingingen. Die statistischen Ergebnisse der Primärstudien wurden in eine einheitliche Effektgröße transformiert, die der Pearson Produkt-Moment-Korrelation entspricht. Es werden die Ergebnisse des Random Effects Models betrachtet.

Die durchschnittliche Korrelation zwischen mathematischer Selbstwirksamkeit und Mathematikleistung liegt bei 0,43, wobei es sich um heterogene ES handelt. Mehrere Moderatorvariablen können identifiziert werden, die jedoch ähnlich wie bei Multon, Brown und Lent keine homogenen Untergruppen bilden. Eine Erfassung der Selbstwirksamkeit über Aufgaben (task-/problem-specific) erzeugt mit 0,47 eine höhere Korrelation zur Leistung als über Fragebögen mit Aussagen (domain-specific), die bei 0,38 liegen. Die Art des Leistungstests bildet einzeln betrachtet keine zueinander heterogenen Untergruppen, jedoch im Zusammenhang mit der Erhebung der Selbstwirksamkeit. Die Höhe der Korrelation hängt nicht nur von der Erfassung der Selbstwirksamkeit ab, sondern u. a. auch von der Nähe zur Leistung. Wird gleichzeitig auch noch der Zeitpunkt der ES (Pretest vs. Posttest) berücksichtigt, so variiert die Korrelation zwischen 0,35 (SWK über Fragebogen, Test als Leistung, Pretest) und 0,64 (SWK über Aufgaben, Test als Leistung, Posttest). Die Moderatoranalysen zeigen, dass mehrere Variablen die Höhe der Effektstärke beeinflussen. Zur gleichzeitigen Berücksichtigung wurden entsprechende Gruppen gebildet, deren Zugehörigkeit als Moderator dient.

Der Selbstwirksamkeit wird eine wichtige Rolle im Lernprozess beigemessen und der positive Zusammenhang zur Leistung ist unumstritten. Die Ergebnisse der Meta-Analyse zeigen, dass die Höhe der Effekte stark variiert und u. a. von der Art der Erhebung der SWK abhängt. Trotz gleichzeitiger Berücksichtigung mehrerer Variablen, konnten keine homogenen Untergruppen gebildet werden. Es ist anzunehmen, dass weitere Charakteristika der Studien bzw. Erhebungsinstrumente von Bedeutung sind, die u. a. aufgrund zu geringer Informationen innerhalb dieser Meta-Analyse nicht berücksichtigt werden konnten (Schwierigkeitsgrad der Aufgaben zur Erfassung der SWK, Exaktheit der eigenen Einschätzung in Bezug auf realisierte Leistung).

E01 Pause mit Poster

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

3

Schröder, Jan; Schwanenberg, Jasmin

Die Entwicklung von Prüfungsangst, Fähigkeitsselbstkonzept und Schülerleistung an Ganztagsgymnasien

Aus der pädagogisch-psychologischen Forschung und Analysen zu Determinanten der Schulleistung lässt sich entnehmen, dass Fähigkeitsselbstkonzepte als „die Gesamtheit der kognitiven Repräsentation eigener Fähigkeiten in akademischen Leistungssituationen“ (Schöne et al., 2002, S. 7) und Prüfungsangst auf verschiedene Weise einen Einfluss auf Lernprozesse haben können. Es konnte bereits gezeigt werden, dass sich das Fähigkeitsselbstkonzept auf das Verhalten in Lernsituationen auswirkt und schulische Leistungen determiniert (vgl. Köller & Baumert, 2001; Marsh & Yeung, 1997). Schulleistungen und das weitere Lernen können darüber hinaus durch Prüfungsangst negativ beeinflusst werden (vgl. Helmke & Schrader, 2010; Pekrun, 1991). Bei Prüfungsangst besteht die Befürchtung, den Leistungsanforderungen nicht entsprechen zu können. Sie stellt einen schulleistungshemmenden Faktor dar (vgl. Rohfs, 2011) und beeinträchtigt in verschiedener Hinsicht die betroffenen Schülerinnen und Schüler in ihrem (Schul-)Alltag. Außerdem kann Prüfungsangst eine Verringerung der Lernmotivation bewirken, das Fähigkeitsselbstkonzept vermindern, aber auch psychosomatische Beschwerden und damit verbundene soziale Probleme zur Folge haben (vgl. ebd). Ängste, Leistung und emotionales Erleben bedingen sich damit gegenseitig. Die Anzahl vorliegender Studien, die Prüfungsangst bei Schülerinnen und Schülern nach dem Übergang in die Sekundarstufe I untersucht haben, ist bisher eher begrenzt.

Auf Grundlage dieser theoretischen Annahmen stehen folgende Fragestellungen im Fokus dieses Beitrags: (1) Wie entwickelt sich die Prüfungsangst und das Fähigkeitsselbstkonzept in der Sekundarstufe I bei Schülerinnen und Schülern an Ganztagsgymnasien? (2) Welche Wirkungsmuster lassen sich in Bezug auf die Prüfungsangst, das Fähigkeitsselbstkonzept und die Leistung im Längsschnitt an Ganztagsgymnasien identifizieren? Als Hypothese lässt sich annehmen, dass die Prüfungsangst negativ und das Fähigkeitsselbstkonzept positiv die Leistung im Längsschnitt beeinflusst.

Mithilfe eines latenten Strukturgleichungsmodells werden die Entwicklung sowie die Wirkungsmuster hinsichtlich der Prüfungsangst, des Fähigkeitsselbstkonzepts und der Schülerleistung analysiert. Als Datengrundlage dienen dafür drei Messzeitpunkte im Schulentwicklungsprojekt „Ganz In – Mit Ganztag mehr Zukunft. Das neue Ganztagsgymnasium NRW“. Im Rahmen von „Ganz In“ werden 31 Gymnasien auf ihrem Weg zum gebundenen Ganztag begleitet und unterstützt. Bezüglich der längsschnittlichen Erhebungen haben etwa 3000 Schülerinnen und Schüler im Schuljahr 2010/11, 2012/13 und 2014/2015 an einer Leistungsmessung und einer Schülerbefragung teilgenommen.

Die Ergebnisse für die ersten zwei Messzeitpunkte haben bereits für das Fach Biologie gezeigt, dass sich eine hohe Prüfungsangst in Klasse 5 negativ auf die Schülerleistung und das Fähigkeitsselbstkonzept in Klasse 7 auswirkt. Zudem wird deutlich, dass bei einer hohen Schülerleistung in Biologie in Klasse 5 die Prüfungsangst geringer und das Fähigkeitsselbstkonzept in Klasse 7 höher wird. Die theoretische Annahme, dass das Fähigkeitsselbstkonzept in Klasse 5 die Schülerleistung in Klasse 7 beeinflusst, konnte jedoch nicht nachgewiesen. Aus diesem Grund sollen nun weiterführende Analysen mit Daten des dritten Messzeitpunktes folgen, um Entwicklungen abbilden zu können. Durch den Beitrag können damit weitere Erkenntnisse über die genauen Mechanismen von Determinanten der Schulleistung im Rahmen eines Längsschnittdesigns gewonnen werden.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Weber, Katrin; Markus, Emden; Elke, Sumfleth

Entwicklung und Validierung einer Learning Progression zum Basiskonzept „Chemische Reaktion“ in der Sekundarstufe I

Es soll eine Learning Progression zum Basiskonzept „Chemische Reaktion“ für die Sekundarstufe I an Gymnasien beschrieben und validiert werden. Desweiteren wird untersucht, inwieweit sich das Lernen zu Inhalten der Basiskonzepte „Chemische Reaktion“ und „Struktur der Materie“ gegenseitig bedingt. Die Erkenntnisse können der Entwicklung optimierter sequenzierter Lehreinheiten bzw. der Curriculumentwicklung dienen.

Theoretischer Hintergrund: Mit Hilfe von Learning Progressions kann die Konzeptentwicklung von Lernenden über einen definierten Zeitraum beschrieben werden (Duschl, Schweingruber & Shouse, 2007). Zwei wichtige Bestandteile einer Learning Progression sind der Lower und der Upper Anchor (Corcoran, Mosher & Rogat, 2009). Der Lower Anchor wird durch die Fähigkeiten zu Beginn einer Lerneinheit beschrieben, während der Upper Anchor die Lernziele dieser Lerneinheit umfasst (Corcoran et al., 2009). Zwischen Lower und Upper Anchor werden mehrere Progressionsstufen durchschritten, die durch für die jeweiligen Progressionsstufen typische Fähigkeiten charakterisiert sind (Corcoran et al., 2009). Auf Basis dieser Fähigkeiten werden Assessments zur Validierung einer Learning Progression entwickelt (Corcoran et al., 2009).

Fragestellung: Es sollen folgende Forschungsfragen beantwortet werden: (1) Kann mit Hilfe einer Learning Progression die Entwicklung konzeptbezogener Fähigkeiten im Chemieunterricht der Sekundarstufe I hinsichtlich der Basiskonzepte „Struktur der Materie“ und „Chemische Reaktion“ valide beschrieben werden? (2) Wie entwickeln sich die Fähigkeiten entlang der Dimension „Chemische Reaktion“ in Abhängigkeit von der Dimension „Struktur der Materie“? (3) Welche Progressionsstufen lassen sich hinsichtlich der Entwicklung Fähigkeiten in den Basiskonzepten identifizieren?

Methode: Während sich der Upper Anchor aus den Bildungsstandards für das Fach Chemie für den Mittleren Schulabschluss (KMK, 2005) und dem nordrhein-westfälischen Kernlehrplan für die Sekundarstufe I an Gymnasien (MSW, 2008a) ableitet, beruht der Lower Anchor auf den nordrhein-westfälischen Richtlinien und Lehrplänen für die Grundschule (MSW, 2008b).

Zur Identifikation von Progressionsstufen wird eine Strand Map für die Basiskonzepte „Struktur der Materie“ und „Chemische Reaktion“ entwickelt (analog zu: AAAS, 2007), die die Zusammenhänge zwischen den Fähigkeiten grafisch visualisiert. Einige Fähigkeiten aus „Struktur der Materie“ scheinen Voraussetzung für den Erwerb von Fähigkeiten im Basiskonzept „Chemische Reaktion“ zu sein.

Auf Grundlage der Strand Map werden folgende Ausprägungen (Progressionsstufen) für die Dimension „Chemische Reaktion“ vermutet: „... als Stoffumwandlung“, „... als Umgruppierung von Atomen“ und „... als Veränderung in der Elektronenhülle“. Folgende Ausprägungen werden für die Dimension „Struktur der Materie“ angenommen: „Phänomenebene“, „Daltonisches Atommodell“ und „Bohrsches Kern-Hülle-Modell“.

Die Validierung der Strand Map erfolgt in einer Expertenbefragung. Die Learning Progression wird in einer Schülerbefragung in den Jahrgangsstufen 7-9 validiert (N ≈ 600). Dazu sollen die Fähigkeiten der Lernenden anhand eines Paper-Pencil-Tests (Multiple-Choice-Single-Select) erfasst werden. Aus der Strand Map geht hervor, dass es in jeder Progressionsstufe eine Fähigkeit gibt, die diese maximal differenziert beschreibt. Für diese Fähigkeiten werden Items zu fünf Beispielen konstruiert. Für die Fähigkeiten, die sich auf dem angenommenen Lernpfad hierhin befinden, werden ebenfalls Items zu genau diesen Beispielen formuliert. Es ergeben sich fünf Itemsets. Die Lernenden bearbeiten jeweils ein vollständiges Itemset und einzelne Items aus den weiteren Itemsets. Die Bearbeitung dieser weiteren Items soll helfen, die Itemschwierigkeiten besser zu schätzen. Bei der Itementwicklung nach Ferber (2014) werden die Komplexität bei der Aufgabenbearbeitung und die kognitiven Prozesse berücksichtigt (vgl. Kauertz, Fischer, Mayer, Sumfleth & Walpuski, 2010).

Die Schätzung der Itemschwierigkeiten erfolgt durch Rasch-Analysen. Zur Überprüfung der Progressionsstufen und Zusammenhänge zwischen den Fähigkeiten werden Korrelationsanalysen durchgeführt. Die Kontrollvariablen aktuelle Motivation (FAM: Rheinberg, Vollmeyer & Burns, 2001), Sprachbeherrschung (C-Test: Baur & Spettmann, 2008) und kognitive Fähigkeiten (KFT: Heller & Perleth, 2000) werden berücksichtigt.

Ergebnisse: Die Pilot-Studie wird im Januar 2015 und die Hauptstudie im Mai 2015 durchgeführt. Zum Zeitpunkt der GEBF-Tagung werden evaluierte Testitems, ein reliabler Test und erste Ergebnisse aus der Pilotstudie vorliegen. Es wird erwartet, dass mit Hilfe von Korrelationsanalysen die Annahmen hinsichtlich der Relationen zwischen den Fähigkeiten und der Progressionsstufen gestützt werden.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Averbeck, Daniel; Strübe, Martina; Sumfleth, Elke

Sachstrukturen im Chemieunterricht – Entwicklung und Erprobung eines Kategoriensystems zur Rekonstruktion von Sachstrukturen

Strukturierung ist ein häufig genanntes Kriterium für Unterrichtsqualität (vgl. Helmke, 2012; Meyer, 2009). Neben dem für Schülerinnen und Schüler nachvollziehbaren Aufbau und Verlauf einer Stunde spielt dabei bezogen auf den Fachunterricht die Sachstruktur (Niedderer, 1974) eine bedeutende Rolle. Der Begriff der Sachstruktur wurde vor allem in der Physikdidaktik weiter ausdifferenziert (vgl. Duit, Häußler & Kircher, 1981; Reinhold, 2006). Heute wird die Sachstruktur definiert als „(...) die sachliche, unter logischen und systematischen Gesichtspunkten gegliederte Struktur der fachlichen Inhalte (...)“ (Brückmann, 2009, S. 57), wobei sich diese sowohl auf die „(...) Begriffe, Konzepte, Modelle und Prinzipien als auch Methoden, Denk- und Arbeitsweisen und die Vorstellungen über die Natur der Naturwissenschaften“ beziehen. Verschiedene Studien (Brückmann, 2009; Müller & Duit, 2004) nutzten erfolgreich Sachstrukturdiagramme, um die Sachstruktur von Unterricht abzubilden. Dabei sind Sachstrukturdiagramme Flussdiagramme, in denen der Inhalt einer Stunde durch Inhaltsblöcke chronologisch aufeinanderfolgend und mithilfe von Einzelpfeilen oder Doppelpfeilen (Rückbezüge) verknüpft und strukturiert dargestellt wird.

In dieser Arbeit wurde ein erster Schritt zur Analyse der Sachstruktur von Chemieunterricht mithilfe von Sachstrukturdiagrammen unternommen und die folgenden Fragestellungen wurden untersucht: FF1a: Inwiefern ist es möglich, reliable Kategoriensysteme zur Erfassung des sachlichen Inhalts von Chemieunterricht für das Thema „Periodensystem der Elemente und Atombau“ und zur Rekonstruktion von Sachstrukturen zu entwickeln? FF1b: Ist es möglich, ausgehend von einem detaillierten Kategoriensystem die Sachstruktur von Chemiestunden in Form von Sachstrukturdiagrammen darzustellen? FF2: Inwiefern können mithilfe der Sachstrukturdiagramme Unterschiede und Gemeinsamkeiten der Sachstrukturen von Unterrichtsstunden untersucht werden?

Das Kategoriensystem wurde in einem deduktiv-induktiven Verfahren (Bortz & Döring, 2006; Brückmann, 2009) entwickelt. In der deduktiven Entwicklungsphase wurde der gymnasiale Lehrplan der Sekundarstufe I (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2008) auf die Inhalte zum Themenfeld Periodensystem der Elemente und Atombau analysiert und mit Fachbüchern weiter ausdifferenziert. Aufgrund dieser Analysen konnte in Anlehnung an Brückmann (2009) ein Kategoriensystem mit zwei Ebenen entwickelt werden. Auf der Ebene A wird zwischen den drei Themenbereichen „Periodensystem der Elemente“, „Atombau & Atomkerne“ sowie „Atomhülle & Radioaktivität“ unterschieden. Auf der Ebene B werden Unterkategorien zu den drei Themenbereichen als eigenständige, ausdifferenzierte Kategorien kodiert.

In der induktiven Phase wurden vier vorliegende Videos (Mutke, in Vorb.; Pollender, in Vorb.) verwendet, um das Kategoriensystem zu prüfen und zu überarbeiten. Für die Bestimmung der Beobachterübereinstimmung und der Reliabilität wurden diese im 10 Sekunden Intervall ($n_{\text{Intervalle}} = 933$) von drei Beobachtern kodiert. Hierbei ist zu beachten, dass nur ein Video von allen drei Beobachtern kodiert wurde. Aus diesem Grund mussten zur Berechnung der Reliabilität zum einen Cohens Kappa und zum anderen Fleiss' Kappa verwendet werden. Für die drei Themenbereiche der Ebene A wurden gute Beobachterübereinstimmungen von über 90 Prozent („Periodensystem der Elemente“: $n_{\text{Intervalle}} = 13$, Beobachterübereinstimmung = 97,5%, $\kappa = .35$ (aufgrund einer unsymmetrischen Randsummenverteilung); „Atombau & Atomkerne“ Beobachterübereinstimmung = 95,1%, $n_{\text{Intervalle}} = 671$, $\kappa = .92$; „Atomhülle & Radioaktivität“ Beobachterübereinstimmung = 93,5 %; $n_{\text{Intervalle}} = 173$, $\kappa = .82$). Für alle 12 beobachtbaren Kategorien der Ebene B ergeben sich Beobachterübereinstimmungen zwischen 95,4 und 100 Prozent. Hier ist jedoch anzumerken, dass nicht alle Kategorien in den Videos vorkamen und dementsprechend nicht kodiert werden konnten. Fast alle der beobachteten Kategorien zeigen nach Wirtz und Caspar (2002) ein mäßig gutes bis sehr gutes Kappa zwischen .50 und 1.00.

Um mithilfe der Sachstrukturdiagramme Aussagen zu den Gemeinsamkeiten und Unterschieden der Sachstruktur zu machen, wurden vier weitere Videos aus der ProwiN-Videostudie Chemie von einem Beobachter kodiert. Bei der Analyse der Sachstrukturdiagramme werden die Anzahl der Vernetzungen zwischen den Inhaltsblöcken, die Anzahl der Inhaltsblöcke und die Inhalte bestimmt.

U.a. werden die Ergebnisse zur Häufigkeit bestimmter beobachtbarer Themengebiete und weiterer sachstruktureller Merkmale aus einem Vergleich der Sachstrukturdiagramme im Hinblick auf die Gemeinsamkeiten und Unterschiede auf dem Poster dargestellt.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Steff, Henning; Sommer, Katrin

Untersuchungen über die Wirkung von Modellexperimenten im Chemieunterricht

Modellexperimente liefern Beiträge zur Erkenntnisgewinnung. Sie erklären und veranschaulichen. In diesem vielschichtigen, modellbasierten Erkenntnisprozess kommt der Analogiebildung eine Schlüsselrolle zu [1, 2], wobei das Modellexperiment als Analogon des Originals fungiert. Eine Formulierung könnte lauten: „Das ist doch wie ...“. Bei Analogien kommt es zur Übertragung von Strukturen, Funktionen bzw. Verhalten sowie Verfahrensweisen im Vorstellen oder Denken von einem Gegenstandsbereich auf einen anderen Gegenstandsbereich [3]. Modellexperimente werden immer dann eingesetzt, wenn das Original aus Gründen des Gefahrenpotentials, der Handhabung, der Komplexität oder der Zugänglichkeit nicht verfügbar ist [4]. Ein konkretes Original wäre beispielsweise die Mehlstaubexplosion, die im Chemieunterricht mittels Modellexperiment nachvollzogen werden kann. Auf Grund der schlechten Zündeigenschaften der Originalsubstanz Weizenmehl greift man bei diesem Modellexperiment allerdings häufig auf reaktivere Modellsustanzen, wie Bärlappsporen, zurück, welche die Originalsubstanzen substituieren sollen. Es gibt allerdings keine Untersuchung darüber, inwieweit Schülerinnen und Schüler Analogien zwischen einem Modellexperiment und dem entsprechenden Original erkennen und ob sich die Auswahl der eingesetzten Modellsustanzen auf den intendierten Analogieschluss auswirkt. Aus diesem Sachverhalt ergeben sich folgende zentrale Fragestellungen: (1) „Erkennen Schülerinnen und Schüler die Analogien zwischen Modellexperimenten des Chemieunterrichts und den entsprechenden Originalen?“ (2) „Hat die Auswahl der eingesetzten Modellsustanzen einen Einfluss auf das Erkennen der Analogien?“

Zur Beantwortung dieser Forschungsfragen wurde eine Interventionsstudie im Pre-Post-Design entwickelt. Die Intervention beinhaltet insgesamt 5 Modellexperimente aus dem „Themenfeld Brände und Brandbekämpfung“ [5]. Die Erhebung wurde mit Fragebögen im offenen und geschlossenen Aufgabenformat durchgeführt. Bei der Stichprobe (N= ca. 350) handelte es sich um Schülerinnen und Schüler 7. Klassen von unterschiedlichen Gymnasien aus Nordrhein-Westfalen. Insgesamt 4 Probandengruppen bestehend aus 3 Schulklassen führten die Modellexperimente durch, wobei als unabhängige Variable ausschließlich die eingesetzte Substanz variierte. Die eingesetzten Substanzen wurden so ausgewählt, dass sie in oberflächlichen Merkmalen wie der Farbigkeit oder ihrer Bezeichnung stärkere oder weniger starke Analogien gegenüber der Originalsubstanz aufwiesen. Die Datenauswertung der eingesetzten Fragebögen erfolgte durch Anwendung der qualitativen Inhaltsanalyse nach Mayring [6]. Erste Ergebnisse lassen die Aussage zu, dass Schülerinnen und Schüler Analogien zwischen Modellexperimenten und den entsprechenden Originalen erkennen und dass dieser Einfluss auf die unterschiedlichen Modellsustanzen zurückzuführen ist.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Reschke, Tim; Koenen, Jenna; Sumfleth, Elke

Mit fiktionalen Geschichten im Chemieunterricht lernen

Im Fach Chemie wird der Regelstandard des Kompetenzbereichs Fachwissen von 41.9 % Schülerinnen und Schülern, die den Mittleren Schulabschluss anstreben, nicht erreicht. Dies zeigten die Ergebnisse des vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) durchgeführten Ländervergleichs. Des Weiteren zeigte dieser, dass das Interesse am Fach Chemie insbesondere im Vergleich zu den Fächern Biologie und Mathematik deutlich geringer ist (Pant et al., 2013). Dies konnte zuvor auch schon in anderen Studien, wie beispielsweise der ROSE-Studie von Sjøberg & Schreiner (2010), gezeigt werden.

Eine Möglichkeit, sowohl dem mangelnden Fachwissen als auch dem geringen Interesse der Schülerinnen und Schüler entgegen zu wirken, könnten Lernmaterialien sein, die den Wissenserwerb und das situationale Interesse gleichzeitig anregen. Wie einige Studien zeigen, könnten fiktionale Geschichten diesen Anforderungen gerecht werden (vgl. Avraamidou & Osborne, 2009). Solche Geschichten werden oft im Rahmen der Methode des Storytellings im Unterricht eingesetzt (z. B. Egan, 1988). Hierbei werden die Geschichten den Schülerinnen und Schülern am Anfang der Unterrichtsstunde mündlich durch die Lehrperson erzählt, wodurch sie für die darauffolgenden Fachinhalte begeistert werden sollen. Geschichten können jedoch auch als Lesegeschichten im Unterricht eingesetzt werden (z. B. Martensen et al., 2007). Diese enthalten in der Regel bereits konkrete zu lernende Fachinhalte (u. a. Kaspar & Mikelskis, 2008).

In fiktionalen Geschichten werden im Vergleich zu Sachtexten andere narrative Elemente genutzt (vgl. Avraamidou & Osborne, 2009). Hier eignen sich vor allem Analogien (z. B. Glynn, 2007) und Personifikationen (z. B. Pütttschneider & Lück, 2004). Analogien können durch Rückgriff auf bereits bekannte Sachverhalte aus dem Alltag eine Brücke zu unbekanntem Fachinhalten darstellen.

Bislang ist die Anzahl empirischer Studien, die die Lernförderlichkeit von fiktionalen Geschichten und ihre Wirkung auf das situationale Interesse der Schülerinnen und Schülern untersuchen, jedoch gering.

Basierend auf den kurz dargestellten Annahmen, ist das Ziel dieser Studie zu untersuchen, inwieweit der Lernerfolg und das situationale Interesse der Schülerinnen und Schüler durch fiktionale Geschichten positiv beeinflusst werden können. Es werden daher je eine fiktionale Geschichte zu den Themen Alkalimetalle und Atombau entwickelt und im Rahmen einer Interventionsstudie mit Prä-Post-Design in der Jahrgangsstufe 8 an Gymnasien im Vergleich zu Sachtexten eingesetzt. Beide Textarten sind als Lernmaterialien in Kombination mit jeweils drei gleichen Aufgaben zur eigenständigen Bearbeitung konzipiert und beinhalten identische Fachinhalte. Der Lernerfolg wird mithilfe eines Fachwissenstests im Multiple-Choice-Single-Select-Format evaluiert. Zusätzlich werden Lückentexte als halboffenes Aufgabenformat und Triaden (Produktion von Zusammenhängen zwischen drei Begriffen) (vgl. Özcan, 2013) als offenes Aufgabenformat eingesetzt. Mit diesen Aufgabenformaten soll überprüft werden, inwieweit die eingesetzten narrativen Elemente einen Einfluss auf den angemessenen Fachspracheneinsatz der Schülerinnen und Schüler haben. Das situationale Interesse wird mithilfe eines chemiespezifischen Fragebogens von Fechner (2009) evaluiert. Als Kontrollvariablen werden die aktuelle Motivation (FAM: Rheinberg et al., 2001), die kognitiven Fähigkeiten (KFT, N2: Heller & Perleth, 2000), die Lesekompetenz (LGVT: Schneider et al., 2007), das Chemieinteresse (Fechner, 2009; Klos, 2008), die Bearbeitungszeit, die Deutsch- und Chemienote, das Geschlecht sowie die zu Hause gesprochene Sprache erhoben. Zudem wird die kognitive Belastung (Kalyuga et al., 1999; Paas, 1992) erhoben, um zu überprüfen, inwieweit die unterschiedliche Gestaltung der Textarten einen Einfluss auf die empfundene Schwierigkeit und die investierte Denkanstrengung der Schülerinnen und Schüler hat. Unterschiedliche Levels an kognitiver Belastung könnten sich darüber hinaus positiv oder negativ auf den Lernerfolg oder auch das situationale Interesse auswirken.

Bevor die Texte im Rahmen der Interventionsstudie eingesetzt werden, sollen sie zunächst von Schülerinnen und Schülern anhand der Methode des Lauten Denkens bearbeitet werden. Im Anschluss daran folgt ein leitfadengestütztes Interview zum Umgang mit dem Text. Auf Grundlage der Ergebnisse dieser qualitativen Untersuchung sollen die entwickelten Texte optimiert werden. Die Ergebnisse dieser Untersuchung können im März im Rahmen der GEBF in Form eines Posters präsentiert werden.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Schüßler, Katrin; Koenen, Jenna; Sumfleth, Elke

Individuell- oder inhaltlich-fokussiert? - Einfluss unterschiedlicher Promptingmaßnahmen beim Lernen mit Lösungsbeispielen

Theoretischer Hintergrund: Lösungsbeispiele entsprechen expertenhaften Musterlösungen für ein Problem (Atkinson, Derry, Renkl, & Wortham, 2000), indem sie eine Problemstellung, die zur Problemlösung notwendigen Lösungsschritte sowie eine abschließende Problemlösung präsentieren (Atkinson, Renkl, & Merrill, 2003). Lösungsbeispiele sind insbesondere während des anfänglichen Wissenserwerbs von Bedeutung (Atkinson et al., 2000; Reimann, 1997). Nach Annahmen der Cognitive Load Theory (u. a. Sweller, 2005) reduzieren Lösungsbeispiele die extrinsische Belastung während des Lernprozesses, wenn sie entsprechend gestaltet sind (u. a. Atkinson et al., 2000), da für Lernende die Suche nach der Problemlösung entfällt. Dadurch stehen den Lernenden beim Lernen mit Lösungsbeispielen mehr kognitive Kapazitäten für das Verständnis der Musterlösung zur Verfügung (Ward & Sweller, 1990). Neben einer angemessenen Gestaltung wird die Generierung von Selbsterklärungen als entscheidender Faktor für erfolgreiches Lernen mit Lösungsbeispielen betrachtet (Renkl, 2005). Selbsterklären wird als dualer Prozess beschrieben, in dessen Verlauf implizite Informationen aufgedeckt und Vorstellungen korrigiert werden (Roy & Chi, 2005). Als Einflussfaktoren des Selbsterklärens werden die Fähigkeit den eigenen Lernprozess zu überwachen und zu steuern und das Vorwissen der Lernenden diskutiert (Chi, 2000; Kross & Lind, 2002; Renkl, 1997; Stark, 1999). Jedoch generiert höchstens die Hälfte der Lernenden spontan lernförderliche Selbsterklärungen (Renkl, 1997; Stark, 1999). Zur Unterstützung der Lernenden beim Generieren von Selbsterklärungen können Prompts eingesetzt werden (Berthold, Eysink, & Renkl, 2009; Gerjets, Scheiter, & Catrambone, 2006).

Fragestellung: Bisher ist jedoch fraglich, welche Promptingmaßnahmen besonders gut geeignet sind. Daher sollen im Rahmen dieser Studie inhaltlich-fokussierte Prompts mit individuellen, inhaltlich-offen formulierten Prompts in Bezug auf ihren Einfluss auf den Lernerfolg beim Lernen mit Lösungsbeispielen zu chemischen Fachinhalten verglichen werden.

Inhaltlich-offene Prompts fordern die Lernenden dazu auf individuell wichtige Ideen eines Textabschnittes zu identifizieren und intensiver zu betrachten. Da mentale Repräsentationen individuell konstruiert werden, sollten sich inhaltlich-offene Prompts als vorteilhaft erweisen (Chi, 2000), weil sie Lernenden die Möglichkeit bieten ihre mentale Repräsentation mit der Darstellung im Lösungsbeispiel abzugleichen (Modellinspektion) und bei Bedarf zu korrigieren (Modellkonstruktion) (Schnotz & Bannert, 2003). Lernenden, denen die metakognitiven Fähigkeiten fehlen, um eigene Verständnisschwierigkeiten zu erkennen und eine entsprechende Reaktion einzuleiten (Renkl, 2001; 2002), könnten mit inhaltlich-offenen Prompts allerdings überfordert sein und sollten stärker von inhaltlich-fokussierten Prompts profitieren. Inhaltlich-fokussierte Prompts lenken die Aufmerksamkeit der Lernenden auf Aussagen zu bestimmten inhaltlichen Aspekten und fordern sie dazu auf sich intensiver mit diesen auseinanderzusetzen.

Neben Basisprompts, die die aktive Beispielbearbeitung unterstützen und Verständnisillusionen vermeiden (Stark, 1999), werden daher inhaltlich-offen bzw. inhaltlich-fokussierte Zusatzprompts eingesetzt. Untersucht wird die Frage, inwieweit sich das Lernen mit den verschiedenen Promptingmaßnahmen positiv auf den Lernerfolg beim Lernen mit Lösungsbeispielen auswirkt. Dabei sollen auch verschiedene Voraussetzungen der Lernenden wie zum Beispiel das Vorwissen mit betrachtet werden.

Methode: Zur Untersuchung der Fragestellung wird momentan eine Interventionsstudie mit Prä-, Post-, Follow up-Design im Chemieunterricht der Jahrgangsstufe 9 an Gymnasien in Nordrhein-Westfalen durchgeführt (N = 175). Eingesetzt werden drei Lösungsbeispiele (je 90 Minuten) zum Thema saure Lösungen, die unterschiedliche Promptingmaßnahmen enthalten. Die Kontrollgruppe erhält Lösungsbeispiele, die lediglich Basisprompts enthalten. Zusätzlich zu den Basisprompts erhalten Lernende der Experimentalgruppen inhaltlich-fokussierte Prompts (EG1) oder inhaltlich-offene Prompts (EG2).

Unterschiede im Lernzuwachs sollen Aufschlüsse über die verschiedenen Wirkungsweisen der Promptingmaßnahmen geben. Neben dem Fachwissen als Indikator für den Lernerfolg werden unter anderem das Interesse (Fechner, 2009; Klos, 2007), das private Leseverhalten, Lesegeschwindigkeit und Leseverständnis (Schneider, Schlagmüller, & Ennemoser, 2007), die kognitiven Fähigkeiten (Heller & Perleth, 2000) und der Cognitive Load (Paas, 1992; Kalyuga, Chandler, Tuovinen, & Sweller, 2001) bei der Bearbeitung der Lösungsbeispiele als Kontrollvariablen erhoben.

Ergebnisse: Ergebnisse der Studie werden im März 2015 vorliegen. Erwartet werden erste Hinweise auf Abhängigkeiten zwischen individuellen Merkmalen der Lernenden und Effektivität der Promptingmaßnahmen.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Bouley, Franziska

Erwerb von Rechnungswesenkompetenz in Abhängigkeit der Unterrichtsmethode – Ein Vergleich der Bilanzmethode und des wirtschaftsinstrumentellen Rechnungswesens

Problemstellung und theoretischer Hintergrund: Die kaufmännische Ausbildung hat zum Ziel, ihre Schülerinnen und Schüler zu befähigen, die Anforderungen des (späteren) Berufes bestmöglich zu bewältigen (KMK, 1991). Dabei nimmt der Buchführungsunterricht eine tragende Rolle ein. Hier sollen Schülerinnen und Schüler nicht nur Fachwissen erwerben, sondern primär unternehmensinterne Prozesse kennenlernen sowie handlungsleitendes ökonomisches Denken entwickeln (Achtenhagen, 1996, Sembill & Seifried, 2005; Sloane, 1996).

Auf dem Weg diese Ziele zu erreichen, begegnen Schülerinnen und Schüler einer Vielzahl von Lernschwierigkeiten (Pawlik, 1979, 1980; Tramm, Hinrichs & Langenheim, 1996) und machen typische Rechnungswesenfehler (Türling, Seifried, Wuttke, Gewiese & Kästner, 2011; Wuttke & Seifried, 2012). In Folge dessen kann festgestellt werden, dass das Wissen von Lernenden im Bereich des Rechnungswesens oft defizitär ist (Lehmann & Seeber, 2007) und der Erwerb von Rechnungswesenkompetenz durchaus beeinträchtigt werden kann.

Wie die oben genannten Ziele am besten erreicht und Lernschwierigkeiten vermieden werden können, d.h. die Frage nach der Gestaltung des Buchführungsunterrichts, wird in der Fachdidaktik kontrovers diskutiert, da sich unterschiedliche Ansätze der didaktischen Sequenzierung gegenüberstehen und jeder den Anspruch erhebt, die Aufgaben des Rechnungswesenunterrichts am besten erfüllen zu können. Dabei nehmen die Bilanzmethode (traditioneller Rechnungswesenunterricht mit Fokus auf der axiomatischen Buchführungssystematik) und das wirtschaftsinstrumentelle Rechnungswesen (WiR; Fokus auf ökonomischer Interpretation der die Unternehmenstätigkeit abbildenden Buchungsvorgänge) die prominenteste Stellung ein. Beide Methoden stehen aus diesem Grund im Mittelpunkt des Dissertationsvorhabens.

Die Diskussion um den besten pädagogischen Ansatz beschränkt sich zum einen auf eine theoretische Beschreibung der Vorteile und Nachteile beider hier fokussierter Methoden (Plink, 2010, 2011, 2013), zum anderen beschreiben die wenigen Forschungsbemühungen in diesem Bereich allgemeine Lernschwierigkeiten im Buchführungsunterricht (Pawlik, 1979, Tramm et al., 1996) oder den Erfolg beider Ansätze in Kombination mit ausgewählten Unterrichtsmethoden (Seifried, 2004). Zusammenfassend lässt sich sagen, dass die Diskussion bislang nur unzureichend anhand von einschlägigen empirischen Studien angereichert wurde (Ernst, 2014). An diese Lücke knüpft das hier vorgestellte Dissertationsprojekt, indem beide didaktischen Ansätze direkt gegenübergestellt und die Frage untersucht wird, welche Methode zu einer höheren Rechnungswesenkompetenz führt und die zentralen Anliegen des Buchführungsunterrichts erfolgreich transportiert.

Methode:

Kompetenzmodell: Zur Bestimmung der Rechnungswesenkompetenz von Schülerinnen und Schülern bedarf es eines zugrundeliegenden Kompetenzmodells, das die Dimensionen der Kompetenz abbildet. Das hier verwendete Modell lehnt sich an Vorarbeiten von KoMeWP (Kompetenzmessung im wirtschaftspädagogischen Studium: Berger et al., im Druck, 2013) sowie OEKOMA (Ökonomische Kompetenzen von Maturandinnen und Maturanden: Schumann, Oepke & Eberle, 2011) an. Kompetenz ist hierbei ein Zusammenspiel aus Motivation und Interesse, Fachwissen sowie Werthaltung und Einstellung.

Design und Stichprobe: Die Durchführung der zugrundeliegenden Studie ist für das Schuljahr 2015/2016 geplant. Der Haupterhebung wird eine Pilotierungsphase im Januar/Februar 2010 vorauslaufen. Um die Entwicklung der Kompetenzen zu messen, wird eine Interventionsstudie durchgeführt. Hierzu werden zwei Klassen ausgewählt, die im Rechnungswesen mittels Bilanzmethode und zwei, die mittels WiR unterwiesen werden. Es handelt sich dabei jeweils um eine Klasse des Wirtschaftsgymnasiums und eine der Berufsfachschule. Somit soll die Adäquanz der Methode hinsichtlich unterschiedlicher Schülerschaft (in Bezug auf Lernvoraussetzungen und Niveau) berücksichtigt werden. Es wird eine Stichprobengröße von ca. 100 Schülern angestrebt. Während eines Pretests sollen die Lernvoraussetzungen (kognitive Fähigkeiten, mathematische Kenntnisse, Interesse an Wirtschaftsthemen sowie Lernmotivation) der Schülerinnen und Schüler erfasst werden. Während der Intervention sollen Prozessdaten die Entwicklung der Motivation, der Lernschwierigkeiten sowie der kognitiven Belastung dokumentieren. Die Intervention schließt mit einem Posttest, der den Wissensstand in den drei zentralen Lerninhaltsbereichen, das Interesse an Wirtschaftsthemen sowie die Lernmotivation erfasst. Ein Follow-up soll einige Zeit später die Entwicklung einer ökonomischen Kompetenz überprüfen.

Ergebnisse: Ergebnisse liegen zum jetzigen Zeitpunkt noch nicht vor. Zur Konferenz werden Ergebnisse der Pilotierung vorgestellt. Hierbei soll bereits gezeigt werden, inwiefern sich das Wissen der Teilnehmer beider Gruppen entwickeln und unterscheiden.

Donnerstag
12. März

10:45 Uhr

Foyer GB - Ost

Marschner, Jessica

Elterntrainings: Wann trainiere ich was? Erste Hinweise aus einer Elternbefragung an Grundschulen

Es steht außer Frage, dass für Schulerfolg neben Lehrer- und Unterrichtsfaktoren vor allem familiäre Einflussfaktoren entscheidend sind (Helmke et al., 1991; Lorenz & Wild, 2007; Sacher, 2012). Eltern fungieren als Modell, Feedbackquelle und „Lehrer“ beim häuslichen Lernen. Sie erhalten allerdings wenig Unterstützung, wie sie häusliche Lehr-Lernsituationen gestalten können (Wild & Gerber, 2009), obwohl viele Eltern daran interessiert sind (Krumm, 1996; Lund et al., 2001). Es bietet sich folglich an Eltern „fortzubilden“. Elterntrainings im Grundschulbereich und der Sekundarstufe I existieren bereits zum Selbstregulierten Lernen oder zu motivationsförderlichem Verhalten (Bruder, 2005; Lund et al., 2001; Otto, 2007; Wild & Gerber, 2009). Vernachlässigt wird bisher die elterliche Fehlerfreundlichkeit (Marschner & Wirth, 2014). Nach Heinrichs et al. (2002) spielen Elterntrainings eine umso größere Rolle je jünger die Kinder sind. Als Hauptproblem sehen sie die Selbstselektion der Eltern und die hohen Drop-out-Quoten. Um möglichst viele Eltern zu gewinnen und sie während des gesamten Trainings beizubehalten, ist es ratsam, das Training Adressaten gerecht zu gestalten und zu schauen, ob bzw. in welcher Jahrgangsstufe Eltern besonders für Elterntrainings ansprechbar wären. Es soll also geprüft werden, in welcher Jahrgangsstufe stärker Probleme in häuslichen Lehr-Lernsituationen wahrgenommen werden bzw. ob sich Eltern unterschiedlicher Jahrgänge unterschiedlich kompetent fühlen die Probleme zu beheben, sich in ihrer Selbstwirksamkeit, ihrem Elternverhalten (kontrollierendes Elternverhalten und Fehlerfreundlichkeit) oder in dem Ausmaß an Informationswünschen unterscheiden.

Befragt wurden N=125 Mütter dreier Grundschulen. 23 Mütter füllten den Fragebogen für Erstklässler aus, 36 Mütter für Zweitklässler, 35 Mütter für Drittklässler und 30 Mütter für Viertklässler. Der Fragebogen erfasste neben demografischen Angaben, spezifische Selbstwirksamkeit ($\alpha=.83$), Fehlerfreundlichkeit ($\alpha=.66$), kontrollierendes Hausaufgabenverhalten ($\alpha=.79$) und für 17 Probleme bei häuslichen Lehr-Lernsituationen die wahrgenommene Problemstärke ($\alpha=.91$) und die eingeschätzte Problembewältigungs-Kompetenz ($\alpha=.96$). Zusätzlich wurde gefragt, ob Eltern gerne Informationen zu häuslichen Lernsituationen bekommen würden, Tipps diesbezüglich von Lehrkräften bekommen haben bzw. gerne bekommen würden. Alle Items waren auf einer vierstufigen Likertskala zu beantworten. Zusätzlich gab es die offene Frage nach weiteren Problemen in häuslichen Lehr-Lernsituationen.

Die eingeschätzte Problemstärke lag bei $M=1,92$ und die eingeschätzte Problembewältigungs-Kompetenz bei $M=2,95$. Der Wunsch nach Informationen zum Gestalten häuslicher Lernsituationen lag dennoch bei $M=2,82$, Eltern bekommen von Lehrern nur wenige Tipps dazu ($M=2,12$), wünschen sich diese aber eher schon ($M=3,07$).

Varianzanalysen zeigen keine signifikanten Unterschiede zwischen den Jahrgängen bezüglich der eingeschätzten Problemstärke, dem Informationswunsch, dem Informationserhalt, der Selbstwirksamkeit, Kompetenzüberzeugung oder dem Elternverhalten. Ein tendenzieller Unterschied zwischen den unterschiedlichen Jahrgängen zeigt sich allerdings in der Fehlerfreundlichkeit. Post-hoc Tests zeigen, dass die Fehlerfreundlichkeit zwischen zweitem und drittem Jahrgang absinkt. Erklärung hierfür könnte die Zensurenvergabe ab der dritten Klasse sein.

Die Problemstärke korreliert erwartungsgemäß positiv mit dem Informationswunsch ($r=.29$). Auch korreliert die Problemstärke theoriekonform positiv mit kontrollierendem Hausaufgabenverhalten ($r=.22$) und negativ mit Fehlerfreundlichkeit ($r=-.22$), Selbstwirksamkeit ($r=-.30$) und Kompetenzüberzeugung ($r=-.25$). Allerdings korreliert das kontrollierende Hausaufgabenverhalten nicht positiv mit dem Wunsch nach Tipps und Informationen. Es gibt die erwarteten negativen Zusammenhänge zwischen Informationswunsch und Fehlerfreundlichkeit ($r=-.22$), Selbstwirksamkeit ($r=-.34$) sowie Kompetenzüberzeugung ($r=-.24$).

Aus den Ergebnissen lässt sich schließen, dass ein Elterntaining in allen Jahrgängen der Grundschule angebracht wäre. Allerdings macht es Sinn gerade vor Beginn der dritten Jahrgangsstufe auf die Komponente Fehlerfreundlichkeit in Trainings Bezug zu nehmen. Zumal sich ein negativer Zusammenhang zwischen Fehlerfreundlichkeit und Problemstärke zeigt. Trotz des positiven Zusammenhangs zwischen Problemstärke und kontrollierendem Hausaufgabenverhalten findet sich kein positiver Zusammenhang zwischen kontrollierendem Hausaufgabenverhalten und dem Wunsch nach Tipps und Informationen. Das spricht für die Annahme, dass viele Eltern die negativen Folgen von kontrollierendem Hausaufgabenverhalten nicht kennen und es für ein angemessenes Elternverhalten halten, ein wichtiger Hinweis für künftige Trainings.

Im nächsten Schritt werden die offenen Antworten danach ausgewertet, welche weiteren Probleme von den Eltern häufig genannt werden, dieses hilft Trainings auf authentische Problemsituationen zu beziehen.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

Schönmoser, Carina; Schmitt, Monja; Relikowski, Ilona

Sozialkompetenzen von Kindergartenkindern – Ein Vergleich der Eltern- und Erzieherperspektive

Soziale Kompetenzen stellen einen wichtigen Aspekt menschlicher Entwicklung dar. Sie nehmen nicht nur entscheidenden Einfluss auf das Sozialverhalten von Kindern und Erwachsenen, sondern gelten auch als Prädiktor für akademische Leistungsfähigkeit oder beruflichen Erfolg (Reinders 2008). Da sich soziale Kompetenzen bereits im Kindesalter ausprägen, sollten sie frühestmöglich erfasst und untersucht werden. Weiterhin sollte diese Erfassung idealerweise in einer Multi-Informanten-Perspektive erfolgen. In verschiedenen Grundlagenbüchern wird betont, wie wichtig die Fremdeinschätzung von mehreren Informanten zu persönlichen Eigenschaften und dem Verhalten der Zielperson ist – gerade wenn die Zielperson noch sehr jung ist. Bekannt ist in diesem Zusammenhang, dass die Einschätzungen der verschiedenen Informanten in den seltensten Fällen übereinstimmen (u.a. Korsch und Petermann 2013), was u. a. auf die unterschiedlichen Beurteiler und die variierenden Kontexte zurückzuführen ist. Welche Variablen allerdings die Differenzen der Einschätzungen in den unterschiedlichen Kontexten erklären, wurde bisher nur in wenigen Studien berücksichtigt (Dinnebeil et al. 2013; Woerner et al. 2004).

Mit Daten der ersten Welle der Startkohorte 2 des Nationalen Bildungspanels (Blossfeld et al. 2011) nimmt vorliegender Beitrag einen Vergleich der Fremdeinschätzungen von Kindergartenkindern bezüglich ihrer sozialen Kompetenz vor. Diese wurde über eine Unterskala des SDQs (Strengths and Difficulties Questionnaire (Goodman (1997)) operationalisiert und sowohl im CATI (Computer Assisted Telephone Interview) für die Eltern als auch im PAPI (Paper and Pencil Interview)) für die Erzieher eingesetzt.

Mittels mehrebenenanalytischer Analyseverfahren sollen folgende Forschungsfragen beantwortet werden: (1) Welche Variablen erklären die Elterneinschätzung bzw. die Erziehereinschätzung der sozialen Kompetenz? Unterscheiden sich diese Variablen in ihrer Relevanz für die beiden Einschätzungen? (2) Inwiefern stimmen die Einschätzungen der Erzieher und der Eltern über die soziale Kompetenz der Kinder überein? Welche Variablen erklären die Unterschiede zwischen der Elterneinschätzung und der Erziehereinschätzung?

Es zeigt sich, dass die Übereinstimmung der Einschätzungen sehr gering ist. Dieses Ergebnis bestätigt die in der Literatur berichtete geringe Korrelation der Urteile. Weiterhin schätzen Erzieher die Sozialkompetenz der Kinder deutlich geringer ein als Eltern. Während die Einschätzungen der Eltern über die verschiedenen berücksichtigten Merkmale hinweg sehr stabil sind, variieren die Einschätzungen der Erzieher stark unter den beobachteten Einflussfaktoren. Dies führt dazu, dass auch die Abstände zwischen den Einschätzungen bei bestimmten Merkmalen unterschiedlich hoch ausfallen.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

Dörr, Lisa; Venitz, Laura; Perels, Franziska

Multimethodale Erfassung des selbstregulierten Lernverhaltens von Vorschulkindern

Ziel der vorliegenden Studie ist es, selbstreguliertes Lernen von Kindern zu fördern und dabei auch die unmittelbaren Bezugspersonen einzubeziehen. Ein besonderer Fokus der durch die Deutsche Forschungsgemeinschaft (DFG) geförderten Studie liegt auf der Erfassung selbstregulierten Lernens bei Kindern im Vorschulalter. Anhand einer Evaluation der eingesetzten Instrumente soll überprüft werden, ob diese geeignet sind, um selbstregulatorische Kompetenzen in der Altersklasse der 5-7 Jährigen zu erfassen.

Selbstreguliertes Lernen bezeichnet unterschiedliche Fähigkeiten, die für das Gelingen einer Lernhandlung eine Rolle spielen. In der aktuellen wissenschaftlichen Diskussion zur Selbstregulation wird hierbei zwischen den Komponenten der Kognition, der Metakognition und der Motivation unterschieden, die stets kontextgebunden betrachtet werden sollten (Zimmerman, 2000). Da gerade der Lernkontext und damit die Lernerfahrungen von Vorschulkindern noch sehr stark durch Bezugspersonen mit bestimmt werden, sollten diese bei der Betrachtung selbstregulierten Lernens im Vorschulalter Berücksichtigung finden.

Die Fähigkeit, das eigene Lernen selbständig, aktiv und strukturiert zu organisieren, wird heutzutage hoch eingeschätzt, da sie dem Lernenden erlaubt, sich lebenslang eigenständig Wissen anzueignen (vgl. Brunstein & Spörer, 2001; Friedrich & Mandl, 1997; Zimmerman, Bonner & Kovach, 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.

Im Rahmen der Untersuchung werden drei verschiedene Trainingsmaßnahmen zur Förderung des selbstregulierten Lernens bei Vorschulkindern kombiniert: Eine direkte Intervention im Rahmen eines Vorschulprojekts unter Anleitung geschulter Projektmitarbeiter sowie zwei indirekte Interventionen in Form von speziellen Eltern- und Erziehertrainings.

Um Maßnahmen der Förderung selbstregulierten Lernens in dieser Altersstufe evaluieren zu können, sind zielgruppenadaptive Verfahren notwendig. So können aufgrund der fehlenden Schriftsprache der Vorschulkinder keine textbasierten Instrumente angewandt werden. Basierend auf einer Untersuchung der im CIndLe-Projekt (Whitebread et al., 2009; Bryce & Whitebread, 2012) eingesetzten Train Track Task werden deshalb in der vorliegenden Studie ein videobasiertes Beobachtungsinstrument sowie Ratingskalen (CHILD Checklist; Anderson et al., 2003; Whitebread et al., 2005) eingesetzt, mit Hilfe derer die ErzieherInnen das kindliche selbstregulierte Verhalten einschätzen (Büttner, Perels & Whitebread, 2011). Im vorliegenden Beitrag wurde die CHILD-Checklist zusätzlich adaptiert, um eine Einschätzung durch die Eltern erfassen zu können.

An der Studie nahmen ca. 150 Kinder, 35 ErzieherInnen sowie 50 Elternteile teil. Vor Beginn und zum Ende der Intervention wurde die abhängige Variable selbstreguliertes Lernverhalten der Kinder durch das videobasierte Beobachtungsinstrument sowie der Einschätzungen der Eltern- und ErzieherInnen mit Hilfe die CHILD-Checklist erfasst. Im aktuellen Beitrag sollen die Ergebnisse herangezogen werden, um die Einschätzskalen für die Erzieherinnen und die Eltern auf ihre testtheoretische Güte hin zu überprüfen.

Im Rahmen des Vortrages möchten wir die Ergebnisse mit dem Fachpublikum diskutieren.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

3

Steffensky, Mirjam; Taskinen, Päivi; Heinig, Elisa; Anders, Yvonne

Frühe naturwissenschaftliche Bildung im Elternhaus

Theoretischer Hintergrund und Forschungsfragen: Vor dem Schuleintritt haben Eltern neben den Kindertagesstätten eine wichtige Funktion in der akademischen Bildung der Kinder. Während es bereits vertieft untersucht wurde, wie Eltern ihre Kinder im Lernen von Mathematik oder Lesen vor dem Schuleintritt unterstützen können, sind über naturwissenschaftsbezogene Bildungsprozesse im Elternhaus kaum Erkenntnisse vorhanden.

Es wird angenommen, dass die frühen Eltern-Kind-Interaktionen das Interesse der Kinder an Naturwissenschaften wecken und ihr Lernen in diesem Bereich fördern können (Alexander, Johnson, & Kelley, 2012); vor allem ihr Verständnis von naturwissenschaftlichen Phänomenen und Denk- und Arbeitsweisen soll hier angesprochen werden. Generell kann bei der elterlichen Unterstützung nach der Quantität und Qualität unterschieden werden. Die Quantität findet ihren Ausdruck vor allem in den Strukturmerkmalen der Elternhäuser, wie beispielsweise in der Anzahl von bildungsrelevanten Materialien oder verschiedenen gemeinsamen bildungsrelevanten Aktivitäten. Die Qualität fokussiert mehr auf die Prozessmerkmale der Elternhäuser: die Weisen, wie Eltern Lernprozesse ihrer Kinder unterstützen (Kluczniok, Lehl, Kuger, & Rossbach, 2013). Aus der Lernforschung ist beispielsweise bekannt, dass Stimulation zum Lernen oder angemessenes Begleiten der Lernprozesse durch Monitoring und gezielte Hilfestellung für eine hohe Qualität von Interaktionen sprechen. Die Forschung im Elementarbereich weist darauf hin, dass Aufrechterhaltung sozial geteilter kognitiver Prozesse - „sustained shared thinking“ - durch kognitive Aktivierung, Scaffolding, Anregung von Dialogen u. ä. für eine hohe Qualität der Interaktionen spricht (Hopf, 2012; Siraj-Blatchford, Sylva, Muttock, Gilden & Bell, 2002).

Das Ziel des Projektes ist es zunächst, im Rahmen einer Pilotierungsstudie Instrumente für die Erhebung von elterlichen Struktur- und Prozessmerkmalen zu entwickeln. Das ermöglicht die Wichtigkeit von verschiedenen naturwissenschaftsbezogenen Eltern-Kind-Interaktionen oder idealtypische Familien im Rahmen weiterer Untersuchungen festzustellen.

Methode: Für die Erstellung der Instrumente werden zunächst Interviews mit Eltern von 5-6 jährigen Kindern (N = 5) zur naturwissenschaftlichen Bildung im Elternhaus durchgeführt. Es folgen einige Beobachtungen von Eltern, während sie mit ihren Kindern in einem vorgegebenen Lernsetting interagieren (N = 3). Die Erkenntnisse fließen in die Erstellung der Instrumente ein. Die zusammengestellten Instrumente werden in einer Pilotstudie (Dezember 2014 –Januar 2015) erprobt und auf ihre Güte hin untersucht (n = 80).

Ausblick auf die Ergebnisse: Die Ergebnisse der Pilotierungsstudie ermöglichen eine Zusammenstellung eines Instruments zur Messung der Struktur- und Prozessmerkmale von naturwissenschaftlichen Bildungsangeboten im Elternhaus. Die Ergebnisse können zum Verständnis der elterlichen Unterstützung beim Lernen von Naturwissenschaften beitragen. Sie liefern Hinweise, welche Praktiken Eltern nutzen, das Lernen ihrer Kinder in Naturwissenschaften zu fördern. Die Daten für die weiterführenden Fragestellungen werden im Rahmen der Hauptstudie erhoben (2. Quartal 2015). Die Anlage der Studie ermöglicht es, die elterlichen Merkmale mit den kindlichen Merkmalen in Beziehung zu setzen sowie die Bildungsqualität des Kindergartens gleichzeitig zu berücksichtigen. Somit ist es möglich, zu genaueren Erkenntnissen zu erlangen, inwieweit bestimmte elterliche Aktivitäten bzw. Maßnahmen für die kindlichen Kompetenzen vorteilhaft sein können.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

Ziegler, Tobias; Hardy, Ilonca; Taskinen, Päivi; Leuchter, Miriam; Steffensky, Mirjam

Naturwissenschaftliche Kompetenzen im Vorschulalter: Erfassung von prozessualen und konzeptuellen Wissenskomponenten

Theoretischer Hintergrund: Naturwissenschaften sind ein Schwerpunkt in den Bildungsplänen für den Elementarbereich (Fthenakis, 2009). Anlehnend an das Konzept Scientific Literacy (Bybee et al., 2009), werden unter naturwissenschaftlicher Kompetenz auch naturwissenschaftliches Wissen und Wissen über die Naturwissenschaften gefasst. Verschiedene empirische Befunde weisen darauf hin, dass schon Kinder im Vorschulalter über spezifische naturwissenschaftliche Kompetenzen verfügen (Zimmermann, 2007). So zeigen entwicklungspsychologische Studien (u.a. Sodian et al., 2006), dass Kinder zwischen Theorie und Evidenz unterscheiden und somit zumindest auf einfachem Niveau Bezüge zwischen Theorie und Evidenz herstellen können (Stephan-Gramberg & Hardy, 2013). Interventionsstudien (Samarapungavan et al., 2009; Carstensen et al., 2011) belegen, dass sich naturwissenschaftliches Wissen von Vorschulkindern in einzelnen Domänen (Biologie, Physik) und Inhaltsgebieten (Aggregatzustände) erfassen lässt. Dieses (anschlussfähige) Wissen ist dabei von hoher Relevanz für das weitere naturwissenschaftliche Lernen in der Primarstufe.

Es finden sich somit empirische Hinweise auf konzeptuelle und prozessbezogene Kompetenzen von Vorschulkindern. Deutlich wird auch, dass über die grundsätzliche Struktur naturwissenschaftlicher Kompetenz bei jungen Kindern bislang wenig bekannt ist (Carstensen et al., 2011). Insbesondere die Breite naturwissenschaftlich relevanter Inhaltsgebiete wurde in bisherigen Studien kaum berücksichtigt. Wissen über die Struktur und inhaltliche Breite naturwissenschaftlicher Kompetenz ist jedoch wichtig, um geeignete Lerngelegenheiten im Elementarbereich konzipieren und deren Wirkung empirisch überprüfen zu können.

Fragestellung: Im Rahmen des Projekts Early Steps Into Science (EASI-Science) werden Kindergartenkinder zu domänen-spezifischen und prozessbezogenen Aspekten naturwissenschaftlicher Kompetenz befragt. Im Mittelpunkt der Konstruktion geeigneter Instrumente steht die altersgemäße, standardisierte und reliable Entwicklung von geeigneten Testformaten sowie die Analyse auf Konstruktvalidität durch korrelative Befunde der eingesetzten Skalen.

Methode: Mit Blick auf die inhaltliche Breite naturwissenschaftlicher Themengebiete wurden insgesamt sieben standardisierte prozess- und domänenbezogene Testinstrumente entwickelt und pilotiert (Sommer und Herbst 2014). In den Pilotierungsstudien wurden bislang 157 Kinder aus 19 Kindertagesstätten in Frankfurt, Jena und Kiel an unterschiedlichen Standorten befragt. Die Testinstrumente bestehen aus Multiple-Choice-Aufgaben, offenen Aufgaben, Richtig-Falsch-Aufgaben und Zuordnungsaufgaben. Das naturwissenschaftliche Wissen wird unter Berücksichtigung altersangemessener Konzepte und Begriffe in den Domänen Aggregatzustände, Magnetismus, Materie und Schwimmen und Sinken erhoben. Die Instrumente zu den prozessbezogenen Kompetenzen fokussieren auf die Prozesse Vorhersagen, Beobachten, Erklären und Vergleichen. Dabei wird erfasst, ob die Kinder zwischen verschiedenen Prozessen (bspw. dem Beobachten und Vorhersagen) unterscheiden können und ob sie innerhalb eines Prozesses qualitativ höherwertige Aussagen erkennen können. Beispielsweise sind Begründungen, die sich auf Gesetzmäßigkeiten beziehen, qualitativ hochwertiger einzuschätzen, als einfache Behauptungen ohne weitere Stützung der Position (Hardy et al., 2010).

Ergebnisse: Die Pilotierungsstudien zeigen, dass das konzeptuelle Wissen in den vier anvisierten Inhaltsbereichen standardisiert und reliabel erfasst werden kann (z.B. Magnetismus: Cronbachs $\alpha = .78$; Prozesswissen: $.77$). Auf Grundlage der ersten Pilotierungsergebnisse zeigt sich zudem ein signifikanter Zusammenhang zwischen prozessbezogenen und konzeptbezogenen Wissenskomponenten. Im Rahmen der Hauptstudie soll der Frage nachgegangen werden, ob sich die weiteren Wissenskomponenten naturwissenschaftlicher Kompetenz empirisch trennen lassen. Hierfür werden auch Instrumente zum Wissenschaftsverständnis von jungen Kindern einbezogen. Als Kontrollvariablen werden zum einen das sprachliche und kognitive Verständnis und soziostrukturelle Merkmale der Kinder erhoben. Zum anderen werden die Ausprägungen der professionellen Kompetenz der betreuenden Fachkräfte sowie struktur- und prozessbezogene Merkmale der Einrichtung berücksichtigt.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

Tresp, Timo; Hardy, Ilona; Steffensky, Mirjam; Ziegler, Tobias

Konzeptualisierung und Erfassung naturwissenschaftlichen Professionswissens frühpädagogischer Fachkräften

Theoretischer Hintergrund: In den letzten Jahren haben sich in Deutschland die Erwartungen an frühkindliche Bildungseinrichtungen sowie damit einhergehend die Anforderungen an die professionelle Kompetenz frühpädagogischer Fachkräfte stark verändert (Anders, 2012; Fröhlich-Gildhoff, et al., 2011). Pädagogische Fachkräfte in Kindertageseinrichtungen übernehmen eine zentrale Rolle in der Gestaltung von Bildungsprozessen. Für die frühe naturwissenschaftliche Bildung wird hierbei das Ziel verfolgt, ein grundlegendes, alltagsnahes Wissen zu entwickeln und anschlussfähige Vorstellungen als Basis für differenzierte wissenschaftliche Konzepte zu fördern (Steffensky et al., 2012).

Vor dem Hintergrund dieser Aufgabe stellt sich die Frage nach der Kompetenz frühpädagogischer Fachkräfte, die für eine angemessene, anschlussfähige Förderung früher naturwissenschaftlicher Bildungsprozesse notwendig ist.

Erkenntnisse der Lehrerprofessionsforschung zeigen, dass das Professionswissen ein wichtiger Einflussfaktor auf die Instruktionsqualität des Unterrichts und somit zentral für erfolgreiche Lernprozesse von Schülerinnen und Schülern ist (Baumert & Kunter, 2006; 2011; Hill et al., 2005; Lipowsky, 2006). In Anlehnung an die Taxonomie Shulmans (1986, 1987) hat sich eine Einteilung des Professionswissens in die drei Dimensionen Fachwissen (content knowledge, CK), fachdidaktisches Wissen (pedagogical content knowledge, PCK) und pädagogisches Wissen (pedagogical knowledge, PK) etabliert. Das fachwissenschaftliche Wissen kann dabei in ein vertieftes Verständnis des Lernstoffs sowie ein übergeordnetes Wissen über das Fach differenziert werden (Ball et al., 2008; Baumert & Kunter, 2011). Für den Bereich des fachdidaktischen Wissens kristallisieren sich das Wissen über Schülerkognitionen sowie das Wissen über Instruktionsstrategien als relevante Aspekte heraus (Depaepe et al., 2013).

Für den Elementarbereich existieren bislang zwar allgemeine Modelle zur professionellen Handlungskompetenz (Fröhlich-Gildhoff et al., 2011), sowohl domänenunspezifisch als auch spezifisch für den naturwissenschaftlichen Bereich gibt es aber kaum empirische Erkenntnisse zur Struktur und Bedeutung des Professionswissens frühpädagogischer Fachkräfte.

In dem Projekt EASI Science (Early Steps Into Science) wird hier angesetzt und die Struktur der naturwissenschaftlichen Kompetenz frühpädagogischer Fachkräfte und deren Wirkung auf die Kompetenz jüngerer Kinder untersucht. Auf dem Poster werden die Konzeptualisierung, Instrumente und erste Ergebnisse bezüglich des fachbezogenen Professionswissens frühpädagogischer Fachkräfte zu den Themen Aggregatzustände, Material, Magnetismus und Schwimmen und Sinken sowie naturwissenschaftliche Arbeitsweisen dargestellt.

Konzeptualisierung des fachbezogenen professionellen Wissens: In Anlehnung an die Forschung zum fachbezogenen Professionswissen von Lehrkräften wird bei der Konzeptualisierung in diesem Projekt für das fachdidaktische Wissen differenziert zwischen 1.) Wissen über Kognitionen der Kinder und 2.) Wissen über Instruktionsstrategien sowie 3.) Wissen über die Unterstützung naturwissenschaftsbezogener Bildungsprozesse. Unter 1. werden z.B. Wissen über spezifische Schwierigkeiten bezüglich einzelner Inhaltsbereiche sowie das Erkennen von nicht oder weniger anschlussfähigen Vorstellungen gefasst. Zu 2. gehören z.B. das Wissen über Situationen, Experimente sowie weitere Zugänge zu bestimmten naturwissenschaftlichen Inhaltsgebieten sowie das Erkennen des Lernpotenzials von Alltagssituationen, während unter 3.) z.B. Wissen über das Aufmerksam machen auf Phänomene, die Anregung zum Nachdenken und die Strukturierung von Bildungsprozessen gefasst wird.

Bezüglich des Fachwissens differenzieren wir das Wissen über die konkreten Inhaltsbereiche auf einem Alltagsniveau sowie ein in Ansätzen übergeordnetes Wissen über z.B. zentrale Konzepte der Naturwissenschaften und deren Bedeutung für die hier im Vordergrund stehenden Inhaltsbereiche.

Fragestellungen: (1) Lassen sich das Fachwissen und das fachdidaktische Wissen frühpädagogischer Fachkräfte für die unterschiedlichen Inhaltsbereiche reliabel erfassen? (2) Wie hängen die beiden Dimensionen des fachspezifischen Professionswissens zusammen?

Vorgehen und Methode: Aktuell werden Aufgaben zu den genannten Inhaltsbereichen und den beiden Wissensdimensionen entwickelt. Dabei werden Multiple-Choice, True-False, sowie offenen Antwortformate verwendet. Fachwissen und fachdidaktisches Wissen sind eng aufeinander abgestimmt. Die entwickelten Aufgaben werden Fachkräften zur Einschätzung des Niveaus und der Relevanz für den Kita-Alltag gegeben (face validity) vorgelegt.

Ausblick: Die Erhebung und Auswertung der Daten erfolgt in den kommenden Wochen an einer Stichprobe an ca. 100 Fachkräften. Auf dem Poster werden ausgewählte Ergebnisse aus den Pilotierungen der Instrumente dargestellt, die einerseits die Testgüte sowie andererseits die Zusammenhänge zwischen Fach- und fachdidaktischem Wissen betreffen.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

Balt, Miriam; Ehlert, Antje; Fritz-Stratmann, Annemarie

Schuleingangsdiagnostik auf dem Prüfstand. Wie sensibel ist die Schuleingangsdiagnostik beim Erkennen von Rechenschwierigkeiten?

Theoretischer Hintergrund: Der Übergang vom Elementar- in den Primarbereich stellt ein normatives kritisches Lebensereignis dar, das von den Kindern in relativ kurzer Zeit Anpassungsleistungen auf verschiedenen Ebenen abverlangt. Zwar stellt der Schulbeginn für viele Kompetenzbereiche wie dem Rechnen keineswegs die „Stunde null“ dar (Selter, 2008), jedoch differieren die Kinder z.T. stark bezüglich ihrer mathematischen Entwicklungsstände (z.B. Wittmann, 2001). Studien wie die SCHOLASTIK- oder LOGIK-Studie (Weinert & Helmke, 1997; Weinert & Schneider, 1999) belegten eine interindividuelle Stabilität in den Leistungen vom letzten Kindergartenjahr über die Schuljahre hinweg und die Bedeutung bereichsspezifischen Vorwissens für die weitere Lernentwicklung. Krajewski und Schneider (2006) wiesen einen hohen Zusammenhang zwischen mengen- und zahlenbezogenem Vorwissen und Mathematikleistungen in der Grundschulzeit nach und stellten das Vorwissen als guten Vorhersagefaktor heraus. Aunola, Leskinen, Lerkkanen und Nurmi (2004) zeigten überdies einen Schereneffekt auf, demzufolge sich Kinder, die bereits vor ihrer Einschulung über ein schwaches Mengen-Zahlen-Verständnis verfügten, im schulischen mathematischen Unterricht deutlich langsamer entwickelten. Diese Befunde legen eine möglichst frühe diagnostische Prüfung der mathematischen Entwicklung nahe. Auf diese Weise könnte rechtzeitig auf Entwicklungsrückstände im mathematischen Bereich reagiert und ein binnendifferenzierter Anfangsunterricht realisiert werden. Kinder, die durch geeignete Maßnahmen im Anfangsunterricht „abgeholt“ wurden, konnten zuvor festgestellte Defizite aufholen (vgl. Hasselhorn & Lohaus, 2008; Hildeschiedt, 1995).

Frage- und Zielstellung: Aufbauend auf diesen Befunden wird in der vorliegenden Studie folgenden Fragen nachgegangen: Wie verbreitet ist die Durchführung schulinterner Eingangsbefragungen zur Feststellung schulnaher Vorläuferkompetenzen (Kammermeyer, 2004), wie z.B. einem grundlegenden Mengenverständnis? Werden bei der Schuleingangsdiagnostik vermehrt informelle Verfahren verwendet oder wird auf standardisierte Tests zurückgegriffen? Wie hoch ist der Anteil an Aufgaben, die mathematisches Vorwissen erfassen? Welche mathematischen Konzepte werden mittels der verwendeten Aufgaben erfasst und entsprechen diese dem altersgerechten Entwicklungsstand?

Aus den Ergebnissen sollen Rückschlüsse gezogen werden, ob die Verfahren, die im Rahmen der Schuleingangsdiagnostik eingesetzt werden, dazu geeignet sind, Rechenschwierigkeiten frühzeitig zu erkennen, sodass Kinder bevor es zu den ersten schulischen Misserfolgen kommt, ihrem Entwicklungsstand entsprechend gefördert werden können.

Methoden: Hierfür wurden Befragungen in drei Bundesländern (Berlin, Brandenburg und Nordrhein-Westfalen) an rund 80 Schulen durchgeführt.

Die von den Schulen zur Verfügung gestellten mathematischen Diagnostikaufgaben wurden ausgehend vom mathematischen Entwicklungsmodell (Fritz & Ricken, 2008; Fritz, Ehlert & Balzer, 2013) theoriebasiert qualitativ analysiert. Dazu wurde zunächst bestimmt, welches mathematische Konzeptwissen benötigt wird, um die Aufgaben zu bewältigen. Entsprechend wurden dann die verschiedenen Aufgaben, wenn möglich, den mathematischen Entwicklungsniveaus des Modells zugeordnet. Diese Zuordnung wird zudem quantifiziert, so dass dokumentiert und deskriptiv beschrieben werden kann, wie viele Aufgaben, welche mathematischen Konzepte abprüfen.

Auf der Grundlage des theoriebasierten standardisierten und normierten mathematischen Vorschultests (MARKO-D von Ricken, Fritz & Balzer, 2013) soll außerdem normativ analysiert werden, inwiefern zwischen den Anforderungen der eingesetzten Instrumente und der altersentsprechenden mathematischen Entwicklungsnorm eine Passung besteht.

Ergebnisse: Der aktuelle Stand der Grundschulbefragung zeigt bereits auf, dass sich der Einsatz schulinterner Eingangsbefragungen bundeslandspezifisch in Abhängigkeit von den bildungspolitischen Rahmenbedingungen unterscheidet. Neben der in allen drei Bundesländern verpflichtenden schulärztlichen Untersuchung führt ein Großteil der Schulen in Brandenburg zusätzlich eigene Untersuchungen durch. In Berlin hingegen scheint dieses Vorgehen eher unüblich und in Nordrhein-Westfalen ist das Verhältnis bisher ausgeglichen.

Aus der Auswertung des bereits vorliegenden diagnostischen Materials kann zunächst geschlussfolgert werden, dass Aufgaben, die mathematisches Vorwissen erfassen, stark unterrepräsentiert sind. Dies gilt sowohl für die standardisierten als auch für die informellen Verfahren. Weiterhin deutet sich bei der Analyse der den Aufgaben zugrundeliegenden mathematischen Konzepte an, dass diese eher Anforderungen enthalten, die sehr frühe Entwicklungsniveaus erfordern (z.B. Aufsagen der Zahlwortreihe oder Zahlen der Größe nach ordnen). Dies wiederum bedeutet, dass die Verfahren nur eine sehr geringe Sensibilität aufweisen, um zum Schulstart bestehende Rechenschwierigkeiten aufzudecken.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

Raddatz, Julia; Kuhn, Jörg-Tobias; Holling, Heinz; Moll, Kristina; Dobel, Christian

Dyskalkulie, Dyslexie und Komorbidität: Unterschiede in den kognitiven Profilen von Kindern mit Lernschwierigkeiten

Eine Dyskalkulie oder Rechenstörung bezeichnet nach ICD-10 eine mathematik- oder rechenspezifische Minderleistung, die nicht auf eine geringe Intelligenz, mangelnde Beschulung oder neurologische Erkrankungen zurückführbar ist. Etwa 3-8% aller Kinder im Grundschulalter sind von dieser spezifischen Lernstörung betroffen (Shalev, Auerbach, Manor & Gross-Tsur, 2000). Sie weisen massive Schwierigkeiten bei den einfachsten mathematischen Aufgaben auf, wie zum Beispiel der Addition und der Subtraktion im Zahlenraum bis 10 sowie dem Einprägen arithmetischer Fakten (Geary, 1993; Jordan, Hanich & Kaplan, 2003). Neben diesen domänen-spezifischen Defiziten in der Zahlenverarbeitung, werden ebenso domänen-spezifische kognitive Defiziten bei Kindern mit Dyskalkulie berichtet, z.B. im Arbeitsgedächtnis (Swanson & Jerman, 2006). Obwohl es in den letzten Jahren eine steigende Anzahl von Publikationen in diesem Forschungsfeld gibt, sind die kausalen Mechanismen von Dyskalkulie bisher nicht vollständig erklärbar. Besonders zeichnet sich diese spezifische Lernstörung durch ihre phänotypische Heterogenität aus (Rubinsten & Henik, 2009). So konnte eine aktuelle Studie (Bartelet, Ansari, Voessen, & Blomert, 2014) bis zu 6 Cluster von unterschiedlichen kognitiven Profilen in einer großen Stichprobe von Kindern mit Dyskalkulie aufzeigen. Die Befunde lassen die Notwendigkeit einer individuellen Diagnostik sowie Intervention erkennbar werden. Deutlich hervorzuheben ist dabei die hohe Komorbidität zwischen Dyskalkulie und Dyslexie, die bei bis zu 47 % liegen kann (Landerl & Moll, 2010). Dieser Umstand stellt eine besonders fundamentale Beeinträchtigung für die betroffenen Kinder dar, die, neben Schwierigkeiten im mathematischen Bereich, zusätzlich Leseschwierigkeiten aufweisen (Landerl & Moll, 2010). Daraus ergeben sich massive Schwierigkeiten im Schulalltag, die es ihnen teilweise unmöglich machen, diesen angemessen zu bestreiten.

Die vorliegende Studie vergleicht die kognitiven Profile von N = 127 Grundschulkindern in den Bereichen basale und komplexe Zahlenverarbeitung (z.B. Mengenvergleich, Transkodieren), Rechnen und Arbeitsgedächtnis. Die untersuchten Kinder verteilen sich dabei auf vier Gruppen, die sich weder bezüglich ihres Alters noch ihres Intelligenzquotienten voneinander unterscheiden: ausschließlich Dyskalkulie (N = 20) oder Dyslexie (N = 40), eine komorbid beeinträchtigte Gruppe (N = 27) sowie eine unbeeinträchtigte Kontrollgruppe (N = 40). Unser Studiendesign ermöglicht es, zu überprüfen, ob sich das kognitive Profil der komorbiden Gruppe additiv aus den Profilen der beiden spezifischen Lernstörungen (Dyskalkulie und Dyslexie) ergibt (sog. Additivität; vgl. auch Moll, Göbel & Snowling 2014).

Die Ergebnisse stehen im Einklang mit aktuellen Befunden (Landerl, Fussenegger, Moll & Willburger 2009; Moll et al., 2014) und deuten auf einen additiven Effekt einer komorbid auftretenden Dyskalkulie und Dyslexie hinsichtlich der Beeinträchtigungsprofile hin. Dies spricht erstens gegen die Annahme, dass den kognitiven Profilen komorbid beeinträchtigter Kinder ein gemeinsames Defizit zugrunde liegt (Unteradditivität). Auch die Hypothese, dass komorbide Kinder über die Summe der beiden Einzeldefizite (Dyskalkulie plus Dyslexie) hinaus beeinträchtigt sind (Überadditivität), kann verworfen werden. Unsere Ergebnisse unterstreichen weiterhin, dass weder ein domänen-spezifisches Arbeitsgedächtnisdefizit noch ein ausschließliches Lesedefizit für die Leistung der komorbid beeinträchtigten Kinder angenommen werden kann. Kinder mit ausschließlich Leseschwierigkeiten weisen eine selektive mathematische Beeinträchtigung im Zählen auf. Ein Defizit im visuell-räumlichen Arbeitsgedächtnis kann bei Kindern mit ausschließlich Dyskalkulie jedoch, entgegen der Annahmen, nicht nachgewiesen werden. Diese Kinder sind allerdings, wie vermutet, in zahlreichen zahlenverarbeitenden Prozessen beeinträchtigt, wie zum Beispiel dem schnellen Erfassen von kleinen Mengen (Subitizing) oder dem präzisen Einordnen von Zahlen auf einem Zahlenstrahl.

Die vorliegende Studie unterstreicht, dass Kinder mit einer Dyskalkulie multiple, heterogene Defizite aufweisen (Dowker, 2005). Kinder mit einer Dyslexie hingegen weisen im mathematischen Bereich nur spezifische Schwierigkeiten in der Zahlenverarbeitung auf, und auch nur dann, wenn die bearbeiteten Aufgaben verbale Fähigkeiten voraussetzen (z.B. Zählen). Die Studie liefert damit wichtige Hinweise in Bezug auf die Entwicklung geeigneter Diagnostikverfahren für Dyskalkulie, Dyslexie oder eine komorbide Lernstörung. Nur eine fundierte Diagnose lässt individuelle, frühestmögliche Interventionsmaßnahmen zu, die es Kindern mit einer spezifischen Lernstörung ermöglicht, ihren Schul- und Lebensweg angemessen beschreiten zu können.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

3

Breier, Sina B.; Kipp, Kestin H.; Arndt, Petra A.

Auswirkungen der Heterogenität mathematischer Fertigkeiten im Vorschulalter auf schulische Leistungen – Ergebnisse aus der wissenschaftlichen Begleitung des Modellprojekts „Bildungshaus 3 – 10“

Am Übergang vom Kindergarten in die Grundschule zeigen Kinder kein einheitliches Leistungsspektrum. Moser et al. (2005) konnten für die Schweiz darlegen, wie heterogen die Leistungen der Kinder beim Übergang sind: während manche Schulanfänger schon Sätze lesen und bis 100 rechnen, kennen andere Kinder weder Buchstaben noch Zahlen. Knapp vier Fünftel der Kinder beherrschen bereits bedeutende Teile des Mathematikstoffs der ersten Klasse, ein Fünftel käme sogar in der zweiten Klasse zurecht. Für Deutschland zeigten Rabenstein et al. (1989, S. 12), dass lernschwache Kinder bereits mit den lehrplanmäßig vorgegebenen Grundleistungen des 1. Schuljahres überfordert sind und leistungsstarke Schüler auch Anforderungen höherer Schuljahre erfüllen könnten. Hasemann (2006) berichtet, dass bereits zu Schulbeginn eine große Heterogenität in den mathematischen Fähigkeiten vorliegt, ein Teil der Kinder schon mit erheblichen Vorkenntnissen im mathematischen Bereich in die Grundschulzeit startet, aber nicht alle Kinder eine gute Basis für den schulischen Mathematikunterricht mitbringen.

Kindertageseinrichtungen und Schulen sind dementsprechend gefordert, dieser Heterogenität bei der Förderung von Kindern Rechnung zu tragen. Vielfältige Projekte wurden initiiert, um entsprechende pädagogische Maßnahmen zu etablieren, und zwar sowohl in Kindertageseinrichtungen (z.B. „Schulreifes Kind“ Hasselhorn et al., 2012, „KidZ“, Roßbach et al., 2010) als auch in der Grundschule (z.B. „Schulanfang auf neuen Wegen“, Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2006). Ein Projekt, das beide Institutionen einbezieht ist das baden-württembergische Landesmodell „Bildungshaus 3 – 10“. Kernelement des Bildungshauses sind regelmäßig stattfindende, gemeinsame Spiel- und Lerneinheiten für Kindergarten- und Grundschulkindern, die von pädagogischen Fachkräften und Lehrkräften gemeinschaftlich vorbereitet und durchgeführt werden. Die Ausschreibung zum Modellprojekt sieht vor, dass die Pädagogen „jedes Kind individuell im gemeinsamen sozialen Rahmen“ fördern (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2007).

Das „Bildungshaus 3 – 10“ wird vom ZNL TransferZentrum für Neurowissenschaften und Lernen, Ulm, gefördert mit Mitteln des Bundesministeriums für Bildung und Forschung und des Europäischen Sozialfonds der Europäischen Union, wissenschaftlich begleitet. An der Evaluation nehmen 33 Modellstandorte sowie 28 Vergleichsstandorte teil. Dieses Experimental-Kontrollgruppen-Design erlaubt zu untersuchen, inwieweit Grundschulen der Heterogenität von Vorschulkindern im mathematischen Bereich Rechnung tragen und ob das „Bildungshaus 3 – 10“ einen Beitrag dazu leistet. Im Einzelnen wurde untersucht, wie stark die Kompetenzen vor Schuleintritt Mathematikleistungen in den verschiedenen Klassenstufen vorhersagen bzw. ob es den Schulen gelingt, die unterschiedlichen Voraussetzungen der Kinder zu berücksichtigen. Zudem wurden mögliche Unterschiede zwischen Kindern aus Bildungshäusern und Vergleichseinrichtungen untersucht.

Mathematische Basiskompetenzen wurden im letzten Quartal vor der Einschulung mit dem Osnabrücker Test zur Zahlbegriffsentwicklung (van Luit, van de Rijt & Hasemann, 2001) erfasst. In der Grundschule wurde am Ende der 2. bis 4. Klasse der Heidelberger Rechentest (Haffner, Baro, Parzer & Resch, 2005) eingesetzt. In die Analyse flossen die Daten von 267 Kindern ein. Zusätzlich wurden Daten zum sozioökonomischen Hintergrund einbezogen (N=230). Die Auswertungen erfolgten anhand von Mehrebenenanalysen.

Die bisherigen Analysen der Daten 2. und 3. Klassenstufen ergaben, dass Kinder, welche bereits vor dem Schulbeginn über bessere mathematische Fertigkeiten verfügen, auch in der 2. und 3. Klasse signifikant bessere mathematische Leistungen zeigen. Ein Ausgleich findet bis zum Ende der 3. Klasse nicht statt, die Daten für die 4. Klasse und Detailanalysen könnten hier weitere Informationen liefern.

Überprüft man den Einfluss des sozio-ökonomischen Hintergrunds der Kinder auf die mathematischen Leistungen, so zeigt sich, dass Kinder aus Familien mit einem niedrigen sozio-ökonomischen Status geringere schulische Leistungen aufweisen und der sozioökonomische Status die mathematischen Fertigkeiten im Kindergarten zusätzlich mediiert. Das bedeutet, dass der sozioökonomische Status der Kinder sowohl auf die im Kindergartenalter erworbenen Basisfertigkeiten wirkt, als auch darüber hinaus zusätzlich auf die Leistungen in der Grundschule. Eine Wirkung der pädagogischen Maßnahmen im „Bildungshaus 3-10“ konnte bisher nicht nachgewiesen werden. In weiteren Analysen sollen die Daten der 4. Klassen sowie institutionelle Qualitätsmerkmale einbezogen werden.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

3

Sturmhöfel, Nicole; Fäsche, Anika; Arndt, Petra A.

Der Einfluss jahrgangs- und institutionsübergreifender Lern- und Spielangebote auf sozial-emotionale Kompetenzen von Kindern im letzten Kindergartenjahr

In den letzten Jahren wird immer häufiger konstatiert, dass Kinder in ihren Fähigkeiten und Lernvoraussetzungen heterogener werden. Von Kindergarten und Grundschule wird deshalb vermehrt ein konstruktiver Umgang mit der Verschiedenheit der Kinder erwartet (z.B. Trautmann & Wischer, 2011). Auch die Gestaltung des Übergangs zwischen den Institutionen soll sich an der individuellen Bildungsbiografie des Kindes orientieren (JFMK & KMK, 2009). An dieser Stelle setzt das Modellprojekt „Bildungshaus 3–10“ an. Der auf sieben Jahre (2007–2014) angelegte Modellversuch des Ministeriums für Kultus, Jugend und Sport Baden-Württemberg besteht aus 32 Bildungshäusern mit jeweils einer Grundschule und mindestens einem Kindergarten. Mit der wissenschaftlichen Begleitung des Modellprojekts ist seit September 2008 das ZNL TransferZentrum für Neurowissenschaften und Lernen in Ulm betraut. Ziel des Modellversuchs ist eine intensiviertere Verzahnung der pädagogischen Arbeit von Kindergarten und Grundschule, die Kindern eine bruchlose Bildungsbiographie ermöglichen soll. Dabei wird die Orientierung an den Voraussetzungen, Potenzialen und dem Tempo des einzelnen Kindes betont.

Studien zeigen, dass Kinder z.B. mit sehr unterschiedlichen Lese- und Rechenkenntnissen eingeschult werden. Über soziale, emotionale und motivationale Kompetenzen liegen bisher wenig Untersuchungsergebnisse vor. Neben ihrem Einfluss auf akademische Leistungen gelten diese als Entwicklungsressourcen und Resilienzfaktoren, welche die psychische und soziale Anpassungsfähigkeit unterstützen (Malti & Perren, 2008). Die vorliegende Studie fokussiert deshalb auf nicht-akademische Kompetenzen im letzten Kindergartenjahr. Es wird von der Hypothese ausgegangen, dass sich die für das „Bildungshaus 3-10“ kennzeichnenden institutions- und jahrgangsübergreifende Lern- und Spielangebote positiv auf entsprechende Kompetenzen der Kindergartenkinder auswirken. Im Rahmen der regelmäßig stattfindenden Angebote bietet sich die Möglichkeit zur Förderung kooperativer Einstellungen und des sozialen Verhaltens sowie zum Lernen am Modell. Bei Untersuchungen im schulischen Kontext konnten in jahrgangsübergreifenden Klassen leichte Vorteile im sozialen Bereich nachgewiesen werden (Kucharz & Wagener, 2007). Die Effekte jahrgangsübergreifenden Lernens und deren Bedeutung sind jedoch umstritten (Roßbach, 1999). Auswirkungen altersgemischten Lernens im Übergang vom Kindergarten in die Grundschule wurden bisher wenig empirisch untersucht.

Die vorliegende Studie geht vor diesem Hintergrund zwei Fragestellungen nach: Zeigen sich Unterschiede zwischen sozial-emotionalen Kompetenzen von Kindern im letzten Kindergartenjahr, die ein „Bildungshaus 3-10“ besuchen im Vergleich zu Kindern, die eine andere Form der Kooperation zwischen Kindergarten und Grundschule erleben (Kontrollgruppe)? Profitieren Kinder im letzten Kindergartenjahr in ihren sozial-emotionalen Kompetenzen, je intensiver sie in Bildungshausangeboten involviert sind?

Zur Erfassung der sozial-emotionalen Kompetenzen der Kindergartenkinder wurden die Einschätzungen von Erzieherinnen genutzt. Diese wurden mit dem Beobachtungsbogen „Positive Entwicklung und Resilienz im Kindergartenalltag“ (Mayr & Ullrich, 2005) erhoben, welcher sechs Skalen umfasst: Kontaktfähigkeit, Selbststeuerung/Rücksichtnahme, Selbstbehauptung, Stressregulierung, Aufgabenorientierung, Explorationsfreude. Die Intensität der institutionsübergreifenden Kooperation (Häufigkeit der Teilnahme an Bildungshausangeboten pro Woche bzw. Stunden pro Woche in Bildungshausangeboten) wurde mit einem selbst entwickelten Fragebogen erfasst. Den Analysen liegen Daten von Kindern im letzten Kindergartenjahr in drei Kohorten (2008/09, 2009/10, 2011/12) zugrunde (Gesamtstichprobe n = 976). Die Auswertung erfolgte anhand hierarchisch linearer Modelle. Als Moderatorvariablen wurden das Geschlecht des Kindes, der Schulabschluss der Mutter, Zuwanderung, Familiensprache und das Familieneinkommen berücksichtigt.

Die Analysen ergeben, dass die Erzieherinnen die Kontaktfähigkeit der Kinder, die ein „Bildungshaus 3-10“ besuchen, negativer bewerten als dies Erzieherinnen in der Kontrollgruppe tun. Dieser Effekt findet sich in allen drei Kohorten. Überprüft man, welchen Einfluss die Intensität der Bildungshausangebote (erhoben für 2009/10 und 2011/12) auf die sozial-emotionalen Kompetenzen der Kinder in den Modelleinrichtungen hat, zeigt sich: Die Häufigkeit der Teilnahme pro Woche beeinflusst die Skalen Kontaktfähigkeit und Explorationsfreude im Erhebungsjahr 2009/10 signifikant positiv. Die mittlere Stundenzahl pro Woche zeigt keinen Einfluss. Ebenso finden sich keine Effekte für das Erhebungsjahr 2011/12. Die Auswertung weiterer moderierender Variablen zeigt u.a. Geschlechtereffekte zugunsten der Mädchen und einen positiven Einfluss eines hohen Schulabschlusses der Mutter.

Donnerstag
12. März

10:45 Uhr

Foyer GB - West

3

Bach, Andreas; Fischer, Thomas; Wolke, Julia; Biedermann, Horst

ISabEL (Intensität und Stabilität berufsbezogener Einstellungen im Lehramtsstudium)

Das ursprünglich aus der Sozialpsychologie stammende Konstrukt der Einstellung ist in den vergangenen Jahren zunehmend im Rahmen der Lehrerbildungsforschung untersucht worden. Neben den Kompetenzfacetten des Wissens und Könnens bilden Einstellungen bzw. Überzeugungen in den derzeit diskutierten lehrerberufsspezifischen Kompetenzmodellen eine eigenständige Kompetenzfacette und damit einen zentralen Bestandteil der beruflichen Handlungskompetenz von Lehrkräften. Es wird davon ausgegangen, dass Einstellungen das Handeln von Lehrkräften bedeutsam beeinflussen (Baumert & Kunter, 2006; Lipowsky, 2006; Richardson, 1996), wobei in empirischen Studien bislang häufiger epistemologische und hier insbesondere domänenspezifische Überzeugungen für das Fach Mathematik in den Blick genommen worden sind (Biedermann, Brühwiler & Krattenmacher, 2012). Weitaus seltener untersucht wurden bislang unterrichts- und schulbezogene Einstellungen. Die unter dem Begriff der „Konstanzer Wanne“ bekannt gewordenen Untersuchungen zu Lehrereinstellungen der 1970er Jahre analysierten Einstellungsänderungen als Sozialisationseffekte des Studiums und der Schule. Aussagekräftige Längsschnittstudien zur Entwicklung berufsbezogener Einstellungen bei Lehramtsstudierenden insbesondere unter Berücksichtigung schulpraktischer Lerngelegenheiten liegen im Anschluss an diese Untersuchungen kaum vor.

Das Forschungsprojekt ISabEL (Intensität und Stabilität berufsbezogener Einstellungen im Lehramtsstudium), das vom Zentrum für Lehrerinnen- und Lehrerbildung der Europa-Universität Flensburg seit 2013 durchgeführt wird, greift dieses Forschungsdesiderat auf und untersucht, wie sich berufsbezogene Einstellungen von Lehramtsstudierenden im Verlauf des Studiums und in schulpraktischen Ausbildungsphasen entwickeln. Folgende Fragestellungen werden verfolgt: (1) Einstellungsänderungen: Wie entwickeln sich berufsspezifische Einstellungen im Verlauf des Lehramtsstudiums und während schulpraktischer Ausbildungsphasen? (2) Differentielle Einstellungsänderungen: Existieren bedeutsame Unterschiede zwischen den Veränderungen in Studien- und Praxisphasen (Praxissemester) sowie zwischen unterschiedlichen Lehramtstypen (Grundschule versus Gemeinschaftsschule)? (3) Bedingungen der Einstellungsänderungen: In welcher Wechselwirkung stehen die Veränderungsprozesse mit individuellen Faktoren (u.a. Studieninteresse, Beanspruchungserleben, Lehrer-Selbstwirksamkeitserwartung der Studierenden)? Welche Einfluss haben pädagogische Vorerfahrungen der Studierenden?

Der Fokus des Forschungsprojekts liegt auf berufsbezogenen Einstellungen zum Unterricht und zur Schule. Bisher fehlt es weitgehend an geeigneten Instrumenten, um jene Einstellungen empirisch quantitativ zu erfassen. Im Rahmen einer Pilotstudie konnten Einstellungsskalen entwickelt und validiert werden, die im Längsschnitt des Projekts eingesetzt werden. Die Skalen beziehen sich auf Einstellungen zur Komplexität von Unterricht, zum Classroom Management, zum Verhältnis von Theorie und Praxis, zur individuellen Entwicklungsbereitschaft sowie zu Funktionen von Schule.

Das Forschungsprojekt ist als Fragebogenstudie im Längsschnittdesign konzipiert und umfasst jeweils drei Messzeitpunkte im Bachelorstudium und Masterstudium, wobei auch das Praxissemester als eigene Phase untersucht wird. Die Stichprobe bilden Studierende des Bachelorstudiengangs Bildungswissenschaften sowie der Masterstudiengänge Lehramt an Grundschulen und an Gemeinschaftsschulen der Europa-Universität Flensburg (N = ca. 720 im Bachelor bzw. im Master N = ca. 400). Die Datenauswertung erfolgt mit Hilfe linearer Strukturgleichungsmodelle mit dem Programm Mplus. Geplant sind Latent-Change-Analysen zur Veränderungsmessung.

Das Forschungsprojekt ISabEL sowie erste Ergebnisse der neu entwickelten Instrumente zur Einstellungsmessung sollen vorgestellt werden. Dabei liegt der Schwerpunkt auf der faktoriellen Validität der Einstellungsskalen.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Rath, Viktoria; Reinhold, Peter

Modellierung situationsspezifischer diagnostischer Kompetenz von Physiklehramtsstudierenden

Fragestellung und Zielsetzung: Im Projekt soll die fachspezifische diagnostische Kompetenz, wie sie beim Beobachten und Analysieren von Präkonzepten bei Gruppenkommunikation im Physikunterricht bedeutend ist, modelliert und operationalisiert werden. Vor dem Hintergrund der bisherigen Forschungsergebnisse fehlt insbesondere für die diagnostische Kompetenz von Präkonzepten sowohl ein Modell als auch eine Operationalisierung guter Diagnoseleistung. Ausgehend von dieser Problematik wurden ein Modell der kognitiven Wissensstrukturen von diagnostischer Kompetenz, die Lehramtsstudierende an der Universität erwerben können, und Merkmale einer guten Diagnose erarbeitet. Auf dieser Grundlage wurde ein videobasiertes und Interviewleitfaden gestütztes, handlungsnahes Testinstrument entwickelt, mit dem die diagnostische Performanz erhoben werden soll. Von dieser qualitativ erhobenen, diagnostischen Performanz soll auf die dahinter liegende situationsspezifische, beobachtungsbezogene diagnostische Kompetenz und mögliche Zusammenhänge dieser Kompetenz mit fachdidaktischem Wissen, pädagogischem Wissen und Fachwissen als Bereiche des Professionswissens geschlossen werden.

Theoretischer Hintergrund und Modellierung: Da für die Domäne Physik nur wenige Ansätze zur diagnostischen Kompetenz (Cappell 2013, Draude 2013) existieren, wurden für die Modellentwicklung zusätzlich angrenzende Domänen (COACTIV, Karst 2012) sowie fachübergreifende Ansätze (Helmke 2009, Schrader 1989 & 2008, Abs 2007) analysiert (vgl. Rath & Reinhold, 2013). Weitere Anhaltspunkte lieferten Modellierungen zum Professionswissens wie beispielsweise Shulmann (1987), Bromme (1997), Baumert & Kunter (2006), Riese (2009), Blömeke et al. (2011), Gramzow et al. (2013). Aus diesen theoretischen Überlegungen ergibt sich, dass für die situationsspezifische diagnostische Kompetenz Wissensfacetten aus allen drei Wissensbereiche des Professionswissens in das Modell einzubeziehen sind: Facetten des Fachwissen werden benötigt, um die Richtigkeit von Schüleräußerungen beurteilen zu können. Zur genaueren Einschätzung der Äußerungen sind Facetten fachdidaktischen Wissens erforderlich. Um das Diagnoseurteil und den -prozess zu reflektieren und einzuordnen, werden Elemente pädagogischen Wissens benötigt.

Das Modell wird empirisch überprüft, indem die Wissensfacetten der drei Wissensbereiche sowie die diagnostische Performanz erhoben werden. Um auf die diagnostische Kompetenz zu schließen, wird anhand der gezeigten Handlungen das dahinter liegende Wissen (Wissen III im Sinne nach Neuweg (2011)) rekonstruiert.

Für die Erfassung und Bewertung der diagnostischen Performanz wurden Qualitätsmerkmale einer guten Diagnose (bezogen auf die obigen geschilderten Situationen) in mehreren Schritten entwickelt. Im ersten Schritt wurden grundlegende Überlegungen aus der Pädagogik (Abs 2007, Helmke et al. 2004), der pädagogischen Psychologie (Ophysen 2010, Bruder et al. 2010) und verschiedener Fachdidaktiken (Cappell 2013, Karst 2012, Kunter et al. 2011) gesichtet. Ebenfalls wurden normative Setzungen zur Lehrerbildung (KMK 2004, Oser 1997) mit hinzugezogen. Nachfolgend wurden die betrachteten Aspekte guter Diagnostik auf die vorliegende Situation bezogen und konkretisiert. Nicht relevante Aspekte wurden gestrichen. Zuletzt wurden Indikatoren entwickelt und den vier übergeordneten Merkmalen Urteilsgenauigkeit, Differenziertheit, Urteilsgüte und Qualität der Reaktionsplanung einer guten Diagnose zugeordnet.

Testinstrument, Ergebnisse und Ausblick: Den Ausgangspunkt für die Entwicklung des Testinstruments zur Erhebung der diagnostischen Performanz stellen videographierte Gruppendiskussionen von 3-4 Lernenden beim Lösen von Physikaufgaben (Mechanik) dar. Hieraus wurden in mehreren Schritten Videosequenzen anhand der folgenden Kriterien Kommunikationsstruktur, fachliche Komplexität und Analysekomplexität der Präkonzepte ausgewählt. Anschließend wurden diese Sequenzen Experten zur inhaltlichen Validierung vorgelegt. Die Ergebnisse dieser Teilstudie werden im Vortrag präsentiert. Zur Erhebung der diagnostischen Performanz sollen die ausgewählten Videosequenzen den Probanden gezeigt werden. Das Testinstrument besteht dann - ausgehend von diesen Videosequenzen - aus einem Leitfaden gestützten Interview und einem Kriterienkatalog für die Inhaltsanalyse der Interviews, der auf den Merkmalen einer guten Diagnose basiert. Zur Durchführung wird nach jeder einzelnen Videosequenz nach den diagnostizierten Präkonzepten, einer möglichen Anschlusshandlung und dem Diagnoseprozess gefragt. Hierzu wurde ein Interviewleitfaden entwickelt und erprobt. Die Daten zur Pilotierung des Testinstruments werden ebenfalls im Vortrag präsentiert. Im Anschluss an die Pilotierungsstudien ist die Haupterhebung mit N=20 geplant.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Klingsieck, Katrin; Müsche, Hanna; Praetorius, Saskia

Entwicklung und Erprobung einer Skala zur Erfassung der Selbstwirksamkeit zum Diagnostizieren lern- und leistungsrelevanter Merkmale im Schulalltag

Gerade im Umgang mit Heterogenität im Kontext der inklusiven Schule kommt der diagnostischen Kompetenz der Lehrkräfte ein besonderer Stellenwert zu. Die diagnostische Kompetenz zählt neben der Sachkompetenz, der didaktischen Kompetenz und der Klassenführungskompetenz zur adaptiven Lehrkompetenz (Beck et al., 2008), welche wiederum für professionelles Handeln in inklusiven Settings essentiell ist (Seitz & Scheidt, 2012). Das Monitoring der Entwicklung der diagnostischen Kompetenz bereits in der universitären Ausbildungsphase, und darüber hinaus, ist demnach äußerst wichtig. Dazu bedarf es ökonomischer Instrumente, die zum Beispiel zur Evaluation und anschließenden Optimierung von Lehrveranstaltungen eingesetzt werden können. Unser Ziel ist die Entwicklung eines solchen Instruments. Aufgrund ihrer Relevanz für den tatsächlichen Handlungserfolg wählen wir dazu (zunächst) den Zugang über die Erfassung der Selbstwirksamkeit (SWK).

Dieser Beitrag stellt die Entwicklung und erste Erprobung einer Skala zur Erfassung der Selbstwirksamkeit zum Diagnostizieren lern- und leistungsrelevanter Merkmale im Schulalltag vor. Ausgehend von der Multidimensionalen Skala der Lehrer-Selbstwirksamkeitserwartung von Schulte und Kollegen (2008, 2011), wurde die darin enthaltene Subskala diagnostische Kompetenz ausdifferenziert. Die Entwicklung der Skala und damit auch die Formulierung sowie Skalierung der Items erfolgte unter Verwendung der Checkliste von Bandura zur Entwicklung von Selbstwirksamkeitsskalen (2006). Die 21 Items umfassen die Einschätzung des Ausprägungsgrads von Schülermerkmalen (kognitiv, motivational, emotional, selbstregulativ und sozial) sowie Kontextmerkmalen (Elternhaus, Klassenklima, Schüler-Schüler-Interaktion, Aufgabenanforderungen). Erfasst wird, inwieweit es sich die Befragten (z. B. angehende Lehrkräfte, Referendare, Lehrkräfte) zutrauen, den Ausprägungsgrad zutreffend zu beurteilen.

Die Skala wurde im Rahmen von pädagogisch-psychologisch orientierten Seminaren im bildungswissenschaftlichen Studium erprobt. Dazu wurde sie als Online-Befragung zusammen mit anderen Skalen vorgelegt. Im Wintersemester 2013 (E1 = Erprobung 1) füllten 340 Lehramtsstudierende (70% weiblich; M[Alter] = 21, M[Semester] = 2) den Fragenbogen komplett aus, im Sommersemester 2014 (E2 = Erprobung 2) waren es 366 Lehramtsstudierende (69% weiblich; M[Alter] = 21, M[Semester] = 2). Zu den für die Erprobung der Skala relevanten Skalen gehörten die Subskala Leistungsbeurteilung (LB; 6 Items; $\alpha = .85$; Bsp-Item: „Ich kenne die Funktionen von Leistungsbeurteilungen und kann sie im Unterricht berücksichtigen“) und die Subskala Diagnostizieren von Lernvoraussetzungen (DL; 6 Items, $\alpha = .78$, Bsp-Item: „Ich kann hochbegabte Schüler identifizieren und kenne Möglichkeiten der Förderung“) der Multidimensionalen Skala der Lehrer-Selbstwirksamkeitserwartung von Schulte und Kollegen (2008, 2011) sowie eine Skala zur Erfassung der Selbstwirksamkeit im Studium (E1: adaptiert nach Schyns & Collani, 2002, 8 Items, $\alpha = .85$; Bsp-Item: „Schwierigkeiten im Studium sehe ich gelassen entgegen, weil ich mich immer auf meine Fähigkeiten verlassen kann“; E2: Schiefele, Moschner & Husstegge, 2002, 10 Items, $\alpha = .82$, Bsp-Item: „Auch wenn ein Kurs sehr schwierig ist, weiß ich, was ich tun muss, um ihn zu bestehen“).

In beiden Semestern zeigten sich gute Reliabilitäten ($\alpha = .95$; Range der Trennschärfen zwischen .48 und .73) sowie eine eindimensionale Faktorenstruktur der neu entwickelten Skala. Die Skala wies zu den Skalen zur Selbstwirksamkeit im Studium mittlere Korrelationen auf (E1: $r = .30$, $p < .01$; E2: $r = .32$, $p < .01$). Die Korrelationen mit den Subskalen der Lehrer-SWK fielen ebenfalls in den mittleren Bereich (LB: $r = .37$, $p < .01$; DL: $r = .40$, $p < .01$). Jedoch wiesen die Subskalen der Lehrer-SWK, anders als die neu entwickelte Skala, niedrige Korrelationen mit den Skalen zur Selbstwirksamkeit im Studium auf.

Aufgrund dieser nicht zufriedenstellenden Befunde zur Validität wird eine gründlich überarbeitete Version der Skala derzeit erneut erprobt. Die Ergebnisse der erneuten Erprobung sollen zur Diskussion gestellt werden. Zum Zeitpunkt der Konferenz wird diese Skala zur Evaluation einer Lehrveranstaltung (Prä-Posttest-Design) eingesetzt worden sein, so dass dann ebenfalls Ergebnisse zum Einsatz in diesem Kontext vorliegen. Neben den Ergebnissen der erneuten Erprobung und der Evaluation sollen Probleme dieses Selbsteinschätzungsinstruments sowie die Weiterentwicklung zur Erfassung der Selbstwirksamkeit zum Fördern im Schulalltag diskutiert werden.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Huber, Stephan; Degen, Sandra; Schwander, Marius; Wolfgramm, Christine

Die Entwicklungen von Selbstkonzepten und Selbstwirksamkeitserwartungen in Abhängigkeit von demographischen und motivationalen Aspekten sowie verschiedener Lehramtsstudien

Infolge der Bologna Reform wurde die Ausbildung für Lehrpersonen auch in der Schweiz an den Pädagogischen Hochschulen in einer modularisierten und standardbasierten Form neu gestaltet. Mit der Formulierung von Professionsstandards geht die Erwartung einher, dass Studierende während ihrer Ausbildung professionelle Kompetenzen sukzessiv entwickeln und diese als Output auch überprüft werden können. Allerdings beschreiben die Professionsstandards nicht, wie Studierende den Standards zugrunde liegende Kompetenzen entwickeln können. Auch für die Ausbildungsinstitutionen in der Schweiz besteht deshalb ein Bedarf an Klärung darüber, welche Lerngelegenheiten, also welche strukturellen Rahmenbedingungen, welche Inhalte und welche methodischen Umsetzung, den Aufbau professioneller Kompetenzen während der Ausbildung optimal unterstützen bzw. wie diese Einflussfaktoren wechselwirkend zusammenhängen. In Anlehnung an das sich in der Lehrerbildungsforschung etablierte Modell der professionellen Handlungskompetenz von Lehrpersonen (Baumert & Kunter, 2006; 2011) werden im Projekt "PaLea - Panel zum Lehramtsstudium" (Bauer & al., 2010) insbesondere Entwicklungen in den vier Kompetenzbereichen Professionswissen, Überzeugungen und Werthaltungen, motivationale Orientierungen sowie überfachliche Kompetenzen durch wiederholte standardisierte Studierendenbefragungen in den Blick genommen.

Im Bereich der motivationalen Orientierungen spielen selbstbezogene Kognitionen wie persönliche Einschätzungen der eigenen Handlungsfähigkeit bei der beruflichen Zielsetzung und deren Umsetzung wie auch bei der Überwachung und Regulation des beruflichen Handelns eine wichtige Rolle (Schulte & al., 2008; Tschannen-Moran, Woolfolk Hoy & Hoy, 1998).

Der schweizerische Beitrag zum Symposium stützt sich auf die von 2011 bis 2014 - parallel zu PaLea in Deutschland – in der Schweiz erhobenen Daten an der PH Luzern, der PH St. Gallen und der PH Zug. Mit den auf den Schweizer Kontext adaptierten Instrumenten von PaLea wurde bei einer Stichprobe von insgesamt 742 Lehramtsstudierenden Daten erhoben. Im Beitrag werden erste längsschnittliche Ergebnisse aktuell laufender Auswertungen zu der Fragestellung präsentiert, wie sich demographische Aspekte wie z.B. Geschlecht, Beruf der Eltern, sozio-ökonomischer Status und motivationalen Aspekte wie z.B. Vorfreude auf das Studium und verschiedene Studieninteressen auf die die Entwicklungen von Selbstkonzepten und Selbstwirksamkeitserwartungen auswirken. Ergebnisse werden auch differenziert bezüglich verschiedener Gruppen von Lehramtsstudierenden (Fächerkombination, Erst-/Zweitstudium & Studiengang).

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Weiland, Laura; Krause, Ulrike-Marie

Kooperative Unterrichtsreflexion im Lehramtsstudium – eine Interventionsstudie

Ziel der Studie ist es, kognitive, einstellungsbezogene, motivationale und affektive Effekte einer Lehrinheit zur kooperativen Unterrichtsreflexion im Lehramtsstudium zu untersuchen. Im Fokus der Untersuchung stehen Reflexionsgespräche zwischen Lehramtsstudierenden nach selbst durchgeführten Unterrichtsstunden. Kooperationsfähigkeit und Reflexivität gelten als wichtige Dimensionen professioneller Kompetenz und können als zentrale Ziele der Lehrer/innenbildung angesehen werden (Oser & Oelkers, 2001; Rahm & Lunkenbein, 2014). Eine Möglichkeit zur Förderung von Kooperationsfähigkeit und Reflexivität können kooperative Unterrichtsreflexionsgespräche sein. Unter kooperativer Unterrichtsreflexion wird hier eine ko-konstruktiv ausgerichtete, kriterienorientierte Unterrichtsanalyse mit einem wechselseitigen Feedback verstanden, die auch eine ressourcenorientierte Standortreflexion des professionellen Selbst ermöglichen soll. Theoretische Grundlage bilden u.a. Konzepte wie reflection on action (vgl. Schön, 1983) und kollegiales Unterrichtscoaching (Kreis & Staub, 2009) sowie Arbeiten zu Feedback und kooperativem Lernen (z.B. Krause, 2007). In der vorliegenden Studie erwerben Lehramtsstudierende innerhalb regulärer Seminare zunächst Wissen für die Planung, Durchführung und Analyse von Unterricht (u. a. Kriterien zur Einschätzung von Unterrichtsqualität). Außerdem erhalten die Studierenden ein Skript (vgl. Fischer, Kollar, Stegmann & Wecker, 2013), das den Inhalt und Ablauf der kooperativen Unterrichtsreflexionsgespräche vorgibt. Anschließend werden Unterrichtsversuche und Reflexionsgespräche durchgeführt. Im Rahmen einer quasi-experimentellen Studie mit Prä-Post-Design und Kontrollgruppe werden kognitive, einstellungsbezogene, motivationale und affektive Effekte der Intervention untersucht. Es werden Tests, Fragebögen und Audiomitschnitte der Reflexionsgespräche eingesetzt, zudem werden Interviews durchgeführt. Die Seminarkonzeption und die Instrumente wurden bereits im Rahmen einer Pilotstudie erprobt. Die Stichprobe der Hauptstudie soll aus ca. 120 Lehramtsstudierenden bestehen. Zum Zeitpunkt der Konferenz können erste Ergebnisse der Hauptstudie präsentiert werden.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Bittner, Andrea; Urhahne, Detlef

Evaluation eines Trainings zum Erwerb von Schlüsselkompetenzen für Studierende

Zu den Schlüsselqualifikationen für den beruflichen Erfolg zählt heutzutage der Erwerb interkultureller Kompetenzen (Dreyer & Hößler, 2011). Thomas (2003, S. 143) definiert das hypothetische Konstrukt wie folgt: „Interkulturelle Kompetenz zeigt sich in der Fähigkeit kulturelle Bedingungen und Einflussfaktoren im Wahrnehmen, Urteilen, Empfinden und Handeln bei sich selbst und bei anderen Personen zu erfassen, zu respektieren, zu würdigen und produktiv zu nutzen [...]“. In der empirischen Bildungsforschung hat die Erforschung interkultureller Kompetenzen jedoch noch wenig Beachtung gefunden. Insbesondere ist nicht geklärt, ob interkulturelle Kompetenzen auf indirektem Wege vermittelt werden können oder ob nur direkte Erfahrungen mit dem Kulturraum zu einem veränderten Bewusstsein beitragen. Die bisherigen Trainingsprogramme zeigen, dass auf kognitiver Ebene Veränderungen erreicht werden können (Kinast, 2005). Dabei scheinen kulturspezifische Inhalte besser vermittelbar zu sein als kulturallgemeine Inhalte (Kölbl & Kreuzer, 2014). Die Erfassung der Evaluationsergebnisse beruht jedoch häufig auf Selbstberichtsdaten und nicht auf objektiven, interkulturellen Tests.

Mit der Evaluation eines Trainings zum Erwerb interkultureller Kompetenzen für den Raum China soll auf der Grundlage eines objektiven Tests überprüft werden, ob sich ein spezielles Trainingsprogramm als wirksam erweist. Es wird erwartet, dass eine spezifisch auf den Kulturraum China trainierte Experimentalgruppe am besten im Wissenstest abschneidet. Eine unspezifisch interkulturell trainierte erste Kontrollgruppe sollte einen höheren Kompetenzerwerb aufweisen als eine zweite Kontrollgruppe ohne interkulturelles Training. Die Resultate sollen auch bei Einbeziehung von Kontrollvariablen stabil bleiben.

In die Studie werden drei Gruppen von Studierenden mit einer Größe von jeweils zwanzig Personen einbezogen. Die Rekrutierung der Gruppen findet in Seminaren des Zentrums für Schlüsselkompetenzen an der Universität Passau statt. Neben Alter und Geschlecht werden Abiturnote, Fachsemester, Studienfach sowie Anzahl und Dauer der Besuche in China als Kontrollvariablen erhoben.

Das Untersuchungsdesign sieht eine Experimentalgruppe und zwei Kontrollgruppen vor, die jeweils vor und nach dem Seminarbesuch einen Fragebogen beantworten. Die Experimentalgruppe besucht ein 16-stündiges Seminar zum Erwerb interkultureller Kompetenzen für den Raum China. Die erste Kontrollgruppe besucht ebenfalls ein 16-stündiges Seminar zum Erwerb interkultureller Kompetenzen für einen anderen Kulturraum. Die zweite Kontrollgruppe besucht ein 16-stündiges Seminar zum Erwerb von Präsentationskompetenzen.

Der Fragebogen zur Messung interkultureller Kompetenzen für den Raum China entstammt einem Trainingsprogramm für Manager, Fach- und Führungskräfte von Thomas und Schenk (2001). In verschiedenen Problemsituationen sollen die richtigen Verhaltensweisen über eine siebenstufige Bewertungsskala angegeben werden. Der Fragebogen enthält fünf Situationen mit jeweils vier Items.

Die Daten werden im Wintersemester 2014/15 erhoben und zu Tagungsbeginn ausgewertet und dokumentiert sein. Die Mittelwerte interkultureller Kompetenzen der drei Gruppen werden mittels Varianz- und Kovarianzanalyse miteinander verglichen. Die Ergebnisse liefern Hinweise darauf, ob in Seminaren zum Erwerb von Schlüsselkompetenzen tatsächliche Kompetenzgewinne erzielt und interkulturelle Kompetenzen auch auf indirektem Wege erworben werden können.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

3

Koch, Helvi; Henke, Thorsten; Spörer, Nadine

Auf die eigene Erfahrung kommt es an. Zur Vermittlung eines schulischen Leseförderprogramms im Lehramtsstudium

Derzeit finden evidenzbasierte Programme im Allgemeinen noch zu selten Anwendung im Regelunterricht (van Ackeren et al., 2011). Hasselhorn, Köller, Maaz und Zimmer (2014) bezeichnen die Lücke, die zwischen empirisch evidenter Forschung, welche theoriebasierte Programme in regelunterrichtlichen Settings erprobt, als Mangel des Theorie-Praxis-Transfers. Evidenzbasierte Lesekompetenzförderprogramme im Speziellen werden noch zu selten in den Unterricht implementiert, unter der Bedingung, dass Regellehrkräfte als Strategieinstruktoren fungieren (Souvignier, 2009; Spörer & Schünemann, 2014).

Die wenigen pädagogisch-psychologischen Studien hierzu zeigen, dass Lehrkräfte zu wenig Unterrichtszeit aufwenden, um Lesestrategien zu instruieren (Artelt et al., 2007; Ness; 2008). Sailors (2008) kam zu dem Befund, dass die meisten Lehrkräfte das Verstehende Lesen von Texten zudem so anleiteten, wie sie es selbst in ihrer Schulzeit erlernt hatten und dass ihnen didaktisches bzw. pädagogisch-psychologisches Strategieinstruktionswissen oftmals fehlte.

Bislang wendet sich die Implementationsforschung zumeist an Regellehrkräfte, die evidenzbasierte Programme im Unterricht anwenden (Koch & Spörer, in Vorbereitung; Souvignier & Mokhlesgerami, 2005). Um evidenzbasierte Fördermaßnahmen jedoch nicht nur punktuell, sondern breitgefächert in der Praxis zu verankern, ist es notwendig, Lehramtsstudierenden ein entsprechendes evidenzbasiertes fachdidaktisches Wissen zu vermitteln. So legen z. B. die Ergebnisse der COACTIV-Studie (Baumert et al., 2011) nahe, dass sich sowohl fachliche als auch fachdidaktische Kompetenzen von Lehrenden positiv auf den Leistungszuwachs von Schülern auswirken (vgl. auch Behrmann & Souvignier, 2013). Darüber hinaus ist zu beachten, dass Studierende eigene positive Erfahrungen mit der Umsetzung und Wirksamkeit einer Fördermaßnahmen machen. Hierdurch kann die allgemeine Akzeptanz evidenzbasierter Programme erhöht und die Zufriedenheit mit einem speziellen Strategieinstruktionsprogramm gesteigert werden (Kline, Desher & Shumaker, 1992).

Ein evidenzbasiertes Programm zur Förderung der Lesekompetenz von Schülern ist das Reziproke Lehren (Palincsar & Brown, 1984). Dieses wurde von Spörer und Kollegen (Schünemann, Spörer & Brunstein, 2013; Spörer & Schünemann, 2014) mit Prozeduren des selbstregulierten Lernens kombiniert (RT+SRL). In diesen Studien konnte gezeigt werden, dass sich die Lesekompetenz von Schülern durch die Anwendung von RT+SRL nachhaltig steigern lässt.

Da sich die Methode des Reziproken Lehrens auch in universitären Settings als effektive Fördermethode erwiesen hat (Freihat & Makhzoomi, 2012), soll vermittelt über die eigenen positiven Erfahrungen der Studierenden mit dem Programm (Guskey, 2002) die Offenheit und Bereitschaft, es im späteren Unterricht selbst einzusetzen, angebahnt werden. Dazu wurde in einem ersten Schritt in der vorliegenden Studie untersucht, ob Studierende hinsichtlich ihrer eigenen Lesekompetenz von der Fördermaßnahme RT+SRL profitieren. Das Programm wurde für diese Zwecke so adaptiert, dass die Inhalte der Lehrtexte, welche in reziproker Gruppenarbeit gelesen wurden, auf spezifische Seminarinhalte ausgerichtet waren. Es wurde angenommen, dass sich die Studierenden, die mit dem Programm RT+SRL arbeiteten, in Effektmäßen der Lesekompetenz deutlich verbessern.

Dazu wurde in einem Experimental-Kontrollgruppendesign mit Pre- und Postmessung die Leseentwicklung von N = 56 Lehramtsstudierenden untersucht. Zur Erfassung der Lesekompetenz wurden standardisierte Verfahren (LGVT 6–12 und WLST 5–7) eingesetzt. Der Erfolg der Interventionsmaßnahme sollte sich in Vorteilseffekten der Treatmentgruppe gegenüber der Kontrollgruppe zeigen. Die Kontrollgruppe nahm an einem Seminar desselben Lehrmoduls innerhalb des Lehramts-Masterstudiums teil. Es wurden die Kriteriumsmaße Leseverständnis, Lesegeschwindigkeit und Lesestrategiewissen erfasst. Kovarianzanalysen der summativen Daten ergaben signifikante Vorteilseffekte für die Treatmentgruppe ($p \leq .05$)."

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Müsche, Hanna; Klingsieck, Katrin; Scharlau, Ingrid; Praetorius, Saskia

Problembasierter versus instruktionsbasierter Erwerb pädagogisch-psychologischen Wissens in den Bildungswissenschaften

Im deutschen Sprachraum ist die Befundlage zur Effektivität problemorientierten (z. B. Gräsel, 1997) oder problembasierten Lernens (z. B. Zumbach, 2003) auf der Basis von Einzelstudien bislang sehr uneinheitlich, insbesondere in Bezug auf die kognitiven Effekte problembasierten Lernens (z. B. Fölling-Albers, Hartinger & Mörtl-Hafizovic, 2005; Krause, Stark & Herzmann, 2011). Die meisten metaanalytischen Befunde qualifizieren das problembasierte Lernen als bewährten Ansatz zur Förderung selbstgesteuerten Lernens und teils auch eines nachhaltigen, anwendungsorientierten Wissenserwerbs (z. B. Strobel & van Barneveld, 2009). In der Tradition einer pädagogischen Kasuistik und unter dem Einfluss des situierten Lernens gewinnt problembasiertes Lernen auch in der Lehrerbildung zunehmend an Popularität. So wird das problembasierte Lernen auch bei einem Teil der pädagogisch-psychologisch orientierten Seminare im bildungswissenschaftlichen Studium an der Universität Paderborn ins Zentrum der Didaktik gestellt.

Im Jahr 2013 wurden die Effekte der problembasierten (PBL) und traditionelleren, d.h. instruktions-basierten, Seminarkonzepte anhand einer Onlinebefragung der Studierenden (N = 414; Frauenanteil: 70%; Lehrämter: 68% Sekundarstufe I, 31% Sekundarstufe II; Fachsemester: M = 1.62, SD = 1.02; Alter: M = 20.98, SD = 2.62) in einem quasi-experimentellen Prä-Post-Design vergleichend überprüft. Zur direkten Veränderungsmessung wurde der selbsteingeschätzte Kompetenzzuwachs nach Paechter, Skliris und Macher (2011) erfasst, zur indirekten Veränderungsmessung wurde u. a. der Fragebogen zur Erfassung von Lernstrategien im Studium (LIST; Wild & Schiefele, 1994) herangezogen. Erstgenanntes Instrument setzt sich aus insgesamt vier Skalen mit vier bis fünf Items und Skalenreliabilitäten zwischen .75 und .89 zusammen: der Skala zur Methodenkompetenz (Beispielitem: „Ich kann mein erworbenes Wissen auf verschiedene Aufgabenstellungen anwenden.“), zu sozial-kommunikativen Kompetenzen (Beispielitem: „Ich konnte mein Wissen und meine Fähigkeiten in die Gruppe einbringen.“), zur Fachkompetenz (Beispielitem: „Ich kann einen guten Überblick über die Inhalte der LV geben.“) sowie zur personalen Kompetenz (Beispielitem: „In der LV habe ich gelernt, meinen Lernfortschritt besser zu prüfen.“). Mit Hilfe des LIST wurde die Nutzungshäufigkeit sowohl kognitiver als auch metakognitiver und ressourcenbezogener Lernstrategien im Selbstbericht erfasst. Die Reliabilitäten der LIST-Skalen lagen bei einer durchschnittlichen Itemanzahl von 7 zwischen .74 und .86. In einer zweiten Evaluationsstudie (N = 372; Frauenanteil: 67%; Lehrämter: 16% Sekundarstufe I, 64% Sekundarstufe II; Fachsemester: M = 2.15, SD = 1.04; Alter: M = 20.97, SD = 2.75) wurde zusätzlich ein Test zur Erfassung veranstaltungsbezogenen Faktenwissens und Anwendungswissens, bestehend aus jeweils fünf MC-Aufgaben, eingesetzt. Auf der Basis bisheriger Befunde wurde erwartet, dass die Studierenden, die problembasiert gearbeitet haben, mit Ausnahme des Faktenwissens vorteilhaftere Lern- und Leistungsparameter (z. B. einen vermehrten Einsatz tiefenorientierter Lernstrategien) zeigen würden.

Im Ergebnis erwiesen sich die PBL-Studierenden den anderen Studierenden im selbsteinge-schätzten Kompetenzzuwachs überlegen ($t(378) = 2.38, p = .02, dz = 0.38; M_{PBL} = 4.05, SD_{PBL} = 0.55, n_{PBL} = 48; M_{ANDERE} = 3.81, SD_{ANDERE} = 0.65, n_{ANDERE} = 332$). Differenziert nach Teilkompetenzen zeigte sich dieses Bild speziell für die Methodenkompetenz und die sozial-kommunikativen Kompetenzen, nicht jedoch für die Fach- und die Personalkompetenz. Anders als erwartet ergaben sich jedoch weder in den objektiven Testleistungen noch hinsichtlich von Selbstregulation und Häufigkeit des Lernstrategieeinsatzes überzufällige Unterschiede zwischen den Studierendengruppen.

Dass sich die Studierenden beispielsweise nicht im Zuwachs ihres Anwendungswissens unterschieden, könnte darauf hindeuten, dass der situativ gebundene Wissenserwerb im Rahmen des problembasierten Lernens die Anwendung des Wissens nicht nur fördert, sondern den Wissenstransfer potenziell auch behindert (z. B. Kaufman, Keselmann & Patel, 2009). Weiterhin werden insbesondere die erwartungswidrigen Ergebnisse und ihre Implikationen für die künftige Veranstaltungsdidaktik diskutiert. So könnte es sich u. a. empfehlen, das bestehende Konzept dahingehend auszuweiten, dass jeder Veranstaltungsinhalt durch multiple Problemsituationen bzw. Falldarstellungen repräsentiert wird. Eine weitere Möglichkeit besteht in der (anfänglichen) Reduzierung der kognitiven Belastung (z. B. Mayer & Moreno, 2003) insbesondere in der Zielgruppe der Erstsemesterstudierenden.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Bergner, Maria; Stark, Robin; Naefgen, Christoph; Krause, Ulrike-Marie

Wissenschaftliches Denken bei Lehramtsstudierenden: Ergebnisse einer Interviewstudie

Im Rahmen von Standards für die Lehrerbildung (KMK, 2004) und Modellen professioneller Kompetenz (z.B. Baumert & Kunter, 2006) wird deutlich, dass wissenschaftliches Denken eine hohe Bedeutung für die Professionalität von Lehrkräften hat und damit ein wichtiges Lernziel im Lehramtsstudium ist (Kunina-Habenicht et al., 2013). Als Teil eines Projektes mit dem Titel „Analyse und Förderung wissenschaftlichen Denkens in der Lehrerbildung“ besteht die Zielsetzung der vorgestellten Interviewstudie in der Erhebung von Stärken und Schwächen im wissenschaftlichen Denken bei Lehramtsstudierenden. Wissenschaftliches Denken wird definiert als die reflektierte Anwendung wissenschaftlicher Erkenntnisse auf Problemstellungen der pädagogischen Praxis sowie als kritische Auseinandersetzung mit Alltagstheorien. Wissenschaftliches Argumentieren basiert auf wissenschaftlichem Denken und liefert einen Ansatzpunkt für die Analyse dieser Kompetenz.

In Anlehnung an eine Interviewstudie von Kuhn (1991) erfolgte die Konzipierung eines Interviewleitfadens, der die Leistungen der Studierenden bei der Generierung und Analyse von Argumenten als Indikatoren für wissenschaftliches Denken erheben sollte. Die Erfassung der Generierungs- und Analyseleistung erfolgte im Kontext der pädagogisch-psychologischen Themen Schulangst, Burnout, Motivationsprobleme und Unterrichtsstörungen. Die Durchführung der Interviewstudie fand im Mai 2014 statt und umfasste eine Stichprobe von 18 Lehramtsstudierenden unterschiedlicher Fachrichtungen und angestrebter Lehramtsabschlüsse (Grundschule, Sonderpädagogik, Gymnasium, Haupt- und Realschule). Das Durchschnittsalter der Befragten lag bei 24 Jahren (SD = 2,06) und die mittlere Anzahl der studierten Semester bei 7,5 (SD = 3,11). Die Auswertung der Interviews erfolgte inhaltsanalytisch hinsichtlich formaler domänenübergreifender Merkmale (Kuhn, 1991) sowie hinsichtlich formaler und inhaltlicher domänenspezifischer Aspekte (Stark & Krause, 2006; Stark, Puhl & Krause, 2009). Zusammenfassend weisen die Ergebnisse darauf hin, dass die Lehramtsstudierenden bei der Generierung von Erklärungen selten auf wissenschaftliche Theorien und empirische Befunde Bezug nehmen. Häufig wird auf Alltagserklärungen, Meinungen oder intuitive Annahmen rekurriert, und es werden Alltagsbeobachtungen und -erfahrungen herangezogen. Diese Ergebnisse decken sich mit den von Stark und Krause (2006) sowie Stark, Puhl und Krause (2009) identifizierten Problemen Studierender beim wissenschaftlichen Argumentieren. Weitere interessante exemplarisch ausgewählte Ergebnisse betreffen zum einen die Probleme Studierender mit der Wissenschaftssprache bzw. mit wissenschaftlichen Fachbegriffen, was zu Schwierigkeiten bei der Analyse präsentierter wissenschaftlicher Erklärungen führt. Viele Studierende hatten Probleme bei der inhaltlich korrekten Wiedergabe der präsentierten wissenschaftlichen Erklärung, und einige konnten bei einem Vergleich von Erklärungen diese bestehenden Schwierigkeiten mit der Wissenschaftssprache direkt benennen. Zugleich gibt es Hinweise darauf, dass die Studierenden die Komplexität und Vielschichtigkeit der präsentierten Themen erkannt haben, was u.a. daran zu erkennen ist, dass zur Erklärung der präsentierten Phänomene in den meisten Fällen mehr als eine Ursache angeführt wurde.

Die in dieser Interviewstudie gewonnenen Erkenntnisse liefern wichtige Hinweise für die im weiteren Projektverlauf angestrebte Förderung wissenschaftlichen Argumentierens bzw. Denkens bei Lehramtsstudierenden.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

3

Pollerhof, Kathleen; Athanassiou, Georgios; Braun, Edith

Entwicklung eines Testsettings zur Erfassung von strategischer oder verständnisorientierter Kommunikation von Studierenden

Kommunikation wird in zahlreichen Qualifikationsrahmen als einer der zentralen Lerninhalte eines Hochschulstudiums beschrieben. Bislang liegen kaum Testverfahren vor um derartige überfachliche Fähigkeiten auf der Individualebene mit anerkannten Gütekriterien zu erfassen. Daher ist das Ziel der hier vorgestellten Forschungsgruppe die Entwicklung eines Testsettings, durch das kommunikative Fähigkeiten mithilfe eines standardisierten Beobachtungsbogens bewertet werden können.

Als Ausgangslage werden die beiden Studiengänge Wirtschaftswissenschaften und Lehramt verwendet. Zum einen handelt es sich um ausbildungsstarke Studiengänge, in denen zusammen 25% aller Studierenden eingeschrieben sind. Zum anderen müssen die Absolvierenden der Studiengänge unterschiedlich in ihren Berufen kommunizieren. Während in den Wirtschaftswissenschaften eher Verhandeln und Überzeugen als kommunikative Anforderungen genannt werden, müssen Lehrer(innen) insbesondere Konflikte lösen, erziehen und lehren.

Als theoretischer Bezugsrahmen zur Definition dieser unterschiedlichen Kommunikationsanforderungen dienen die von Habermas (1981) entwickelten Typen der Kommunikation. Habermas unterscheidet in seiner Theorie des kommunikativen Handelns zwischen strategischer und verständigungsorientierter Kommunikation. Unter der strategischen Kommunikation ist im Allgemeinen ein an partikularen Nutzenkalkülen orientierter Versuch der Beeinflussung von Meinungen, Einstellungen und Haltungen zu verstehen, während in der verständigungsorientierten Kommunikation die Herstellung eines rational motivierten Einverständnisses auf Basis gemeinsamer Überzeugungen im Vordergrund steht. Diese in der Soziologie verbreitete Theorie wird mit den Methoden der psychometrischen Testentwicklung verbunden.

Um das Testverfahren entwickeln zu können, müssen zwei Forschungsfragen vorab bearbeitet werden: (1) Welche konzeptuellen Komponenten kommunikativer Fähigkeiten können identifiziert werden, die zur beobachtungs-basierten Ermittlung und Bewertung von strategischer oder verständnisorientierter Kommunikationsfähigkeiten verwendet werden können. (2) Welche spezifischen Situationen sind typisch für die Kommunikation im Kontext der Berufstätigkeit von Hochschulabsolvierenden der genannten Studienrichtungen (Lehramt und Wirtschaftswissenschaften). Wie können solche Situationen simuliert werden, um ein standardisiertes Setting für die beobachtungs-basierte Ermittlung der Ausprägung der modellrelevanten kommunikativen Fähigkeiten bereitzustellen.

Ausgehend vom Ziel, sowohl strategische als auch verständigungsorientierte Kommunikationskomponenten zu erfassen, wird ein eigenes Modell der kommunikativen Fähigkeiten entwickelt. Für diesen Zweck wurden auf Basis einer umfänglichen Literaturrecherche einzelne Komponenten kommunikativer Fähigkeit extrahiert (Gartmeier et al. 2011, Grice 1975, Hinsch und Pflingsten 2007, Rogers 1981; Schulz von Thun 2000, Lasswell-Formel nach Traut-Mattausch und Frey (2005) Hargie und Dickson 2004 nach Röhner und Schütz 2012, Watzlawick 1969).

Die identifizierten Modellkomponenten sind: Intention, Sachinhalt, Beziehung (ggf. Selbstoffenbarung), Sprache/Sprechweise, Interpunktion der Kommunikation, soziale Rolle. Auf Grund dieser Komponenten können die Kommunikationstypen charakterisiert werden, wie es im Folgenden beispielhaft für die Intention dargestellt ist. In strategischen Gesprächskontexten besteht die Intention häufig über die Gesprächssituation hinaus. Das Gesprächsziel ist schwierig zu erkennen, da eine direkte Ansprache meist nicht zu einer nutzenkalkulierten Erreichung des Ziels beiträgt. In der verständigungsorientierten Kommunikation steht eine kooperative Problemlösung im Vordergrund, indem eine gegenseitig akzeptierte Zielvereinbarung angestrebt wird. Die Beziehung und die Selbstoffenbarung spielen möglicherweise eine Art Antagonisten bei den Kommunikationstypen.

Es wird erwartet, dass eine differenzierte Ausprägung dieser Komponenten auf das Vorliegen unterschiedlicher Kommunikationstypen hinweisen und deren Ermittlung ermöglichen wird.

Zur Beantwortung der zweiten Fragestellung wurde eine Online-Befragung (Frühjahr 2014) von ca. 10.000 Absolvent(inn)en (deutschlandweite repräsentative Stichprobe des DZHW Panels) genutzt. Die Befragten wurden gebeten, typische berufliche Situationen als qualitative Angaben anzugeben, in denen sie mit anderen Personen kommunizieren müssen. Die am häufigsten genannte Situation war die „Präsentation von Arbeitsergebnissen“ (941), gefolgt von den Situationen „Austausch über Aufgaben oder Ideen“ (784), und „Unterrichten“ (726) sowie „Abstimmung mit Beteiligten“ (468).

Das Modell der Kommunikationskomponenten auf der einen Seite soll genutzt werden, um beobachtbare Merkmale zu identifizieren, die als Grundlage eines Beobachtungsbogens dienen. Die berufsspezifischen Kommunikationssituationen auf der anderen Seite sollen als thematische Grundlage für die Konzeption standardisierter Rollenspiele genutzt werden. Die herausgearbeiteten Kommunikationskomponenten werden dabei berücksichtigt, um Rahmenbedingungen und Instruktion für jeweils strategische und verständnisorientierte Gespräche zu standardisieren.

Donnerstag
12. März

10:45 Uhr

Foyer GC - Ost

Hemker, Nele; Hertel, Silke

Effektivität von Trainings zu verschiedenen Facetten der Selbstregulationskompetenz in der ersten Phase der Lehrerbildung

Lehrkräfte stehen im Schulalltag komplexen Anforderungen gegenüber, auf die sie schnell, effizient und kompetent reagieren müssen. Folglich wird Selbstregulation als essenzieller Bestandteil der professionellen Lehrerkompetenz betrachtet, die notwendig ist, um die Aufgaben des Lehrerberufes erfolgreich zu bewältigen und zugleich das eigene Wohlbefinden aufrechterhalten zu können (Klusmann, 2011; KMK, 2004; Kunter et al., 2011).

Selbstregulation kann auf emotionaler, kognitionaler, motivationaler sowie auf Verhaltensebene erfolgen (vgl. Zeidner, Boekaerts, & Pintrich, 2000). Bisher gibt es nur wenig Forschung zur Vermittlung von Selbstregulationskompetenzen in der Lehreraus- und Weiterbildung. Beispielsweise hat Mattern (2012) ein Training zur Handlungs- und Motivationsregulation während der Unterrichtsvorbereitung mit dem Fokus auf die Überwachung und Steuerung zielgerichteten Verhaltens entwickelt und positive Effekte auf volitionale Kontrollstrategien und Umgang mit Belastungen gefunden.

Inwiefern Selbstregulationskompetenzen auf allen vier Ebenen, d.h. auf emotionaler, kognitionaler, motivationaler und Verhaltensebene, bereits in der ersten Phase der Lehrerbildung erworben werden können, wurde bisher kaum untersucht. Ziel dieser Studie ist es daher, die Wirkung zweier Lehrveranstaltungen zu selbstregulatorischen Fähigkeiten differenziert zu analysieren.

Im Rahmen von zwei wöchentlichen Veranstaltungen mit 14 Seminareinheiten á 90 Minuten wurde zu den Themen „Lehrergesundheit“ (N=16) und „Selbstreguliertes Lernen“ (N=17) Wissen vermittelt, reflektiert und angewandt. Während das Seminar zur Lehrergesundheit breit ausgelegt verschiedene Elemente der Selbstregulation thematisierte, lag der Schwerpunkt des Seminars zu selbstreguliertem Lernen ganz spezifisch auf selbstregulatorischen Kompetenzen, die für effektives Lernen notwendig sind. Mit einem Fragebogen wurden zu Beginn und am Ende der Veranstaltungsreihen die Einschätzung zu verschiedenen Facetten der Selbstregulation erfasst. Die Formulierung der Items basiert auf Kuhl und Fuhrmann (1997), Kaluza (2004), sowie Schmitz und Salisch (2002). Die Studierenden wählten sich in die Lehrveranstaltungen selbst ein. Zur Auswertung der Daten wurde eine MANOVA mit Messwiederholung berechnet (PASW 22). Als unabhängige Variable wurde der Faktor Treatment eingeführt. Es lagen keine signifikanten Prätest-Unterschiede zwischen den Gruppen vor ($p > .25$).

Auf multivariater Ebene zeigten sich eine signifikante Interaktion von Zeit x Treatment (Wilks-Lambda, $F(7/25)=3.817$, $p=.006$, $\eta^2=.517$).

Die univariaten Kontraste ergaben für die Aspekte Zielsetzung ($F(1/31)=1.283$, $p=.266$), Planung ($F(1/31)=.277$, $p=.602$), Selbstmotivation ($F(1/31)=2.388$, $p=.132$), Emotionsausdruck ($F(1/31)=1.51$, $p=.419$), sowie Selbstberuhigung ($F(1/31)=.171$, $p=.602$) keine signifikanten Zeit x Treatment Interaktionen. Für die Aspekte Prokrastination ($F(1/31)=6.221$, $p=.018$, $\eta^2=.167$) und Denkmuster ($F(1/31)=10.368$, $p=.003$, $\eta^2=.251$) ergeben sich signifikante Interaktionen von Zeit x Treatment.

Für Studierende des Seminars Lehrergesundheit war die Veränderung in den Bereichen Prokrastination und Denkmuster positiver als für Studierende des Seminars zu selbst-reguliertem Lernen: Nach eigener Einschätzung planen sie Dinge nicht nur sondern gehen sie direkt an und überwinden dabei potenzielle Hindernisse. In unangenehmen Situationen bleiben sie mit ihren Gedanken eher bei sich und grübeln nicht darüber nach, was andere Personen in ihrer Situation denken würden.

Für die Aspekte Selbstmotivierung ($F(1/31)=4.402$, $p=.044$, $\eta^2=.124$), Selbstberuhigung ($F(1/31)=4.205$, $p=.049$, $\eta^2=.119$) und Planung ($F(1/31)=9.506$, $p=.004$, $\eta^2=.235$) ergaben sich signifikante Effekte für den Faktor Zeit. In diesen Bereichen haben Studierende in beiden Seminaren an Selbstregulationskompetenzen gewonnen.

Insgesamt weisen die Befunde darauf hin, dass sich wichtige Aspekte der Selbstregulation für den beruflichen Kontext bereits im Rahmen der ersten Phase der Lehrerbildung vermittelt werden können. Sie unterstützen somit die Ergebnisse von Mattern (2012).

Es wurde sehr deutlich, dass es lohnenswert ist, unterschiedliche Aspekte und Facetten des Konstrukts „Selbstregulation“ differenziert zu betrachten – sowohl aus theoretischer Perspektive als auch vor dem Hintergrund der Weiterentwicklung der Lehreraus- und -weiterbildung.

Einschränkend ist anzumerken, dass die Stichprobengröße sehr klein ist und keine randomisierte Zuweisung von Studierenden zu Treatmentgruppen erfolgen konnte. Zudem basieren die Daten auf selbstberichteten Angaben der Studierenden zu ihren Selbstregulationsstrategien. In Folgestudien ist geplant, Zusammenhänge verschiedener Facetten der Selbstregulation von angehenden Lehrkräften zu Leistungs- und Gesundheitsindikatoren zu untersuchen, sowie objektivere Maße zum Vergleich von subjektiven Maßen heranzuziehen.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Hüfner, Christiane; Wilde, Matthias

subjektive Vorstellungen (angehender) Biologielehrer

Mit der „kognitiven Wende“ (Mischke & Kiper 2008) und dem einhergehenden Übergang von behavioristischen zu kognitionspsychologischen Sichtweisen auf das Lernen, rückt der Lernende als aktiver Teilnehmer am Lernprozess (sgL) zunehmend in das Blickfeld der Lernforschung. Hierbei ist es vor allem der didaktische Diskurs um das sgL, der diese Lernform akzentuiert (Kiper & Mischke 2008). Nach Schiefele und Pekrun (1996), Zimmerman (2002) und Boekaerts und Corno (2005) entscheidet der Lernende dabei auch über den Einsatz metakognitiver Lernstrategien innerhalb seines Lernprozesses. Diesen metakognitiven Lernstrategien weisen Schraw, Crippen und Hartley (2006) eine zentrale Rolle innerhalb naturwissenschaftlicher Bildung zu. Teil der naturwissenschaftlichen Bildung ist das hypothesengeleitete Experimentieren im Biologieunterricht (KMK 2004). Für ein zielgerichtetes Ausführen von Experimenten benötigen Schülerinnen und Schüler metakognitives Wissen und sollten in der Lage sein, dieses auch anwenden zu können. In welcher Form sgL und damit auch metakognitive Lernstrategien im Unterricht eingesetzt werden, hängt von den individuellen Vorstellungen der jeweiligen Lehrperson ab. Demnach besteht nach Hartinger et al. (2006) und Dubberke et al. (2008) ein Zusammenhang zwischen diesen individuellen Vorstellungen zum Lernen und der unterrichtlichen Gestaltungen durch die Lehrperson.

Da die individuellen Vorstellungen zum sgL als Teil des Professionswissens (Shulman 1987) oder als beliefs über den Lerner und das Lernen (Calderhead 1996) gesehen werden können, werden diese beiden Begriffe in Anlehnung an Magnusson et al. (1999) und Hartinger et al. (2006) hier unter „subjektive Vorstellungen“ zusammengefasst.

Nach Shulman (1986, 1987) entwickelt sich das Professionswissen von (angehenden) Biologielehrern mit ihrer Aus- und Weiterbildung bzw. ihrer Berufserfahrung. Sind sie in ihren Anfängen des Lehramtsstudiums noch Novizen des Lehrens, können sie ihr biologisches Fachwissen und biologiepädagogisches Wissen weiterentwickeln und so zu Experten des Lehrens werden. Dieser Studie liegt die Annahme zugrunde, dass sich die Entwicklung des Professionswissens zum sgL mit Hilfe der Aussagen von angehenden Lehrpersonen und praktizierender Lehrpersonen rekonstruieren lässt. Die Forschungsfrage lautet:

Unterscheiden sich die subjektiven Vorstellungen zum sgL im Biologieunterricht von Studierenden und Biologielehrern gemäß der postulierten Entwicklung von fachdidaktischem Wissen während der Ausbildung zum Biologielehrer?

Mit Hilfe eines halbstandardisierten, leitfadengestützten Interviews (Flick 2011) wurden die subjektiven Vorstellungen von 12 Biologiestudierenden (Lehramt), 4 Referendaren und 13 Biologielehrenden erhoben. Nach Lamneck (2005) kann es als ein Interview beschrieben werden, durch das die komplexen Wissensbestände der Lehrer zum sgL sichtbar gemacht werden. Als Grundlage dient das Leitfadenterview nach Scheele und Groeben (1988). Das durchgeführte Interview gliedert sich in zwei Abschnitte. Im ersten Teil werden die Interviewpartner dazu angehalten, über ihre eigenen Erinnerungen, Vorstellungen und Ideen zum Biologieunterricht, mit Bezug zum sgL, zu sprechen. Die zweite Phase konfrontiert die Befragten mit den theoretischen Konstrukten des sgL nach Schiefele und Pekrun (1996). Es erfolgte eine qualitative Inhaltsanalyse, in Form einer inhaltlichen Strukturierung nach Mayring (2003) und Schilling (2006). Die quantitative Analyse der Ergebnisse erfolgte mittels deskriptiver Statistik. Die Übereinstimmung zwischen Erst- und Zweitkodierer kann als sehr gut eingeschätzt werden (Cohens Kappa = .83).

Studierende wie Lehrende besetzten alle Oberkategorien zum sgL (Probanden pro Kategorie; Studierenden/Lehrperson; Lernstrategien: 100%/100%, Volition: 50%/47%, Motivation: 83%/82%, externe Lernsteuerung: 100%/100%, Lernprodukte: 41,6%/64%, Schülvoraussetzungen: 83%/100%). Erste Analysen ergaben, dass in der übergeordneten Kategorie „sgL im Biologieunterricht“ verschiedene Sichtweisen auf den Lernprozess eingenommen werden. Zum einen wird sgL unter der kognitionspsychologischen Perspektive betrachtet, wobei die Probanden unter anderem von einer tieferen und nachhaltigeren Verarbeitung der Lerninhalte durch sgL sprechen. Zum anderen ist sgL eine Methode, die im Biologieunterricht eingesetzt werden kann. Die für den Biologieunterricht wohl spezifische Sicht auf das sgL ergibt sich aus den Annahmen, dass das sgL mithilfe des „induktiv-hypothetischen“ Wegs der Erkenntnisgewinnung umgesetzt wird. Letztere Perspektive wird vor allem von den Lehrpersonen eingenommen, wohingegen sgL von 28 der 29 Probanden auch als Methode zur Strukturierung des Biologieunterrichts wahrgenommen wird.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Roth, Anne; Kellermann, Anne; Schmitz, Bernhard

„Thinking about thinking“ – die deutschsprachige Adaption des Metacognitive Awareness Inventory

Lernende, die über umfangreiches metakognitives Wissen verfügen und metakognitive Strategien in hohem Maße einsetzen, sind besser in der Lage, das Lernen zu planen, Informationen zu organisieren, sich selbst beim Lernen zu beobachten, Probleme zu lösen und ihr Lernen zu evaluieren (Schraw & Dennison, 1994). In diesem Sinne fanden beispielsweise Young und Fry (2008) signifikante Zusammenhänge zwischen Metakognition und den Leistungen am Ende eines Kurses sowie signifikante Unterschiede hinsichtlich des metakognitiven Strategieeinsatzes zwischen Studierenden höherer und niedrigerer Semester. Trotz der Relevanz von Metakognition für akademische Leistungen liegt derzeit noch kein deutschsprachiges Instrument zur Erfassung dieses Konstruktes bei Erwachsenen vor. Ein etabliertes und validiertes Instrument zur Erfassung metakognitiven Wissens und metakognitiver Regulation im angloamerikanischen Raum ist das „Metacognitive Awareness Inventory“ (MAI, Schraw & Dennison, 1994). Dieses Inventar wurde bereits vielfach genutzt (z.B. Kleitman & Stankov, 2007; Teo & Lee, 2012), übersetzt (Akin, Abaci & Cetin, 2007) und als Junior- Version für Kinder adaptiert (Sperling, Howard, Miller & Murphy, 2002). Bedauerlicherweise liegt keine validierte deutschsprachige Version des MAI vor.

Das MAI orientiert sich an der Definition der Metakognition von Flavell (1976) und bildet die beiden Komponenten metakognitives Wissen und metakognitive Regulation sowie insgesamt acht Subkomponenten mit 52 Items ab.

Ziel der vorliegenden Arbeit ist die Entwicklung einer deutschsprachigen Adaption des MAI. Im Rahmen dessen wird überprüft, ob das adaptierte Inventar eine ausreichende Reliabilität und die Items eine angemessene statistische Güte aufweisen. Im Hinblick auf die Konstruktvalidierung wird geprüft, ob sich Studierende in verschiedenen Studienabschnitten hinsichtlich ihrer Selbstauskunft im MAI unterscheiden.

Die Übersetzung des Fragebogens wurde durch zwei Bilingualisten mit vertieften pädagogisch-psychologischen Kenntnissen im Rahmen eines Backtranslation-Verfahrens durchgeführt. Auftretende Unklarheiten wurden in einem Expertengremium mit den Autoren dieses Beitrages diskutiert und gelöst. Anschließend wurde eine Online-Befragung mit 13 Bilingualisten, welche über exzellente Deutsch- und Englischkenntnisse verfügten, durchgeführt, in der die Angemessenheit der Übersetzung beurteilt werden sollte. Eine als inadäquat eingeschätzte Übersetzung wurde durch das Expertengremium erneut überarbeitet.

An der anschließenden Online-Pilotierung des MAI nahmen insgesamt 227 Studierende (♀70%) verschiedener Universitäten und Fachbereiche mit einem mittleren Alter von $M=23,33$ Jahren ($SD=3,81$) teil. Im Rahmen eines Crossover-Designs wurden den Studierenden sowohl die englische als auch die deutsche Version des MAI randomisiert in zweiwöchigem Abstand vorgelegt. 154 Studierende füllten beide Inventar-Versionen aus, 204 Studierende ausschließlich die deutsche Version.

Die Ergebnisse der Untersuchung verdeutlichen die Vergleichbarkeit der deutschen und der englischen Inventarversion. Hinsichtlich des Gesamtscores zeigen sich Zusammenhänge von $r=.80^{**}$.

Die interne Konsistenz der Gesamtskala (Cronbach's α) liegt bei .93. Die Itemkennwerte der adaptierten Inventarversion sind ebenfalls vielversprechend: Die Schwierigkeitsindizes und Trennschärfekoeffizienten liegen überwiegend im gewünschten Bereich. Im Rahmen der Konstruktvalidierung zeigen sich signifikante Unterschiede hinsichtlich des metakognitiven Strategieeinsatzes bei Studierenden höherer und niedrigerer Semester, $t(200)=2.36$, $p<.05$.

Abschließend werden Hinweise auf mögliche weitere Adaptionsschritte sowie eine konvergente und diskriminante Kriteriumsvalidierung gegeben.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Lohbeck, Annette; Mühlig, Andrea; Rohde, Julia; Moschner, Barbara

Bedingungen des Studierverhaltens: Passive akademische Prokrastination und Selbstregulation im Studium

Theoretischer Hintergrund: Das Lernverhalten im Studium hat einen starken Einfluss auf Studienleistungen und Studienerfolg (Klingsieck, Fries, Horz, & Hofer, 2012; Steel, 2007). Einige Studien zeigen dabei, dass insbesondere selbstregulatorische Fähigkeiten wie der Gebrauch von Lernstrategien zum Studienerfolg beitragen (Stumm, Thomas & Dormann, 2010). Ein häufig zu beobachtendes Phänomen unter Studierenden ist jedoch passive akademische Prokrastination. Dieses beschreibt das dysfunktionale Aufschieben von leistungsbezogenen Aufgaben und ist vor allem durch eine mangelnde Selbstregulation gekennzeichnet (Grund, Schmid, Klingsieck & Fries, 2012). Ausgehend von der Selbstbestimmungstheorie von Deci und Ryan (2000) sowie dem Erwartungs-mal-Wert-Modell von Eccles und Wigfield (2002) lässt sich Prokrastination auf aversive, external motivierte und weniger autonome Handlungen zurückführen, die neben negativen Konsequenzen wie z. B. unzureichende akademische Leistungen (Steel, 2007) die Handlungsausführung blockieren (Patzek, Grunschel, König & Fries, 2014). Trotz der weiten Verbreitung von passiver akademischer Prokrastination liegen bislang jedoch nur wenige Studien zu den Bedingungsfaktoren von passiver akademischer Prokrastination und Selbstregulation bei Studierenden im deutschen Sprachraum vor (z. B. Rustemeyer & Rausch, 2007). Um Anhaltspunkte für Interventionen zur Verringerung von passiver akademischer Prokrastination und zur Steigerung von Selbstregulation bei Studierenden zu gewinnen, sind deshalb solche Studien von hoher Relevanz.

Fragestellung: Ausgehend von der bisherigen Forschungslage geht die vorliegende Studie folgender Fragestellung nach: Welche individuellen und studienrelevanten Umweltmerkmale bedingen passive akademische Prokrastination und Selbstregulation im Studium?

Methode: Die Daten basieren auf einer Studierendenbefragung, die im Sommersemester 2014 an verschiedenen deutschen Hochschulen in mehreren Lehrveranstaltungen durchgeführt wurde. Befragt wurden 908 Studierende der Erziehungswissenschaften und eines Lehramtsstudiums (Männer: n = 245, Frauen: n = 663; Alter: M = 22.62, SD = 3.98) zu verschiedenen individuellen Merkmalen (z. B. Selbstwirksamkeit, Studienzufriedenheit) und studienrelevanten Umweltmerkmalen (z. B. Leistungsdruck, Belastungen). Zur Überprüfung der Vorhersagekraft dieser Merkmale wurden hierarchische Regressionsanalysen jeweils für die passive akademische Prokrastination und Selbstregulation als Kriterium herangezogen. Als Kontrollvariablen dienten die Abitur- und Semesterdurchschnittsnoten, Geschlecht und Alter der Studierenden.

Ergebnisse: Die Regressionsanalysen weisen darauf hin, dass passive akademische Prokrastination durch Anstrengungsbereitschaft, Selbstwirksamkeit, Selbstregulation, Arbeitsvermeidung, Neurotizismus, Extraversion, Gewissenhaftigkeit, Ungewissheitstoleranz, soziale Eingebundenheit von Mitstudierenden und Semesterdurchschnittsnoten erklärbar ist. Bedeutsame Prädiktoren der Selbstregulation stellen ebenso Anstrengungsbereitschaft, Selbstwirksamkeit, Neurotizismus, Extraversion und zusätzlich Resignationstendenz, Prokrastination, das Alter sowie die Abiturdurchschnittsnoten dar. Interventionsmaßnahmen sollten demzufolge insbesondere die Anstrengungsbereitschaft und Selbstwirksamkeit von Studierenden fördern.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Funk, Alexandra; Wichmann, Astrid; Rummel, Nikol

Steigerung der Fähigkeit Textprobleme zu erkennen?

Das Verfassen eigener Texte fällt Studierenden schwer. Beim Erlernen von Schreibfähigkeiten, haben sich Peer-Feedback-Verfahren als wirksam erwiesen, da Lernende durch Rückmeldung von Mitstudierenden auf Fehler in ihren Texten aufmerksam gemacht werden und diese daraufhin entsprechend überarbeiten können. Neben der Fähigkeit der Problemüberarbeitung ist es wichtig Probleme im eigenen Text zu erkennen (Hayes & Flowers, 1986). Die Fähigkeit zur Problemerkennung kann auf unterschiedliche Weise gefördert werden. Zum einen müssen Lernende in die Lage versetzt werden, den eigenen Text anhand der in der Aufgabenstellung enthaltenen Schreibkriterien und Ziele zu bewerten. Durch das Einnehmen einer Gutachterrolle kann es Lernenden erleichtert werden, Schreibkriterien und Ziele zu reflektieren. In dieser Rolle können Wissenslücken erkannt und Schreibkriterien internalisiert werden (Topping, 1998; Cho & Cho, 2011). Zum anderen sollten Lernende befähigt werden, fehlerhafte Textstellen selbst zu erkennen. Dies kann beispielsweise durch Rückmeldung zum eigenen Text unterstützt werden, bei der nicht jede falsche Textstelle markiert wird, sondern Schreibfehler nur exemplarisch hervorgehoben werden. Exemplarisches Feedback kann Suchprozesse anregen, die zur selbstständigen Erkennung weiterer Fehler im eigenen Text führen. Im Gegensatz dazu lässt eine ausführlichere Rückmeldung von Schreibfehlern wenig eigene Fehlererkennungsaktivität zu.

Ziel einer experimentellen Studie war die Verbesserung der Fähigkeit zur Problemerkennung von Studierenden beim Verfassen wissenschaftlicher Texte. Zwei Faktoren wurden variiert: Das Einnehmen einer Gutachterrolle (ja vs. nein) und die Menge der Rückmeldung (niedrig vs. hoch). Insgesamt 83 Studierende der Geisteswissenschaften wurden zufällig einer von vier Bedingungen zugewiesen: Gutachterrolle und Erhalten einer niedrigen Menge von Feedback (G+/5), nur Erhalten einer niedrigen Menge von Feedback (G-/5), Gutachterrolle und Erhalten einer hohen Menge von Feedback (G+/10) oder nur Erhalten einer hohen Menge von Feedback (G-/10). Die Studie fand im Rahmen einer zweiwöchigen Online-Schreibübung statt. Darin verfassten die Teilnehmer einen Aufsatz von 600 Worten, den Sie nach dem Erhalt von Feedback überarbeiteten. Nach dem Schreiben des Aufsatzes erhielten alle Teilnehmer und Teilnehmerinnen eine Beschreibung zu verschiedenen Schreibkriterien. Zusätzlich gaben die Gruppen mit Gutachterrolle Feedback zu einem Aufsatz. Die Rückmeldung wurde mit Hilfe der Kommentierungsfunktion von WORD an den problematischen Textstellen platziert. Teilnehmer der Gruppen mit einer niedrigen Menge an Feedback erhielten fünf Feedbackkommentare und die Gruppen mit einer hohen Menge an Feedback erhielten zehn Feedbackkommentare. Vor und nach den Schreibaktivitäten wurde die Überarbeitungskompetenz mit einem zweiteiligen Test zur Problemerkennung und Problemkorrektur erfasst. Abschließend erfolgte ein Anwendungstest, der aus einem fehlerhaften Aufsatz eines/r Kommilitonen/in bestand, in dem Fehler gesucht, markiert und die Schreibkriterien zugeordnet werden sollten. Zusätzlich erfassten wir die Wahrnehmung des Feedbacks mit einer adaptierten Version des Feedback Perception Questionnaire von Srijbos et al. (2010) und das Frustrationslevel der Teilnehmer bezüglich des Feedbacks mit einem selbst entwickelten Fragebogen ($\alpha=.65$). Die Studie wurde im Sommer durchgeführt und die Daten werden zur Zeit ausgewertet. Es liegen erste Ergebnisse zur Wahrnehmung des Feedbacks vor. Eine zweifaktorielle ANOVA zeigte, dass Studierende in den Bedingungen die eine niedrige Menge an Feedback erhielten, das Feedback als positiver empfanden ($MG+/5=7.20$, $SD=1.35$; $MG-/5=7.04$, $SD=.87$), als Studierende in den Bedingungen die eine höhere Menge an Feedback erhielten ($MG+/10=6.21$, $SD=1.19$; $MG-/10=6.10$, $SD=1.25$; $F=19.73$, $p=.00$). Zudem empfanden Studierende in den Bedingungen mit einer hohen Menge an Feedback, das Feedback als frustrierender ($MG-/10=3.50$, $SD=1.48$; $MG+/10=2.65$, $SD=1.17$), als Studierenden in den Bedingungen mit einer niedrigen Menge an Feedback ($MG+/5=2.67$, $SD=1.37$; $MG-/5=2.19$, $SD=.86$; $F=10.12$, $p=.00$). Eine geringere Menge an Feedback wirkt sich also insgesamt positiv auf die Wahrnehmung von Feedback aus. Weitere Ergebnisse zur Wahrnehmung des Feedbacks und zur Überarbeitungsfähigkeit werden auf der Konferenz präsentiert.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

3

Schönfeld, Tim; Wichmann, Astrid; Rummel, Nikol

Wie nehmen Studierende Feedback wahr? Eine empirische Untersuchung der Wahrnehmung von Peer- und Experten-Feedback in der universitären Lehre

Im Lehramtstudium trifft eine große Anzahl von Studierenden unterschiedlichster Disziplinen aufeinander, um gemeinsam auf den Lehrerberuf vorbereitet zu werden. Studierende mit sehr heterogenen Voraussetzungen werden mit gleichen Anforderungen konfrontiert. Hierfür kann Feedback entscheidend zum Erfolg der Studierenden beitragen. Allerdings müssen sich Lehrende aus Zeitmangel oft auf kurzes, ergebnisorientiertes Feedback beschränken (Narciss & Huth, 2004). Um die Studierenden dennoch durch elaborierte und individualisierte Rückmeldungen zu unterstützen, können Peer-Feedback-Verfahren eingesetzt werden, welche aktuell große Popularität erfahren (Kollar & Fischer, 2010). Hierbei begutachten Studierende die Arbeit ihrer Mitstudierenden (Peers) und geben sich gegenseitig Feedback (Topping, Smith, Swanson & Elliot, 2000). Obwohl sich Peer-Feedback positiv auf das Lernen auswirken kann (Topping, 1998), führt es nicht notwendigerweise zu positiveren Lernergebnissen. Ein wichtiger Erfolgsfaktor ist unter anderem wie Peer-Feedback (auch im Vergleich zu Experten-Feedback) von den Studierenden wahrgenommen wird (Kaufmann & Schunn, 2010). Darüberhinaus stellt sich die Frage, inwiefern sich die Erwartungshaltung der Studierenden ändert, nachdem sie an einem Peer-Feedback-Verfahren teilgenommen haben.

Diese Fragen wurden im Rahmen einer Studie mit Studierenden (N=28) der Erziehungswissenschaft untersucht. Über den Zeitraum eines Semesters hinweg haben Studierende Produkte (z.B. Handouts) erstellt, die in einem Feedback-Zyklus überarbeitet wurden. Innerhalb eines Feedback-Zyklus haben sich die Studierenden zunächst gegenseitig Feedback gegeben (Peer-Feedback), konnten daraufhin ihre Produkte überarbeiten und haben anschließend Feedback von der Dozentin (Experten-Feedback) erhalten. Mithilfe eines adaptierten Fragebogens basierend auf dem Feedback Perception Questionnaire (FPQ) von Strijbos, Narciss & Dünnebier (2010), wurde die Wahrnehmung (hinsichtlich Nützlichkeit, Fairness, Motivation, Verbesserungswille und das Vertrauen in den Assessor) einmal zu Beginn und jeweils einmal nach jedem Feedback-Zyklus erhoben. Die Reliabilitäten (interne Konsistenz) der einzelnen erhobenen Skalen waren zufriedenstellend ($\alpha > 0.7$).

Die Ergebnisse zu den Messungen vor dem Feedback-Zyklus zeigten eine positivere Erwartungshaltung gegenüber Experten-Feedback im Vergleich zu Peer-Feedback in allen Konstrukten. Auch nach dem Durchlaufen des ersten Feedback-Zyklus gaben die Studierenden an, dass sie das Peer-Feedback ($M = 3.88$, $SD = .79$) als weniger nützlich empfanden, als das Experten-Feedback ($M = 4.40$, $SD = .52$, $p = .004$). Ebenso hatten die Studierenden mehr Vertrauen in die Experten ($M = 4.56$, $SD = .50$) als in Peers ($M = 3.80$, $SD = .63$, $p = .000$) und gaben eine höhere Bereitschaft an, das Feedback der Experten ($M = 4.27$, $SD = .83$) als das der Peers umzusetzen ($M = 4.00$, $SD = .65$, $p = .051$). Interessanterweise hielten die Studierenden ihre Peers ($M = 4.28$, $SD = .56$) aber für ebenso fair wie die Experten ($M = 4.30$, $SD = .72$, $p = .461$).

Da die Studie im Feld durchgeführt wurde und dementsprechend an den Seminarkontext angepasst werden musste, wurde von einem between-subject design abgesehen und ein within-subject design ohne counter-balancing durchgeführt. Die Wahrnehmung wurde nach jedem Zyklus sowohl für Peer-Feedback als auch für Experten-Feedback einmal erhoben. Es ist davon auszugehen, dass die Wahrnehmung des Experten-Feedback von der vorausgehenden Erfahrung des Peer-Feedback beeinflusst wurde. Ein Vergleich zwischen Peer und Experten-Wahrnehmung ist deshalb nur mit Einschränkung möglich.

Die Ergebnisse bestätigen Annahmen aus der Peer-Assessment Forschung (Kaufmann & Schunn, 2010). Wenn es um die Einstellung zu Feedback geht, sind Studierende gegenüber Experten-Feedback positiver eingestellt als gegenüber Peer-Feedback, was sich, bis auf wenige Ausnahmen nicht ändert nachdem sie einen Feedback-Zyklus durchlaufen haben. Eine Ausnahme stellt Fairness dar. Interessanterweise verbesserte sich die Wahrnehmung des Peer-Feedback signifikant im Hinblick auf Fairness nach dem Durchlaufen des Peer-Expert-Feedback Zyklus (Wahrnehmung vor dem Zyklus: $M = 3.58$, $SD = .73$; Wahrnehmung nach dem Zyklus: $M = 4.18$, $SD = 0.56$, $p = .001$). Damit wurde das Peer-Feedback als genauso fair wahrgenommen wie das Experten-Feedback. Für zukünftige Studien stellt sich die Frage, welche Maßnahmen ergriffen werden können, damit Studierende Peer-Feedback (in Relation zu Experten-Feedback) positiver wahrnehmen.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Lammerding, Sandra; König, Johannes

Erfassung von Lerngelegenheiten angehender Englischlehrkräfte – Instrumentenvalidierung durch Expertenurteile

Theoretischer Hintergrund: Mit dem aktuell laufenden DFG-Projekt ""Professionelle Kompetenz von Englischlehrkräften: Fachdidaktisches Wissen angehender Englischlehrkräfte – Konzeption, Messung, Validierung"" (PKE, König, Rothers & Nold, 2013) wird derzeit, aufbauend auf die Arbeiten im Kontext von TEDS-M, ein Test zur standardisierten Erfassung des fachdidaktischen Wissens von Englischlehrkräften entwickelt. Damit werden im Sommer 2015 angehende Englischlehrkräfte der ersten und zweiten Ausbildungsphase getestet. Ziel ist es unter anderem, einen Beitrag zu leisten zu den Diskursen, Positionen und Konzepten der Fremdsprachenlehrausbildung in der Englischdidaktik, denen bislang eine empirische Prüfung fehlt.

Bei der geplanten Testung angehender Englischlehrkräfte unterschiedlicher Ausbildungsstadien (Master-Studierende, Referendare) ist ferner die Erfassung von Merkmalen ihrer Lerngelegenheiten entscheidend, um den Einfluss ausbildungsbezogener Variablen als Bedingungen der Kompetenzentwicklung zu prüfen. Im Rahmen des PKE-Projekts wurde hierfür in Anlehnung an Vorarbeiten zur standardisierten Untersuchung von Lerngelegenheiten im Lehramtsstudium (u.a. Darge et al., 2012; König et al., 2014) ein entsprechender Fragebogen entwickelt und an angehenden Englischlehrkräften aus beiden Ausbildungsphasen pilotiert. Dabei werden Lehramtsstudierende der Sekundarstufe I mit dem Fach Englisch systematisch u.a. zu den inhaltlichen Schwerpunkten ihrer Lerngelegenheiten an der Universität befragt.

Fragestellung: Zur Prüfung der curricularen Validität dieses neu entwickelten Fragebogens zur Erfassung von Lerngelegenheiten angehender Englischlehrkräfte werden derzeit die Items des Fragebogens in einer Experten-Befragung Fachdidaktikern an deutschen Hochschulen vorgelegt. Diese schätzen mittels eines Online-Fragebogens die Items auf standortbezogene curriculare Relevanz ein. Ziel dieser Instrumentenvalidierung durch Expertenurteile ist es, die Relevanz der zur Erhebung verwendeten Items von Expertinnen und Experten zu erfassen. Im Vortrag gehen wir zwei analytischen Fragestellungen nach, die sich im Zusammenhang dieser Expertenbefragung stellen: (1) Gibt es Unterschiede in der Relevanz der eingeschätzten Inhalte, die im Zusammenhang mit curricularen Schwerpunkten der englischdidaktischen Lehrerbildung an Hochschulen stehen? Wir vermuten, dass disziplinbezogene Inhalte der Fachdidaktik (z.B. Sprachdidaktik) eine höhere Relevanzzuschreibung erhalten als schulpraxisbezogene (z.B. konkrete Unterrichtsmethoden), da letztere stärker Lerngegenstand der zweiten Ausbildungsphase sind. (2) Gibt es Unterschiede in der eingeschätzten Relevanz in Abhängigkeit der Berufsqualifikationen der Experten? Wir vermuten, dass forschungsstarke Experten (u.a. Habilitation, hohe Publikationsanzahl innerhalb der vergangenen fünf Jahre) die disziplinbezogenen Inhalte der Fachdidaktik stärker gewichten als Experten mit primär schulpraktischem Werdegang.

Methode: Aktuell werden bundesweit 45 Fachdidaktiker als Experten angeschrieben und um Teilnahme an der Online-Befragung bis 30.11.2014 gebeten. Deckungsgleich mit der Itemanzahl im Lerngelegenheitsfragebogen für die angehenden Lehrkräfte werden in unserer Studie Einschätzungen der Experten auf insgesamt 115 Items zu Inhalten fachdidaktischer Lehre vorliegen. Folgende Inhalte werden abgedeckt: Sprachdidaktik/ Spracherwerb (32 Items), Literaturdidaktik (12 Items), Cultural Studies/ Kulturdidaktik (11 Items), Mediendidaktik (13 Items), Methodik des Englischunterrichts (18 Items), Curriculum und Assessment/ Beurteilung (16 Items), Fachdidaktische Forschung (13 Items). Experten lokalisieren dabei ihr Urteil auf einer fünfstufigen Skala (sehr wichtig – wichtig – weniger wichtig – unwichtig – keine Angabe). Ferner werden personale Merkmale der Experten erfragt (u.a. wiss. Qualifikation, berufliche Position, Anzahl der Publikationen innerhalb der letzten fünf Jahre), welche eine Unterscheidung voraussichtlich in zwei Expertengruppen erlaubt (forschungsstark vs. praxisnah). Die mit unseren Fragen vermuteten Unterschiede werden varianzanalytisch untersucht und auf statistische Signifikanz geprüft.

Ergebnisse: Im Vortrag wird zunächst die Konzeptualisierung der Inhalte englischdidaktischer Lehre, wie sie mit dem Erhebungsinstrument strukturiert werden, dargelegt. Da die Datenerhebung bei Fertigstellung dieses Abstracts noch nicht abgeschlossen ist, können an dieser Stelle noch keine Ergebnisse genannt werden. Die zur Tagung final vorliegenden und zu präsentierenden Ergebnisse tragen jedoch nicht nur zur Validierung der Erfassung von Lerngelegenheiten im Rahmen des PKE-Projekts durch Expertenurteile bei, sondern geben einen Einblick in den Stellenwert einzelner Themen und Studienschwerpunkte innerhalb der fachdidaktischen Ausbildung im Fach Englisch an deutschen Hochschulen. Dies soll zur Diskussion gestellt werden.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Raser, Marcus; Hense, Jan

Was geschieht nach der Rückmeldung? Entwicklung eines Instruments zur Erfassung der Einflussmechanismen von Lehrveranstaltungsevaluationen

Problemstellung: Zur Verbesserung der Lehrqualität an Hochschulen werden vielfach Verfahren zur Lehrveranstaltungsevaluation (LVE) eingesetzt, die zwar breitflächig etabliert sind, oft aber als folgenlose Belastung wahrgenommen werden (Schmidt, 2007). Durch Optimierung dieser Verfahren könnte diese wahrgenommene Folgenlosigkeit reduziert werden. Dazu ist eine differenzierte Analyse der Wirkungen von LVE auf Seiten der Lehrenden notwendig, um untersuchen zu können, inwiefern verschiedene Variationen des LVE-Verfahrens (z.B. Modifikationen der Ergebnismrückmeldungen) zu einer verbesserten Wirksamkeit der Rückmeldungen führen. Bisher fehlt für diese differenzierte Analyse ein geeignetes Instrument.

Theoretischer Hintergrund: Die reine Rückmeldung von LVE-Ergebnissen an Lehrende scheint kaum Einfluss auf die Verbesserung der Lehrqualität zu haben, wenn die Ergebnismrückmeldung nicht mit Beratungsangeboten verknüpft wird (z. B. Rindermann, 2003; Marsh, 2007). Daneben ist es aufgrund vieler konfundierender Variablen und des hohen methodischen Aufwands schwierig, direkte Verhaltensänderungen der Lehrenden zu messen und sichtbar zu machen. Die Theory of Evaluation Influence von Mark & Henry (2004) beschreibt neben einer solchen (direkten) Verhaltensänderung weitere Einflussmechanismen bzw. intraindividuelle Reaktionen von Evaluationsrezipienten (z. B. Lehrpersonen), die als Vorläufer für Wirkungen von LVE-Verfahren betrachtet werden können. Beispiele sind etwa Salienz (subjektive Wichtigkeit der Lehre verändert sich in Folge der Rückmeldung) oder Elaboration (LVE als Anlass über Lehre nachzudenken).

Fragestellung: Ziel des Forschungsprojekts ist eine differenzierte Analyse der Rezeptions- und Wirkungsprozesse von LVE auf Seiten der Lehrenden zur gezielten Optimierung von LVE-Verfahren. Als erster Schritt wird in dieser Studie die Erstellung eines Instruments zur Erfassung der Einflussmechanismen von Lehrveranstaltungsevaluationen (ELVE) vorgestellt, das solche intraindividuellen Veränderungen auf motivationaler, kognitiver und behavioraler Ebene erfassen soll.

Methode: Zunächst wurden theoriegeleitet folgende (möglichen) Einflussmechanismen identifiziert: Elaboration, Priming, Salienz, Fertigungsaneignung, Verhaltensänderung (Mark & Henry, 2004) sowie Verbesserungshinweise, Nachweis der Lehrbefähigung, Äußere Kontrolle und Selbstbild (Viebahn, 2004).

Im Rahmen einer qualitativen Vorstudie wurden anschließend mit neun Lehrenden gleicher Hierarchiestufe an deutschen Hochschulen leitfadengestützte Interviews geführt, in denen diese zu jedem der neun Einflussmechanismen befragt wurden, ob durch die LVE in diesem Bereich eine Veränderung bei ihnen stattgefunden hat. Ziel war es, zu testen, ob die theoretisch postulierten Einflussmechanismen tatsächlich von Lehrpersonen im Rahmen der Rezeption von LVE-Ergebnissen berichtet werden.

Anschließend wurden zu jedem der neun Einflussmechanismen drei bis max. fünf Items formuliert. Jedes Item ist als Aussage formuliert, die auf einer 5-stufigen Zustimmungsskala bewertet werden soll. Diese erste Version des Instruments wird zu Beginn des Wintersemesters 2014/15 online an alle Lehrpersonen einer großen deutschen Universität versandt. Anschließend sollen Itemanalysen durchgeführt und die Faktorstruktur mittels konfirmatorischer Faktorenanalyse geprüft werden. Außerdem werden Hinweise auf Differenzen sowohl in der Faktorstruktur als auch in den absoluten Werten abhängig von der Statusgruppe sowie der Lehrerfahrung der Lehrpersonen analysiert.

Nach einer Itemreduktion anhand der Item- und Faktorkennwerte, soll am Ende der Vorlesungszeit des Wintersemesters 2014/15 mit der Validierung des Instruments begonnen werden.

Ergebnisse: Durch die qualitative Vorstudie konnte gezeigt werden, dass der Unterstellung von Folgenlosigkeit von LVEs nicht pauschal zustimmen ist. So wurde zu jedem der neun Einflussmechanismen (mit Ausnahme der Verhaltensänderung) von einer bis acht Lehrpersonen beschrieben, dass durch LVE diese Veränderung stattgefunden habe. Erste Ergebnisse lassen erwarten, dass die in der qualitativen Vorstudie gefundenen Effekte sich auch mit einem quantitativen Instrumente empirisch erschließen lassen.

Ausblick: Insofern liegt mit dem ELVE eine wichtige methodische Grundlage vor, um die Wirkungen von Evaluationsrückmeldungen in Lehrkontexten differenziert erschließbar zu machen. Dies ist eine wichtige Basis für die Erhöhung der Wirkungen von LVE insgesamt und sollte langfristig - trotz steigender Heterogenität der Studierendenschaft - zu einer Verbesserung der Lehrqualität an Hochschulen beitragen.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Weißerrieder, Jochen

Effektive Evaluation – Eine quasi-experimentelle Studie zum Nutzen und zur Nutzung von Evaluationen durch Fortbildner_innen

Motivation: Die Evaluation von Fortbildungen geht häufig mit einem hohen Aufwand für Lehrende und Teilnehmende einher, selten jedoch mit einem hohen Nutzen in Form von Weiterentwicklung des Designs der Fortbildungsmaßnahme.

Theoretische Einbindung: Wissenschaftlichen Studien zu Evaluationen in Schule und Hochschule zeigen, dass Evaluationen ineffektiv bleiben, wenn die Evaluationsergebnisse nur vorgelegt werden (im Überblick: Rindermann, 2001). Aus der Forschung zur Wahrnehmung von Schulleistungsstudien entstand z.B. das „Zyklusmodell von der Rückmeldung zur Innovation“ (Hosenfeld, 2005): Lehrende müssen demnach die Ergebnisse für relevant befinden (wahrnehmen) und reflektieren, um die Maßnahme weiterzuentwickeln. Marsh (1993) wiesen in experimentellen Feldstudien nach, dass die Lehre dann verbessert wird, wenn die Lehrenden den Erhebungsinstrumenten eine praktische Bedeutung zumessen und zu den Ergebnissen beraten werden. Laut qualitativen Studien von Koch (2011) lässt sich auch ein gesteigerter Nutzen nach der Teilnahme von Fortbildungen feststellen.

Im Bereich von Fortbildungen kann dieser Aufwand nur selten realisiert werden. Es braucht andere Wege, um die inhaltliche Relevanz für die Lehrenden und deren Fähigkeit, die Ergebnisse zu reflektieren, zu erhöhen. Die Evaluationsforschung zeigt, dass der Nutzen der Evaluation durch eine wahrgenommene Nähe zwischen Evaluierendem und Evaluation gefördert werden können (Weiss, 1980)

Forschungsfragen: Die in diesem Beitrag vorgestellte Dissertationsstudie überträgt diese theoretischen Ansätze auf die Evaluation von 60 Berliner Lehrerfortbildungen im November und Dezember 2014. Dabei ist die folgende Fragestellung richtungsweisend: Lassen sich durch eine enge Einbindung der Lehrenden in den Evaluationsprozess einerseits und eine verbesserte Führung durch den Reflexionsprozess andererseits die Relevanz der Evaluation und die Reflexion der Ergebnisse verbessern?

Design: Auf einen Pretest für die Referentinnen und Referenten (RuR) folgt eine Fortbildungsevaluation, der sowohl ein Feedbackbogen der Fortbildungsteilnehmenden als auch eine schriftliche Reflexion durch die RuR enthält. Abschließend wird ein Posttest mit den RuR durchgeführt. Als abhängige Variablen werden die Veränderung der Einstellung gegenüber Evaluation (vgl. Weiss, 1980; Schloemerkerper, 2002) (Pre-Posttest-Vergleich), die Befragungsergebnisse selbst (TN-Feedback), Reflexionsgenauigkeit sowie Reflexionstiefe (schriftliche Reflexion) sowie Veränderungsabsicht und -intensität (Posttest) herangezogen.

Untersucht werden drei Variationen der Fortbildungsevaluation: Einbindung der RuR, Zeitpunkt der Rückmeldung sowie Art und Weise der schriftlichen Reflexion. Das Design und die Variationen wurden nach einer Pilotstudie mit 16 Fortbildungen überarbeitet.

Einbindung: Die Hälfte der RuR kann den Fragebogen für die Teilnehmenden durch Auswahl von Fragen und Formulierung eigener Ziele mitgestalten. Ergebnisse der Pilotstudie zeigen, dass dies zwar zu besseren Befragungsergebnissen führt (Prozessnutzen), weder Reflexionsgenauigkeit oder -tiefe, noch Veränderungsabsichten oder -intensität wurden beeinflusst.

Zeitpunkt der Rückmeldung: Die Hälfte der RuR erhält die Ergebnisse der TN nach der eigenen schriftlichen Reflexion. Ergebnisse der Pilotstudie legten die Vermutung nahe, dass RuR die Ergebnisse zwar erhielten, für die Reflexion aber auf die eigenen Erfahrungen zurückgriffen. Daher wurde diese Variation aufgenommen, um zu überprüfen, inwieweit die Vorlage der Ergebnisse einen Einfluss auf spätere Prozesse hat.

Art und Weise der schriftlichen Reflexion: Ein Teil der RuR erhält Standardfragen (z.B. Was lief gut?). Die andere Hälfte der RuR erhält spezifische und getrennte Fragen zur Wahrnehmung (insbesondere auch von Minderheitenmeinungen) und Interpretation, sowie zu Erklärungstheorien. Die Ergebnisse aus der Pilotstudie führten zu einer grundsätzlichen Überarbeitung der Reflexionsanleitungen, da kein Einfluss der Variation auf abhängige Variablen nachgewiesen werden konnte.

Auswertung: Durch drei Variationen mit je zwei Möglichkeiten ergibt sich ein Design mit acht Subgruppen. Auswertungen in Form von Regressionsanalysen lassen Rückschlüsse auf die jeweils einzelne Variation als auch auf verschiedene Kombinationen zu. Der Beitrag wird zeigen, wie und inwieweit durch Variationen der Evaluationspraxis eine effektive - im Sinne von Nutzung und Nutzen - Evaluation von Fortbildung möglich ist. Wäre die Rückmeldung der Ergebnisse jedoch irrelevant für weitere Prozesse, muss die Praxis in anderer Weise überdacht werden.

E04 Pause mit Poster

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

3

Strippel, Christian Georg; Sommer, Katrin

Chemische Forschung für die Öffentlichkeit – Public Understanding of Research für Chemie-Ausstellungen

Wissenschaftskommunikation mit der breiten Öffentlichkeit findet häufig in einem von hohem Maß an Freiwilligkeit geprägten Umfeld statt, z.B. in Ausstellungen, Vorträgen, Medien (National Research Council, 2009; Waidacher, 2005). Ein solches Umfeld beeinflusst sowohl das Lernverhalten der TeilnehmerInnen als auch die Auswahl von Forschungsmethoden (ebda.). Einen Rahmen für die Betrachtung von Wissen über Forschung in der Öffentlichkeit bietet Public Understanding of Research. Hierzu zählen die Vorstellungen von Personen über Zukunftsperspektiven, Beweggründe, Akteure und den Prozess des Erkenntnisgewinns in der Forschung (Field & Powell, 2001). Die hier vorgestellte Studie befasst sich damit, wie Wissen über den Prozess des naturwissenschaftlichen Erkenntnisgewinns in der Öffentlichkeit gemessen werden kann und über welches Wissen diese Zielgruppe verfügt. Damit wird einer der vier Kompetenzbereiche der KMK-Bildungsstandards – für den bereits Tests für den schulischen Bereich bestehen – auch für eine weitere Öffentlichkeit in den Blick genommen (KMK, 2004; Lederman et al., 2014; Wellnitz, Fischer, Kauertz, Neumann, & Pant, 2012).

Für den Ausstellungsbereich ist bereits bekannt, dass Lernen über naturwissenschaftliches Wissen stattfindet und dass dies in Abhängigkeit vom Vorwissen der Lernenden geschieht (J. H. Falk & Adelman, 2003; J. Falk & Storksdieck, 2005). Dabei scheinen insbesondere Probanden mit geringem Vorwissen überdurchschnittlich an Wissen zu gewinnen (ebda.). Wissen über den Prozess des naturwissenschaftlichen Erkenntnisgewinns wird insbesondere dann erworben, wenn eine explizite Vermittlung stattfindet und zur Reflexion angeregt wird (Khishfe & Abd-El-Khalick, 2002).

Von dieser Grundlage ausgehend lauten die Fragestellungen in der hier vorgestellten Studie wie folgt: 1. Wie bildet ein Interview zu ausgewählten Aspekten des Prozesses des naturwissenschaftlichen Erkenntnisgewinns (naturwissenschaftliche Fragen, Methodenauswahl) das Wissen von Probanden im Vergleich zu einem schriftlichen Fragebogen ab? 2. Welches Wissen besitzen Personen in der Öffentlichkeit (insbesondere Ausstellungsbesucher) über ausgewählte Aspekte des Prozesses des naturwissenschaftlichen Erkenntnisgewinns (naturwissenschaftliche Fragen, Methodenauswahl)?

Diesen Fragestellungen wird mittels eines empirischen Ansatzes nachgegangen. Hierzu wird zunächst ein bestehendes Interview für SchülerInnen zum naturwissenschaftlichen Erkenntnisprozess für Erwachsene und in deutscher Sprache weiterentwickelt und durch Vergleich mit einem schriftlichen Fragebogen validiert (Lederman et al., 2014). Darauf aufbauend werden mittels des Interviews die Vorstellungen der Öffentlichkeit über ausgewählte Aspekte des naturwissenschaftlichen Prozesses der Erkenntnisgewinnung erhoben – in naturwissenschaftlichen Ausstellungen und kontrastierend an anderen öffentlichen Orten. Die Analyse der Daten erfolgt mittels qualitativer Inhaltsanalyse (Mayring, 2010). Die Ergebnisse dieser Untersuchung werden auch in die Konzeption einer Ausstellung zu den Arbeiten des RESOLV Cluster of Excellence EXC 1069 einfließen. Der Beitrag präsentiert das Forschungsvorhaben und erste Ergebnisse.

Donnerstag
12. März

10:45 Uhr

Foyer GC - West

Drechsel, Barbara; Roder, Jana; Breunig, Katharina

Theorie-Praxiserfahrungen zur Leseförderung - Eine Lerntagebuchstudie

In der aktuellen Diskussion wird immer wieder die systematische und strukturierte Verknüpfung von Theorie und Praxis im Lehramtsstudium gefordert (z.B. Hascher, 2012). Gelegenheiten, Gelerntes anzuwenden und zu erproben sind aber nach wie vor rar und die Qualität von Lerngelegenheiten in Praktika ist häufig dem Zufall überlassen (Gröschner & Seidel, 2012). Das Projekt LuPe (Leseförderung und Praxiserfahrung) versucht diesen Anforderungen an Praxiserfahrungen, die mit dem Bedürfnis vieler Studierender nach „mehr Praxis“ in Einklang stehen, zu entsprechen. Am Beispiel der Lesekompetenzförderung, die als bedeutsame Aufgabe von Lehrerinnen und Lehrern aller Fächer und Schularten angesehen werden kann (vgl. z.B. Artelt & Dörfler, 2010), wird eine Gelegenheit zum Lernen in der Praxis angeboten, die zunächst Anforderungen stellt, die die Komplexität des Lehrerhandelns reduziert, um einen engen Rückbezug zu theoretischen Grundlagen zu ermöglichen (Hascher, 2012). In einem universitären Psychologie-Seminar der ersten Ausbildungsphase werden Studierende in einem ersten, theoriebezogenen Baustein mit wichtigen Grundlagen der Leseförderung (z.B. kognitionspsychologischen Prozessen, Lesestrategien und deren Vermittlung) und der Gestaltung von Lehr-Lernprozessen (z.B. Feedback richtig geben, Fördermaßnahmen adaptiv gestalten) vertraut gemacht. In der zweiten, der Praxisphase des Seminars führen die Studierenden in fünf Tutoring-Sitzungen Leseförderung mit Kindern der fünften Klasse einer Mittelschule durch. Lerntagebücher mit Prompts zur Reflektion des Gelernten, ein individuelles Videofeedback und eine gemeinsame Feedbackrunde im Seminar sollen in einem dritten Baustein dazu beitragen, den Transfer von der Theorie in die Praxis zu erleichtern, zu beobachten und zu reflektieren, um so die gewonnenen Erkenntnisse zu konsolidieren. Ein ausschließlich theoriebasiertes Seminar zum selben Thema wird diesem Konzept gegenübergestellt. Die Auswertung der Lerntagebücher aus beiden Seminarformen steht mit folgenden Fragen im Mittelpunkt des Beitrags: Im Vergleich der Theorie- und der Praxisphase mit einem theoriebasierten Seminar werden folgende Fragen gestellt (1) Welche theoretischen Inhalte werden in den Lerntagebüchern aufgegriffen? (2) Über welche Lernprozesse und –ergebnisse berichten Studierende aus ihrer Theorie- und / oder Praxis-Erfahrung? Welche Bereiche professionellen Handelns werden fokussiert? (3) Für die Praxisphase: Gibt es Hinweise auf evidenzbasiertes Handeln, das in direkter Beziehung zur vermittelten Theorie steht? Greifen Studierende die Inhalte der Theoriebausteine auf?

30 Studierende der Lehrämter für Grund-, Haupt- und Realschulen sowie Gymnasien, absolvierten je eine Variante des Seminars zur Leseförderung und bearbeiteten währenddessen ein Lerntagebuch. Darin werden Studierende u.a. nach Lernprozessen und –schwierigkeiten während des Lernens in Theorie und Praxis gefragt. Die inhaltsanalytische Auswertung der Aussagen in den Tagebüchern gibt (1) Hinweise darauf, welche Inhalte des Seminars aufgegriffen werden, (2) mit welchen Bereichen Studierende neue Erfahrungen gewonnen haben, die für sie von besonderer Bedeutung sind und ggf. (3) ob und in welcher Weise sie die theoretisch vermittelten Kenntnisse in der Praxis aufgreifen.

Die Befunde geben Hinweise darauf, dass die Studierenden in beiden Seminarvarianten ähnliche Inhalte als bedeutsam ansehen, jeweils bedeutsame unterrichts- und leseförderungsrelevante Erkenntnisse berichten, die sich mit der Vorbereitung und Gestaltung des Unterrichts, der Person des Schülers / der Schülerin und mit der eigenen Rolle als Lehrperson beschäftigen. In der Seminarvariante, die Theorie und Praxis verknüpft, bezieht sich der größte Anteil der Nennungen sich auf das eigene unterrichtliche Handeln in der Leseförderung, was als Hinweis darauf gedeutet wird, dass Studierende die Praxiserfahrung nutzen, um sich selbst und ihr Wissen zu erproben. Aus den Ergebnissen werden Überlegungen für die Weiterentwicklung der Seminarvarianten angestellt und Schlussfolgerungen bezüglich einer evidenzbasierten Handlungsweise im praktischen Teil des Seminars gezogen. Vor dem Hintergrund der vorliegenden Befunde werden weitere Forschungsfragen diskutiert.

Donnerstag
12. März

11:15 Uhr

GA 03/49

Chair(s)

Drechsel, Barbara

Kleinbub, Iris

Berücksichtigung heterogener Lernvoraussetzungen bei Texterschließungsaktivitäten im Leseunterricht

Das Lesen zählt zu den am besten erforschten Lernbereichen der Unterrichtsforschung. Seit es Lernstandserhebungen gibt, steht auch die Fragen nach der Unterrichtsqualität im Zentrum des Interesses, da hier wesentliche Faktoren liegen, die für das Zustandekommen von Lernergebnissen eine Rolle spielen (vgl. Helmke 2009). Der Leseunterricht in der Grundschule, der durch die Diversität der Lernvoraussetzungen der Schülerinnen und Schüler stark geprägt ist, verdient dabei besondere Aufmerksamkeit, da er kompensatorische Aufgaben hat (vgl. Bos et al. 2012). Ein Defizit besteht jedoch bei beobachtungs-basierten Daten des Verlaufs von Lehr-Lern-Prozessen. Hierüber weiß man in Deutschland z.Zt. noch wenig.

Der vorliegende Beitrag geht diesem Desiderat nach und gibt Einblick in Teile der Studie „VERA – Gute Unterrichtspraxis“, der ersten Videostudie zur Lesekompetenz, die in deutschen Grundschulklassen durchgeführt wurde (Leitung: Prof. Dr. A. Helmke, Universität Koblenz-Landau). In 41 vierten Klassen in Rheinland-Pfalz wurde der Leseunterricht videografiert, um Hinweise auf kompetenzförderliche Unterrichtsangebote zu erlangen. In diesem Rahmen wurde untersucht, ob im Leseunterricht durch gezielte Unterrichtsangebote der Lehrpersonen Neigungs-, Leistungs- und Sprachenheterogenität der Schülerinnen und Schüler bei der Förderung von Lesekompetenz Berücksichtigung finden, um divergierenden familiären Lesesozialisierungen entgegenzuwirken. Des Weiteren wurde analysiert, welche Unterrichtsangebote allen Schülerinnen und Schülern zur Unterstützung der Texterschließung gemacht werden. Im Mittelpunkt dieses Beitrags stehen folgende Forschungsfragen: (1) Mit welchen Unterrichtsangeboten werden die Lerner auf die Lektüre vorbereitet? (2) Welche Themen werden zum Gegenstand der Textreflexion? (3) Mit welchen Unterrichtsangeboten werden unterschiedliche Leistungsniveaus, sprachliche Voraussetzungen und Neigungen der Lerner berücksichtigt? (4) Mit welchen Mitteln werden die Schülerinnen und Schüler im Lernprozess von der Lehrkraft unterstützt?

Darüber hinaus wurde der Frage nach der angemessenen Durchführung der Unterrichtsangebote im Hinblick auf eine Kompetenzorientierung nachgegangen. Zur Erfassung und Analyse relevanter Aspekte wurde ein hochinferentes, konzept-orientiertes Rating-Instrument entwickelt, mit dessen Hilfe Perspektiven auf die Unterrichtsangebote der Lehrenden an die Lerner eröffnet werden (Langer & Schulz von Thun, 2007). Experten-Rater wurden geschult, um den videografierten Unterricht mit Hilfe des Rating-Inventars zu beobachten und einzuschätzen. Es wurden Items zu didaktischen Inhalten wie Lektürevorbereitung und Textreflexion, zum gezielten Umgang mit Heterogenität durch Aufgabenqualität und -quantität und zu Unterstützungsangeboten entwickelt, anhand derer die Unterrichtsangebote beschrieben sowie ihre Kompetenzorientierung von Ratern eingeschätzt wurden.

Die Unterrichtsbeobachtung zeigt, dass die Lerner nur rudimentär auf die Textlektüre vorbereitet werden: Der Aufbau von Weltwissen sowie die Vermittlung von Lesestrategien und Arbeitstechniken erfolgen nicht explizit. Merkmale der Textoberfläche werden kaum zur Erschließung der Textbedeutung genutzt. Relevante Reflexionsgegenstände (bspw. formale und sprachliche Textmerkmale, kommunikative Textfunktionen, Unterscheidung von Realität und Fiktion, Austausch über subjektive Textverständnisse, Textvergleich), die über den reinen Textinhalt hinausgehen, sind unterrepräsentiert. In einem Großteil der untersuchten Unterrichtsstunden werden keine differenzierenden Maßnahmen in Form von gesonderten Arbeitstechniken, -materialien oder Aufgaben zum gezielten Umgang mit Heterogenität eingesetzt. Gleiches gilt für die Berücksichtigung der Neigungen der Lerner. Verhältnismäßig ausgeprägt sind hingegen die Angebote zur persönlichen Unterstützung der Schülerinnen und Schüler durch die Lehrpersonen (Klären von Aufgaben, Lern- und Verständnisschwierigkeiten).

Donnerstag
12. März

11:15 Uhr

GA 03/49

Chair(s)

Drechsel, Barbara

Hoya, Fabian; Hellmich, Frank

Feedback zu Leseprozessen und wahrgenommenes Klassenklima aus der Sicht von Grundschulkindern im dritten und vierten Schuljahr

Feedback zu Leseprozessen gilt als eine bedeutsame Maßnahme für die Entwicklung der Lesekompetenzen von Schülerinnen und Schülern im Leseunterricht der Grundschule. Ergebnisse von Tarelli, Lankes, Drossel und Gegenfurtner (2012) aus dem letzten Durchlauf der Internationalen Grundschul-Lese-Untersuchung (IGLU) verdeutlichen, dass ein Großteil der befragten Kinder in Deutschland Leseunterricht erhält, der im Klassenverband mit nur wenig individuellen Unterstützungsmomenten erteilt wird. Der Leseunterricht ist damit vermutlich nur wenig mit dem Blick auf individuelle Strukturierungen von Leseprozessen und wenig individuellem Feedback zu Leseleistungen ausgerichtet. Unter Rückmeldungen bzw. Feedbacks werden in diesem Zusammenhang Informationen verstanden, die Grundschullehrkräfte Kindern in der Folge von Leseleistungen bereit stellen, um Diskrepanzen zwischen aktuell erbrachten Lern- oder Leistungsergebnissen und den zu erreichenden Lernzielen aufzuzeigen (vgl. Dinnen & Collopy, 2009).

Während Hattie (2009) zwar eine hohe Effektstärke von $d=,73$ für die Erteilung von Feedback durch Lehrerinnen und Lehrer für das Lernen von Schülerinnen und Schülern im Unterricht berichtet, ist gegenwärtig noch weitgehend ungeklärt, wie Kinder Rückmeldungen in Abhängigkeit der jeweils zugrunde liegenden Rahmenmerkmale des Leseunterrichts wahrnehmen. Theoretisch naheliegend ist, dass sie Rückmeldungen in Abhängigkeit des von ihnen wahrgenommenen Klassenklimas und der perzipierten Lehrer-Schüler-Beziehung deuten. So konnte Burnett (2002) beispielsweise Zusammenhänge zwischen positiven und negativen Rückmeldungen zu Lernprozessen sowie dem Klassenklima bzw. den wahrgenommenen Lehrer-Schüler-Beziehungen im Rahmen einer Untersuchung, an der Kinder der dritten bis sechsten Jahrgangsstufe beteiligt waren, in Aspekten nachweisen. Ebenso konnte Baker (1999) im Rahmen einer qualitativ angelegten Studie mit Dritt- bis Fünftklässlern Hinweise in Bezug auf einen Zusammenhang zwischen der Lehrer-Schüler-Beziehung und perzipierten Rückmeldungen belegen.

Ungeklärt ist allerdings gegenwärtig, ob und inwiefern Unterschiede in Lesekompetenzen von Kindern durch das Zusammenspiel von Klassenklima, Lehrer-Schüler-Beziehung und Rückmeldungen erklärt werden können. Im Speziellen wird angenommen, dass die Wahrnehmung von positiven und negativen Feedbacks zu Leseprozessen der Grundschulkindern und -schüler durch das von ihnen perzipierte Klassenklima und die Lehrer-Schüler-Beziehung vorhergesagt werden können. Die Wahrnehmung positiver und negativer Feedbacks dient des Weiteren zur Prädiktion von Unterschieden in Leseselbstkonzepten, Lesemotivation und Leseverständnis von Kindern im Grundschulalter.

In unserer Studie wurden vor diesem Hintergrund $N=684$ Kinder aus dem dritten und vierten Schuljahr im Rahmen eines Fragebogens zu wahrgenommenem Klassenklima und perzipierter Lehrer-Schüler-Beziehung sowie zu den positiven und negativen Rückmeldungen, die sie im Leseunterricht erhalten, befragt. Zusätzlich nahmen sie Stellung zu ihren Leseselbstkonzepten sowie ihrer Lesemotivation und bearbeiteten den Leseverständnistest ELFE 1-6 (Lenhard & Schneider, 2006).

Die Ergebnisse aus einem Strukturgleichungsmodell – mit dem Klassenklima und der Lehrer-Schüler-Beziehung als exogene Variablen, positiven und negativen Rückmeldungen, Lesemotivation sowie Leseselbstkonzept als intervenierende Variablen und der Leseleistung als endogene Variable – verdeutlichen einen engen Zusammenhang zwischen Klassenklima und Lehrer-Schüler-Beziehung. Weiterhin zeigt sich, dass die Wahrnehmung des positiven Feedbacks über die Lehrer-Schüler-Beziehung erklärt werden kann, allerdings nicht über das Klassenklima. Dagegen können Unterschiede im perzipierten negativen Feedback über Unterschiede im wahrgenommenen Klassenklima erklärt werden. Weiterhin wird deutlich, dass der Effekt des wahrgenommenen positiven Feedbacks auf das Leseselbstkonzept durch die perzipierte Lehrer-Schüler-Beziehung moderiert wird, während derjenige des wahrgenommenen negativen Feedbacks auf das Leseselbstkonzept durch das perzipierte Klassenklima moderiert wird. Unterschiede in der Lesemotivation der Kinder werden durch Unterschiede in ihren lesebezogenen Selbstkonzepten deutlich. Dabei werden die Effekte der Leseselbstkonzepte auf die Lesemotivation der Kinder durch die von ihnen wahrgenommenen positiven wie negativen Rückmeldungen im Leseunterricht moderiert. Ein enger Zusammenhang besteht außerdem zwischen der Lesemotivation und dem Klassenklima aus der Sicht der Kinder. In Bezug auf die Leseleistung erweisen sich das Leseselbstkonzept der Kinder sowie die perzipierten negativen Rückmeldungen als signifikante Prädiktoren.

Die Ergebnisse der Studie geben zusammenfassend Hinweise darauf, dass Kinder in der zweiten Hälfte der Grundschulzeit positive wie negative Feedbacks vor dem Hintergrund der von ihnen wahrgenommenen Lehrer-Schüler-Beziehung und dem durch sie perzipierten Klassenklima deuten.

Donnerstag
12. März

11:15 Uhr

GA 03/49

Chair(s)
Drechsel, Barbara

3

Nonte, Sonja; Stubbe, Tobias

Mehrgruppenanalysen zur Beschreibung von geschlechtsspezifischen Unterschieden in der LeseEinstellung von Grundschülerinnen und Grundschülern in den Jahren 2001, 2006 und 2011

Allgemein wird die Einstellung zum Lesen als ein Bewusstseinszustand beschrieben, der von Gefühlen und Emotionen begleitet wird, der Lesen mehr oder weniger wahrscheinlich werden lässt (Smith, 1990). In verschiedenen Studien wurde eine konsistente Abnahme der positiven LeseEinstellungen während der Grundschulzeit beobachtet (McKenna et al., 1995; Kush & Watkins, 1996). Darüber hinaus zeigen sich geschlechtsspezifische Unterschiede zu Gunsten der Mädchen (Smith, 1990; Kush & Watkins, 1996). Die Einstellung zum Lesen steht in Deutschland in einem moderaten Zusammenhang mit den Leseleistungen (Bos et al., 2012) und gilt darüber hinaus als relevanter Prädiktor für die Vorhersage der Häufigkeit von Leseaktivitäten und der Ausdauer beim Lesen (McKenna et al., 1995).

Bos, Bremerich-Vos, Tarelli und Valtin (2012) berichten anhand der Daten aus der „Internationalen Grundschul-Lese-Untersuchung“ (IGLU) einen positiven Trend für die Lesemotivation und das Leseselbstkonzept von Grundschülerinnen und Grundschülern im Zeitraum von 2001 bis 2011, wobei sich auch hier deutliche Geschlechtsunterschiede zu Gunsten der Mädchen zeigen.

Es wird der Frage nachgegangen, ob der berichtete positive Trend der LeseEinstellung in den Jahren 2001, 2006 und 2011 auch unter Kontrolle von verschiedenen individuellen und schulspezifischen Einflussfaktoren beobachtet werden kann. Darüber hinaus wird analysiert, wie sich insbesondere die Einstellung der Jungen in diesem Zeitraum verändert hat und ob sich Unterschiede in den Zusammenhangsmustern gegenüber denen der Mädchen zeigen.

Die vier Items der internationalen Skala „Students' Attitude Toward Reading“ bildeten die Basis für die Schätzung der latenten Modelle im Strukturgleichungsansatz. Anhand der IGLU-Daten aus den Jahren 2001 (n = 7633), 2006 (n = 7899) und 2011 (n = 4000) wurden Unterschiede zwischen den Erhebungszeitpunkten mit Hilfe von Mehrgruppenanalysen in Mplus 7.11 (Muthén & Muthén, 1998-2013) ermittelt. Darauf aufbauend wurden Mehrgruppen-Mehrebenenanalysen geschätzt, um Einflussfaktoren auf Individual- und Schulebene zu identifizieren. Darüber hinaus konnten Mehrgruppenmodelle für die Gruppierungsvariable Geschlecht realisiert werden, welche die Analyse von Geschlechtsdisparitäten ermöglichen. Fehlende Werte wurden mit Mplus 7.11 multipel imputiert (m=10).

Für die vier Items umfassende Skala der LeseEinstellung kann das Vorhandensein von strikter Messinvarianz sowohl für die einzelnen Kohorten als auch für das Geschlecht bestätigt werden. Darüber hinaus zeigt sich ein signifikant positiver Trend der LeseEinstellung für die Jahre 2001, 2006 und 2011. Jungen weisen jedoch zu allen Erhebungszeitpunkten geringere Werte auf als Mädchen. Diese Unterschiede bleiben über die Zeit stabil. Der Zusammenhang zwischen den frühen Leseaktivitäten in der Familie und der Einstellung der Jungen zum Lesen (im Vergleich zu den Mädchen) ist besonders stark in den Jahren 2001 und 2006 ausgeprägt, wohingegen im Jahr 2011 eher die Anzahl an heimischen Büchern für eine positive LeseEinstellung der Jungen relevant ist.

Donnerstag
12. März

11:15 Uhr

GA 03/49

Chair(s)

Drechsel, Barbara

Wolf, Katrin; Schroeders, Ulrich; Kriegbaum, Katharina

Metaanalyse zur Wirksamkeit deutschsprachiger Trainingsprogramme zur Förderung der phonologischen Bewusstheit

Die Bedeutung der phonologischen Bewusstheit als Vorläuferfähigkeit des Schriftspracherwerbs wurde umfassend untersucht und empirisch wiederholt bestätigt. Lundberg, Frost und Petersen (1988) konnten überdies in einer international viel beachteten Studie zeigen, dass sich die phonologische Bewusstheit vor Schuleintritt fördern lässt und diese Förderung nachhaltige Effekte auf das Lesen und Schreiben bis zum Ende der Grundschulzeit besitzt. Seitdem wurden vergleichbare Förderprogramme in vielen Ländern und Sprachen entwickelt und die ursprünglichen Befunde von Lundberg und Kollegen wiederholt repliziert. Auch in Metaanalysen konnte die Wirksamkeit der Programme bestätigt werden (Bus & van Ijzendoorn, 1999; Ehri et al., 2001). Der Fördereffekt auf die Lesekompetenz wurde dabei mit einer Stärke von durchschnittlich $d = 0,70$ (Bus & van Ijzendoorn, 1999) beziehungsweise $d = 0,53$ (Ehri et al., 2001) angegeben. Nach bis zu 36 Monaten nach dem Training waren die Trainingseffekte mit $d = 0,16$ beziehungsweise $d = 0,23$ jedoch deutlich geringer ausgeprägt. Auch in Deutschland wurde eine Reihe von Programmen zur Förderung der phonologischen Bewusstheit entwickelt, um günstigere Voraussetzungen für den Schriftspracherwerb von Schulanfängern zu schaffen (z.B. „Hören, lauschen, lernen“ von Küspert & Schneider, 2006).

Aufgrund von Unterschieden in der Konsistenz der Graphem-Phonem-Korrespondenzen und der Phonem-Graphem-Korrespondenzen zwischen der deutschen und englischen Sprache können die Befunde der internationalen Metaanalysen nicht auf die deutschen Förderprogramme übertragen werden. Es wurde angenommen, dass im Kontext des deutschen Schriftsystems, einer transparenten Orthografie mit konsistenten Graphem-Phonem-Korrespondenzen, die phonologische Bewusstheit für Leseanfänger weniger bedeutend ist als in intransparenten Orthografien wie z.B. dem Englischen (vgl. Ziegler & Goswami, 2005). Deshalb wurde erwartet, dass eine Förderung der phonologischen Bewusstheit im Kontext der deutschen Orthografie signifikant geringere Effekte auf die frühe Lesekompetenz hat, als dies in den internationalen Metaanalysen berichtet wurde.

Mit Hilfe einer Metaanalyse wurde die Wirksamkeit deutschsprachiger Programme zur Förderung der phonologischen Bewusstheit, auch hinsichtlich der Transfereffekte auf die frühe Lese- und Schreibkompetenz, untersucht. Zur Ermittlung der mittleren Effektstärken wurden 25 größtenteils nicht randomisierte Primärstudien mit unbehandelten Kontrollgruppen in einer Random-Effekt-Metaanalyse berücksichtigt. Zusätzlich wurde mittels Moderatorenanalyse der Einfluss folgender Variablen auf die Ausprägung der Effektstärke abgeschätzt: (a) Förderzeitpunkt (letztes Kitajahr vs. erstes Schuljahr), (b) Förderinhalte (rein phonologisch vs. phonologische und andere Inhalte), (c) Sprachhintergrund (deutsch vs. nicht-deutsch) und (d) Ausgangskompetenzen der Stichprobe (schwach vs. nicht schwach). Die einzelnen Effektstärkemaße wurden als Leistungsunterschiede zwischen Förder- und Kontrollgruppe zum Posttest und den Folgeuntersuchungen mittels standardisierter Mittelwertsdifferenz (Hedges g) kalkuliert und für die Integration in ein Random-Effekt-Modell mit der inversen Varianz gewichtet. Mögliche Unterschiede von Förder- und Kontrollgruppen bezüglich ihrer Ausgangskompetenzen wurden durch den Einbezug eines Moderators zur Vergleichbarkeit der Gruppenkompetenzen kontrolliert. Die Homogenität der Effektmaße wurde durch entsprechende Statistiken geprüft (Cochrans Q-Test, 1954).

Die Analysen belegen einen moderaten, unmittelbaren Effekt der Förderung auf die phonologische Bewusstheit ($d = 0,55$). Dagegen konnten weder in der ersten noch in der zweiten Jahrgangsstufe Effekte der phonologischen Fördermaßnahmen auf die Dekodierfähigkeit festgestellt werden ($d = 0,05$ bzw. $d = 0,06$). Der durchschnittliche Effekt auf die Rechtschreibkompetenz betrug in der ersten Jahrgangsstufe $d = 0,23$ und in der zweiten Jahrgangsstufe $d = 0,20$. Trotz teilweise nachgewiesener Heterogenität konnten die untersuchten Moderatoren keinen bedeutsamen Beitrag zur Erklärung der Heterogenität leisten.

Die vorliegende Metaanalyse stellt eine systematische, quantitative Zusammenfassung der Ergebnisse von Studien zur Förderung der phonologischen Bewusstheit im deutschsprachigen Raum dar. Die identifizierten mittleren Effekte sind deutlich geringer als die in der internationalen Literatur berichteten Fördereffekte. Dies stützt die Annahme, dass in einer transparenten Orthografie wie der deutschen Sprache mit konsistenten Graphem-Phonem-Korrespondenzen die Bedeutung der phonologischen Bewusstheit als Vorläuferfähigkeit, insbesondere des Leseerwerbs, geringer ist als in intransparenten Orthografien wie dem Englischen. Die Ergebnisse werden vor dem Hintergrund der Implikationen für die Forschung und für die pädagogische Praxis diskutiert.

Donnerstag
12. März

11:15 Uhr

GA 03/49

Chair(s)
Drechsel, Barbara

Kramer, Jochen; Trautwein, Ulrich

Studienverläufe von Abiturientinnen und Abiturienten allgemeinbildender und beruflicher Gymnasien im Vergleich

In Baden-Württemberg werden ca. ein Drittel aller allgemeinbildender Hochschulzugangsberechtigungen (HZB) an beruflichen Gymnasien (BG) erworben. Die Schülerschaft an BG unterscheidet sich in mehrerer Hinsicht von derjenigen an allgemeinbildenden Gymnasien (AG). Beispielsweise ist der Ausländeranteil an BG höher als an AG (Landesinstitut für Schulentwicklung & Statistisches Landesamt Baden-Württemberg, 2011), der Anteil an Schülern und Schülerinnen aus bildungsferneren Familien und weniger privilegierter Sozialschichten ist höher, ebenso zeigen sich Leistungsunterschiede zu Gunsten des AG beim Abitur (vgl. Köller, Watermann, Trautwein & Lüdtke, 2004). Berufliche Gymnasiastinnen und Gymnasiasten haben jedoch die Möglichkeit in der Oberstufe fachspezifisches Wissen in verschiedenen Fachbereichen zu erwerben (beispielsweise in Technik oder Wirtschaftswissenschaft), das bei entsprechender Studienfachwahl Startvorteile verschaffen kann. Außerdem stehen verschiedene Hochschultypen zum Studium zur Verfügung (z.B. Universitäten, Hochschulen für angewandte Wissenschaften, Duale Hochschule Baden-Württemberg), die sich insbesondere in ihrer Praxis- bzw. Theorieorientierung voneinander unterscheiden und – den Schultypen AG bzw. BG vergleichbar – unterschiedliche Studierendenschaften aufweisen (Kramer et al., 2011).

Im Beitrag wird die Frage beantwortet, wie sich die Studienverläufe von Abiturientinnen und Abiturienten, die ihre HZB an einem BG erworben haben von denjenigen unterscheiden, die ein AG besuchten. Sind Studienabbrüche bzw. Studienfachwechsel zwischen beiden Gruppen unterschiedlich wahrscheinlich?

Zur Überprüfung der Fragestellung werden Paneldaten aus der TOSCA-Studie (vgl. Köller et al., 2004; Trautwein, Neumann, Nagy, Lüdtke & Maaz, 2010) genutzt, in der zwei Kohorten von Abiturientinnen und Abiturienten baden-württembergischer AG und BG (Abiturjahrgänge 2002 und 2006) alle zwei Jahre wiederholt befragt werden. Betrachtet werden jene TOSCA-Teilnehmenden, die spätestens im auf das Abitur folgenden Kalenderjahr ein Studium an einer Universität, HaW oder der DHBW aufgenommen haben und sich an min. einer der Folgebefragungen zwei, vier und sechs Jahre nach dem Abitur beteiligt haben ($n_{AG} = 2586$, $n_{BG} = 1371$). Die Daten werden regressionsanalytisch ausgewertet. Tatsächlich zeigen sich signifikante Unterschiede zwischen AG und BG in der Wahrscheinlichkeit, mit der ein Studium abgebrochen bzw. fortgeführt wird zugunsten der AG-Gruppe (Cox Regression, Waldtest $T = 21.56$, $p < .001$). Die Wahrscheinlichkeit, den Studiengang zu wechseln, ist hingegen zwischen beiden Gruppen nicht signifikant unterschiedlich ($T = 0.67$, $p > .400$).

In weiteren Analysen wird untersucht, welche Variablen sich als Prädiktoren für die unterschiedlichen Abbruchquoten eignen (bsp. Leistungsfähigkeit, familiärer Hintergrund) und ob sich die Abbruchquoten je nach gewähltem Hochschultyp unterscheiden. Antworten auf diese Fragen können als Hinweise darauf interpretiert werden, inwiefern es Absolventinnen und –Absolventen von AG und BG gelingt, die Studiengänge zu wählen, in denen sie erfolgreich sein können.

Donnerstag
12. März

11:15 Uhr

GA 03/142

Chair(s)

Kramer, Jochen

Braun, Edith; Brachem, Julia-Carolin

Anforderungsprofile des Arbeitsmarktes an Hochschulabsolvent(inn)en

Mittlerweile nehmen über fünfzig Prozent eines Jahrgangs ein Hochschulstudium auf (Orr et al. 2011), so dass Hochschulen einen Auftrag hinsichtlich der Ausbildung für den Arbeitsmarkt übernehmen. Daher erscheint es bedeutsam, die Anforderungen des Arbeitsmarktes an Hochschulabsolvent(inn)en zu kennen.

Derartige Anforderungen werden in diversen, größtenteils internationalen Studien, meist über Selbsteinschätzungen erworbener Kompetenzen, erfasst (Allen und van der Velden 2007; Fabian et al. 2013; Schomburg und Teichler 2012). Obgleich die ökonomischen Gründe solcher Erhebungen überzeugend sind, ist die Aussagekraft selbsteingeschätzter Kompetenzen eingeschränkt. Der Job Requirements Approach (JRA) versucht dieser Kritik insofern gerecht zu werden, dass nicht mehr eigene Kompetenzniveaus eingeschätzt werden, sondern Aussagen dazu gemacht werden, wie oft bestimmte berufliche Tätigkeiten verrichtet werden. Dabei variieren das Niveau und die Bereiche der abgefragten Tätigkeiten.

Es gibt eine Reihe von Studien, die die Methode des JRA verwenden: Generalized Work Activities Questionnaire (GWA) (O*NET 2012; Peterson et al. 2001; Toolsema 2003), UK Skills Survey (BMRB 2006; Felstead et al. 2007) oder Programme for the International Assessment of Adult Competencies (PIAAC) der OECD (OECD 2013a, b).

Bislang fehlte für den deutschsprachigen Raum eine systematische Analyse, was von Hochschulabsolvent(inn)en auf dem Arbeitsmarkt gefordert wird und ein entsprechendes Instrumentarium, das auch aktuelle Entwicklungen erfasst (z. B. Social Media, Akademisierung, Führungsaufgaben).

In diesem Beitrag sollen die psychometrischen Kennwerte eines Fragebogens vorgestellt werden, der generische, berufliche Tätigkeiten und Anforderungen von Hochschulabsolvent(inn)en auf dem Arbeitsmarkt erfasst. Weiterhin werden erste Ergebnisse unterschiedlicher Anforderungsprofile von Hochschulabsolvent(inn)en unterschiedlicher Studienrichtungen berichtet.

Um möglichst alle relevanten beruflichen Tätigkeiten und Anforderungen zu identifizieren, wurde eine umfassende Literaturrecherche vorgenommen. Zudem konnten die Ergebnisse qualitativer Arbeitgeberbefragungen unserer Kooperationspartner(innen) (Prinz et al. 2014) berücksichtigt werden. Darauf aufbauend wurden, angelehnt an den JRA, 49 Items entwickelt, die die Häufigkeit der ausgeführten beruflichen Tätigkeiten und der erfahrenen Anforderungen abfragen.

In einer Online Befragung mit einer deutschlandweit repräsentativen Stichprobe von Hochschulabsolvent(inn)en wurden drei Kohorten (Prüfungsjahrgänge 2001, 2005, 2009) des DZHW-Absolventenpanels (vormals HIS) befragt. Es haben sich 11.504 Hochschulabsolvent(inn)en beteiligt.

Um Vergleiche verschiedener Studienrichtungen vornehmen zu können, wird auf die Kategorisierung des Statistischen Bundesamtes zurückgegriffen: (1) Sprach- und Kulturwissenschaften (n=2.102); (2) Sport (n=53); (3) Rechts-, Wirtschafts- und Sozialwissenschaften (n=2.826); (4) Mathematik, Naturwissenschaften (n=1.926); (5) Humanmedizin/Gesundheitswissenschaften (n=570); (6) Veterinärmedizin (n=141); (7) Agrar-, Forst- und Ernährungswissenschaften (n=464); (8) Ingenieurwissenschaften (n=1.891); (9) Kunst, Kunstwissenschaft (n=281). Aufgrund der sehr großen Stichprobe werden nahezu alle Vergleiche signifikant. Daher werden in diesem Beitrag lediglich Differenzen von mehr als einem halben Punkt auf einer fünfstufigen Likertskala berichtet.

Aufgrund der Literaturrecherche wurden acht Tätigkeits- und Anforderungsbereiche angenommen. Ziel war möglichst viele der identifizierten Tätigkeiten und Anforderungen beizubehalten. In einer konfirmatorischen Faktorenanalyse konnten acht Bereiche generischer, beruflicher Tätigkeiten und Anforderungen zufriedenstellend bestätigt werden ($\chi^2(751) = 30.316$; CFI = .737; RMSEA = .073; SRMR = .081): (1) Planung (8 Items, $\alpha = .75$); (2) (Personal) Entwicklung (3 Items, $\alpha = .81$); (3) Beratung (5 Items, $\alpha = .72$); (4) Selbstständigkeit bei herausfordernden Aufgaben (9 Items, $\alpha = .79$); (5) Informationsaufbereitung (6 Items, $\alpha = .79$); (6) Umgang mit Zahlen (3 Items, $\alpha = .62$); (7) Interkulturelle Kommunikation (4 Items, $\alpha = .74$); (8) Körperliches Arbeiten (3 Items, $\alpha = .77$).

Ein erster explorativer Vergleich der Studienrichtungen zeigt, dass in den Bereichen Planung, Beratung, Selbstständigkeit bei herausfordernden Aufgaben und Interkulturelle Kommunikation kaum Variation zwischen den Studienrichtungen vorhanden ist und diese Bereiche somit als gleichbedeutend für alle Studienrichtungen betrachtet werden können, während sich in den Bereichen (Personal) Entwicklung, Informationsaufbereitung, Umgang mit Zahlen und Körperliches Arbeiten Unterschiede zwischen den Studienrichtungen zeigen (Alle Werte lagen den Gutachtern vor).

Die Abfrage von Häufigkeiten ausgeführter Tätigkeiten und erfahrener Anforderungen ist möglicherweise weniger verzerrt als Selbsteinschätzungen von Kompetenzniveaus. Neben der Bereitstellung eines Instruments für Absolvent(inn)enbefragungen, hoffen wir die Diskussion, welche Kompetenzen von Hochschulabsolvent(inn)en auf dem Arbeitsmarkt erwartet werden, auf eine empirische Basis stellen zu können.

Donnerstag
12. März

11:15 Uhr

GA 03/142

Chair(s)
Kramer, Jochen

3

Bergann, Susanne; Blüthmann, Irmela; Watermann, Rainer

Bildungsaufsteiger im Studium - Empirische Befunde zum Zusammenhang von akademischem Bildungshintergrund, Lernschwierigkeiten und dem Übergang in ein Masterstudium

Mit der Umstellung auf die neue Studienstruktur war die Erwartung verbunden, verstärkt Studienberechtigte aus nicht-akademischen Elternhäusern zur Aufnahme eines Studiums zu motivieren. Forschungsbefunde zeigen, dass dies bisher nur unzureichend gelungen ist (Maaz, 2006; Schindler & Reimer, 2010), wobei insbesondere beim Zugang zum Masterstudium soziale Disparitäten zu bestehen scheinen (z.B. Bargel & Bargel, 2010). Nur wenig Aufmerksamkeit wurde zudem der Frage gewidmet, ob sich soziale Unterschiede auch im Hinblick auf den Studienerfolg beobachten lassen. Der vorliegende Beitrag widmet sich der Frage, ob und inwieweit Zusammenhänge zwischen der Bildungsherkunft von Bachelorstudierenden, Lernschwierigkeiten und der Absicht, ein Masterstudium aufzunehmen, bestehen.

In Anlehnung an Boudon (1974) werden bei der Erklärung von Bildungsungleichheiten primäre und sekundäre Effekte der sozialen Herkunft unterschieden. Primäre Effekte bezeichnen alle Faktoren, die dazu führen, dass Personen unterschiedlicher sozialer Schichten unterschiedliche Kompetenzen erlangen. Solche interindividuellen Leistungsunterschiede werden auf eine unterschiedliche Ressourcenausstattung zurückgeführt (Boudon, 1974; Erikson & Jonsson, 1996). Nach Bourdieu (1982, 1992) spielt dabei nicht nur die Ausstattung mit ökonomischem und sozialem Kapital eine Rolle, sondern auch die Verfügbarkeit von kulturellem Kapital, welches auch durch die familiäre Sozialisation erworbene Fähig- und Fertigkeiten (z.B. Lernstrategien, wissenschaftlicher Duktus) umfasst. Sekundäre Effekte entstehen hingegen, wenn trotz gleichen Leistungsniveaus unterschiedliche, von der sozialen Herkunft abhängige Entscheidungen an Bildungsübergängen getroffen werden.

Vor dem Hintergrund dieser theoretischen Überlegungen sollen folgende Fragestellungen beantwortet werden: (1) Berichten Studierende aus nicht-akademischen Elternhäusern größere Lernschwierigkeiten und inwieweit sind diese auf ihre im Vergleich mit Akademikerkindern schlechteren Eingangsvoraussetzungen zurückzuführen? (2) Welche Rolle spielen Lernschwierigkeiten für die Entstehung von Gruppenunterschieden in Bezug auf die Studienabbruchneigung bzw. in Bezug auf die Absicht, ein Masterstudium aufzunehmen? (3) Lassen sich Hinweise auf sekundäre Effekte des Bildungshintergrunds beim Zugang zum Masterstudium identifizieren?

Die Fragestellungen wurden, basierend auf einem für Bachelorstudierende an der Freien Universität Berlin repräsentativen Datensatz aus dem Jahr 2008 und unter Berücksichtigung der geschachtelten Datenstruktur, mithilfe von Strukturgleichungsmodellen und Mediationsanalysen untersucht. Die Stichprobe umfasst N=2353 Studierende. Der Anteil sog. „Bildungsaufsteiger“, d.h. Studierender, deren beide Elternteile nicht studiert haben, beträgt rund 38 Prozent (n=883). Es wurden zwei verschiedene Formen von Lernschwierigkeiten (allgemeine Lernschwierigkeiten [z.B. kognitiv, Zeitmanagement] und Schwierigkeiten mit wissenschaftlichen Arbeits- und Kommunikationsformen [z.B. Referate, Verfassen schriftlicher Arbeiten]) sowie die Studienabbruchneigung und die Absicht, ein Masterstudium aufzunehmen, erfasst.

Die Ergebnisse der Analysen zeigen, dass Studierende aus nicht-akademischen Elternhäusern auch nach Kontrolle relevanter soziodemografischer Merkmale sowohl größere allgemeine Lernschwierigkeiten als auch größere Schwierigkeiten mit wissenschaftlichen Arbeits- und Kommunikationsformen aufweisen als Studierende aus einem akademischen Elternhaus. Während die Gruppenunterschiede in allgemeinen Lernschwierigkeiten fast vollständig über Unterschiede in den Zugangsvoraussetzungen (Abiturnote) vermittelt sind, lässt sich in Bezug auf Schwierigkeiten mit wissenschaftlichen Arbeits- und Kommunikationsformen auch nach Kontrolle der Zugangsvoraussetzungen und der Belastung durch Erwerbsarbeit ein signifikanter Effekt der Bildungsherkunft belegen. Des Weiteren zeigt sich, dass Studierende aus nicht-akademischen Elternhäusern eine höhere Studienabbruchneigung sowie eine geringere Intention aufweisen, ein Masterstudium aufzunehmen als Studierende aus einem akademischen Elternhaus. Der Effekt des akademischen Bildungshintergrunds auf die Studienabbruchneigung kann dabei zu 31% und der Effekt des akademischen Bildungshintergrunds auf die Absicht, ein Masterstudium aufzunehmen, zu 80% über Gruppenunterschiede in Lernschwierigkeiten erklärt werden. Auch nach Kontrolle relevanter soziodemografischer Merkmale, der Zugangsvoraussetzungen und der Lernschwierigkeiten ist jedoch ein signifikanter Effekt des akademischen Bildungshintergrunds auf die Absicht, ein Masterstudium aufzunehmen, zu beobachten.

„Bildungsaufsteiger“ weisen im Vergleich zu ihren Kommilitonen also größere (Lern-) Schwierigkeiten im Studium auf, insbesondere mit wissenschaftlichen Arbeits- und Kommunikationsformen, wobei dies mit einer im Vergleich erhöhten Studienabbruchneigung und einer geringeren Absicht, ein Masterstudium aufzunehmen, einhergeht. Die Ergebnisse weisen darauf hin, dass im Zuge der „sozialen Öffnung der Hochschulen“ der Unterstützungsbedarf dieser Studierendengruppe verstärkt berücksichtigt werden sollte.

Donnerstag
12. März

11:15 Uhr

GA 03/142

Chair(s)
Kramer, Jochen

Kretschmann, Julia; Vock, Miriam; Gronostaj, Anna; Schulze, Annelie; Pannier, Birthe

Wenn sich die Masterfrage stellt: Analyse sozialer Determinanten der Übergangsentention nach dem Bachelorstudium

Theoretischer Hintergrund: Das deutsche Bildungssystem ist durch zahlreiche Übergänge gekennzeichnet und das Vorhandensein sozialer Disparitäten bei der Bewältigung von Übergängen im Bildungsverlauf in Form von primären und sekundären Herkunftseffekten wurde bereits in einer Reihe von Studien und insbesondere für den Übergang von den Grundschule in die weiterführenden Schulen umfassend belegt (z.B. Maaz, Baumert, Gresch & McElvany, 2010; Stubbe, Bos & Euen, 2012). Dabei werden insbesondere sekundäre Herkunftseffekte in den Fokus gerückt - solche Effekte des familiären Hintergrundes, die nicht über Unterschiede in schulischen Leistungen vermittelt werden, sondern lediglich auf Unterschiede im Entscheidungsverhalten der beteiligten Akteure zurückzuführen sind (für einen Überblick s. Dumont, Maaz, Neumann & Becker, 2014). Im Zuge des Bologna-Prozesses wurde durch die landesweite Einführung des zweistufigen Bachelor-Master-Systems von Studienabschlüssen in Deutschland ein weiterer Übergang geschaffen, an welchem potentiell selektive Mechanismen greifen. Vor dem Hintergrund der bislang nur geringen und z.T. widersprüchlichen empirischen Befundlage besteht jedoch nachwievor Unsicherheit darüber, ob und in welcher Art dabei Einflüsse der Herkunft auf das Übergangsverhalten der Studierenden eine Rolle spielen (Falkenhagen, 2013; Gärtner & Himpele, 2009). Bisherige Befunde deuten zudem darauf hin, dass der Effekt des familiären Hintergrundes auf die Intention, ein weiterführendes Studium aufzunehmen, bedeutsam zwischen den Studiengängen variiert (Auspurg & Hinz, 2011; Kronfeld, 2013).

Fragestellung: Die vorliegende Untersuchung geht daher der Frage nach, ob und in welcher Weise die Intention zur Aufnahme eines Masterstudiums bei Bachelorstudierenden an deutschen Hochschuleinrichtungen mit Merkmalen des familiären Hintergrundes in Zusammenhang gebracht werden kann. Neben Migrationsstatus, sozio-ökonomischen Merkmalen der Studierenden und Angaben zur Studienfinanzierung sollen weitere Variablen Berücksichtigung finden, die für die Übergangsentcheidung von Bedeutung sein können: demographische Angaben, erbrachte Leistung im Abitur, Studienfach, familiäre Situation und die Zufriedenheit der Studierenden mit der Studienfachwahl und der Studiensituation.

Methode: Die wenigen bislang verfügbaren empirischen Studien, die den Übergang vom Bachelor- ins Master-System betrachten, gehen nur selten über die Anwendung deskriptiver Analysen hinaus. Werden methodisch anspruchsvollere Methoden angewandt, so mangelt es zudem an repräsentativen Daten aus Deutschland (Auspurg & Hinz, 2011). In der vorliegenden Untersuchung wird daher auf umfangreiche repräsentative Daten des Nationalen Bildungspanels (NEPS) zurückgegriffen. Befragt wurden ca. 14.000 Studierende im ersten Studienjahr im Wintersemester 2010/11. Die Bachelorstudierenden wurden dabei gebeten, Auskunft darüber zu geben, ob sie die Absicht haben, nach erreichtem Bachelorabschluss ein Masterstudium zu absolvieren. Unter Anwendung logistischer Regressionsanalysen und unter Kontrolle wichtiger Kovariaten (z.B. Alter, Geschlecht, erreichte Punktzahl im Abitur, Angaben zur Studienzufriedenheit) werden Effekte familiärer Hintergrundmerkmale und Einflüsse der Studienfinanzierung auf die berichtete Intention, ein Masterstudium aufzunehmen, geprüft. Studienfachspezifische Wirkgrößen werden zudem im Rahmen von Mehrebenen-Logitmodellen berücksichtigt, wobei das jeweilige Studienfach als Gruppierungsvariable dient.

Ergebnisse: Die Ergebnisse deuten darauf hin, dass familiäre Hintergrundmerkmale und die Art der Studienfinanzierung für die Intention, ein Masterstudium aufzunehmen, lediglich eine untergeordnete Rolle spielen, sowohl mit als auch ohne Kontrolle von Leistungsmerkmalen. Eine wesentliche Rolle spielt hingegen das Alter der Befragten. Jedoch bestätigen die Befunde der Multi-Level-Analysen die Abhängigkeit der Bedeutsamkeit sozialer Herkunftsmerkmale vom jeweiligen Studienfach.

Donnerstag
12. März

11:15 Uhr

GA 03/142

Chair(s)

Kramer, Jochen

Schumann, Stephan; Jüttler, Michael

Mit guten Wirtschaftskenntnissen zum Wirtschaftsstudium? Zum Zusammenhang zwischen ökonomischem Wissen und der Studienaspiration am Ende der Sekundarstufe II

Theoretischer Hintergrund: Die Vergabe von Hochschulberechtigungen an Lernende in der Sekundarstufe II ist in Ländern wie der Schweiz, Österreich 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). Gerade in Systemen mit berechtigungsorientierten Zugangsregelungen zur Hochschule (dazu Orr, 2013) kommt der Passung von Fähigkeiten und Studienanforderungen eine große Bedeutung zu. Dass diese lückenlose Studierfähigkeit für einen Teil der Absolventinnen und Absolventen nicht gegeben ist, zeigen die zwar tendenziell rückläufigen, jedoch weiterhin hohen Studienabbruchzahlen, aber auch Untersuchungen wie z.B. die Studie EVAMAR II für die Schweiz (Eberle et al., 2008). Dieser Befund wird durch das Argument entschärft, dass auch eine Selbstselektion der Studieninteressierten wirksam wird (Ramseier et al., 2004). Dabei spielen bei der Studienfachwahl neben Merkmalen wie Interessen, beruflichen Orientierungen, Karrierezielen, sozialer Herkunft etc. besonders auch die individuellen domänenspezifischen Fähigkeiten eine bedeutsame Rolle. Typischerweise rekurren einschlägige Studien zur Erfassung dieser Fähigkeiten auf die Schul- und Zeugnisnoten und selten auf mit deutlich größeren Aufwand erhobene Testergebnisse, die zudem nur für Kerndomänen wie z.B. Mathematik oder Englisch vorliegen (Bornkessel & Asdonk, 2011). Für das im vorliegenden Beitrag interessierende Wissen und Können in der Domäne „Wirtschaft“ (nachfolgend ökonomisches Wissen) liegen zumindest für den deutschsprachigen Raum keine Befunde zu diesem Zusammenhang vor. Dies überrascht auch angesichts der Tatsache, dass betriebswirtschaftliche bzw. wirtschaftswissenschaftliche Studiengänge zu den meistgewählten Studienfächern gehören.

Fragestellung: Vor diesem Hintergrund wird der Frage nachgegangen, inwieweit das ökonomische Wissen von schweizerischen Lernenden im letzten Jahr des Gymnasiums bzw. der Berufsmaturitätsschule mit der Studienintention korrespondiert. Dabei wird zwischen einer allgemeinen Studienaspiration und der Aspiration, Wirtschaft zu studieren, unterschieden. Erweitert wird die Fragestellung nach dem Nettoeffekt des ökonomischen Wissens, d.h. nach dem Effekt unter Kontrolle von Kovariaten.

Methode: Als Datengrundlage der Untersuchung dient die querschnittlich angelegte Studie OEKOMA mit einer für die Deutschschweiz repräsentativen Stichprobe von N = 2.328 Lernenden aus Gymnasien und Berufsmaturitätsschulen des Schuljahres 2010/11 (vgl. Schumann & Eberle, 2014). Damit die interessierenden Parameter möglichst zuverlässig geschätzt werden können und zuverlässige Vergleiche von interessierenden Subpopulationen möglich sind, wurden vier explizite Strata gezogen: (1) Gymnasialklassen mit Schwerpunktfach „Wirtschaft und Recht“, (2) Gymnasialklassen mit anderen Schwerpunktfächern, (3) kaufmännische Berufsmaturitätsklassen und (4) nicht-kaufmännische Berufsmaturitätsklassen. Das Instrumentarium setzte sich aus Tests und einem Fragebogen zusammen. Die Tests wurden im Multi-Matrix-Design eingesetzt. Zur Erfassung des ökonomischen Wissens und Könnens wurden sechs Testhälften gebildet (Eigenentwicklung, Schumann & Eberle, 2014). Das Testletdesigns zur Erhebung der Deutsch- und Mathematikleistung (jeweils aus Eberle et al., 2008) sowie der kognitiven Grundfähigkeit (Kurzversion des KFT 4–12R, Heller & Perleth, 2000) erfolgte nach derselben Logik bei jeweils vier Testheften. Die EAP/PV-Reliabilitäten der Tests liegen zwischen 0.75 und 0.83. Die Daten der Tests wurden auf der Basis der Item-Response-Theorie mittels des Programms „ConQuest“ (Wu et al. 2007) ausgewertet. Die Erfassung der Studienaspiration erfolgte mittels der Frage: „Beabsichtigen Sie zu studieren?“ (4-stufige Antwortskala: ganz sicher, wahrscheinlich, wahrscheinlich nicht, ganz sicher nicht).

Ergebnisse: Die Analysen dokumentieren zunächst einen starken Profileffekt: Lernende des gymnasialen Schwerpunktfachs Wirtschaft bzw. der kaufmännischen Berufsmaturität wählen deutlich häufiger einen wirtschaftswissenschaftlichen Studiengang. Die weiteren, innerhalb der vier Straten vorgenommenen Auswertungen zeigen, dass die Lernenden mit höheren ökonomischen Kompetenzen eher zu einem Hochschulstudium, vor allem zu einem wirtschaftswissenschaftlichen Studium, tendieren. Unter Kontrolle der Mathematik- und Deutschleistung sowie der kognitiven Grundfähigkeit „verschwindet“ der Effekt auf die allgemeine Studienaspiration, jedoch nicht auf die Intention, Wirtschaft zu studieren (signifikante Effekte innerhalb der Straten zw. 0.11 und 0.19). Inwiefern sich dieser Zusammenhang auch im faktischen Studienerfolg niederschlägt, ist Gegenstand einer aktuell startenden Follow-up-Studie.

Donnerstag
12. März

11:15 Uhr

GA 03/142

Chair(s)

Kramer, Jochen

Praetorius, Anna-Katharina; Drexler, Katharina; Nitsche, Sebastian; Janke, Stefan; Dickhäuser, Oliver; Dresel, Markus
Lehrermotivation und Unterrichtsqualität: Längsschnittuntersuchungen zu ihrem Zusammenhang

Die Motivation von Lehrkräften gilt als wichtiger Teilbereich der professionellen Kompetenz von Lehrkräften. Es wird angenommen, dass Merkmale der Lehrermotivation das Erleben und Verhalten von Lehrkräften beeinflussen (z. B. persönliches Wohlbefinden, Unterrichtsqualität). Zeitgleich existieren aber auch Hinweise darauf, dass Lehrermotivation durch persönliche und kontextuelle Variablen beeinflusst wird (s. z. B. das zyklische Modell der Lehrerselbstwirksamkeit von Tschannen-Moran, Woolfolk Hoy & Hoy, 1998). Unterrichtsqualität – als Kerngeschäft von Lehrkräften – kann dementsprechend sowohl als Prädiktor als auch als Kriterium von Lehrermotivation gesehen werden (s. auch Holzberger, Philipp & Kunter, 2013). Bislang existieren nur wenige empirische Studien, die Einflüsse und Wirkungen von Lehrermotivation simultan betrachten. In der vorliegenden Studie wird daher der Frage nach den längsschnittlichen Beziehungen zwischen Merkmalen der Lehrermotivation und Unterrichtsqualität nachgegangen.

Zur Beantwortung der Forschungsfrage wurden die Daten von zwei Messzeitpunkten (Abstand: 6 Monate) einer Längsschnittstudie verwendet. Insgesamt gingen in die Berechnungen Daten von 90 Mathematiklehrkräften und deren 1537 Schüler/-innen der sechsten Jahrgangsstufe ein. In Bezug auf die Lehrermotivation wurden die folgenden Konstrukte erhoben: Selbstwirksamkeit (9 Items; Schwarzer & Schmitz, 1999), Enthusiasmus für das Unterrichten (6 Items; Baumert et al., 2008), extrinsische Motivation (4 Items; Müller, 2010) und berufliche Zielorientierungen (4 Skalen mit 36 Items; Nitsche, Dickhäuser, Fasching & Dresel, 2011). Zur Erfassung von Unterrichtsqualität wurden Schülereinschätzungen zu den drei Basisdimensionen von Unterrichtsqualität (s. Klieme, Schümer & Knoll, 2001) eingesetzt. Klassenführung wurde mit 3 Items (Ramm et al., 2006), unterstützendes Unterrichtsklima mit 5 Items (Wild, 1999) und kognitive Aktivierung mit 6 Items (Ramm et al., 2006) erfasst. Die Daten wurden mittels Zwei-Ebenen-Cross-Lagged-Panel-Modellen analysiert. Für jede Basisdimension in Kombination mit jedem der untersuchten Lehrermotivationsmerkmale wurden getrennte Modelle berechnet.

Neben signifikanten autoregressiven Pfaden fanden sich in vier Modellen ausschließlich Effekte der Basisdimensionen von Unterrichtsqualität auf Merkmale der Lehrermotivation (unterstützendes Unterrichtsklima auf Lehrereнтуasmus: $\beta = .23$, $p = .01$; unterstützendes Unterrichtsklima auf Selbstwirksamkeit: $\beta = .24$, $p = .02$; unterstützendes Unterrichtsklima auf Annäherungsziele: $\beta = -.35$, $p = .004$; kognitive Aktivierung auf Annäherungsziele: $\beta = .24$, $p = .04$). Effekte von Merkmalen der Lehrermotivation auf Unterrichtsqualitätsdimensionen existierten in einem Fall (extrinsische Motivation auf kognitive Aktivierung: $\beta = .25$, $p = .02$). Effekte in beide Richtungen wurden in einem Fall gefunden (kognitive Aktivierung auf Selbstwirksamkeit: $\beta = .25$, $p = .04$; Selbstwirksamkeit auf kognitive Aktivierung: $\beta = .16$, $p = .05$).

Insgesamt zeigen die Ergebnisse, dass die Wirkmechanismen zwischen Merkmalen der Lehrermotivation und Unterrichtsqualität komplexer sind als ursprünglich angenommen. Basierend auf den Befunden kann vermutet werden, dass Effekte von Lehrermotivation auf Unterrichtsqualität in geringerem Ausmaß existieren als die umgekehrten Effekte. Dieses Ergebnis steht in Einklang mit den wenigen bislang existierenden Studien zu dieser Thematik (s. z. B. Holzberger et al., 2013). Die Befunde zeigen, dass in zukünftigen Studien in stärkerem Ausmaß die mit Lehrermotivation verbundenen kausalen Mechanismen untersucht werden sollten (s. auch Klassen, Tze, Betts, & Gordon, 2011).

Krüger, Marvin; Korneck, Friederike; Oettinghaus, Lars; Kunter, Mareike

Erfassung von Unterrichtsqualität in Unterrichtsminiaturen: Perspektiven und deren Zusammenhänge

Theoretischer Hintergrund: Für eine umfassende und aussagekräftige Erfassung von Unterrichtsqualität wurden zuletzt vermehrt mehrperspektivische Ansätze in Studien implementiert (z. B. Kunter & Baumert, 2006; Waldis et al., 2010). Das Spektrum reicht dabei von einfachen Schülerbefragungen bis zu aufwendigen (Video-)Beobachterratings. Es hat sich gezeigt, dass Unterrichtsqualität häufig perspektivenspezifisch ausgeprägt zu sein scheint, da nur in wenigen Kombinationen überhaupt substanzielle Korrelationen zwischen den unterschiedlichen Perspektiven berichtet werden (Clausen, 2002; De Jong & Westerhof, 2001; Kunter & Baumert, 2006; Waldis et al., 2010; Fauth et al., 2014). Daher wurde zuweilen eine entsprechend perspektivenspezifische Operationalisierung der unterschiedlichen Basisdimensionen von Unterrichtsqualität gefordert (Clausen, 2002).

Optionen, um eine Annäherung der Perspektiven zu erreichen, sind die zeitliche Begrenzung eines für die unterschiedlichen Perspektiven identischen Beobachtungszeitraums, die mehrfache Beobachtung sowie eine Vergrößerung der Beobachtungsstichprobe (ebd.). Die Studie Φ actio (Korneck, Oettinghaus & Kunter, in Arbeit), in deren Kontext dieser Beitrag entstanden ist, greift diese Forderungen in ihrem speziellen Setting auf, indem die Unterrichtsqualität nicht im Regelunterricht, sondern in zwei kurzen Unterrichtsminiaturen mit Hilfe von Schülerratings sowie einer großen Anzahl von beobachtenden Peers (d.h. anderen Lehramtsstudierenden bzw. Referendaren) und der Möglichkeit zur anschließenden Videokodierung erhoben wird.

Fragestellungen: In diesem Beitrag stehen die Schüler- und Peereinschätzungen im Fokus. Es ergeben sich die folgenden Fragestellungen: (1) Sind die drei Unterrichtsqualitätsmerkmale kognitive Aktivierung, konstruktive Unterstützung und effektive Klassenführung auch in Unterrichtsminiaturen mit Schüler- und Peerratings reliabel messbar? (2) Welche Zusammenhänge gibt es zwischen den Urteilen von Schülern und hospitierenden Peers?

Design und Methode: Die vorliegende Studie erfasst Unterrichtsqualität angehender Physiklehrkräfte durch eine Erhebung im Rahmen 12- bis 15-minütiger Unterrichtssequenzen, welche in einer studiengangübergreifenden fachmethodischen Lehrveranstaltung entwickelt werden und in deren Mittelpunkt jeweils ein einzelnes physikalisches Freihandexperiment steht. Die Lehrkräfte sind Studierende des Haupt- und Realschul- sowie des gymnasialen Lehramts im Hauptstudium und Lehrkräfte im ersten Modul des Vorbereitungsdienstes.

Die Stichprobe umfasst 172 Unterrichtsminiaturen, die von insgesamt 86 Lehrkräften (zwei pro Lehrkraft) gestaltet wurden. Es handelt sich um eine Vollerhebung der Physiklehramtsstudierenden am Studienort Frankfurt seit Wintersemester 2012/13. Jede Sequenz wird von durchschnittlich 10 Peers sowie den durchschnittlich 12 Schülern einer Klassenhälfte bewertet.

Die Erfassung der Unterrichtsqualität erfolgt mit Hilfe vierstufiger Likert-skalierten Fragebögen, deren Items von der COACTIV-Studie (Baumert et al., 2009) stammen, an Setting sowie die zweiminütige Erhebungsdauer adaptiert wurden und für Schüler und Peers explizit parallel formuliert sind. Das Instrument basiert auf der Einteilung in die drei Basisdimensionen kognitive Aktivierung, Klassenführung sowie konstruktive Unterstützung, die mit 7 bis 8 Items sowie jeweils 3 Subskalen vertreten sind. Die zwei verwendeten Testhefte enthalten jeweils 18 der 22 Items, von denen wiederum 14 als Ankeritems geführt sind.

Ergebnisse: Um die geschachtelte Struktur der Daten zu berücksichtigen, werden zur Auswertung Mehrebenenstrukturgleichungsmodelle in Mplus (Muthén & Muthén, 2010) verwendet. Als Indikatoren für die drei Basisdimensionen werden dabei auf beiden Ebenen ihre drei Subskalen in Form entsprechender WLE-Schätzer verwendet, die mit Hilfe von ConQuest (Wu et al., 2007) einem jeweils dreidimensionalen Partial Credit Model entnommen wurden. Die geschätzten Modelle zeigen entsprechend allgemeiner Cutoff-Kriterien angemessene Modellfits (Geiser, 2011).

Zur Bestimmung der Reliabilität sind in Mehrebenenanalysen insbesondere die Intraklassenkorrelationen geeignet (Wagner, 2008). In diesem Fall zeigen sie mit $ICC(2) > .70$ eine reliable Erfassung aller drei Unterrichtsqualitätsmerkmale an (Lüdtke et al., 2006).

Zwischen Schüler- und Peerperspektive können ferner für alle drei Merkmale signifikante substanzielle Korrelationen gefunden werden.

Diskussion: Neben dem erneuten erfolgreichen Einsatz von Schülerratings, zeigten sich die Peers als weitere geeignete Perspektive zur Erfassung von Unterrichtsqualität. Die hohe Konvergenz beider Perspektiven kann dabei vermutlich auf die hohe Standardisierung der zugrunde gelegten Unterrichtsminiaturen zurückgeführt werden.

Donnerstag
12. März

11:15 Uhr

GA 03/149

Chair(s)

Praetorius, Anna-Katharina

Lindl, Alfred

Professionalität im Unterrichtsfach Latein

Um für einen modernen altsprachlichen Unterricht eine fruchtbare Symbiose aus dem Erfahrungsschatz einer mehr als 2000-jährigen Vermittlungstradition und aktuellen Erkenntnissen der Unterrichtsforschung zu erreichen, ist eine eigenständige Evaluation des Schulfaches Latein unter gegenwärtigen Forschungsfragen unerlässlich, die jüngst die Person des Lehrers, deren Kompetenzen „im Sinne eines kohärenten Ensembles von Wissen und Können“ (Bromme, 1997, S. 187) zunehmend als Schlüsselvariablen für erfolgreichen Unterricht hervortreten, in den Fokus rücken. Während vor allem für das Fach Mathematik durch die COACTIV-Studie (Kunter et al., 2011) und für eine Auswahl an „gering strukturierten Domänen“ durch TEDS (Blömeke et al., 2011; 2013) bereits Operationalisierungen für das Expertenwissen einer Lehrkraft vorliegen sowie Zusammenhänge zwischen diesem und dem Unterrichtserfolg messbar gemacht werden können, stellt eine Konzeptualisierung und Erhebung des fachspezifischen Professionswissens von Lateinlehrkräften bisher ein Desiderat dar. Diese Forschungslücke greift die Fachgruppe Latein innerhalb des Regensburger FALKO-Projektes auf und stellt im vorliegenden Beitrag Überlegungen einer domänenspezifischen Modellierung des Professionswissens sowie den Testentwicklungsprozess für das Fach Latein vor.

In Anlehnung an die theoretische Fundierung der COACTIV-Studie (Krauss et al., 2011) und der Wissenstaxonomie nach Shulman (1986; 1987) wurde ein Professionswissensmodell für Lateinlehrkräfte entwickelt, welches sich in die zwei Hauptkomponenten Fachwissen (CK), das als vertieftes Hintergrundwissen über den Schulstoff operationalisiert wird, und fachdidaktisches Wissen (PCK), das die drei domänenspezifischen Subfacetten „Wissen über Erklären und Repräsentieren“, „Wissen über lateinbezogene Schülerkognitionen“ und „Wissen über das Potenzial von Materialien im Lateinunterricht“ umfasst, gliedert. Bei der Konzeptualisierung dieser Komponenten galt es, die fachspezifischen Besonderheiten adäquat zu berücksichtigen, da zum einen Latein als „Gymnasialfach par excellence“ (Westphalen, 1992, S. 60) a priori auf die Fachlehrergruppe und Anforderungen dieser Schulform beschränkt ist, zum anderen das dem Unterrichtsfach zugrunde liegende Multivalenzprinzip eine inhaltliche Reduktion erfordert, die sich auf die Kernfacetten des Lateinunterrichts als Sprachfach konzentriert. In Ermangelung wissenschaftlicher Vorarbeiten sowie nationaler Bildungsstandards und der fehlenden Erfassung im Gemeinsamen europäischen Referenzrahmen für Sprachen wurde bei der inhaltlichen Ausgestaltung einerseits auf bundesländerspezifische Lehrpläne, die einheitlichen Prüfungsanforderungen für das Abitur, andererseits auf die jüngst etablierten Lehrerbildungsstandards sowie ein Extrakt aus universitären Lehrerbildungscurricula zurückgegriffen, um ein angemessenes Niveau abschätzen und eine optimale Passung der Testitems realisieren zu können.

Darauf fußend wurde ein Papier-und-Bleistift-Test entwickelt, der analog in einen fachdidaktischen und fachwissenschaftlichen Abschnitt zerfällt und eine Mischung geschlossener wie offener Itemformate mit einer handlungsorientierten Ausrichtung bietet. Die Items wurden von Experten übereinstimmend als inhaltlich valide erachtet und deren Urteil von den hohen Akzeptanzwerten der Testpersonen bei Pilotierungsrunden bestärkt, indem diese auf einer vierstufigen Skala (1 = Trifft gar nicht zu., 4 = Trifft zu.) die Berufsrelevanz ($M = 3,53$; $\sigma = 0,38$), die Angemessenheit der Schwierigkeit der Items ($M = 3,44$; $\sigma = 0,40$) und deren Bedeutung für die Lehrerausbildung ($M = 3,11$; $\sigma = 0,75$) einschätzen sollten. Erste Ergebnisse von Trennschärfe- und Schwierigkeitsanalysen machten Adaptionen des Testentwurfs nötig, die unter interdependenter Berücksichtigung des theoretischen Konstruktes und psychometrischer Kriterien vollzogen wurden. Zudem erfolgte die Erstellung und Ausdifferenzierung eines Kodiermanuals, so dass eine Auswertung der Interraterreliabilität der offenen Items durch geschulte Rater zu akzeptablen Werten führt (Spearman's $\rho > .64$; $p < .001$). Trotz der thematischen Heterogenität der Testitems, die aus dem Spannungsverhältnis einer angemessenen Bearbeitungszeit von maximal 90 Minuten und einer adäquaten Erfassung der komplexen Professionswissensstruktur resultiert, gelingt es, in beiden Testpartien ein hinreichendes Maß der internen Konsistenz zu erzielen (PCK: Cronbach's $\alpha = .66$; CK: $\alpha = .68$), wobei beide Teile auf niedrigem Niveau miteinander korrelieren ($r = .49$; $p < .001$). Außerdem lassen sich sowohl im Fachwissen als auch im fachdidaktischem Wissen Unterschiede zwischen den Leistungen von Lehramtsstudierenden ($N = 35$) und Lehrkräften ($N = 135$) zugunsten der Berufstätigen nachweisen (PCK: Cohen's $d = .77$; CK: $d = .71$).

Donnerstag
12. März

11:15 Uhr

GA 03/149

Chair(s)

Praetorius, Anna-Katharina

Minkley, Nina; Rest, Miriam; Terstegen, Susanna; Kirchner, Wolfgang; Wolf, Oliver

Mehr Stress durch G8? Akute und chronische Stressbelastung von Abiturienten mit regulärer und verkürzter Gymnasialzeit

Seit ihrer Einführung löst die Schulzeitverkürzung am Gymnasium vom neunjährigen (G9) zum achtjährigen (G8) Abitur gesellschaftliche Debatten aus und steht besonders aufgrund überfrachteter Stundenpläne in der Kritik (Bölling 2008). Dadurch bedingt, werden für die G8-Schüler eine höhere Stressbelastung sowie gesundheitliche Beschwerden angenommen (Kühn et al. 2013).

Die Befunde systematischer Untersuchungen hierzu sind jedoch nicht eindeutig. So zeigt die Untersuchung von Milde-Busch et al. (2010) zwar, dass Gymnasiasten allgemein überdurchschnittlich viel Stress empfinden, jedoch besteht kein signifikanter Unterschied zwischen den G8- und G9-Schülern. Auch in Bezug auf die gesundheitliche Belastung (z.B. Kopfschmerzen) zeigt sich kein Unterschied.

In unserer Studie untersuchten wir daher die akute und langfristige Stressbelastung von G8- und G9-Abiturienten, die 2013 gleichzeitig ihr Abitur (NRW) absolvierten, hinsichtlich evtl. vorhandener Unterschiede. Die beiden Schülergruppen wurden zuvor in der Unter- und Mittelstufe getrennt und in der Oberstufe gemeinsam unterrichtet (Ministerium für Schule und Weiterbildung NRW 2009, 2010).

Als Indikatoren für die akute Stressbelastung im Schulalltag dienten die Cortisolkonzentration im Speichel und die subjektive Stresseinschätzung (Kirschbaum & Hellhammer 2000). Dazu gaben die Probanden an drei Tagen einer Schulwoche (Montag, Mittwoch, Freitag) an jeweils sechs verschiedenen Zeitpunkten (Aufwachen, 30 min. nach dem Aufwachen, Schulbeginn, jeweils nach der 2., 4. und 6. Stunde) Speichelproben ab, deren Cortisolkonzentration in unserem Labor gemessen wurde. Zeitgleich wurde die subjektive Stresseinschätzung mittels einer visuellen Analogskala erfasst (Folstein & Luria 1973). Anhand der Cortisolkonzentration im Haar wurde zudem die langfristige Stressbelastung (Karlén et al. 2011; van Uum et al. 2008) während der Phase der Abiturprüfungen und davor betrachtet.

Dazu wurden je 5 cm lange Haarsträhnen von den Probanden entnommen und halbiert, so dass die kopfhautnahen 2 cm die zweimonatige Prüfungsphase und die folgenden 2 cm die zweimonatige Vorprüfungsphase repräsentierten. Die Analyse der Cortisolkonzentration erfolgte schließlich am Lehrstuhl für Biopsychologie der TU Dresden (Kirschbaum et al. 2009).

Dabei zeigte sich, dass die Cortisolkonzentration im Speichel der beiden Schülerinnengruppen im Tagesverlauf sehr ähnlich ist: Beide Gruppen zeigen das typische diurnale Profil, mit der morgendlichen Aufwachreaktion und dem Absinken der Cortisolkonzentration im Tagesverlauf (Steptoe 2007). In der Tendenz zeigen die G8-Schülerinnen jedoch eine leicht höhere Cortisolkonzentration (ANOVA, Innersubjektfaktor ZEIT (6 Messzeitpunkte), Zwischensubjektfaktor SCHULBIOGRAPHIE (G8, G9) = Faktor ZEIT: $F(2,29; 43,51) = 39,61$; $p < 0,001$; $\eta^2 = 0,68$; Faktor SCHULBIOGRAPHIE: $F(1; 19) = 2,95$; $p = 0,10$; $h^2 = 0,13$). Dieser leichte Unterschied zeigt sich insbesondere bei den Messungen bis zum Schulbeginn – während des Unterrichtes sind die Konzentrationen quasi gleich.

Ein solch ähnliches hormonelles Stressprofil spricht für einen Gewöhnungseffekt: Zum Zeitpunkt der Untersuchung wurden die Schülerinnen bereits seit zwei Jahren gemeinsam unterrichtet und waren dementsprechend denselben Stressoren ausgesetzt, so dass evtl. ehemals vorhandene Unterschiede in der Stressbelastung u.U. nicht mehr vorhanden sind.

In Bezug auf das subjektive Stressempfinden zeigen beide Schülerinnengruppen ein nahezu identisches Profil, mit einem Anstieg vom Aufwachen bis zum Schulbeginn und einem relativ gleichbleibenden Empfinden während des Unterrichtes (Faktor ZEIT: $F(2,32; 41,78) = 3,77$; $p < 0,05$; $h^2 = 0,17$; Faktor SCHULBIOGRAPHIE: n.s.).

In Bezug auf die langfristige Stressbelastung zeigt sich ebenfalls kein Unterschied zwischen den beiden Gruppen: Bei beiden besteht ein signifikanter Anstieg der Cortisolkonzentration im Haar von der Vorprüfungs- zur Prüfungsphase (ANOVA, Innersubjektfaktoren ZEIT (Vorprüfungsphase, Prüfungsphase), Zwischensubjektfaktor SCHULBIOGRAPHIE (G8, G9) = Faktor ZEIT: $F(1; 41) = 14,19$; $p < 0,001$; $h^2 = 0,26$; Faktor SCHULBIOGRAPHIE : n.s.). Dies deutet darauf hin, dass die Prüfungsphase einen langfristigen Stressor darstellt.

Die Ergebnisse unserer Untersuchung widersprechen somit der Befürchtung vieler Eltern, dass mit der verkürzten Schulzeit ein höheres Belastungserleben einhergeht (Kühn et al. 2013).

Donnerstag
12. März

11:15 Uhr

GB 03/49

Chair(s)

Minkley, Nina

Hübner, Nicolas; Wagner, Wolfgang; Kramer, Jochen; Trautwein, Ulrich

Die G8-Reform in Baden-Württemberg – Vergleichende Analysen zum Wohlbefinden, zur Inanspruchnahme von Nachhilfe und zum Freizeitverhalten

Theoretischer Hintergrund: Seit Beginn des 21. Jahrhunderts lässt sich in der deutschen Bildungslandschaft ein Trend zur systematischen Verkürzung der gymnasialen Schulzeit beobachten (Trautwein & Neumann, 2008). Mit Ausnahme der Bundesländer Sachsen und Thüringen, in denen das in der DDR übliche Abitur nach acht Jahren am Gymnasium nach der Wende beibehalten wurde, hatten zunächst auch die ostdeutschen Bundesländern das G9 eingeführt. In Sachsen-Anhalt lässt sich (seit 2003) eine vergleichsweise frühzeitige Reform des Bildungswesens zurück zu G8 konstatieren. Einzig Brandenburg und Mecklenburg-Vorpommern finden sich in der großen westdeutschen G8-Reformwelle wieder, die in den Jahren 2004 bis 2006 stattfand. In diesem Zeitraum wurde die gymnasiale Schulzeit in insgesamt sieben westdeutschen Bundesländern von neun auf acht Jahre reduziert. Neben den aktuell fünfzehn Ländern in der Bundesrepublik Deutschland, die G8-Modelle flächendeckend eingeführt haben, hielt Rheinland-Pfalz bisher am G9-Modell fest und führte G8 ausschließlich im Rahmen eines Modellversuchs ein (KMK, 2014).

Trotz der zunächst politisch überwiegend einvernehmlichen und großflächigen Umstrukturierungsmaßnahmen der gymnasialen Bildungsgänge fast aller Bundesländer hin zu G8, wird die Diskussion um Schulzeitverkürzungen in jüngster Vergangenheit zunehmend kritisch geführt. Vor diesem Hintergrund sind auch aktuelle Aufweichungstendenzen von G8-Strukturen in verschiedenen Bundesländern zu interpretieren. So kehrt beispielsweise Niedersachsen als bisher einziges Bundesland mit dem Schuljahr 2015/2016 flächendeckend zu G9 zurück (KMK, 2014). Erstaunlich ist, dass der Forschungsstand zu den intendierten und nicht-intendierten Reformeffekten der Schulzeitverkürzung derzeit als vergleichsweise unbefriedigend bezeichnet werden muss (vgl. Kühn, van Ackeren, Bellenberg, Reintjes & im Brahm, 2013).

Fragestellung: Der Beitrag behandelt die Frage, ob sich Schülerinnen und Schüler aus G8- und G9-Jahrgängen in Baden-Württemberg im Hinblick auf die folgenden Bereiche unterscheiden: Belastung und Gesundheit, Teilnahme an Nachhilfe und Freizeitverhalten am Ende der Sekundarstufe II. Diesbezüglich wurden auf Basis von Daten des Nationalen Bildungspanels (NEPS Zusatzstudie Baden-Württemberg; Blossfeld, Roszbach & Maurice, 2011; Wagner et al., 2011) zwei Abschlussjahrgänge eingehend miteinander verglichen (Ebenfalls sollen die Daten einer weiteren reinen G8-Kohorte aus dem Abschlussjahr 2013 berücksichtigt werden (voraussichtliche Veröffentlichung: Dezember 2014). Im Hinblick auf die Konstrukte „Belastung“ und „Gesundheit“ wurden ebenfalls Geschlechterunterschiede in Abhängigkeit der jeweiligen Kohortenzugehörigkeit näher untersucht.

Methode: Bei den Kohorten handelt es sich zum einen um den G9-Abschlussjahrgang aus dem Jahr 2011 und zum anderen um den aus G8- und G9-Schülerinnen und Schülern bestehenden Abschlussjahrgang aus dem Jahr 2012 (jeweils etwa 1200 Schülerinnen und Schüler aus insg. 48 Schulen). Die Schülerinnen und Schüler der zwei Jahrgänge wurden unter Verwendung von MIMIC-Modellen (Jöreskog & Goldberger, 1975) und multiplen logistischen Regressionsmodellen bezüglich der verschiedenen Konstrukte miteinander verglichen. Latente Variablenmodelle ermöglichen in diesem Zusammenhang einerseits die messfehlerfreie Betrachtung von Zusammenhängen und explizite Berücksichtigung der Ordinalität der Indikatoren der latenten Variablen sowie andererseits die umfangreiche Testung von Messinvarianz (Meredith, 1993; Millsap & Yun-Tein, 2004). Die hierarchische Struktur der Daten wurde durch Verwendung von robusten Schätzverfahren berücksichtigt. Fehlende Werte wurden mittels Full-Information-Maximum-Likelihood-(FIML)Verfahren berücksichtigt.

Ergebnisse: Es zeigten sich beim Belastungserleben und gesundheitlichen Wohlbefinden mittlere Effekte (Cohen, 1988) zuungunsten der untersuchten G8-Kohorte sowie deutliche Geschlechterunterschiede: Mädchen wiesen unabhängig von der jeweiligen Kohorte eine höhere Belastung und ein geringeres gesundheitliches Wohlbefinden auf als Jungen. Im Hinblick auf die Teilnahme an außerschulischer Nachhilfe und hinsichtlich des Freizeitverhaltens liegen uneinheitliche Befunde vor. Die Bedeutung der Ergebnisse und mögliche Implikationen werden diskutiert.

Donnerstag
12. März

11:15 Uhr

GB 03/49

Chair(s)

Minkley, Nina

Kärner, Tobias; Warwas, Julia; Sembill, Detlef

Analyse von Person-Situation-Interaktionen im Berufsschulunterricht am Beispiel schülerseitiger Belastungsreaktionen.

Schulischer Unterricht steht mit unterschiedlichen Stressoren wie beispielsweise zeitlichem Druck, Erfolgsdruck, Unsicherheit oder empfundener Stoffschwierigkeit in Verbindung (z. B. Kouzma & Kennedy 2004; Böhm-Kasper & Weishaupt 2002). Individuelle Belastungsreaktionen sind hierbei als mehrdimensional anzusehen, da sowohl physiologische als auch psychische Dispositionen und zeitlich variierende Zustände involviert sind. So ist beispielsweise eine Erhöhung der individuellen Cortisol-Baseline mit chronischer Belastung assoziiert (Takahashi et al. 2005), das sympathische Nervensystem dominiert unter situativer Stresseinwirkung gegenüber dem parasympathischen Nervensystem (Baert et al. 2012) und individuelle Bewältigungsressourcen sind hinsichtlich der Bewertung potenzieller Stressoren von großer Bedeutung (Goetz et al. 2008). Da Belastungsreaktionen aus einer Interaktion zwischen Person und Situation resultieren, müssen bei ihrer empirischen Analyse Interaktionseffekte adäquat berücksichtigt werden (Lazarus 1966 et passim). Entsprechende forschungsmethodische Ansätze für den schulischen Kontext sind u. a. schon in der Aptitude-Treatment-Interaction-Theory (Snow 1989) zu finden und werden im vorliegenden Beitrag um einen prozess- und mehrbenenanalytischen Ansatz erweitert.

Erste Befunde zu physiologische und psychologischen Effekten auf das zeitbezogene Stresserleben von Schülern wurden von uns bereits berichtet (Kärner, Golyszny & Sembill 2014). Darauf aufbauend liegt unser aktuelles Forschungsanliegen in der Analyse von Effekten relativ zeitstabiler Eigenschaften auf den Zusammenhang zwischen situationalen Unterrichtsmerkmalen und dem subjektiven Belastungserleben von Berufsschüler(innen). Hiermit soll der Heterogenität individueller Dispositionen und Prozesse entsprechend Rechnung getragen werden.

Der Analyse liegen Daten von insgesamt 27 Schüler(inne)n einer bayerischen Berufsschule zugrunde ($M = 19.3$, $SD = 3.9$ Jahre), welche über einen Zeitraum von 3 Wochen (9×45 Min.) begleitet wurden. Die Videostudie wurde als Produkt-Prozess-Design realisiert, wobei zunächst dispositionelle Merkmale (Traits) erhoben wurden, bevor situationsabhängige States in 10-minütigen Zeitintervallen mittels der Continuous-State-Sampling-Methode während des Unterrichtsgeschehens erfasst wurden (vgl. Sembill, Seifried, & Dreyer 2008). Hierbei ergeben sich für jeden Lernenden im Durchschnitt ca. 28 State-Messungen. Auf Basis der videografierten Unterrichtseinheiten wurden anschließend situationsspezifische Unterrichtsmerkmale auf Basis 15-sekündiger Zeitintervalle kodiert, welche sich auf Schülerarbeitsphasen (z. B. Aufgaben bearbeiten, Recherchieren) beziehen und für die nachfolgenden Analysen auf 10-minütige Zeitintervalle zusammengefasst wurden. Somit konnte auf insgesamt 753 synchrone State-Messungen und aggregierte Beobachtungsintervalle zurückgegriffen werden.

In den korrelativen Analysen zeigt sich, dass das situationsabhängige Belastungserleben der Lernenden signifikant negativ mit dem schulischen Fähigkeitsselbstkonzept der Schüler sowie mit deren subjektivem Erleben eigener Bewältigungsfähigkeit im Unterrichtsprozess zusammenhängt. Überzufällige positive Korrelationen zeigen sich zwischen dem Belastungserleben und der individuellen Cortisol-Baseline, Phasen schülerzentrierten Arbeitens sowie sympathischer und parasympathischer Aktivität (LF/HF-ratio). Mittels SPSS® MIXED wurde ein 2-Ebenen-Modell mit autoregressiver Kovarianzstruktur geschätzt, welches die in den Individuen geschachtelten, wiederholten State-Messungen sowie Interaktionseffekte berücksichtigt. Entsprechend der korrelativen Analysen zeigen sich auch im Mehrebenenmodell signifikante Effekte dispositioneller und aktueller Bewältigungsressourcen (Fähigkeitsselbstkonzept, Bewältigungserleben) sowie der sympathovagalen Balance (LF/HF-ratio) auf das subjektive Belastungserleben. Weiterhin zeigen sich tendenzielle Effekte sowohl des Fähigkeitsselbstkonzepts ($p = .098$) als auch der Cortisol-Baseline ($p = .097$) auf den Zusammenhang zwischen Belastungserleben und objektiven Umgebungsmerkmalen in Form von Schülerarbeitsphasen. Die sich andeutenden Moderatoreffekte sind dergestalt zu interpretieren, dass ein Ansteigen des subjektiven Belastungserlebens mit zunehmendem Anteil schülerzentrierter Arbeitsphasen sowohl vom individuellen Niveau des schulischen Fähigkeitsselbstkonzepts als auch von der individuellen Ausprägung der Cortisol-Baseline abhängt.

Die Befunde unterstreichen einmal mehr, dass unterrichtsbezogene Stressoren keine homogenen Belastungsreaktionen der Lernenden bedingen, sondern dass diese von individuellen Schülervoraussetzungen mit beeinflusst werden (vgl. Sembill 2008; Sembill & Zilch 2010; Sembill 2012).

Donnerstag
12. März

11:15 Uhr

GB 03/49

Chair(s)
Minkley, Nina

Crede, Julia; Wirthwein, Linda; Steinmayr, Ricarda

Zum reziproken Zusammenhang zwischen Schulleistung, Leistungsängstlichkeit und subjektivem Wohlbefinden bei Jugendlichen

Als relevante Determinanten des Wohlbefindens Jugendlicher werden in letzter Zeit zunehmend diverse schulische Variablen untersucht (Gilman & Huebner, 2006; Proctor, Linley & Maltby, 2010). So konnten Kirkcaldy, Furnham und Siefen (2004; siehe auch Suldo, Shaffer, & Riley, 2008) Zusammenhänge zwischen schulischer Leistung und subjektivem Wohlbefinden aufzeigen. Ungeklärt ist jedoch die Frage der Wirkrichtung, obwohl Mortimore (1991) oder Samdal, Wold und Bronis (1999) von einer wechselseitigen Beeinflussung zwischen Schulleistung und Wohlbefinden ausgehen. Eine weitere für schulische Emotionen relevante Variable stellt die Leistungsängstlichkeit dar (Pekrun, Goetz, Titz & Perry, 2002; Ringeisen & Buchwald, 2010), die jedoch als Determinante subjektiven Wohlbefindens bislang noch nicht untersucht wurde. Auch die Frage nach der Wirkrichtung dieses Zusammenhangs ist ungeklärt. Aufgrund der Annahmen des Transaktionsmodells im Kontext von Leistungsängstlichkeit (vgl. Ringeisen & Buchwald, 2010) lässt sich jedoch vermuten, dass sich vor allem die Leistungsängstlichkeit negativ auf das subjektive Wohlbefinden auswirkt. Das Ziel der vorliegenden Studie ist es zu untersuchen, inwiefern sich die berücksichtigten Variablen (Schulleistung, Leistungsängstlichkeit, subjektives Wohlbefinden) über die Zeit gegenseitig beeinflussen. Eine Stichprobe von N = 290 Schülern/innen der 11. Klasse von drei Gymnasien (n = 138 weiblich; Alter: M = 16.54 Jahre, SD = .57) machten im Abstand von sechs Monaten Angaben zur Emotionalitäts- und Besorgniskomponenten der Leistungsängstlichkeit (Hodapp, 1991) sowie zur emotionalen und kognitiven Komponenten des subjektiven Wohlbefindens (Dalbert, 2003). Die durchschnittliche Zeugnisnote diente als Schulleistungskriterium. Mittels SEM cross lagged-Modellen wurden die reziproken Beziehungen zwischen Schulnoten, Leistungsängstlichkeit und subjektivem Wohlbefinden analysiert. Entsprechende Modelle bieten sich an, um die Veränderung einer Variablen über die Zeit auf eine andere Variable zurückzuführen. Der Modellfit war zufriedenstellend. Die Ergebnisse zeigen, dass die Besorgniskomponente der Leistungsängstlichkeit Veränderungen in der kognitiven Komponente des subjektiven Wohlbefindens (Lebenszufriedenheit), aber nicht in der affektiven Komponente des subjektiven Wohlbefindens von Jugendlichen vorhersagte. Des Weiteren sagte die Besorgniskomponente Veränderungen der Schulleistung vorher. Die Schulleistung sagte positiv Veränderungen in der kognitiven Komponente der Lebenszufriedenheit, aber nicht in der affektiven Komponente des subjektiven Wohlbefindens vorher. Effekte des subjektiven Wohlbefindens auf die Emotionalitäts- oder Besorgniskomponente der Leistungsängstlichkeit wurden nicht gefunden. Weitere Forschungsperspektiven und die unterschiedlichen prädiktiven Qualitäten von Leistungsängstlichkeit und Schulleistung werden hinsichtlich ihrer Auswirkungen auf das subjektive Wohlbefinden Jugendlicher und ihrer praktischen Bedeutsamkeit diskutiert.

Donnerstag
12. März

11:15 Uhr

GB 03/49

Chair(s)
Minkley, Nina

Dörsam, Michael; Lauber, Verena

The Faster, the Worse? Evidence From a German High School Reform

Reducing the time spent in school has some obvious advantages: Individuals benefit from a decrease in the opportunity costs of schooling due to a reduction in the expected school leaving age, while at the same time fiscal costs are reduced. However, these gains may be outweighed by a decrease in average labor productivity caused by the reduced investment in educational production. Thus, to design an optimal educational policy it is crucial to understand to what extent shortening the time spent in school reduces effective human capital of high school graduates. To date, the empirical evidence has provided mixed results (see Morin, 2013; Patai et al, 2010; Pischke, 2007). A recent high school reform in Germany that reduced the duration of secondary schooling from nine to eight years while keeping the curriculum constant may shed new light on this important question.

Striving to provide first causal evidence on the effects of this reform on academic outcomes at the tertiary level, we exploit a unique data set on university students, collected at the University of Konstanz. In total, the data set comprises more than 200.000 examination results of all students enrolled at the university since the winter term 2007. The fact that the reform was not implemented in every German state at the same time allows us to apply a difference-in-differences identification strategy to isolate the reform effect from confounding factors varying across federal states and over time. More precisely, the reform was conducted in the federal state of Bavaria one year earlier than in Baden-Wuerttemberg. As a consequence, the first eight-year high school students (G8 students) from Bavaria were already enrolled at the University of Konstanz while the majority of students still had a nine-year high school degree (G9 students). Other German states, such as Hesse, North Rhine-Westphalia and Rhineland-Palatinate, conducted the reform even later, providing us with an additional control group. Furthermore, the reform was only implemented in academic high schools ("Gymnasien"), allowing us to apply a within-state difference-in-differences approach with students from other types of high schools in Baden-Wuerttemberg as control group.

We find that G8 double-cohort students, i.e., the first G8 students who graduated high school simultaneously with the last G9 students of a particular state, perform at the beginning of their studies significantly worse in terms of academic achievement and failure rates relative to their G9 counterparts of the same cohort. Further, a significantly lower fraction of G8 double-cohort students' obtained the best possible grade. However, the G8 double-cohort students caught up quickly and the negative effects already vanish after the first semester. Considering the first pure G8 cohorts, i.e., the cohorts which graduated after the double-cohorts, no significant performance differences are found right from the first semester. Robustness checks with respect to maturity and age do not alter these findings.

Our study is the first one that investigates the effects of the German high school reform on academic outcomes at the tertiary level. However, other studies on this reform with a different focus, for example on students' high school grades, or students' stress level and health, were conducted and came to similar conclusions (see Anger et al, 2014; Blumentritt et al, 2014). Thus, although there is a lot of resistance in Germany against the reform, especially among parents, no objective arguments can be found so far. Instead, the empirical evidence indicates that there was indeed scope for a more efficient high school system, and thus, major public cost savings.

Donnerstag
12. März

11:15 Uhr

GB 03/49

Chair(s)
Minkley, Nina

Lüdtke, Oliver

Methodische Herausforderungen bei der Durchführung und Auswertung von Large-Scale Assessment Studien

Large-Scale Assessments im Bildungsbereich, wie z.B. große Schulleistungsstudien, sind populationsbeschreibende Untersuchungen, in denen die Heterogenität eines gesamten Schülerjahrgangs hinsichtlich verschiedener Merkmale (z.B. Kompetenzen, Fachinteresse) abgebildet werden soll. Dies ist mit besonderen methodischen Herausforderungen verbunden, da naturgemäß sehr unterschiedliche Schülerinnen und Schüler teilnehmen, die sich sowohl hinsichtlich ihres sozialen Hintergrunds, ihren Fähigkeiten, aber auch ihrer Testmotivation teilweise massiv unterscheiden. Um diese Heterogenität adäquat abbilden zu können, bedarf es ausgefeilter Untersuchungsinstrumente und geeigneter statistischer Auswertungsverfahren. Das vorliegende Symposium widmet sich den Herausforderungen, die mit der Durchführung und Auswertung von Large-Scale Assessments verbunden sind. Die Beiträge wählen unterschiedliche methodische Zugänge (experimentelle Designs, Simulationen, Analyse von realen Datensätzen, mathematisch-analytische Betrachtungen) und sind in verschiedenen Disziplinen verortet (Statistik, Psychologie, Erziehungswissenschaft). Gemeinsames Ziel der Arbeiten ist es, eine valide Erfassung von schulischen Kompetenzen sicherzustellen.

Der erste Beitrag (Hacker, Goldhammer & Kröhne) befasst sich mit Effekten des Zeitverhaltens bei der Bearbeitung von Lesekompetenztests. Da Schüler sich teilweise stark in ihrer Fähigkeit unterscheiden, sich die zur Verfügung stehende Testzeit effektiv einzuteilen, können Unterschiede im Zeitmanagement mit der gezeigten Testleistung konfundiert sein. In einem experimentellen Design werden verschiedene Ansätze der Rückmeldung der verbleibenden Testzeit während der Testbearbeitung miteinander verglichen.

Ein weiterer Beitrag (Robitzsch & Lüdtke) greift die Problematik nicht bearbeiteter Items in Leistungstests aus psychometrischer Perspektive auf. Nicht bearbeitete Items können in Testmodellen als falsch bewertet oder wie fehlende Daten behandelt werden. Es wird ein Ansatz zur Skalierung von Leistungstests vorgestellt, der diese beiden Vorgehensweisen als Spezialfälle enthält und es ermöglicht abzuschätzen, wie sich die Behandlung der fehlenden Antworten auf die Interpretation der Testwerte, auch in Hinblick auf vergleichende Länderreihungen, auswirkt.

Ein dritter Beitrag (Ünlü) gibt einen konzeptionellen Rundblick über eine Auswahl an methodisch-heterogenen Forschungsergebnissen im ZIB und widmet sich den Grundlagen des in PISA eingesetzten psychometrischen Skalierungsverfahrens. Im Mittelpunkt stehen computationale Aspekte der verwendeten Algorithmen und Reproduzierbarkeit der Ergebnisse, der Umgang mit lokalen Abhängigkeiten im Testdesign sowie die Rolle des Hintergrundmodells bei der Erzeugung der Plausible Values als den Schülerleistungswerten. Alternative Ansätze zur Optimierung der Auswertungsmethoden werden diskutiert.

Im vierten Beitrag (Nagy, Lüdtke, Köller, Heine & Mang) wird über die methodischen Herausforderungen bei der Skalierung der Längsschnittstudie PISA Plus berichtet, in der das deutsche PISA-Sample ein Jahr später erneut in den Bereichen Mathematik, Naturwissenschaften und Lesen getestet wurde. Die Bedeutung von Item-Positionseffekten auf die Abschätzung der Leistungszuwächse wird diskutiert und verschiedene IRT-Modelle verglichen, die eine Berücksichtigung von Positionseffekten ermöglichen.

Donnerstag
12. März

11:15 Uhr

HGB 20

Chair(s)

Goldhammer, Frank
Lüdtke, Oliver
Ünlü, Ali**Methodische Herausforderungen bei der Durchführung und Auswertung von Large-Scale Assessment Studien**

Hacker, Miriam; Goldhammer, Frank; Kröhne, Ulf

Heterogenität im Zeitverhalten bei der Bearbeitung von Kompetenztests in LSA-Settings

Hintergrund und Fragestellung: In traditionellen Large-Scale Assessments (LSAs), z.B. PISA, bearbeiten die TestteilnehmerInnen Kompetenztests im Rahmen eines Zeitlimits auf Testebene. Wie die TestteilnehmerInnen die zur Verfügung stehende Zeit für die Aufgabenbearbeitung einsetzen, bleibt ihnen überlassen und hängt von Personen- sowie Situationsmerkmalen ab. Sie können unter Annahme einer motivierten Testbearbeitung (high-stake) bspw. versuchen, die Zeit gleichmäßig zu verteilen, oder sie bearbeiten die Aufgaben so, als gäbe es keine Zeitbegrenzung. Im Bereich von low-stake LSAs kann nicht durchgehend von einer hohen Motivation zur Testbearbeitung ausgegangen werden (Wise & DeMars, 2005). Häufig werden aufgrund mangelnder Konsequenzen für den Testteilnehmer Aufgaben übersprungen oder die Teilnehmer zeigen Rateverhalten. Daraus folgen Antwortmuster, die sich hinsichtlich der individuellen Rate fehlender Werte in Form nicht erreichter Items wie auch in der Rate schnell beantworteter bzw. erratener Items unterscheiden können. Somit sind die Antwortmuster als Indikator der interessierenden Kompetenzausprägung zwischen Personen nicht mehr ohne weiteres vergleichbar. Dieses Problem kann bei der Kompetenzmodellierung statistisch behandelt oder durch geeignete Testprozeduren von vornherein vermieden werden (Goldhammer & Kröhne, 2014). Die vorliegende Studie verfolgt primär letzteren Ansatz, indem durch die Art und Weise der Itemadministrierung die Heterogenität im Zeitverhalten verringert bzw. das Zeitmanagement der Testpersonen durch Zeitfeedback und Zeitlimits auf Aufgabenebene unterstützt wird (Walczyk et al., 1999). Im Hinblick auf low-stake Assessments soll untersucht werden, ob die verschiedenen Arten der Unterstützung des Zeitmanagements einen Einfluss auf die Testmotivation bzw. das Rateverhalten haben.

Methode: In der Studie ReTiCo (Response Time Control) bearbeiteten SchülerInnen am Computer parallele Lesekompetenztests unter verschiedenen Administrierungsbedingungen. Neben einer Kontrollbedingung mit Zeitbegrenzung auf Testebene wurde eine Bedingung mit Feedback über die verbleibende empfohlene Zeit auf Aufgabenebene sowie Bedingungen mit unterschiedlichen Zeitbegrenzungen auf Aufgabenebene realisiert. Die TestteilnehmerInnen bearbeiteten 15 Aufgaben zur Lesekompetenz in jeder der Bedingungen. Die SchülerInnen beantworteten außerdem Fragen zu testbezogener Motivation sowie Persönlichkeitsdispositionen (z.B. Selbstkonzept) und bearbeiteten zusätzliche Tests aus dem Bereich Lesen (Worterkennung, Satzverifikation, Lesegeschwindigkeit). An der Studie nahmen 1065 SchülerInnen aus berufsbildenden Schulen teil (Alter M=20.59 Jahre; 66.01 % männlich). Die Datenerhebung wurde im Sommer 2014 abgeschlossen.

Analysen/ erste Ergebnisse: Die vermutete Heterogenität des Zeitverhaltens in der Kontrollbedingung wird anhand der Reaktionszeiten untersucht. Es wird erwartet, dass in der Kontrollbedingung stärkere Unterschiede in Bearbeitungszeiten auftreten als in den Feedback- und Zeitlimitbedingungen (auch bei Bedingung auf das gezeigte Fähigkeitsniveau). In ersten Analysen der Antwortdaten wurde zunächst die Reliabilität (interne Konsistenz) der Lesekompetenztests ermittelt, die für Kontroll- und Feedbackbedingung ausreichend hoch waren, für die Zeitlimitbedingungen jedoch niedriger ausfielen. Die gezeigte Lesekompetenzleistung variierte zwischen den experimentellen Bedingungen. Erwartungsgemäß fiel die Leistung in den Bedingungen mit milden Zeitlimits schwächer als in der Kontrollbedingung aus, während bei strengem Zeitlimit die gezeigte Leistung noch weiter absank. Im Gegensatz dazu führte das Feedback über die verbleibende empfohlene Zeit auf Aufgabenebene zu einem signifikant höheren Leistungsniveau. In Korrelationsanalysen auf Basis der Testwerte zeigt sich erwartungsgemäß, dass die Feedback- und Zeitlimitbedingungen untereinander höher korrelieren als mit der jeweiligen Kontrollbedingung. Der Zusammenhang der Testwerte mit zusätzlichen Lesemaßen (z.B. Leseverständnistest) fällt in den Feedback- und milden Zeitlimitbedingungen größer aus als in der Kontrollbedingung. Dies spricht für eine höhere konvergente Validität. Erste Ergebnisse zu Effekten auf die Testmotivation weisen darauf hin, dass Zeitfeedback und Zeitlimits einen Einfluss auf die aktuelle Testmotivation haben. Sowohl die Feedbackbedingung als auch die milde Form von Zeitlimits wirken sich demnach im Vergleich zur Kontrollbedingung motivationssteigernd aus. Strenge Zeitlimits hingegen haben einen negative Einfluss. Anhand der Reaktionszeiten wird zusätzlich ein Index zum Lösungsverhalten gebildet, der schnelles Rateverhalten bzw. motiviertes Testbearbeitungsverhalten anzeigen soll (Wise & Kong, 2005). Es wird erwartet, dass in der Feedbackbedingung sowie den milden Zeitbedingungen weniger Rateverhalten gezeigt wird als in der Kontrollbedingung.

Donnerstag
12. März

11:15 Uhr

HGB 20

Robitzsch, Alexander; Lüdtke, Oliver

Behandlung fehlender Item Responses in Leistungstests: Nichtignorierbare Konsequenzen verschiedener Modellierungsansätze

Fehlende Antworten bei Fragen in Leistungstests stellen eine besondere methodische Herausforderung dar, da häufig nicht klar ist, ob die fehlende Antwort z.B. eine geringe Fähigkeit widerspiegelt oder nur Ausdruck einer geringen Testmotivation ist. In der methodischen Literatur wurden verschiedene Ansätze zur Modellierung fehlender Item Responses in Leistungstests vorgeschlagen. So wurde kürzlich in einer Reihe von Arbeiten (z.B. Pohl, Gräfe & Rose, 2014; Pohl & Carstensen, 2013) auf Basis von Simulationen und testtheoretischen Überlegungen argumentiert, dass fehlende Item Responses nicht wie eine falsche Antwort gewertet werden sollten, sondern es angemessener wäre, die fehlenden Item Responses wie fehlende Daten (missing data) zu behandeln und sie somit zu ignorieren.

Im vorliegenden Beitrag werden zunächst kritisch die testtheoretischen Argumente (Rose, 2013) diskutiert, die für eine Behandlung fehlender Item Responses als falsch vorgebracht wurden. Oft wird dabei die unrealistische Annahme getroffen, dass fehlende Item Responses nur von der Fähigkeit, nicht aber von den unbeobachtbaren Responses selbst abhängen. Außerdem wird für die Argumentation angenommen, dass Itemschwierigkeiten für die beiden Gruppen von Schülern, die ein Item beantworten bzw. nicht beantworten, invariant sind, was zu einer zirkulären Definition der Fähigkeit führt (vgl. auch Rohwer, 2013).

Des Weiteren werden die Implikationen der verschiedenen Item-Response-Modelle (siehe Rose, von Davier und Xu 2010) für das Scoring von Items in Leistungstests aufgezeigt. Anhand eines kleinen Datenbeispiels wird illustriert, wie sich jeweils der Umgang mit den fehlenden Items Responses (Bewertung als falsch vs. missing data) auf die Interpretation der Fähigkeitsschätzungen auswirkt.

In einem nächsten Schritt werden zwei alternative Item-Response-Modelle zur Behandlung der fehlenden Item-Responses vorgeschlagen (Robitzsch, 2014). Der erste Ansatz basiert auf einem Pseudo-Maximum-Likelihood-Ansatz, der die Fälle der Bepunktung fehlender Item Responses als falsch und als ignorierbar als Spezialfälle enthält. Der zweite Ansatz modelliert die Wahrscheinlichkeit einer fehlenden Antwort auf einem Item als Funktion der unbeobachtbaren Itemantwort. Beide Ansätze besitzen eher den Status von Sensitivitätsanalysen und liefern eine quantitative Abschätzung für die Stabilität inhaltlicher Schlussfolgerungen (z.B. Unterschiede zwischen Gruppen) gegenüber unterschiedlichen Behandlungen von fehlenden Item-Responses.

Abschließend wird auf Basis des Ländervergleiches in PIRLS 2011 die Verwendung verschiedener IRT-Modelle zum Umgang mit fehlenden Item Responses diskutiert. Die Konsequenzen der verschiedenen Modellierungsansätze werden dabei im Hinblick auf Länderreihenfolgen, Mittelwerte und Quantile der Länderverteilungen, Zusammenhänge zu Kovariaten und Item-DIF untersucht. Die Ergebnisse zeigen, dass verschiedene IRT-Modelle zur Behandlung von fehlenden Item Responses zu deutlichen Änderungen in Länderreihenfolgen führen und somit einen erheblichen Einfluss auf die Validität der Ergebnisse besitzen. So erreichen z.B. Länder, mit einem relativ geringen Anteil fehlender Item Responses, niedrige Ränge bei einer Behandlung der fehlenden Antworten als missing data und somit nicht falsch. Dagegen erhalten Länder mit relativ hohem Anteil fehlender Antworten höhere Rangplätze bei diesem Vorgehen.

Der Beitrag zeigt insgesamt, dass weder theoretisch noch empirisch eindeutig entschieden werden kann, wie fehlende Item Responses optimal zu behandeln sind. Die Wahl einer bestimmten Skalierungsmethode jedoch bedeutsame Konsequenzen für die Interpretation der Testwerte besitzen kann.

Donnerstag
12. März

11:15 Uhr

HGB 20

Chair(s)

Goldhammer, Frank
Lüdtke, Oliver
Ünlü, Ali

Methodische Herausforderungen bei der Durchführung und Auswertung von
Large-Scale Assessment Studien

Ünlü, Ali

Ein methodisch-heterogener Rundblick über statistische Measurementforschung zu PISA im Zentrum für internationale Vergleichsstudien

Dieser Vortrag gibt einen Rundblick über eine persönliche Auswahl an Forschungsergebnissen, die im Rahmen eines ZIB-Teilprojekts entstanden sind. Details auslassend, werden die verschiedenartigen Themenbereiche und Resultate vielmehr konzeptionell abgehandelt. Die methodisch-heterogenen Inhalte zu PISA (OECD, 2012) oder Large-Scale Assessment umfassen hierbei die Grundlagen und Reproduzierbarkeit der IRT-Skalierung, lokale Abhängigkeit und klassische faktorenanalytische Verfahren unter Abweichung von Modellannahmen.

Reproduzierbarkeit von Analysen ist eine wichtige Forderung in der Wissenschaft, insbesondere wenn es um den Vergleich zu anderen Verfahren oder um Re-Analysen geht. Dem Thema der psychometrischen Skalierung in PISA und ihrer Reproduzierbarkeit widmet sich der erste Teil (Ünlü et al., 2014). Nach einer kurzen Rezension der Hauptskalierungsschritte in PISA, kann festgehalten werden, dass aus computationaler Sicht eine vollständige Reproduktion der PISA-Ergebnisse nicht möglich ist. Es werden Maßnahmen vorgeschlagen wie diese Lücken geschlossen werden können.

Verletzungen der zentralen lokalen Unabhängigkeitsannahme in dem PISA-Skalierungsmodell sind neuerlich und mit Recht kritisiert worden. Im zweiten Teil (Trendtel et al., 2014; Ünlü, 2006) wird eine Vorgehensweise vorgeschlagen, die es erlaubt, Mehrinformation unter lokaler Abhängigkeit zu nutzen. Im Rahmen von latenten Klassen mit Zufallseffekten wird ein verallgemeinerter statistischer Abhängigkeitsbegriff eingeführt, der lokale Unabhängigkeit wie in PISA angenommen als einen Spezialfall umfasst. Beispielsweise erscheint ein 2-Klassen Latent-Class-Modell mit Zufallseffekten als statistisch ausreichend zur Beschreibung von Abhängigkeiten in den vorliegenden PISA-Item-Units.

Die Bestimmung der Plausible Values in PISA basiert unter anderem auf Hintergrundmerkmalen. Diese fließen auch in Form von aggregierten Kennwerten einer angewendeten Hauptkomponentenmethode ein. Der Überprüfung der Sensitivität der Hauptkomponentenmethode und verwandter faktorenanalytischer Verfahren widmet sich der dritte Teil (Kasper und Ünlü, 2013). Die Normalverteilungsannahme, die auch in der PISA-Skalierung zentral ist, kann verletzt sein, insbesondere in den Hintergrundmerkmalen von Schülerinnen und Schülern. Es kann festgestellt werden, dass Verletzungen dieser zentralen latenten Verteilungsannahme in diesen klassischen faktoranalytischen Berechnungen die Schätzung der Hintergrundmerkmalswerte von Schülerinnen und Schülern oder ihrer Eigenschaften deutlich verzerren, während hingegen die faktorielle Struktur in den Hintergrundvariablen relativ robust bei Vorliegen schiefer Verteilungen geschätzt werden kann.

Donnerstag
12. März

11:15 Uhr

HGB 20

Nagy, Gabriel; Lüdtke, Oliver; Köller, Olaf; Heine, Jörg-Henrik; Mang, Julia

IRT Skalierung der Leistungstests in der PISA-Längsschnittstudie 2012/2013: Konsequenzen von Positionseffekten auf die Abschätzung der Leistungsentwicklung

Die Erfassung der Leistungsentwicklung setzt besondere Anforderungen an die Qualität der Leistungsmessung. Insbesondere gilt, dass die Zuwächse in den Testleistungen nicht durch zeitverändernde Effekte von Merkmalen des Testdesigns auf die Itemantworten der Schülerinnen und Schüler verzerrt sein sollten. Zeitverändernde Artefakte der Leistungsmessung können in realen Situationen niemals vollständig ausgeschlossen werden. Sie können jedoch (1) durch die Konstanthaltung des Testmaterials und des Testdesigns über die Erhebungswellen hinweg oder (2) durch die Adjustierung der Testleistungen um zeitspezifische Effekte von Designmerkmalen minimiert werden. Obwohl die designbasierte Konstanthaltung von Testmerkmalen in vielerlei Hinsicht vorteilhaft erscheint, lässt sich diese Maßnahme in den meisten großen Längsschnittstudien aus forschungspraktischen Gründen nicht umsetzen. Die Auswertung der Leistungstests muss sich somit auf die Adjustierung der Einflüsse spezifischer Designmerkmale beschränken, wobei sich die Frage nach deren adäquaten Modellierung stellt.

Gegenstand des vorliegenden Beitrags ist die Erfassung und Modellierung der Auswirkungen von (Item-)Positionseffekten auf die Abschätzung der Leistungsentwicklung in der PISA-Längsschnittstudie 2012/2013 in den Bereichen Mathematik, Naturwissenschaften und Lesen. Positionseffekte manifestieren sich in einem Anstieg der Itemschwierigkeiten in Abhängigkeit von deren Position in einem Test und scheinen im Bereich der Schulleistungsmessung eher die Regel als die Ausnahme darzustellen (z.B., Meyers, Miller & Way, 2009). Als Antwort auf die im Rahmen der PISA-Assessments wiederholt ermittelten Positionseffekte, werden in PISA Multi-Matrix-Designs mit balancierten Itemclusterpositionen verwendet. Die aufgrund dieses Testdesigns ermittelten Itemschwierigkeiten und Populationsmittelwerte sind hinsichtlich der „mittleren“ Itemposition definiert (Harris, 1991). Diese Eigenschaft gilt auch für die Leistungsmessung in der ersten Welle der PISA-Längsschnittstudie 2012/2013. In der zweiten Erhebung musste das balancierte Testdesign zugunsten einer Erweiterung des Testmaterials aufgegeben werden, wobei jedoch zwei Testbooklets unverändert in beiden Erhebungen eingesetzt wurden. Die Änderung des Testdesigns kann je nach Inhaltsdomäne eine verstärkte oder abgeschwächte Akzentuierung von Positionseffekten nach sich ziehen, die sich verzerrend auf die Abschätzung der Leistungsentwicklung auswirken kann (Mazzeo & von Davier, 2009).

Die Auswirkungen von Positionseffekten auf die Abschätzung der Leistungsentwicklung wurden mittels einer Serie alternativer IRT-Modelle für Längsschnittdaten untersucht (Andersen, 1985). Die Auswertungen basieren auf einem Sample von N = 6974 Schülerinnen und Schüler aus Schulen, die an der Retesterhebung teilgenommen haben. Die verwendeten IRT-Modelle sind eine multidimensionale Erweiterung des Ein-Parameter logistischen Modells, in denen die domänenspezifischen personenseitigen Testleistungen zu jedem Messzeitpunkt durch jeweils eine latente Variable dargestellt wurden. Ausgehend von einem längsschnittlichen IRT-Modell mit zeitinvarianten Itemschwierigkeiten, wurden alternative Modelle spezifiziert. Eine Gruppe von Modellen beinhaltet (Cluster-)Positionseffekte auf die Itemschwierigkeiten und ermöglicht die Abschätzung von Leistungszuwächsen, die in Referenz zu ausgewählten Clusterpositionen definiert sind. Eine zweite Gruppe von Modellen umfasst zeitspezifische Bookleteffekte. Diese Modelle erlauben die Abschätzung von Leistungszuwächsen, die in Referenz zu einem wiederholt eingesetzten Booklet definiert sind.

Die Ergebnisse der Auswertungen liefern Hinweise für die Bedeutung von Positionseffekten bei der Abschätzung von längsschnittlichen Leistungstrends. Sie dokumentieren graduell unterschiedliche Trendschätzungen, die von der Modellspezifikation und der Inhaltsdomäne abhängen.

Donnerstag
12. März

11:15 Uhr

HGB 20

Chair(s)

Goldhammer, Frank
Lüdtke, Oliver
Ünlü, Ali

Methodische Herausforderungen bei der Durchführung und Auswertung von
Large-Scale Assessment Studien

Sumfleth, Elke

Das Professionswissen von Lehrkräften in den Naturwissenschaften - Ein Vergleich der Befunde aus KiL und ProwiN

In den letzten zehn Jahren hat sich das Professionswissen von Lehrkräften in den Didaktiken der Mathematik und Naturwissenschaften zu einem bedeutenden Forschungsschwerpunkt entwickelt. In der Folge der COAKTIV-Studie in Mathematik (u. a. Kunter et al., 2011) sind für die Naturwissenschaften die Projekte Professionswissen in den Naturwissenschaften (ProwiN; u. a. Tepner et al., 2012) an der Universität Duisburg-Essen und Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen (KiL; Kleickmann et al.) am IPN in Kiel durchgeführt worden. Beide Projekte wurden in Biologie, Chemie und Physik (KiL auch in Mathematik) auf der Basis ähnlicher Konzeptualisierungen des Professionswissens, aber separater Testentwicklung durchgeführt. Im Zentrum des Kieler Projekts stehen Lehramtsstudierende, im Zentrum des Essener Projekts Lehrkräfte an Gymnasien und Haupt- sowie Realschulen.

Im beantragten Symposium sollen die unabhängig voneinander erreichten Ergebnisse miteinander verglichen werden, um aus Gemeinsamkeiten und Unterschieden Schlussfolgerungen für weitere Studien – insbesondere auf die Entwicklung des Professionswissens - zu gewinnen. Leitfragen für die Diskussion sollen sein: (1) Welche Erkenntnisse ergeben sich aus dem Vergleich der Ergebnisse aus KiL und ProwiN für die Struktur des Professionswissens? (2) Inwieweit sind diese möglicherweise durch Unterschiede in der Operationalisierung des Professionswissens bedingt? (3) Welche Konsequenzen ergeben sich für die Lehrerbildung?

Mit Blick auf das GEBF Disziplinencluster werden in diesem Symposium Inhalte der Disziplinen Erziehungswissenschaft, Lehrerbildung, Psychologie und Didaktik der Naturwissenschaften und Technik adressiert.

Im ersten Beitrag werden von Hans Fischer und Knut Neumann der gemeinsame theoretische Ansatz, die verschiedenen Rahmen der Testentwicklung und die Durchführung der beiden Studien vorgestellt und verglichen. Danach werden die Ergebnisse der ersten Projektphasen präsentiert und verglichen, und zwar die für Biologie von Ute Harms und Birgit Neuhaus, die für Chemie von Mirjam Steffensky und Oliver Tepner und die für Physik von Andreas Borowski und Knut Neumann. Dabei werden in den drei Präsentationen sowohl inhaltlich unterschiedliche Schwerpunkte gesetzt als auch die in jedem Fall als dreidimensional beschriebene Struktur des Professionswissens im Sinne von CK, PCK und PK thematisiert. Zum Abschluss werden die Beiträge mit Blick auf verallgemeinerbare Gemeinsamkeiten und Unterschiede diskutiert.

Donnerstag
12. März

11:15 Uhr

HGB 30

Chair(s)
Sumfleth, Elke**Das Professionswissen von Lehrkräften in den Naturwissenschaften - Ein Vergleich der Befunde aus KiL und ProwiN**

Neumann, Knut; Hans Ernst, Fischer

Das Professionswissen von Lehramtsstudierenden und Lehrkräften der Naturwissenschaften - Theoretischer Hintergrund und Methode

Das Professionswissen von Lehrkräften gilt als zentraler Faktor für den Lernerfolg (z.B. Baumert et al., 2010; Magnusson et al., 1999). In Anlehnung an Shulman (1987) wird das Professionswissen dabei üblicherweise in Fachwissen (CK), fachdidaktisches Wissen (PCK) und pädagogisches Wissen (PK) gegliedert. Doch während bezüglich der Bedeutung des Professionswissens und dessen Gliederung in die genannten drei Komponenten Einigkeit besteht, ist bisher weitgehend unklar, welches Fach-, fachdidaktische und pädagogische Wissen Lehrkräfte für die Gestaltung erfolgreichen Unterrichts benötigen und wie sie dieses entwickeln (vgl. Abell, 2007). Abell (2007) kommt entsprechend in ihrem einflussreichen Review der Forschung zum Professionswissen von Lehrkräften in den Naturwissenschaften zu dem Schluss, dass mehr Studien zur Struktur des Professionswissens, seiner Entwicklung und seiner Nutzung im Unterricht benötigt werden. Zwei Studien, die sich in den vergangenen Jahren der Struktur des Professionswissens von Lehrkräften in den Naturwissenschaften gewidmet haben, sind das Projekt „Professionswissen in den Naturwissenschaften“ (ProwiN) und das Projekt „Messung professioneller Kompetenzen in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen“ (KiL).

Im Rahmen des vom BMBF geförderten Projekts ProwiN wurde das Professionswissen von Lehrkräften in den Naturwissenschaften untersucht. In der ersten Projektphase wurden Instrumente zur Erfassung des Professionswissens in den Disziplinen Biologie, Chemie, Physik sowie Bildungswissenschaften konzipiert. Die Instrumente bauen auf einem gemeinsamen Rahmenmodell des Professionswissens auf, dass die Komponenten CK, PCK und PK ausdifferenziert. Zur Aufgabenkonstruktion wurde ein Modell verwendet, in dem zwischen deklarativem, prozeduralem und konditionalem Wissen zu verschiedenen Themenbereichen der jeweiligen Fächer unterschieden wurde, in erster Linie, um die Varianz der Aufgaben zu erhöhen. Die Gesamtstichprobe besteht aus 518 Personen, davon sind 279 Physiklehrkräfte, davon 144 aus NRW, 64 aus Bayern und 71 aus anderen Bundesländern. Zur Validierung wurden 31 Gymnasiallehrkräfte anderer Fächer, 103 Lehramtsstudierende der Physik, 62 Physikreferendare sowie 43 Physik-Studierende analysiert. Die entwickelten Instrumente werden derzeit in einer Videostudie (zweite Projektphase von 2012 bis 2015) genutzt, um Zusammenhänge zwischen den Dimensionen des Professionswissens und Lehrerhandeln im Unterricht zu untersuchen sowie differenzierte Vorhersagen über die Wirkung des Professionswissens auf den Lernzuwachs bei Schülern abzuleiten.

Ziel des von der Leibniz-Gemeinschaft geförderten Projekts KiL war es, Instrumente zur Erfassung des Professionswissens von Lehrkräften der Fächer Biologie, Chemie, Physik und Mathematik bereitzustellen. In der ersten Phase des Projekts wurden, ausgehend vom Stand der Forschung, den Standards für die Lehrerbildung (KMK, 2008) sowie einer Analyse der Modulpläne verschiedener Universitäten aus den 16 Bundesländern Modelle für PK sowie fachspezifische Modelle für CK und PCK entwickelt. Diese Modelle unterscheiden sich sowohl bezüglich inhaltlicher Schwerpunktsetzungen (z.B. unterschiedliche Inhaltsbereiche in der Physik als Grundlage für eine Operationalisierung des Fachwissens, Basiskonzepte in der Chemie), als auch in der Charakterisierung des Wissens (z.B. Wissensarten in der Physik, verschiedene Arten des Umgangs mit Wissen in der Biologie). Die Modelle waren Basis für die Entwicklung entsprechender Testinstrumente zur Erfassung des CK, PCK und PK, die nach umfangreicher Pilotierung N=1058 Lehramtsstudierenden von 12 deutschen Universitäten zur Bearbeitung vorgelegt wurden. Zusätzlich zu den genannten Konstrukten wurden weitere persönliche Merkmale der Befragten (z.B. Interesse, Selbstkonzept, usw.) sowie studienbezogene Merkmale (z.B. Lerngelegenheiten) erhoben.

Basierend auf den gewonnenen Daten wurden in beiden Projekten mit Hilfe von Methoden der Item Response Theory und des Structural Equation Modeling Erkenntnisse über die Struktur des Professionswissens gewonnen. Im Symposium wird in dieser Einführung ausgehend von einem kurzen Überblick über den Stand der Forschung zum Professionswissen von Lehrkräften die Anlage der beiden Projekte – einschließlich des theoretischen Rahmens, des Designs und der Vorgehensweise bei der Entwicklung der Testinstrumente – vorgestellt. Dabei wird insbesondere auf die Spezifika beider Projekte im Vergleich zu einander eingegangen. In den weiteren Beiträgen werden dann die spezifischen Modelle für CK und PCK, Beispielaufgaben sowie Ergebnisse zu Struktur und Entwicklung des Professionswissens vorgestellt.

Donnerstag
12. März

11:15 Uhr

HGB 30

Chair(s)

Sumfleth, Elke

Das Professionswissen von Lehrkräften in den Naturwissenschaften - Ein Vergleich der Befunde aus KiL und ProwiN

Neuhaus, Birgit; Harms, Ute; Großschedl, Jörg

Professionswissen von Lehrkräften und Lehramtsstudierenden in den Naturwissenschaften - Ergebnisse aus der Biologie

Theoretischer Hintergrund: Das Professionswissen einer Lehrkraft gilt als zentrales Element ihrer professionellen Kompetenz (Baumert & Kunter, 2006) und ist für alle Phasen der Lehrerbildung wesentlich. In Anlehnung an Baumert und Kunter (2006) werden in den Projekten KiL und ProWiN drei Bereiche des Professionswissens untersucht: das biologische Fachwissen (CK), das biomedizinische Wissen (PCK) und das allgemein pädagogisch-psychologische Wissen (PK). Ziel beider Projekte ist es (1) objektive, reliable und valide Testinstrumente für die Messung des Professionswissens von Biologielehramtsstudierenden (KiL) bzw. Biologielehrern (ProWiN) bereitzustellen, sowie (2) zur Strukturaufklärung des Professionswissens beizutragen.

Methoden: KiL: Auf Basis (a) der Lehrerbildungsstandards (KMK, 2004, 2008) und (b) einer Analyse aktueller Modulkataloge verschiedener Universitäten aller Bundesländer wurde ein curricular-valider Itempool mit unterschiedlichen kognitiven Anforderungen entwickelt (78 CK-, 66 PCK- und 36 PK-Items). Dieser wurde einer Stichprobe aus 274 Lehramtsstudierenden (64,7% gymnasiales Lehramt) in einem Multimatrixdesign zur Beantwortung vorgelegt (mittlerer Studienfortschritt = 4,7 Semester; 28% männlich). Zur Überprüfung der diskriminanten Validität erfolgte eine konfirmatorische Faktorenanalyse (Mplus). Die Modellierung der Bereiche des Professionswissens erfolgte über jeweils sechs Itempakete für CK und PCK sowie zwei Itempakete für PK. Um weitere Validitätshinweise zu finden, wurden latente Regressionen zwischen den drei Bereichen des Professionswissens und z.B. ausgewählten Skalen des KFT (Heller & Perleth, 2000) oder dem angestrebten Lehramtszugang durchgeführt.

ProWiN: Basierend auf einem Fachwissenstest für Schüler (N=461) wurden Schülerfehler identifiziert, die als Grundlage für Test der Lehrkräfte dienen. Der finale PCK-Test besteht aus 24, der CK-Test aus 21 Items (Jüttner et al., 2013). Die Tests wurden von 158 Biologielehrkräften ausgefüllt, von zwei Ratern ausgewertet und die Objektivität (ICC) bestimmt. Die Tests wurden mittels Winsteps Rasch-skaliert (Partial Credit Model). Zur Überprüfung der Validität wurde die Tests zudem von Biologen und Psychologen bearbeitet (n(Biologen)=20; n(Psychologen)=11) und die Testleistungen dieser Subpopulationen mittels ANOVA mit den Ergebnissen der Lehrkräfte verglichen. Die Inhaltsvalidität wurde durch Protokolle lauten Denkens (N=5) bestimmt, die von zwei Beobachtern ausgewertet wurden. Um Zusammenhänge zwischen CK und PCK zu ermitteln, wurde eine Pearson-Korrelation mit den einheitlich intervallskalierten Raschwerten der Personenfähigkeiten berechnet.

Ergebnisse: KiL: Eine Faktorenanalyse der Ergebnisse des KiL Projekts zeigt, dass CK, PCK und PK zwar korreliert, aber doch empirisch trennbare Konstrukte darstellen. PCK korreliert hoch mit CK ($r=0,68$; $p<0,01$) und moderat mit PK ($r=0,35$; $p<0,01$). Zwischen CK und PK besteht kein signifikanter Zusammenhang ($r=0,11$; $p>0,05$). Die latenten Regressionsanalysen weisen insgesamt auf eine gute Validität der entwickelten Messinstrumente hin. Beispielsweise übertreffen Studierende des gymnasialen Lehramts ihre Kommilitonen aus dem nicht-gymnasialen Lehramt erwartungsgemäß hinsichtlich des CK und PCK. Als Maß für die Reliabilität der Skalen wurde der Faktor α Koeffizient bestimmt, der auf eine gute Messgenauigkeit der Instrumente hinweist (CK: $\alpha=0,85$; PCK: $\alpha=0,88$; PK: $\alpha=0,75$).

ProWiN: Die Objektivität bei der Auswertung der offenen Items ist hoch (CK: ICC(unjust) = 0,83; PCK: ICC(unjust)=0,92). Bezüglich der Reliabilität wurden für den PCK-Test (0,97/0,58) und den CK-Test (0,98/0,76) zufriedenstellende Item- und Personen-Reliabilität gefunden. Die Infit und Outfit MNSQ-Werte, sowie die standardisierten t-Werte, sind für alle Items zufriedenstellend (0,5-1,5). Im Bereich der Konstruktvalidität wurde deutlich, dass die Pädagogen im CK-Test signifikant schlechter abschneiden als die beiden anderen Gruppen ($F_{2,186}=22,80$, $p<0,001$, $\omega^2=0,045$) und im PCK Test die Biologielehrkräfte signifikant besser abschneiden ($F_{2,186}=33,82$, $p<0,001$, $\omega^2=0,092$). Die Protokolle lauten Denkens weisen auf eine hohe Inhaltsvalidität hin. Zwischen CK und PCK wurde eine geringe Korrelation gefunden ($r=0,22$; $p=0,006$; $N=158$).

Diskussion/Ausblick: Die Instrumente aus beiden Projekten verfügen über gute Gütekriterien. Es konnte gezeigt werden, dass CK und PCK zwar korrelieren, aber eindeutig eigenständige Wissensbereiche darstellen. Zurzeit wird der Test aus dem Projekt ProWiN in einer Videostudie eingesetzt, um Zusammenhänge zwischen dem Professionswissen und der Unterrichtsgestaltung herauszuarbeiten, im Projekt KiL wird in einem Längsschnitt überprüft wie sich das Professionswissen von Lehramtsstudierenden im Verlauf des Studiums entwickelt.

Donnerstag
12. März

11:15 Uhr

HGB 30

Chair(s)
Sumfleth, Elke

Das Professionswissen von Lehrkräften in den Naturwissenschaften - Ein Vergleich der Befunde aus KiL und ProWiN

Tepner, Oliver; Steffensky, Mirjam; Parchmann, Ilka; Sumfleth, Elke

Fachspezifisches Professionswissen von Chemielehrkräften - Vergleich der Projekte KiL und ProwiN

Ausgangslage und theoretischer Hintergrund: Während Theorien zur Konzeptualisierung des Professionswissens von Lehrkräften seit geraumer Zeit existieren (z. B. Elbaz, 1983; Shulman, 1987), werden Tests zur empirischen Erforschung des Professionswissens erst seit ca. einem Jahrzehnt entwickelt (Baumert & Kunter, 2006). Dabei werden vorrangig das Fachwissen (CK), das pädagogische Wissen (PK) und das fachdidaktische Wissen (PCK) in der aktuellen Forschung betrachtet (Abell, 2007; Blömeke et al., 2009; Kunter et al., 2013).

Um Erkenntnisse über die Struktur, die Entwicklung und die Förderung des fachbezogenen Professionswissens zu erhalten, ist die Verwendung von Testinstrumenten sinnvoll, welche Ausprägungen einzelner Wissensdimensionen quantitativ abbilden können (Baumert et al., 2010). Für das Fach Chemie stehen vergleichsweise wenige Instrumente zur Verfügung – eine Ausnahme stellt hier z.B. das Instrument von Rollnick und Mavhunga (2014) dar.

Um das bestehende Desiderat zu schließen, wurden in den Projekten KiL und ProwiN Testinstrumente entwickelt, die u.a. der Untersuchung folgender Fragestellungen dienen. Während das Projekt ProwiN auf das Professionswissen von Lehrkräften abzielt, geht es im Projekt KiL um das Wissen von Lehramtsstudierenden.

Fragestellungen: (1) Lässt sich das fachspezifische Professionswissen im Fach Chemie mit unterschiedlichen Konzeptualisierungen reliabel und valide erfassen? (2) Welche Unterschiede im Zusammenhang zwischen den einzelnen Wissensdimensionen zeigen sich in den Tests in den Projekten KiL und ProwiN? (3) Inwiefern unterscheiden sich in beiden Tests Chemielehrkräfte und angehende Chemielehrkräfte im fachspezifischen Professionswissen bzgl. ihrer Schulform?

Methode: Für die Konzeption der Testitems wurden in beiden Projekten auf eine ähnliche Konzeptualisierung des Fachwissens und des fachdidaktischen Wissens zurückgegriffen und die Items modellbasiert entwickelt (Kleickmann et al., 2014; Tepner et al., 2012).

Während im Projekt KiL Paper/Pencil-Tests sowohl mit offenem als auch geschlossenem Aufgabenformat genutzt werden, sind im Projekt ProwiN-Chemie (1. Phase) geschlossene Aufgaben im Multiple-Choice- und Rating-Scale-Format konzipiert. Hier wurden Items zu den Inhaltsbereichen ‚Atombau und Periodensystem‘, ‚Chemische Bindungen‘ und ‚Chemische Reaktionen am Beispiel von Säuren und Basen‘ entwickelt (Tepner & Dollny, 2014), während die Items des KiL-Projekts den Fokus auf Struktur-Eigenschafts-Beziehungen im Kontext verschiedener Inhaltsbereiche legen (Herzog, Taskin, Bernholt, Steffensky & Parchmann, 2013).

Im Bereich des fachdidaktischen Wissens wird sich in beiden Projekten auf die von Shulman (1987) vorgeschlagenen Facetten des Wissens über bestimmte fachspezifische Schülerfehler und Schülervorstellungen sowie fachspezifische Instruktionsstrategien bezogen. Letztere werden über den Einsatz von Experimentierstrategien und Modellen (inklusive Repräsentationsformen) konzeptualisiert (Herzog et al., 2013; Witner & Tepner, 2011).

Die Stichproben bestehen im KiL-Projekt aus 322 Lehramtsstudierenden aus verschiedenen Semestern an 12 Universitätsstandorten und im ProwiN-Projekt aus 98 Gymnasial- und 12 Hauptschullehrkräften.

Ergebnisse: Die im KiL-Projekt konzipierten Instrumente bestehen aus 36 CK-Items und 24 PCK-Items mit zufriedenstellenden Reliabilitätskennwerten (CK=.87; PCK=.77). Die ProwiN-Testinstrumente setzen sich aus insgesamt 44 Items (NCK=25; NPC=19) zusammen und haben mit CK=.87 und PCK=.73 vergleichbare Reliabilitätskennwerte. Hinweise auf die Validität geben Expertenratings und Kontrastgruppenvergleiche (ProwiN) sowie Zusammenhänge zu vielfältigen externen Kriterien wie Studienfortschritt oder Chemie-Abiturnote (KiL).

Die Korrelationen zwischen CK und PCK sind im KiL-Projekt mit $r=.546$ ($p<.001$; $N=322$) etwas höher als im ProwiN-Projekt ($r=.419$; $p<.001$; $N=110$).

Zwischen deklarativem PK und PCK sind im ProwiN-Projekt keine signifikanten Zusammenhänge erkennbar, während dies zwischen konditional-prozeduralem PK und PCK der Fall ist ($r=.501$; $p<.001$; $N=110$). Im KiL-Projekt lässt sich ein signifikanter Zusammenhang zwischen PCK und PK (deklaratives Wissen) feststellen ($r=.361$; $p<.001$; $N=322$).

Hinsichtlich der unterschiedlichen Schulformen zeigen sich im ProwiN-Projekt bezüglich des CK ($t(12,2)=5.76$; $p<.001$; $d=2.4$) und des PCK ($t(108)=3.43$; $p=.001$; $d=1.0$) signifikante Unterschiede zugunsten der Gymnasiallehrkräfte. Im KiL-Projekt finden sich die entsprechenden Unterschiede für die beiden Gruppen der angehenden Lehrkräfte (CK: Mann-Whitney-U=4297.50; $z=-3.34$; $p=.001$; PCK: Mann-Whitney-U=4265.50; $z=-3.40$; $p=.001$).

In dem Vortrag werden die Ergebnisse der beiden Projekte miteinander in Beziehung gesetzt und hinsichtlich der unterschiedlich konzipierten Testinstrumente diskutiert.

Donnerstag
12. März

11:15 Uhr

HGB 30

Chair(s)
Sumfleth, Elke

Das Professionswissen von Lehrkräften in den Naturwissenschaften - Ein Vergleich der Befunde aus KiL und ProwiN

Borowski, Andreas; Neumann, Knut

Die Struktur des Professionswissens von Lehramtsstudierenden und Lehrkräften der Physik im Vergleich

In der Folge der zunehmenden Diskussion um das Professionswissen von Lehrkräften hat sich in den letzten Jahren eine zunehmende Zahl von Studien dem Professionswissen von Physiklehrkräften in verschiedenen Stadien ihrer Ausbildung bzw. ihres Berufslebens gewidmet (Kirschner, 2013; Kröger et al., 2014; Olszewski, 2010; Riese, 2009). Dabei wird ein Vergleich der Ergebnisse durch unterschiedliche Modelle und unterschiedliche Stichproben erschwert. So lassen sich Unterschiede in den Befunden zur Struktur des Professionswissens kaum interpretieren. Auch die Projekte „Kompetenzmessung im Lehramtsstudium“ (KiL) und „Professionswissen in den Naturwissenschaften“ (ProwiN) untersuchen das Professionswissen in unterschiedlichen Stichproben – Lehramtsstudierenden (KiL) und Lehrkräften (ProwiN) der Physik. Beide Studien bauen jedoch auf einem weitgehend identischen Modell des Fach- (CK) und fachdidaktischen Wissens (PCK) auf. Das Modell des CK unterscheidet dabei zwischen verschiedenen Arten von Wissen in verschiedenen Inhaltsbereichen der Physik. Das Modell des PCK verschiedene Arten von Wissen bezüglich verschiedener physikdidaktischer Inhaltsbereiche und physikalischer Inhaltsbereiche (vgl. Kirschner 2013; Kröger et al., 2014). Unterschiede betreffen vor allem die physikalischen Inhaltsbereiche. Während ProwiN auf die Mechanik fokussiert, werden in KiL weitere Inhaltsbereiche der universitären Physik berücksichtigt.

Die Fragestellungen beider Projekte, zielen auf die Struktur des Professionswissens in Physik in den jeweiligen Ausbildungsabschnitten, für die entsprechende Tests konzipiert wurden. Daraus ergeben sich folgende Fragestellungen für den vorliegenden Beitrag: Wie unterscheidet sich die Struktur des Professionswissens von Lehramtsstudierenden und Lehrkräften in Physik? Und: Inwieweit lassen sich daraus Erkenntnisse über die Entwicklung des Professionswissens ableiten?

Zur Untersuchung der Fragestellung liegen aus KiL Daten von insgesamt $N = 201$ Studierenden des Lehramts Physik vor (Durchschnittliches Alter: 23,5 Jahre, Standardabweichung 3,3 Jahre; Durchschnittliches Studiensemester: 5,6 Semester, Standardabweichung 2,7 Semester) vor. Die Studierenden erhielten jeweils einen Test zum CK (Reliabilität $\alpha_{CK} = .85$), PCK (Reliabilität $\alpha_{PCK} = .74$) sowie pädagogischem Wissen (PK; Reliabilität $\alpha_{PK} = .85$). Die Struktur des Professionswissens wurde mittels Strukturgleichungsmodellen untersucht. Der Vergleich des 1-dimensionalen Modells mit dem 3-dimensionalen Modell legt dabei eine 3-dimensionale Struktur des Professionswissens mit den Dimensionen CK, PCK und PK nahe (3-dimensional: CFI = .97, TLI = .95, SRMR = .06, RMSEA .06; 1-dimensional: CFI = .90, TLI = .89, SRMR = .09, RMSEA .10). Die Korrelationen zwischen den Dimensionen betragen $\square_{CK-PCK} = .79$, $\square_{CK-PK} = .49$, und $\square_{PK-PCK} = .77$ (all $p < .001$).

Aus dem Projekt ProwiN liegen Daten von $N = 279$ Physiklehrkräften (28% weiblich, Alter 44 Jahre, Standardabweichung 10 Jahre; Unterrichtserfahrung 12,2 Jahre, Standardabweichung 10,7 Jahre) vor. Die Reliabilitäten der Tests (EAP/PVCK = .77; E-AP/PVPCK = .86; EAP/PVVK = .71) waren zufriedenstellend bis gut (Kirschner, 2013). Eine mehrdimensionale Raschanalyse ergab, dass auch hier das 3-dimensionale Modell dem 1-dimensionalen Modell überlegen ist, das Professionswissen also auch hier eine 3-dimensionale Struktur aufweist (3-dimensional: $df = 131$, deviance = 20786, BIC = 25270, CAIC = 25080; 1-dimensional: $df = 126$, deviance = 25388, BIC = 26004, CAIC = 25822). Die latenten Korrelationen zwischen den Dimensionen fallen jedoch durchgehend niedriger aus als in KiL: $\square_{CK-PCK} = .60$, $\square_{CK-PK} = .37$ und $\square_{PK-PCK} = .23$.

Ein Vergleich dieser Ergebnisse legt nahe, dass sich die Struktur des Professionswissens über die Zeit verändert. In welcher Form ist bisher jedoch offen und Bedarf einer detaillierten Analyse sowohl der verwendeten Instrumente als auch der Ergebnisse im Vergleich. Mit dem Ziel Hypothesen für eine weitere Untersuchung dieser Veränderung im Konstrukt Professionswissen bei Physiklehrkräften über das Berufsleben hinweg zu generieren, werden im Rahmen des Symposiums zunächst die Instrumente ausführlich vorgestellt und auf dieser Grundlage die Ergebnisse diskutiert.

Donnerstag
12. März

11:15 Uhr

HGB 30

Chair(s)
Sumfleth, Elke

Das Professionswissen von Lehrkräften in den Naturwissenschaften - Ein Vergleich der Befunde aus KiL und ProwiN

Biedermann, Horst

Nutzung neuer Medien im Jugendalter: Licht und Schatten

Der Alltag vieler (junger) Menschen gestaltet sich heutzutage derart, dass der Tag mit einem Blick auf die sozialen Netzwerke beginnt, der Tagesverlauf durch die mit neuen Medien verbundene Unmittelbarkeit (mit-)geleitet wird und der Tagesabschluss durch einen letzten Eintrag in den Netzwerken Ritualisierung gefunden hat. Kurz: die heutigen Lebenswelten junger Menschen sind von verschiedenen Medien durchdrungen. Sowohl Eltern, Lehrkräfte, politische Bildungsverantwortliche als auch Forschende fragen sich, welche Auswirkungen ein derartiges Medienverhalten für die Individuen und die Gesellschaft hat. Auch wenn zu Mediennutzung und Medienwirkung in den letzten Jahren intensiviertere Forschungsanstrengungen unternommen wurden, zeigen sich noch immer deutliche Kenntnislücken (Porsch & Pieschl, 2014). Die Beiträge in diesem Symposium nehmen sich derartiger Lücken an, indem Wirkungen der Mediennutzung junger Menschen aus unterschiedlichen Perspektiven (schulpädagogisch, soziologisch und [entwicklungs-]psychologisch) hinsichtlich unterschiedlicher Facetten analysiert werden – und dadurch (Mosaikstücke von) Licht und Schatten neuer Medien herausgearbeitet werden.

Arvid Nagel geht in einem ersten Beitrag der Frage nach, inwieweit kognitiv unterrichtmeidendes Verhalten von Schüler/innen durch ihre Mediennutzung erklärt werden kann. Dabei zeigen sich für die Geschlechter unterschiedliche Ergebnisse. Während sich bei Jungen insbesondere die Computerspielabhängigkeit als Erklärungsvariable der kognitiven Unterrichtsmeidung zeigt, ist es bei Mädchen stärker die Internetabhängigkeit.

Albert Düggeli und Katja Kinder richten den Blick auf die Bedeutung der Internetnutzung im Berufswahlprozess. Konkret interessiert sie, wie häufig Heranwachsende das Internet zum berufswahlspezifischen Wissensaufbau nutzen, wie nützlich sie die dabei gewonnenen Informationen bewerten und inwieweit diese beiden Aspekte durch IT-Kompetenz und familiäre Hintergrundvariablen Erklärung finden können. Dabei zeigt sich, dass Nutzung und Nützlichkeitsbewertung je nach der individuellen Statusform im Prozess der Berufswahl unterschiedlich ausfallen.

Das Lästern in neuen Medien steht im Mittelpunkt des Beitrags von Horst Biedermann, Julia Wolke und Fritz Oser. Sie gehen den Fragen nach, wie verbreitet das Lästern bzw. Cyber-Mobbing bei Schüler/innen ist und welche Einstellungen sie gegenüber derartigem Verhalten zeigen. Mehr als ein Viertel der Schüler/innen berichten, dass sie bereits Erfahrungen mit Mobbing gesammelt haben – viele „nur“ als Beobachtende, aber auch 9% als Opfer und 6% als Täter/innen. Während Cyber-Mobbing vom Großteil der Schüler/innen abgelehnt wird, rechtfertigt die Täterschaft ein derartiges Handeln jedoch nicht (eindeutig) verwerflich.

Fritz Oser und Horst Biedermann interessiert im letzten Beitrag die Frage, inwieweit sich neue Medien auf Freundschaftsverhältnisse und -konzepte auswirken. Konkret interessiert sie, ob sich Freundschaften in der virtuellen und realen Welt ausschließen oder doch Ergänzung finden können und wie sich die jeweiligen Qualitäten zeigen. Die Ergebnisse lassen erkennen, dass sich Freundschaften in den beiden Welten (eher) entgegenstehen.

Donnerstag
12. März

11:15 Uhr

HGB 40

Chair(s)

Biedermann, Horst

3

Nagel, Arvid

Mediennutzungsbezogene Einflussfaktoren auf kognitiv unterricht-meidendes Verhalten von Schülerinnen und Schülern

Bisherige Forschungen zur Unterrichtsmeidung konzentrieren sich verstärkt auf die Einflussbedingungen der Sozialisationsinstanzen Familie, Schule und Peers. Als ein zentrales Desiderat lässt sich festhalten, dass der Einfluss von Mediennutzung im Zusammenhang mit (kognitiver und physischer) Unterrichtsmeidung bis anhin gänzlich vernachlässigt wird, was jedoch dem gegenwärtigen Diskurs einer „Jugend der Medien“ (Friedrichs & Sander, 2010, S. 26) entgegensteht. Aufgrund gesamtgesellschaftlicher Veränderungen und der „Medialisierung des Alltags“ (Ferchhoff, 2007, S. 370) prägt sich gegenwärtig eine immer stärker zu beobachtende „digitale Jugend“ (Hugger, 2010, S. 21) aus. Interaktive Medien nehmen innerhalb der jugendkulturellen Lebenswirklichkeiten eine zentrale Sozialisationsinstanz ein und bestimmen ausschlaggebend die Lebenswelt der heutigen Jugendlichen mit (vgl. u.a. Feierabend, Karg & Rathgeb, 2014; Mößle, Kleimann & Rehbein, 2007; Morgenroth, 2012; Treumann et al., 2007). Die empirische Klärung auf die Frage nach der Wirkung der Neuen Medien auf Unterrichtsmeidung steht jedoch noch offen und wird von Ricking, Schulze und Wittrock (2009, S. 34f.) eingefordert.

In Anlehnung an dieses Forschungsdesiderat soll im Rahmen dieses Beitrages die forschungsleitende Fragestellung untersucht werden: Lassen sich mediennutzungsbezogene Faktoren als Einflussbedingungen für kognitive Unterrichtsmeidung von Schüler/innen unter Berücksichtigung individueller, familiärer, schulbezogener und peer-kultureller Bedingungsfaktoren belegen?

Generell fehlt es bisher an geeigneten Instrumenten, um kognitiv unterrichtmeidendes Verhalten von Schüler/innen während der physischen Anwesenheit in dem Wirkungsbereich Unterricht empirisch-quantitativ zu erfassen. Das Instrument kognUm (Skala zur Unterrichtsunaufmerksamkeit) ist generiert worden, um kognitive Unterrichtsmeidung von jungen Menschen zu untersuchen. Das Instrument wurde in einer ersten Schüler/innenbefragung (N = 166) genutzt und die psychometrischen Daten zeigen eine 1-faktorielle Struktur, wie sie auch theoretisch postuliert wurde. Die Eindimensionalität des Konstruktes der kognitiven Unterrichtsmeidung wurde mit Hilfe konfirmatorischer Faktorenanalysen (CFA) mittels Maximum-Likelihood-Methode zur Überprüfung der Skala empirisch geprüft und konnte auch bestätigt werden. Das Modell weist einen sehr guten Modell-Fit auf ($\chi^2 = 1.76$; $df = 2$; $p = .42$; CFI = 1.000; RMSEA = .000; SRMR = .013). Die ersten Ergebnisse zeigen, dass kognUm für die Erhebung der Unterrichtsunaufmerksamkeit geeignet ist und zur Erforschung von unterrichtsmeidenden Verhaltensformen von Schüler/innen beitragen kann.

Die vorliegende Untersuchung ist eine Pilotstudie und umfasst einen Messzeitpunkt. Untersucht wurden Schüler/innen aus den Jahrgangsstufen 7, 8 und 9 einer Gemeinschaftsschule in Schleswig-Holstein (N = 170). Zur Erfassung der (problematischen) Mediennutzung wurde die Computerspielabhängigkeitsskala KFN-CSAS-II von Rehbein, Kleimann und Mößle (2010) sowie die Skala zur Internetabhängigkeit (ISS) von Hahn und Jerusalem (2001) verwendet. Die Datenauswertung erfolgte mit Hilfe linearer Regressionsmodelle (OLS-Regressionen).

Die Ergebnisse zeigen einen signifikanten geschlechtsspezifischen Unterschied in der kognitiven Unterrichtsmeidung ($t(163) = -2.12$, $p < .05$; MMädchen = 2.1, SD = .80; MJungen = 2.3, SD = .65; $d = .33$). Weiterhin lassen sich bei den Mädchen statistisch bedeutsame Einflüsse des akademischen Selbstkonzeptes, der negativen Schüler/innen-Schüler/innen-Beziehung (Konkurrenz und soziale Desintegration) und der Internetabhängigkeit von mittlerer Effektstärke belegen. Bei den Jungen identifiziert sich ein mittelstarker Netto-Effekt der Computerspielabhängigkeit als mediennutzungsbezogener Einfluss bezüglich der kognitiven Unterrichtsmeidung.

Zusammenfassend bedeutet es, dass im Fokus des Beitrags einerseits die Vorstellung einer Skala zur Messung kognitiver Unterrichtsmeidung (Unterrichtsunaufmerksamkeit) steht und andererseits werden Ergebnisse einer Pilotstudie bezüglich der Untersuchung kognitiver Unterrichtsmeidung von Schüler/innen im Fokus ihrer Mediennutzung präsentiert.

Donnerstag
12. März

11:15 Uhr

HGB 40

Chair(s)

Biedermann, Horst

Düggeli, Albert; Kinder, Katja

Das Internet und das Berufswahlwissen: Oft genutzt und auch nützlich?

Gegen Ende der obligatorischen Schulzeit müssen sich Jugendliche für einen berufsbezogenen weiteren Bildungsweg entscheiden. Die gelingende Bewältigung dieser Entwicklungsaufgabe (Havighurst, 1972; Dreher & Dreher, 1985) ist für ihr Erwachsenwerden zentral, weil sie Aspekte der gesellschaftlichen Inklusion genauso tangiert, wie die Konzeption einer persönlichen Berufsidentität. Informationsbeschaffung und Wissensaufbau zu persönlichkeits- und berufsbezogenem Wissen gelten als zentrale Bestimmungsgrößen entscheidungstheoretischer Berufs- und Laufbahnkonzeptionen (Brown & Brooks, 1994). Ihre Bedeutung lässt sich auch in lern- bzw. in entwicklungstheoretisch basierten Berufswahlmodellen finden (Lent, Brown & Hackett, 2002; Super, 1994). Wissen kann helfen, schlecht strukturierte Probleme besser einordnen, miteinander in Konflikt stehende Ziele differenzierter analysieren oder Konsequenzen von Entscheidungen besser abschätzen zu können. Umgekehrt wird mangelndes oder defizitäres Wissen mit beruflicher Entscheidungsunsicherheit in Verbindung gebracht (Hellberg, 2009; Häfeli & Schellenberg, 2009). Persönlichkeits- und berufsbezogenes Wissen können Heranwachsende auf unterschiedliche Weise aufbauen, z.B. indem sie personale Ressourcen nutzen wie Eltern, Lehrpersonen, Berufsberatende oder auch Peers (Häfeli & Schellenberg, 2009; Beinke, 2008). Daneben steht ihnen als weitere Möglichkeit das Internet als Informationsquelle zur Verfügung (Kanning, 2013). Aktuelle Studien zeigen, dass für Jugendliche das Internet eine zentrale Nutzungsform für die Informationssuche ist (Feierabend, Karg, Rathgeb & Südwest, 2013; Willemse, Waller, Süss, Genner & Huber, 2012). Auch im Berufswahlprozess wird das Internet als Informationsquelle genutzt. Dazu berichten Herzog et al. (2006) von etwa 40% Lernender der achten und neunten Klassenstufe (N=512), die zur Informationsbeschaffung Internetseiten aufgerufen haben. Bei Beinke (2008) finden sich aus einer Untersuchung zur berufsorientierenden Internetnutzung Hinweise darauf, dass der Grossteil der befragten Jugendlichen (69%) die Qualität der erhaltenen Informationen aus dem Internet als gleich wichtig taxiert, wie anderweitig erhaltene Informationen (N=480). Inwiefern diese Befunde über den jeweiligen Forschungskontext hinausgehend interpretiert werden können, ist noch weitgehend offen, und vor allem stellt sich die berufswahltheoretisch relevante Frage, inwiefern die Nutzungshäufigkeit oder die Nützlichkeits einschätzung berufswahlstatus-spezifisch sind. Aus diesem Grund stehen in diesem Beitrag drei Fragen im Zentrum: (1) Wie häufig nutzen Heranwachsende das Internet zum berufswahl-spezifischen Wissensaufbau? (2) Wie nützlich werden die recherchierten Informationen eingeschätzt? (3) Erklären die IT-Kompetenz, die soziale Herkunft, der Migrationsstatus, das kulturelle Kapital und das schulisch angeeignete Wissen über die Nutzung des Internets die Nützlichkeits einschätzung?

Diese drei Fragen werden auf die in Anlehnung an Herzog (2006) vorgenommene Einteilung der Heranwachsenden in vier Statusformen der Berufswahl bearbeitet (diffuse Orientierung, Konkretisierung, Suche nach einer Anschlusslösung, Konsolidierung). Auf der Basis von Daten, die im Schweizer Ergänzungsteil der laufenden IEA ICIL-Studie (8. Klassenstufe; N ca. 3100) erhoben wurden, zeigte sich, dass im Durchschnitt 13% der Lernenden aus dem 8. Schuljahr das Internet zum Wissensaufbau "oft" nutzen. Ein differenzierter Blick auf jene, die diffus orientiert sind, zeigt, dass im Durchschnitt 7% sagen, das Internet "oft" zu nutzen, dagegen aber über 30% angeben, das Internet zum Wissensaufbau "nie" zu nutzen. Tendenziell ist zu sehen, dass je differenzierter die Vorstellung der Heranwachsenden darüber wird, wie sie sich beruflich oder schulisch weiter orientieren wollen, desto nützlicher sie die im Internet recherchierten Informationen einschätzen. Die Vorhersage dieser Nützlichkeits einschätzung ist in den einzelnen Gruppen aber unterschiedlich. In der Gruppe der diffus orientierten Lernenden erweist sich der Migrationsstatus und das schulisch angeeignete Wissen über die Nutzung des Internets als bedeutsam. Bei Lernenden, die sich im Status der Konkretisierung bzw. der Suche nach einer Anschlusslösung befinden, fallen hingegen eher Aspekte der erfassten IT-Kompetenz und des kulturellen Kapitals ins Gewicht. Bei Lernenden in der Gruppe der Konsolidierung wird keiner der genannten Determinanten bedeutsam. Über alle Gruppen hinweg trägt der Migrationsstatus (HISEI beider Eltern) nicht zur Erklärung der Nützlichkeits einschätzung bei.

Donnerstag
12. März

11:15 Uhr

HGB 40

Chair(s)

Biedermann, Horst

Biedermann, Horst; Wolke, Julia; Oser, Fritz

Neue Medien und virtuelles Lästern – Cyber-Mobbing

Cyber-Mobbing stellt ein nicht zu unterschätzendes Problem in Zeiten der Sozialen Medien dar. Diese Problematik aufnehmend interessiert in diesem Beitrag, (1) wie verbreitet Erfahrungen mit Cyber-Mobbing bei Schüler/innen der 8. Klassenstufe der Schweiz sind, und (2) welche Einstellungen diese gegenüber Cyber-Mobbing haben.

Cyber-Mobbing wird in Anlehnung an gängige Mobbing-Definitionen als eine absichtliche, wiederholte, aggressive und schädigende Handlung gegenüber einer anderen Person mithilfe von Informations- und Kommunikationstechnologie bestimmt (Petermann & Marées, 2013; Langos, 2012; Smith et al., 2008). Besonderheiten des Cyber-Mobbings entstehen aus der räumlichen und zeitlichen Entgrenzung durch die Kommunikationsmedien. Es werden verschiedene Formen von Mobbing unterschieden (Willard, 2007) – wobei sich dieser Beitrag auf die Formen des Bloßstellens, des Ausschließens und der Verunglimpfung (das Verbreiten gehässiger, unwahrer oder verletzender Inhalte) bezieht.

Hinsichtlich des Ausmaßes von Cyber-Mobbings unter Jugendlichen besteht keine einheitliche Befundlage (z.B. Willemse et al., 2012; Petermann & Marées, 2013). Es stellt sich daher zunächst die Frage, wie verbreitet Cyber-Mobbing unter den Schüler/innen der 8. Klassenstufe ist. Hierbei gilt es zwischen unterschiedlichen Formen der Erfahrung zu unterscheiden: Es interessiert, (a) wie viele Achtklässler/innen Cyber-Mobbing in ihrem Umfeld bereits beobachten konnten (sog. Bystander), (b) wie viele schon einmal selbst Opfer von Cyber-Mobbing geworden sind, (c) wie viele Täter/innen es gibt und (d) wie viele bereits sowohl Opfer als auch Täter/innen waren. Unter Differenzierung des Erfahrungshintergrunds wird in einem zweiten Schritt der Frage nachgegangen, wie die Jugendlichen ein Lästern bzw. Bloßstellen, Ausschließen und Verunglimpflichen (über) andere/r Menschen im Netz bewerten und (gegebenenfalls) rechtfertigen.

Die Stichprobe der vorliegenden Studie entstammt der IEA-Studie ICILS in der Schweiz (nationales Modul) und umfasst 3200 Schülerinnen und Schüler der 8. Klassenstufe. Die Erfassung der Erfahrungen mit Cyber-Mobbing geschah in Form einer Fragebogen-Vignette. Die Einstellung gegenüber Cyber-Mobbing wurde anhand von fünf Items erhoben, zu welchen die Hauptkomponentenanalyse (PCA) eine eindimensionale Lösung mit guten Kennwerten extrahierte (Eigenkonstruktion).

Bezogen auf die Gesamtstichprobe berichten 27 Prozent der Jugendlichen von Erfahrungen mit Mobbing, sei es als Beobachtende oder als direkt Beteiligte. Diese Erfahrungen teilen sich dermaßen auf, als dass 18 Prozent nur Beobachtende waren, 6 Prozent nur Täter/innen, 9 Prozent nur Opfer und 2 Prozent sowohl Opfer als auch Täter/innen. Insbesondere die Gruppe der Schüler/innen, die sowohl Erfahrungen als Opfer als auch als Täter/innen haben erscheint mit 2 Prozent klein. Spiegelt man diese Gruppe jedoch nur an jenen Schüler/innen, die bereits aktiv an (Cyber-)Mobbing beteiligt waren, so zeigt sich, dass 27 Prozent aller Täter/innen auch schon einmal Opfer waren, und dass knapp ein Fünftel aller Opfer (19 Prozent) auch schon selbst gemobbt haben.

Die Bewertung bzw. Einstellungen der einzelnen Gruppen gegenüber Cyber-Mobbing fallen unterschiedlich aus. Die Einstellung aller Schüler/innen liegt mit einem Mittelwert von 1.81 (bei einer theoretischen Skalenmitte von 2.5) deutlich im ablehnenden Bereich. Die geringste Ablehnung findet sich für die Schüler/innen, welche nur Erfahrungen als Täter/innen besitzen (2.19). Mit einem ähnlichen Mittelwert von 2.09 zeigt sich die Gruppe der Schüler/innen, welche sowohl Erfahrungen in Täterschaft als auch als Opfer besitzen. Demgegenüber fallen die Ablehnungen von Cyber-Mobbing bei den Gruppen „Unbeteiligte“ (1.71), „nur Beobachtende“ (1.69) sowie „nur Opfer“ (1.67) wesentlich deutlicher aus. Es schälen sich also zwei zu unterscheidende Fraktionen heraus: zum einen die Täter/innen, einschließlich derjenigen, die bereits unter Cyber-Mobbing gelitten haben, und zum anderen die Unbeteiligten, Beobachtenden und Opfer. Die Ergebnisse lassen somit erkennen, dass Cyber-Mobbing durchaus in beachtlichem Ausmaß von Schüler/innen der 8. Klassenstufe erfahren wird und dass die dabei auftretende Täterschaft ein derartiges Handeln als nicht (eindeutig) verwerflich rechtfertigt. Die Ergebnisse werden im Rahmen des Beitrags in vertiefender Analyse diskutiert und im Lichte des schulischen Auftrags hinsichtlich Einführung in und Umgang mit Informations- und Kommunikationstechnologien reflektiert.

Donnerstag
12. März

11:15 Uhr

HGB 40

Chair(s)

Biedermann, Horst

Oser, Fritz; Biedermann, Horst

Neue Medien und Freundschaft in der realen und virtuellen Welt

Peer-Beziehungen – und dabei insbesondere Freundschaften – basieren auf Gleichberechtigung, Gegenseitigkeit, Kooperation und Vertrautheit. Im Umgang mit Gleichaltrigen bzw. Peers vollziehen Kinder und Jugendliche zentrale kognitive, soziale und emotionale Entwicklungen. Gerade die Fähigkeit der sozialen Perspektivenübernahme kann sich nur in engen (Freundschafts-)Beziehungen vollziehen. Nach Selman (1984) verläuft diese ausgehend von der undifferenzierten egozentrischen Perspektivenübernahme, über die differenzierte und subjektive Perspektivenübernahme, der Fähigkeit der wechselseitigen Perspektivenkoordination hin zur Fähigkeit der Übernahme der Perspektive der sozialen Bezugsgruppe. Als Bedingung einer derartigen Entwicklung sieht Selman (a.a.O.) im ausgehenden Kindes- und angehenden Jugendalter das Eingehen von engen Freundschaften im Sinne eines (stabilen) intimen gegenseitigen Austausch gefolgt von einer engen Freundschaft als Autonomie und Interdependenz.

Im Fokus heutiger mediatisierter Lebenswelten, in welchen soziale Onlinenetze (wie Facebook, Instagram, WhatsApp), webbasierte digitale Kommunikationsmedien (wie Foren, Wikis, Web-Blocks) und Video-Games fester Bestandteil des Freizeitverhaltens und der der alltäglichen Kommunikation von Jugendlichen geworden sind – wodurch Lee (2013) auch von einer „Net-Generation“ spricht –, stellt sich die Frage, wie sich dies auf reale soziale Beziehungen bzw. Freundschaften auswirkt (Trost, 2013). So interessiert in diesem Zusammenhang beispielsweise, was es bezüglich Freundschaften bedeutet, wenn Jugendliche soziale Netzwerke für Ihre Kontakte nutzen? In welchem (positiven oder negativen) Verhältnis Freundschaften der realen und der virtuellen Welt stehen? Ob ein (intensives) mediales Handeln Veränderungen in Bezug auf Werte, Verhaltensweisen und Bedeutungszuschreiben – auch in Bezug auf das eigene Verständnis von Freundschaft – mit sich bringt? Welche Bedeutung ein einseitiges Ansiedeln von Freundschaften im virtuellen Raum für die sozial-kognitive Entwicklung hat? Inwieweit ein intensives Nutzen neuer Medien mit Prozessen der Vereinsamung in Beziehung steht?

Im Rahmen dieses Beitrags soll aus diesem Fragenkatalog der Aspekt der Beziehung zwischen virtuellen Freundschaften (z.B. Über das Internet [z.B. Facebook, Twitter, MySpace] lerne ich Freunde kennen, was mir ansonsten schwer fällt – 4 Items, Cronbach $\alpha = .80$) und realem (Freundschafts)Netz (z.B. Wenn ich es brauche, sind meine Freunde für mich da – 4 Items, Cronbach $\alpha = .86$) aufgenommen werden. Die Datenerhebung geschah in Form eines nationalen Moduls (Schweiz) im Rahmen der internationalen Ländervergleichsstudie ICILS (International Computer and Information Literacy Study) der IEA (International Association for the Evaluation of Educational Achievement). Insgesamt wurden dabei im Jahr 2013 eine zufällige Auswahl von 3200 Schülerinnen und Schülern der 8. Klassenstufe der Schweiz befragt.

Die ersten Ergebnisse lassen erkennen, dass es einem Großteil der jungen Menschen (eher) einfacher fällt, in der virtuellen Welt Bekanntschaften und Freundschaften als in der realen Welt zu finden und zu pflegen. Demgegenüber fallen die Aussagen hinsichtlich stabiler und verlässlicher Freundschaften in der realen Welt (eher) skeptisch aus. Die beiden Betrachtungen von Freundschaften in Beziehung setzend zeigt sich eine negative Wechselwirkung, je höher virtuelle Freundschaften bewertet werden, desto tiefer fallen Bewertungen zu realen Freundschaften aus ($r = -.30$).

Diese beiden Freundschaftswelten werden im Rahmen des Beitrags in vertiefender Analyse sowohl einzeln als auch in deren Beziehung betrachtet und diskutiert sowie hinsichtlich der Bedeutung der sozial-kognitiven Entwicklung und damit verbunden des Freundschaftskonzepts reflektiert.

Donnerstag
12. März

11:15 Uhr

HGB 40

Chair(s)

Biedermann, Horst

Gretsch, Petra

Visualisierungen im Deutsch- und Mathematikunterricht (VisDeM)

Das Promotionskolleg VisDeM (Visualisierungen im Deutsch- und Mathematikunterricht) ist interdisziplinär ausgerichtet und bezieht neben den Fachdidaktiken Deutsch und Mathematik auch die kognitive und pädagogische Psychologie in Konstrukt- bzw. Modellbildungen mit ein. Visualisierungen übernehmen im didaktischen Zusammenhang unterschiedliche Funktionen, wobei wir uns primär auf informationstragende Visualisierungen fokussieren. Die Interpretation, Verwendung und Konstruktion von Visualisierungen erfordert Visualisierungskompetenzen auf Seiten aller am Bildungsgeschehen Beteiligten. Die erste Annäherung an unterrichtsbezogene Visualisierungen konzentriert sich innerhalb unseres Promotionskollegs auf die Schülerperspektive und zunächst auf Fragen der Rezeption.

Folgende Punkte werden uns näher beschäftigen:

- (1) Wie muss ein Modell zur Visualisierungskompetenz -- hier fokussiert auf den schulischen Kontext für den Deutsch- und Mathematikunterricht -- beschaffen sein? Welche Kompetenzfacetten lassen sich identifizieren? (Beitrag von S. Wafi & M. Wirtz)
- (2) Welche Zusammenhänge zeigen sich bei Lernenden zwischen Präferenzen, Selbstwirksamkeitsüberzeugungen und der Leistung im Umgang mit Repräsentationen? (hier beschränkt auf die Repräsentation von linearen Funktionen; Beitrag U. Dreher, T. Leuders, L. Holzäpfel)
- (3) Wie lassen sich Überzeugungen von Studierenden und SchülerInnen durch den Einbezug angemessener graphischer Repräsentationen in einem spezifischen Wissensbereich fördern? (hier bezogen auf eine Interventionsstudie, welche Baumdiagramme und Einheitsquadrat in der Wahrscheinlichkeitsrechnung nutzt; Beitrag von A. Sturm & A. Eichler)
- (4) Wie bewerten Lernende Bilder, die narrative Texte begleiten? Korrelieren Bildpräferenz und Lesefähigkeit in kompensatorischer Weise (schwache Lesefähigkeit korreliert mit hoher Bildpräferenz)? Welche Funktionen können dekorativen und darstellenden Bildern im narrativen Kontext generell zugeschrieben werden? (Beitrag M. Kist)

Die Untersuchungen konzentrieren sich um Sekundarstufenschüler_innen (Sek I) und weisen unterschiedliche Mixed-Method-Designs auf.

Donnerstag
12. März

11:15 Uhr

HGB 50

Chair(s)

Gretsch, Petra
Holzäpfel, Lars

Dreher, Ulrike; Leuders, Timo; Holzäpfel, Lars

Zusammenhänge beim Umgang von Lernenden mit graphischen und numerischen Repräsentationen von Funktionen im Mathematikunterricht (Sek I)

Das Teilprojekt im Promotionskolleg VisDeM (Visualisierungen im Deutsch- & Mathematikunterricht) fokussiert die aufgaben- und individuumsabhängigen Faktoren der adaptiven Nutzung von Repräsentationen zur Lösung von Aufgaben im Bereich Funktionen im Fach Mathematik bei Schülerinnen und Schülern der 8. Klasse an Realschulen. Dabei soll der Zusammenhang der Präferenzen der Lernenden und der Selbstwirksamkeitsüberzeugungen (vgl. Gagatsis et al. 2009) mit ihren repräsentationalen Fähigkeiten analysiert werden.

In den letzten Jahren wurden zahlreiche Studien zur Erfassung und Beschreibung von Kompetenzen zum Umgang mit Funktionen und deren verschiedenen Repräsentationen vorgelegt (Bayrhuber et al. 2010; Acevedo Nistal et al. 2012). Während Bayrhuber et al. (2010) die Übersetzungsfähigkeit zwischen verschiedenen Repräsentation bei Siebt- und Achtklässlern erfassten und in Kompetenzdimensionen einteilen konnten, ging es Acevedo Nistal et al. (2012) um die Erfassung des flexiblen Umgangs mit Repräsentationen von Neunt- bis Elftklässlern. Beide Studien zeigen, dass sich der Umgang mit Repräsentationen über die Schuljahre hinweg verändert und dass verschiedene Fähigkeitsfacetten messbar und trennbar sind. In einer Interviewstudie von Acevedo Nistal et al. (2013) mit linearen Funktionen konnten zudem aufgrund selbstberichteter Urteile der Schülerinnen und Schüler verschiedene Wirkfaktoren (individuums-, aufgaben-, kontextspezifisch) auf die Nutzung herausgestellt werden, aber noch nicht systematisch mit den tatsächlichen Leistungen in Beziehung gesetzt werden. Als weiteren Einflussfaktor auf die Leistungen im Umgang mit Repräsentationen konnten Gagatsis et al. (2009) im Themengebiet Bruchrechnung die Bedeutung von Selbstwirksamkeitsüberzeugungen herausstellen und auch zeigen, dass sich die Selbstwirksamkeitsüberzeugungen bezüglich der verschiedenen Repräsentationen von Primar- zu Sekundarstufe tendenziell rückläufig entwickeln. Dieser Aspekt soll als weiterer Moderator im Wirkgefüge repräsentationaler Kompetenzen untersucht werden. Das Konstrukt der „Metarepräsentationalen Kompetenzen“ (diSessa et al. 2000) ist dazu geeignet, den Umgang mit Repräsentationen umfassend zu beschreiben. Unklar ist bisher, in welchem Zusammenhang die verschiedenen Kompetenzfacetten (Übersetzungsfähigkeiten, adaptive Kompetenzen, Metarepräsentationale Kompetenz, Präferenzen, Selbstwirksamkeit) stehen, dies gilt es in der vorliegenden Studie aufzudecken. Die Hauptfragestellung gilt der Klärung des Zusammenhangs zwischen den genannten Faktoren:

Welcher Zusammenhang besteht bei Lernenden zwischen ihren Kompetenzen bezüglich der Übersetzung verschiedener Repräsentationen, ihren Präferenzen für einzelne Repräsentationen und der flexiblen Nutzung mehrerer Repräsentationen im Bereich Funktionen?

Diese Frage soll mit Hilfe eines Mixed-Methods-Designs aufgeklärt werden. In der Vorstudie wurden aufgabenbasierte Interviews zu unbewussten und bewussten Präferenzen für einzelne Repräsentationen erhoben und systematisiert sowie metarepräsentationales Wissen (vgl. diSessa et al. 2000) erfragt. Skalen zu Selbstwirksamkeitsüberzeugungen bezüglich der Repräsentationen von Funktionen und zu Präferenzen der Lernenden wurden in diesem Schritt entwickelt und optimiert (n = 95).

In der Hauptstudie werden zunächst die quantitativen Instrumente eingesetzt und anschließend eine vertiefende Analyse durch den Einsatz qualitativer Instrumente sichergestellt. Im quantitativen Teil der Studie wird in 15 Schulklassen an 6 Realschulen sowohl ein Testinstrument zur Messung der Übersetzungsfähigkeit zwischen Repräsentationen (HEUREKO-Test von Bayrhuber et al. 2010) und offene Items mit Reflexionsfragen eingesetzt. Des Weiteren dienen zwei Fragebogenerhebungen dazu, die bewussten Präferenzen für einzelne Repräsentationen und Selbstwirksamkeitsüberzeugungen bezüglich der Repräsentationen von Funktionen und allgemeine mathematische Selbstwirksamkeitsskalen (Ramm et al. 2006) zu erheben. In einem zweiten Schritt werden die Testergebnisse miteinander in Beziehung gesetzt. In einer Clusteranalyse werden differente Profile bzw. deren Vertreter herausgestellt. Im anschließenden qualitativen Teil werden Einzelinterviews mit Extremfällen geführt, um vertiefende Analysen der einzelnen Profile liefern zu können.

Der Schwerpunkt des Vortrags bilden die Vorstudienenergebnisse und ein erster Ausblick in die derzeit durchgeführte Hauptstudie.

Donnerstag
12. März

11:15 Uhr

HGB 50

Wafi, Sammy; Wirtz, Markus

Strukturmodellierung von Visualisierungskompetenz in der Orientierungsstufe: Eine literatur- und expertenbasierte Herleitung

Die Interpretation, Verwendung und Konstruktion visuell abbildbarer Informationen ist für schulische Lernprozesse von hoher Bedeutung (Schnotz 2010; Mayer 2005; Weidenmann, 1994). Neben Anforderungen an die Gestaltung der Visualisierungen selbst bedarf es zielgerichteter Untersuchungen zu den erforderlichen und vorhandenen Fähigkeiten auf Seiten der Schülerinnen und Schüler bei der Arbeit mit Visualisierungen - kurz Visualisierungskompetenz.

Während Kompetenzen im Umgang mit Texten (Literacy) und die dazu erforderlichen mentalen Modelle Gegenstand vielfältiger Forschungstätigkeiten sind, ist der Kenntnisstand zum Umgang mit Visualisierungen nicht gleichermaßen ausgeprägt (Ploetzner et al. 2013). Inzwischen liegt hinreichende empirische Evidenz vor, dass Visualisierungskompetenz wegen grundlegender Unterschiede in der Verarbeitung und in den mentalen Systemen bei Text- und Bilddarstellungen in Abgrenzung zu Literacy aufgefasst werden muss (Schnotz & Bannert 2003). Beispielsweise ist bei Visualisierungen im Gegensatz zu Texten für die semantische Verarbeitung im Allgemeinen keine bestimmte Bearbeitungsreihenfolge vorgesehen und es werden gemäß der Theorie der dualen Kodierung nach Paivio (1990) unterschiedliche Verarbeitungskanäle angesprochen.

Obwohl Visualisierungen zunehmend in der Bildungswissenschaftlichen Forschung berücksichtigt werden, existiert bisher keine pädagogisch-psychologische und fachdidaktische Modellbasis, welche die erforderlichen Visualisierungskompetenzen zur lernförderlichen Verarbeitung in disziplinübergreifender Anwendungsperspektive beinhaltet. Die längere, multidisziplinäre Forschungstradition im anglo-amerikanischen Raum zum eng verwandten Konzept der Visual Literacy (Debes 1970, Wileman 1993, Avgerinou 2001) kann nur in Teilen dazu dienen, da die Thematik selten konzeptionell und empirisch mit einem Fokus auf den Schulkontext untersucht wurde.

Aus diesem Grund fokussiert das Projekt Kompetenzen zum lernförderlichen Umgang mit externalen, statischen Visualisierungen im schulischen Kontext. Es strebt die Entwicklung eines Strukturmodells von Visualisierungskompetenz an, das (1) den internationalen erziehungswissenschaftlichen Diskurs zum Thema Visualisierungskompetenz integriert und (2) die Sichtweisen von Expertinnen und Experten im Fokus auf Deutsch- und Mathematikunterricht berücksichtigt. Die übergreifende Forschungsfrage lautet: Welche Kompetenzstruktur lässt sich für den Umgang mit Visualisierungen in den Fächern Deutsch und Mathematik in der Sekundarstufe I (mit Fokus auf der Orientierungsstufe) aus theoretischen und empirischen Erkenntnissen identifizieren?

Als Ausgangspunkt zur Konstruktion des Strukturmodells diente eine literaturbasierte Identifikation relevanter Kompetenzfacetten mit synoptischen Vergleichen. Die folgende Expertenbefragung in Form leitfadenbasierter Einzelinterviews mit Expertinnen und Experten im Bereich der Lehr-/Lernforschung und fachdidaktischen Visualisierungsforschung sowie mit Lehrpersonen mit fachdidaktisch-spezifischem Erfahrungswissen (N=14) führte in einem mehrstufigem Vorgehen zur Entwicklung eines fächerübergreifenden Strukturmodells von Visualisierungskompetenz. Methodisch wurden dabei Impulsfragen, maximale Offenheit in der Befragung, Strukturlegetechnik und Think-Aloud-Methode eingesetzt. Das resultierende mehrdimensionale Strukturmodell - bestehend aus den vier Hauptdimensionen des Erkennens, Verstehens, Verknüpfens und Generierens von Visualisierungen - trägt zur konzeptuellen Klärung von Visualisierungskompetenz im Schulkontext bei und liefert die Grundlage zur Entwicklung eines empirisch fundierten Diagnoseinstrumentariums.

Donnerstag
12. März

11:15 Uhr

HGB 50

Chair(s)

Gretsch, Petra
Holzapfel, Lars

Sturm, Alexandra; Eichler, Andreas

Förderung der Überzeugungen von Studierenden und SchülerInnen zur Anwendbarkeit ihres statistischen Wissens mittels grafischer Repräsentationen

Medientexte und andere Informationsquellen, die statistische Informationen enthalten, sind in unserem Alltag nahezu omnipräsent. Daher wird etwa vom Mathematical Science Education Board (1990) der Umgang mit quantitativen Daten als unabdingbare Kompetenz gesehen: „Citizens who cannot properly interpret quantitative data are, in this day and age, functionally illiterate“ (S. 8). Tatsächlich scheinen aber Erwachsene diese Kompetenz oftmals nicht zu besitzen (Gal, 2004) und SchülerInnen der Bedeutung dieser Kompetenz nicht bewusst zu sein. In einer Studie mit 20 SchülerInnen (Sek II) fand Eichler (2008) eine erhebliche Diskrepanz zwischen einem Anstieg des statistischen Wissens und einer Förderung der statistikbezogenen Überzeugungen. So nahm das Wissen der SchülerInnen innerhalb eines halben Jahres deutlich zu, während die statistikbezogenen Überzeugungen davon weitgehend unberührt blieben. Die meisten von Ihnen maßen der Statistik weder eine Bedeutung für gesellschaftliche Entscheidungen noch für ihr eigenes Leben bei.

Im Rahmen dieses Forschungsprojekts liegt daher der Fokus auf der Förderung der Statistical Literacy in dem Sinne, dass sowohl Dispositionen als auch Wissens Elemente dieses Konstrukts beachtet werden (Gal, 2004), wobei der Schwerpunkt auf der dispositionalen Facette liegt. Es wird insbesondere die Frage untersucht, ob sich die Überzeugungen von Studierenden und SchülerInnen zur Anwendbarkeit ihres statistischen Wissens durch eine kurze Intervention fördern lassen. Dabei wurde als inhaltliches Thema der Satz von Bayes ausgewählt. Drei Gründe sprechen für diese Wahl: So existieren zum Satz von Bayes Anwendungen, welche die Bedeutung der Statistik deutlich machen können. Des Weiteren scheint es möglich, in Kurzinterventionen, die sich auf graphische Repräsentationen des Baumdiagramms (Bea, 1995) und des Einheitsquadrats (Sedlmeier & Gigerenzer, 2001) stützen, erhebliche Lernerfolge zu ermöglichen. Schließlich lässt sich zusätzlich untersuchen, ob sich die Wahl der Repräsentation auf die Art des geförderten Wissens auswirkt. So wird angenommen, dass das Baumdiagramm hinsichtlich des Aufbaus von prozeduralem Wissen überlegen ist, während durch das Einheitsquadrat eher konzeptionelles Wissen gefördert wird (Böcherer-Linder & Eichler, 2014). Daraus ergibt sich die zweite Forschungsfrage, nämlich welche dieser zwei konkurrierenden graphischen Repräsentationen in dieser Kurzintervention hinsichtlich des Lernerfolgs überlegen ist.

Im Rahmen der Hauptstudie wird an insgesamt ca. jeweils 60 Studierende aus unterschiedlichen Disziplinen sowie ca. 60 SchülerInnen die Wirkung einer Intervention zum Satz von Bayes in einem Pre-Post-Follow up-Design untersucht. Die Stichprobenauswahl begründet sich darin, dass ein breites Spektrum mit dieser Intervention adressiert werden soll. Es sollen nicht nur Studierende mit direktem Mathematikfokus (Mathematikstudierende), sondern auch mathematikferne Studierende (Deutschstudierende) sowie Studierende, die in ihrem späteren Berufsfeld Statistik potentiell anwenden werden (Studierende der Gesundheitspädagogik), in den Blick genommen werden.

Sowohl im Pre-Test als auch im Post- bzw. Follow up-Test werden die Überzeugungen der Studierenden und Schüler mittels Fragebogen erhoben. Dabei werden neben den statistikbezogenen Überzeugungen (vgl. u.a. Schau et al., 1995) auch mathematikbezogene Überzeugungen erhoben (Grigutsch, 1996). Als Kovariaten werden weiterhin das Mathematikinteresse und das Selbstkonzept sowie schließlich das Wissen der Studierenden und Schüler zum Satz von Bayes erhoben. Um der zweiten Forschungsfrage nachzugehen, werden die Studierenden randomisiert zwei Untersuchungsbedingungen zugeteilt, die sich lediglich in der Wahl der Visualisierung (Baumdiagramm, Einheitsquadrat) unterscheiden.

Im Rahmen des Vortrags wird der Fokus auf ersten Ergebnissen der Hauptstudie liegen.

Donnerstag
12. März

11:15 Uhr

HGB 50

Chair(s)

Gretsch, Petra
Holzapfel, Lars

Kist, Magdalena

Bilder als „verführerische Details“ oder Verstehenshilfe für narrative Texte? Einstellungen von Schülerinnen und Schülern bezogen auf Text-Bild-Kombinationen

Die Theorie der dualen Kodierung (Paivio, 2004) sowie das integrative Modell des Text- und Bildverständnisses (Schnotz, 2005) zeigen, dass sprachliche Informationen aus Sachtexten mithilfe von instruktionalen, d.h. vornehmlich informativen Bildern, leichter verarbeitet und erinnert werden können und demnach kognitive Lernprozesse positiv beeinflussen.

Aber auch narrative Texte, wie zum Beispiel Fabeln und Märchen, werden häufig mit Bildern kombiniert. Eine in der Vorstudie bereits durchgeführte Analyse von Lehrwerken des Faches Deutsch aus mehreren Bundesländern hat ergeben, dass es sich dabei um Zeichnungen, Fotografien und Bildabfolgen handelt, welche überwiegend eine rein dekorative oder eine darstellende Funktion übernehmen. In darstellenden Bildern werden Figuren, Objekte und Aktivitäten aus narrativen Handlungsszenen gezeigt. Daher stehen sie besonders häufig in Verbindung mit narrativen Texten, treten aber vereinzelt auch bei expositorischen Texten auf (Levin et al., 1987).

Dekorative Bilder zeigen, was den kognitiven Bereich anbelangt, negative Effekte. Es werden kognitive Ressourcen beansprucht, ohne dass relevante Informationen im Bild enthalten sind, weshalb sie auch als „seductive details“ bezeichnet werden (Harp & Mayer 1997; 1998). Vor allem im affektiven Bereich konnten aber positive Effekte bestätigt werden, wie z. B. eine positive Stimmung der Lernenden oder ein geringer wahrgenommener Schwierigkeitsgrad des Lernmaterials (Lenzner, 2009; Lenzner, Schnotz & Müller, 2013).

Wie die aktuelle Schulbuchforschung bestätigt, sind dekorative Bilder und Illustrationen in immer größerer Anzahl in Lernmaterialien vorhanden (Pettersson, 2010). Es ist anzunehmen, dass diese eingesetzt werden, um narrative Texte ästhetisch ansprechend zu gestalten und so die Lesemotivation und Lesefreude der jungen Leser zu steigern. Ob dies erfolgreich gelingt, wird mittels der ersten Forschungsfrage „Wie bewerten Schülerinnen und Schüler Bilder, die narrative Texte begleiten?“ empirisch überprüft.

Inwiefern jedoch das Verstehen narrativer Texte durch dekorative und vor allem durch darstellende Bilder beeinflusst wird, blieb bisher ungeklärt. Die bisherigen Forschungsergebnisse beziehen sich auf das Lernen mit Text-Bild-Kombinationen aus instruktionalen Bildern und Sachtexten. Daraus ergibt sich die zweite Forschungsfrage dieser Studie: „Welche Funktion(en) haben dekorative und darstellende Bilder, die in Verbindung mit narrativen Texten stehen?“.

In einer Vorstudie wurden Schülerinnen und Schüler der 4. bis 6. Klasse, im Rahmen eines Mixed-Methods-Designs mittels Fragebögen (N=173) und halbstrukturierten Interviews (N=42), zu ihrer allgemeinen Einstellung in Bezug auf verschiedene Bildtypen und zu ihren Bildpräferenzen befragt. Einstellungen liegen implizit vor und können daher nicht direkt erfasst werden. Sie äußern sich jedoch immer in einer Bewertung des Einstellungsobjektes (Eagly & Chaiken 1993). Demnach bewerteten Schülerinnen und Schüler Bilder aus aktuellen Lehrwerken auf Basis einer Sortieraufgabe. Sie erläuterten ihre Präferenzen in Bezug auf Bildtypen und Bewertungskriterien in einem Interview. Der Fragebogen diente dazu, einzelne Facetten der Einstellung zum Lesen von narrativen Texten mit Bildern zu erfassen.

In einer zweiten, derzeit laufenden Hauptstudie werden auch die kognitiven Effekte der verschiedenen Bildtypen, die mit narrativen Texten gemeinsam auftreten, untersucht. Ziel ist es, die Text-Bild-Integration beim Lesen narrativer Texte und die Effekte der verschiedenen Bildtypen auf das Leseverstehen zu untersuchen. Hierbei wird auch die allgemeine Lesefähigkeit der Studienteilnehmer mittels eines standardisierten Tests zum Leseverständnis erfasst. Dies ist vor allem deshalb interessant, da Forschungsergebnisse darauf hindeuten, dass vor allem schwache Leser mehr von Bildern profitieren als starke Leser (Levie & Lentz, 1982).

Im Vortrag wird der Fokus auf die Ergebnisse der Vorstudie gelegt und über erste Ergebnisse der Hauptstudie berichtet.

Donnerstag
12. März

11:15 Uhr

HGB 50

Chair(s)

Gretsch, Petra
Holzapfel, Lars

Schmidt-Hertha, Bernhard

Heterogene Informationsumwelten und individuelle Statusübergänge

Informationsumwelten Studierender sind typischerweise aus familiären und schulischen Netzwerken zusammengesetzt, die mit fachspezifisch akademischen und studentischen "peer groups" und deren Wissens- und Beurteilungsstandards konfrontiert werden. Diese Erfahrungen verändern potentiell das Informations- und Entscheidungsverhalten in zunehmend digitalisierten Informationsumwelten. Die Analyse der individuellen Konsequenzen dieser beiden Entwicklungen, deren Abhängigkeit von sozialen Kontexten sowie deren Rückwirkungen auf soziale Ungleichheiten stehen im Zentrum der Arbeiten einer interdisziplinären Forschergruppe, die Statuspassagen in Bildungskontexten mit Blick auf die Relevanz moderner Informationsumwelten untersucht (vgl. Hillmert et al., 2014). Im Mittelpunkt steht die Frage nach der Bedeutung dieser Informationsumwelten bzw. ihrer individuellen Nutzung und Bewertung für Übergänge im Kontext eines Studiums, wobei sowohl die Studieneingangsphase wie auch Übergänge im Verlauf des Studiums und die Vorbereitung auf den Berufseintritt in den Blick genommen werden. In den einzelnen Beiträgen des Symposiums geht es erstens um die Frage nach der Bewertung dieser Informationen im Hinblick auf Gerechtigkeitsevaluationen studienbezogener Leistungsbeurteilungen und es werden in diesem Kontext auch Dropout-Risiken in den Blick genommen (Groß/Burger). Zweitens kann Dropout im Studium auch eine Frage der Identifikation bzw. Disidentifikation mit dem Studierendenstatus sein, wobei sich u.a. die Frage stellt inwieweit Disidentifikation mit der Nutzung von Informationsressourcen interagiert (Matschke/de Vreeze). Schließlich wird davon ausgegangen, dass die effiziente Nutzung moderner Informationsumwelten spezifische Kompetenzen voraussetzt, wobei sich die diesbezüglichen Anforderungen studiengangsspezifisch differieren. Daher fragt das Symposium drittens nach der Entwicklung von Medienkompetenzfacetten im Studium und deren Anschlussfähigkeit an Anforderungen des avisierten Berufsfelds (Schmidt-Hertha/Rott). Viertens wird auf Basis einer anderen aktuellen Studie die Heterogenität von Studienstandorten in verschiedenen sozialwissenschaftlichen Disziplinen diskutiert, und nach der Bedeutung dieser Heterogenität für Statusübergänge gefragt (Tafertshofer/Werner).

Die verschiedenen Fragen werden auf Basis einer gemeinsamen Befragung von über 2500 Studierenden sowie verschiedenen weiteren Erhebungen bearbeitet.

Das Thema Heterogenität spielt im Symposium eine zentrale Rolle und wird aus mindestens vier Perspektiven aufgegriffen: (a) Die Heterogenität der Voraussetzungen, die Studierende mitbringen (b) Die Heterogenität der Nutzung von Informationsumwelten, d.h. die Frage nach Ursachen und Wirkungen eines immer selektiven und mit Wertungen verbundenen Informationsverhaltens (c) Die Heterogenität von Studienfächern und disziplinären Informationsumwelten Die Heterogenität disziplinärer Perspektiven auf einen Forschungsgegenstand, die in der Zusammensetzung der aus Soziologen, Psychologen und Erziehungswissenschaftlern bestehenden Forschergruppe verankert ist.

In dem Symposium werden theoretische Bezugspunkte und empirische Ergebnisse zu den vier Teilprojekten vorgestellt und diskutiert, wobei die Projekte nicht nur hinsichtlich sich aufeinander beziehender Fragestellungen, sondern in drei Fällen auch durch eine gemeinsame Datenbasis miteinander verbunden sind.

Donnerstag
12. März

11:15 Uhr

HGC 20

Chair(s)

Schmidt-Hertha, Bernhard

Burger, Roland; Groß, Martin

Heterogenität in der Bewertung von Studienbedingungen und ihre Konsequenzen für den Studienerfolg

Das Projekt „Gerechtigkeit im Studienverlauf“ beschäftigt sich mit der Frage, welchen Einfluss die Gerechtigkeitsevaluationen von Studierenden hinsichtlich ihrer Studienbedingungen auf ihre Bindung an die Universität, ihre Zufriedenheit mit den Studienbedingungen und damit auf ihren Studienerfolg (hier operationalisiert mit Dropout-Intentionen) haben.

In der vorliegenden Studie steht die prozedurale Gerechtigkeit im Fokus (vgl. Thibaut & Walker, 1975; Leventhal, 1980). Diese Gerechtigkeitsdimension nimmt Bezug auf die Verfahren, welche DozentInnen bei der Bewertung von Studienleistungen verwenden. Aus Sicht der prozeduralen Gerechtigkeit sind Benotungsverfahren gerecht, wenn die DozentInnen bei der Benotung unvoreingenommen sind, einheitliche Standards anwenden sowie versuchen, den Wissensstand der Studierenden so gut wie möglich zu erfassen (vgl. Colquitt, 2001).

Ein Verfehlen dieser Kriterien führt dazu, dass Studierende systematisch daran gehindert werden, eine für ihre Anstrengungen angemessene Benotung zu erhalten. Wenn Noten ungeachtet guter Leistungen schlecht ausfallen, ist das Erreichen angemessener Noten außerhalb der Kontrolle der Studierenden. So entsteht ein Investitionsrisiko. Da die Chance eines erfolgreichen Studiums so verringert wird erwarten wir, dass die Erfahrung mangelnder prozeduraler Gerechtigkeit zu erhöhten Abbruchsintentionen führt.

In diesem Zusammenhang wird Heterogenität in zwei Formen relevant. Dabei handelt es sich erstens um die Heterogenität des Umfeldes, innerhalb dessen die Gerechtigkeitsevaluationen getätigt werden. Hier zeigt sich, dass die Gerechtigkeitsbeurteilung der Studierenden zwischen universitären Kontexten variiert: Strukturelle Rahmenbedingungen von Fächern einerseits, sowie in den Verhaltensweisen der Lehrenden zum Vorschein kommende Fachkulturen andererseits erzeugen kontextspezifische „Justice Climates“ (vgl. Naumann & Bennett, 2000). Studierende tauschen sich mit ihren KommilitonInnen über ihre Erfahrungen aus, was zu einer Reevaluation der Erfolgchancen führen kann. Das weist auf die Notwendigkeit hin zu überprüfen, inwiefern der Einfluss individueller Gerechtigkeitsurteile von dem fachspezifischen Justice Climate moderiert wird. Zweitens ist zu prüfen, inwiefern der Einfluss der prozeduralen Gerechtigkeit zwischen Studierenden mit unterschiedlicher sozialer Herkunft Differenzen aufweist. Das mit einer ungerechten Behandlung einhergehende Investitionsrisiko wirkt sich insbesondere für solche Studierenden negativ aus, welche die mit einem Scheitern verbundenen Kosten schlechter kompensieren können. Deshalb untersuchen wir, ob die Abbruchsintentionen von Studierenden aus Elternhäusern mit geringem sozialem Status sowie von Studierenden mit Migrationshintergrund in größerem Umfang von Gerechtigkeitsevaluationen beeinflusst sind.

Daten: Die Studie beruht auf Daten des CampusPanels. Dabei handelt es sich um die erste Welle einer Onlinebefragung, die im Wintersemester 2013/14 unter den Studierenden einer der größten deutschen Universitäten durchgeführt wurde. Die Umfrage entstand aus einer Kollaboration von vier Teilprojekten eines interdisziplinären Forschungsclusters. Insgesamt haben n=2564 Studierende die Umfrage vollständig ausgefüllt. Aufgrund des großen Samples ist es uns möglich, differenzierte Analysen auch auf Fachebene vorzunehmen. Neben den Gerechtigkeitsevaluationen enthalten die Daten detaillierte Informationen zu Studien- und Informationsverhalten sowie zur Identifikation mit der Rolle als Studierende. Ebenso liegen umfassende Angaben über die soziale Herkunft vor.

Methode: Die Bildung der zentralen Untersuchungskonstrukte erfolgt auf Basis explorativer und konfirmatorischer Faktorenanalysen sowie darauf aufbauenden regressionsanalytischen Verfahren.

Ergebnisse: Selbst unter Kontrolle weiterer relevanter Einflussgrößen (insbesondere der Zufriedenheit mit der akademischen Leistung) verringert sich das Risiko, Abbruchsintentionen zu haben, wenn die prozedurale Gerechtigkeit höher eingeschätzt wird. Diese Beziehung wird von dem fachspezifischen Justice Climate moderiert: In Studienfächern, in welchen prozedurale Gerechtigkeit im Schnitt als niedrig eingeschätzt wird, wirken die individuellen Urteile sich stärker auf die Abbruchsintentionen aus. Ebenso konnte gezeigt werden, dass die Abbruchsintentionen von Studierenden aus sozial schwachen Elternhäusern sowie von Studierenden mit Migrationshintergrund deutlich stärker von deren Gerechtigkeitsevaluationen beeinflusst. Die Ergebnisse stützen damit die Annahme der hohen Bedeutung der Heterogenität des Umfeldes, innerhalb dessen Gerechtigkeitsevaluationen entstehen sowie der Heterogenität der Reaktionen, welche die Evaluationen hervorrufen.

Donnerstag
12. März

11:15 Uhr

HGC 20

Chair(s)

Schmidt-Hertha, Bernhard

Matschke, Christina; de Vreeze, Jort

„Ich gehöre nicht hierher!“ Der Einfluss von Inkompatibilität zwischen sozialem Hintergrund und der Universität auf Disidentifikation, Diskriminierung und den akademischen Werdegang

Der Statusübergang zur Universität bedeutet häufig nicht nur einen Ortswechsel, sondern auch eine Veränderung der Bezugsgruppe. Neben Familie und Freunden aus der Schulzeit erschließen sich neue Netzwerke aus Studierenden an der Universität. Im besten Fall entwickeln Erstsemester eine soziale Identifikation mit der Gruppe der Studierenden, d.h. sie beschreiben, verhalten und fühlen sich im Sinne der Gruppe (Tajfel & Turner, 1986). Im schlechtesten Fall entwickelt sich eine sog. Disidentifikation, d.h. sie beschreiben, verhalten und fühlen sich im Gegensatz zur Gruppe (Elbach & Bhattacharya, 2001; Matschke & Sassenberg, 2010a). In drei Studien wird erstmals untersucht, wie eine Disidentifikation zur neuen Bezugsgruppe der Studierenden entsteht und welche Konsequenzen diese für Informationsverhalten und den akademischen Werdegang hat.

Die Forschung hat gezeigt, dass insbesondere Studierende mit niedrigem Bildungshintergrund Inkompatibilität zwischen ihrer Herkunft und dem akademischen Leben erleben. Diese Inkompatibilität wirkt sich erschwerend auf die Ausbildung einer gesunden Beziehung zur neuen Bezugsgruppe der Studierenden aus (Iyer, Jetten, Tsivrikos, Postmes, & Haslam, 2009). Auf der Basis dieser Befunde wird angenommen, dass Inkompatibilität die Disidentifikation mit Studierenden erhöht. Eine soziale Identifikation stärkt den Einfluss der Gruppe sowie die Einhaltung der Normen der Gruppe (z.B. Sassenberg, Matschke, & Scholl, 2011), während eine Disidentifikation dazu führt, dass das Gegenteil der Gruppennorm angestrebt wird (Matschke & Sassenberg, 2010b). Daher wird angenommen, dass Disidentifikation mit Studierenden den Informationseinfluss anderer Studierender herabsetzt und die Leistung im Studium mindert. Frühere Ergebnisse zeigen außerdem, dass Disidentifikation die Herabsetzung der eigenen, aber auch anderer status-niedriger Gruppen erhöht (Ikegami & Ischida, 2007). Es wird daher angenommen, dass Disidentifikation mit Studierenden die Präferenz für herabsetzende Informationen über status-niedrige Gruppen erhöht.

Diese Annahmen wurden anhand von Fragebogenmaßen mit einer großen Stichprobe von Studierenden der Universität Tübingen geprüft und an zwei unabhängigen Studierendenstichproben validiert. Die Daten wurden anhand von Regressionsanalysen und Strukturgleichungsmodellen analysiert. Die Analysen zeigen, dass ein niedriger sozialer Hintergrund wie erwartet zu einer stärkeren Inkompatibilität zwischen Herkunftsidetität und studentischer Identität führt. Diese begünstigt eine Disidentifikation mit Studierenden. Der Zusammenhang zwischen Status und Disidentifikation wird durch Inkompatibilität mediiert. Disidentifizierte Studierende zeigen wiederum eine geringere Präferenz für andere Studierende als Informationsquelle sowie insgesamt schlechtere Leistungen im Studium. Eine starke Disidentifikation geht außerdem mit der Präferenz für diskriminierende Informationen über status-niedrige Gruppen, wie z.B. übergewichtige Menschen einher. Die Befunde zeigen, dass insbesondere Studierende mit niedrigem sozialen Hintergrund Gefahr laufen, durch erlebte Inkompatibilität Leistungseinbußen in der Universität zu erleiden, einen schlechten Zugang zu universitätsinternen Informationen zu haben und diskriminierendes Verhalten zu zeigen. Die praktischen Implikationen der Befunde werden diskutiert.

Donnerstag
12. März

11:15 Uhr

HGC 20

Chair(s)

Schmidt-Hertha, Bernhard

Rott, Karin; Schmidt-Hertha, Bernhard

Medienkompetenz als Schlüsselqualifikation im Studium: Ergebnisse eines Fächervergleichs

Medienkompetenz zählt in der Informationsgesellschaft als berufliche Schlüsselkompetenz (Schmidt-Hertha et al., 2011), insbesondere für Angestellte mit akademischem Hintergrund. Dennoch wird Medienkompetenz meist nicht berufsfeldspezifisch definiert. Zudem ist über berufsfeldspezifische Anforderungen zum Umgang mit digitalen Technologien sowie die Entwicklung beruflich relevanter Medienkompetenz im Studium bislang wenig bekannt. Bisherige Forschungsergebnisse konzentrieren sich meist auf die Mediennutzung Studierender im universitären und privaten Kontext (z.B. Margaryan et al., 2011). Zur Frage, inwieweit Mediennutzungsmuster und verschiedene Facetten von Medienkompetenz von Studierenden mit den Anforderungen in potentiellen Berufsfeldern kompatibel sind, liegen bislang lediglich für den Bereich der Lehrerbildung einzelne Studien vor (z.B. Blömeke, 2000). Das Projekt Aufbau von Medienkompetenz und beruflich relevanten Informationsverhalten im Studium nimmt sich diesen Forschungsfragen mit einem multi-methodischen Forschungsdesign an, welches Experteninterviews, eine standardisierte Studierendenbefragung zu mehreren Messzeitpunkten, Mediennutzungstagebücher und qualitative Interviews mit Studierenden umfasst. Im Zentrum steht dabei die Frage nach der fachspezifischen Entwicklung von Mediennutzungsmustern und Medienkompetenzen im Studienverlauf sowie deren Anschlussfähigkeit an das avisierte Berufsfeld.

In diesem Beitrag wird der Frage nachgegangen, inwiefern die aktuellen medienbezogenen Fähigkeiten und Kompetenzen von Studierenden der Universität Tübingen an die Nutzungsanforderungen von Medien in verschiedenen Berufsfeldern anschlussfähig sind. Dazu wird zunächst auf Basis von Experteninterviews mit Personalverantwortlichen mittels qualitativer Inhaltsanalyse (Gläser & Laudel, 2010) die Relevanz medialer Ressourcen in typischen Berufsfeldern für Studierende der Betriebswirtschaft, der Erziehungswissenschaft sowie der Medizin eruiert. Dabei zeigen sich berufsfeldspezifische Anforderungsprofile an die Studierenden der verschiedenen Fächer, die in allen Berufsfeldern über die Handhabung von ‚Standardprogrammen‘ hinausgehen. Im Bereich Betriebswirtschaft wurde u.a. auf die gestiegene Bedeutung der professionellen Tools zur Kommunikation und Kooperation in multilokalen (oft auch internationalen) Projektteams verwiesen.

Anschließend werden die im Forschungscluster erhobenen Daten der ersten Erhebung von Tübinger Studierenden aller Fachbereiche mittels eines Onlinefragebogens (n=2576) zum Informationsverhalten, zu Kenntnissen in verschiedenen Computeranwendungen und Einstellungen zu verschiedenen Medien der Studierenden vorgestellt. Aus diesen Daten lassen sich mittels einer Faktorenanalyse (Hauptkomponentenanalyse, Varimax-Rotation) fünf Faktoren von studienbezogener Medienkompetenz identifizieren, die wie folgt benannt werden können: spezielle Anwendungskennnisse, studienbezogene Anwendungskennnisse, kritische Einstellungen gegenüber Medien, Onlineaktivität und kritische Informationsbewertung (vgl. Schmidt-Hertha & Rott, 2014). Diese Faktoren weisen heterogene Ausprägungen bei Studierenden verschiedener Fachrichtungen auf. Dabei können die studienfachbezogenen Unterschiede auf unterschiedliche Einflüsse zurückgeführt werden, wie etwa auf Studieninhalte oder die Standardisierung des Studienfaches, aber auch auf verschiedene Fachbereichs- und Fachkulturen. Zum Beispiel weisen Studierende der Geistes- und Sozialwissenschaften meist höhere Werte in studienbezogenen Anwendungskennnissen auf, welche Kenntnisse über Textverarbeitungsprogramme und verschiedene Recherche-Tools beinhalten, als etwa Studierende der Naturwissenschaften. Aus den Modulhandbüchern der verschiedenen Fächer wird ersichtlich, dass in den Geistes- und Sozialwissenschaften häufiger das Schreiben einer Hausarbeit als schriftliche Prüfungsleistung verlangt wird als in naturwissenschaftlichen Fächern, wofür diese speziellen studienbezogenen Anwendungskennnisse essenziell sind.

Ein Vergleich der Anforderungsprofile aus den Experteninterviews mit den Daten der ersten Studierendenbefragung zeigt auf der einen Seite Übereinstimmungen der Fähigkeiten und Medienkompetenzfacetten mit den in jeweils typischen potentiellen Berufsfeldern geforderten Fähigkeiten von Studierenden, auf der anderen Seite gibt es aber auch Abweichungen. In jedem Berufsfeld wurden Fähigkeiten identifiziert, die bei einem Großteil der Studierenden (noch) nicht gut ausgebildet sind. In längsschnittlicher Perspektive ist mit den weiteren Erhebungen zu prüfen, inwieweit sich die identifizierten Facetten von Medienkompetenz während des Studiums verändern und inwiefern eine Annäherung an berufsfeldspezifische Anforderungen festzustellen ist.

Donnerstag
12. März

11:15 Uhr

HGC 20

Chair(s)

Schmidt-Hertha, Bernhard

Tafertshofer, Lorenz; Schmidt-Hertha, Bernhard

Vereinbarkeit von Exzellenz in Forschung und Lehre in sozialwissenschaftlichen Hochschuleinrichtungen

Hochschulen definieren sich traditionell über die Verbindung von Forschung und Lehre als zentrale Aufgabengebiete, die als komplementäre, sich ergänzende Aufgaben verstanden werden können, aufgrund begrenzter zeitlicher, personeller und materieller Ressourcen im hochschulischen Alltag aber auch in Konkurrenz treten können (vgl. Jacob & Teichler, 2011). Bisher nehmen Studien beide Bereiche (Forschung und Lehre) zumeist getrennt in den Blick (z.B. Thole u.a. 2012), während in Studienrankings Indikatoren für Forschungs- und Lehrqualität in der Regel miteinander kombiniert werden. Die Studie geht daher der Frage nach, inwieweit es einzelnen sozialwissenschaftlichen Fachbereichen an verschiedenen Standorten gelingt, sowohl in der Forschung als auch in der Lehre Spitzenleistungen zu erbringen oder inwieweit eine Fokussierung auf eine der beiden Kernaufgaben Voraussetzung für Spitzenleistungen ist.

Dem hier vorgestellten ersten empirischen Zugriff liegt eine Typologisierung von sozialwissenschaftlichen Einrichtungen (Soziologie, Politologie, Erziehungswissenschaften, Psychologie, Wirtschaftswissenschaften und Sozialwesen) an öffentlichen Hochschulen in Deutschland hinsichtlich der Forschungs- bzw. Lehrintensivität des jeweiligen Fachs zugrunde. Basis ist die Analyse eigens recherchierter und über bestehende Ressourcen verfügbarer Daten zu zentralen Indikatoren der Forschungs- und Lehrintensität, mit dem Ziel, eine indikatorengestützte Charakterisierung der Lehr- und Forschungskultur für die jeweilige Disziplin an den verschiedenen Standorten zu liefern. Im Zentrum steht eine clusteranalytische Kontrastierung der verschiedenen Standorte unter Einbezug vorhandener Daten u.a. zu Veröffentlichungen, Drittmittelaufkommen, Studiendauer, Lehrpreisen, Promotionen und Habilitationen.

Hierzu wurde für jedes zu untersuchende Fachgebiet eine hierarchische Clusteranalyse (Ward-Verfahren) durchgeführt. Um fachspezifischen Besonderheiten hinsichtlich der Forschungskultur gerecht zu werden, erfolgte die Wahl der Clustervariablen ebenfalls fachspezifisch. So wurde z.B. für die Fächer Wirtschaftswissenschaften und Psychologie der Indikator „Zitationen pro Publikation“ aufgrund der Relevanz von Zitationskennzahlen im fachspezifischen Forschungsalltag als Clustervariable herangezogen.

Die Ergebnisse zeigen eine – je nach Fachgebiet – unterschiedlich stark ausgeprägte Heterogenität in den Profilen der verschiedenen Standorte bzgl. der untersuchten Variablen. In den Erziehungswissenschaften (Clustervariablen: Anzahl Promotionen, Habilitationen und Publikationen je Professor, Drittmittel pro Professor, Anteil Absolventen innerhalb der Regelstudienzeit, Betreuungsrelation; Untersuchungszeitraum 2010-2012; Untersuchungseinheiten: 52 erziehungswissenschaftliche Einrichtungen an öffentlichen Universitäten in Deutschland) ergibt sich beispielsweise folgende 5-Cluster-Lösung: Cluster 1 (n=10) zeichnet sich durch die niedrigste Betreuungsrelation aus, liegt aber auch bei allen anderen Variablen unter dem Durchschnitt. Cluster 2 (n=21) weist neben einer niedrigen Betreuungsrelation einen überdurchschnittlich hohen Anteil an Absolventen in der Regelstudienzeit auf, während die übrigen Variablen unterrepräsentiert sind. Dieses Cluster scheint damit eher lehrorientierte Standorte zu vereinigen. Cluster 3 (n=9) stellt (trotz nur durchschnittlichem Drittmittelaufkommen) das publikationsstärkste Cluster dar. Auch hier fällt der Anteil an Absolventen in der Regelstudienzeit hoch aus. Das vierte Cluster (n=8) ist durch ein hohes Drittmittelaufkommen bei durchschnittlichem Publikationsoutput sowie einer hohen Promotions- und Habilitationsquote gekennzeichnet. Trotz höchster Betreuungsrelation liegt der Anteil der Absolventen innerhalb der Regelstudienzeit nur leicht unter dem Durchschnitt. Dieses Cluster lässt sich als drittmittellorientiert und stark in der Nachwuchsförderung charakterisieren. Auch Cluster 5 (n=4) zeigt überdurchschnittliche Promotions- und Habilitationsquoten, der Publikationsoutput und das Drittmittelaufkommen liegen aber deutlich unter dem Durchschnitt, während der Anteil der Absolventen innerhalb der Regelstudienzeit am höchsten ausfällt. Die in diesem Cluster vereinigten Hochschulen (u.a. Pädagogische Hochschulen) können als lehrorientierte Nachwuchsförderer umschrieben werden, wobei die Forschungsorientierung nur gering ausgeprägt ist.

Diese und die Analysen der weiteren Fachgebiete machen deutlich, dass sich die Hochschullandschaft durch ein hohes Maß an Heterogenität auszeichnet und eine einseitige Orientierung an einer „Einheitsuniversität“ – wie sie in gängigen internationalen Hochschulrankings gebräuchlich ist – den bestehenden Unterschieden zwischen den Profilen nicht gerecht wird (vgl. Federkeil, 2013) und dabei übersieht, dass gute Leistungen in Forschung und Lehre auf vielfältige Weise erzielt werden können. Die Heterogenität der fachspezifischen Lehr- und Forschungsprofile spiegelt sich schließlich auch auf Studiengangsebene wider und steht damit in direkter Wechselwirkung zu den Studierenden, die sich mit unterschiedlichen fach- und studiengangspezifischen Informationsumwelten und den damit verbundenen Anforderungen konfrontiert sehen.

Donnerstag
12. März

11:15 Uhr

HGC 20

Chair(s)

Schmidt-Hertha, Bernhard

Nienkemper, Barbara

Heterogenität und Teilhabe der erwachsenen Bevölkerung im Hinblick auf ihre Kompetenzniveaus unter besonderer Berücksichtigung der Weiterbildungsbeteiligung

In diesem Symposium wird die gemeinsame Fragestellung bearbeitet, welchen erweiternden Beitrag die Ergebnisse aus aktuellen Studien mit Kompetenzmessungen in der erwachsenen Bevölkerung (PIAAC, LEO-Level-One Studie) leisten, um die ungleiche Verteilung von gesellschaftlicher Teilhabe zu erklären. Ökonomische Interpretationen auf Basis der Humankapitaltheorie werden zu diesem Zweck kritisch rezipiert und soziologischen Ansätzen zur sozialen Vererbung von Bildungsarmut (Allmendinger 2003) gegenübergestellt. Ein besonderer Fokus wird auf die Gruppe der Personen mit niedrigen Lesekompetenzwerten (bis Stufe 1) und ihre im Vergleich zu Personen mit hohen Lesekompetenzwerten (Stufe IV/V) deutlich geringere Beteiligung an Weiterbildung (70% vs. 28%) (Maehler et al. 2013, S. 114f.) gelegt.

Am Symposium sind Wissenschaftler/innen unterschiedlicher Disziplinen beteiligt.

Im ersten Beitrag des Symposiums wird das soziodemographische Profil der Personen mit niedriger Lesekompetenz mit Hilfe der PIAAC-Daten beschrieben. Da sich die Gesamtgruppe der Personen mit niedriger Lesekompetenz als relativ heterogen erweist, nimmt der Soziologe Bönisch eine eigenständige Betrachtung der unterschiedlichen Kompetenzstufen vor (Personen mit mangelnder Lese- bzw. Sprachfähigkeit, Personen in Kompetenzstufe unter 1 bzw. 1) und bestimmt die relevanten Einflussfaktoren (formaler Bildungsabschluss, soziale Herkunft und Sprache) auf niedrige Lesekompetenz.

In einer Kooperation des GESIS –Leibniz-Instituts für Sozialwissenschaften und des erziehungswissenschaftlichen Fachbereichs der Universität Hamburg präsentieren Perry, Rammstedt, Nienkemper und Grotlüschen im zweiten Beitrag einen Forschungsansatz zur Identifikation von weiteren Zielgruppen für die Grund- und Weiterbildung. Sie schlagen vor, die in der PIAAC Studie verwendeten Reading Components ergänzend zum Lower-Rungs-Approach der LEO – Level-One Studie zu nutzen, um innerhalb der heterogenen Gruppe von Personen mit niedrigen Lesekompetenzen weiter differenzieren zu können.

Kastner und Schlögl modellieren dem Konstrukt des „Basisbildungsbedarfs“ entsprechend im dritten Beitrag eine Risikogruppe, die in allen drei gemessenen Testdomänen niedrige Kompetenzwerte erzielt hat und beschreiben deren Sozialprofil auf Basis der österreichischen PIAAC-Daten. Aus Perspektive der Erwachsenen- und der Berufsbildung untersuchen sie die Beteiligung der Risikogruppe in den relevanten gesellschaftlichen Teilhabedimensionen und diskutieren kritisch, inwieweit der gewählte Forschungszugang des large-scale-assessment der Heterogenität der Personen bzw. der Zielgruppen mit niedrigen Kompetenzen in Bezug auf ihre Bildungsbedarfe/-bedürfnisse, Erreichbarkeit, auf Ansprache und Angebotsformate sowie Weiterbildungsbarrieren gerecht wird.

Die empirischen Bildungsforscherinnen Kaufmann und Widany integrieren im vierten Beitrag Faktoren der Angebots- und der Nachfrageseite, um die Bedingungen beruflicher Weiterbildungsbeteiligung zu untersuchen. Sie unterscheiden drei Finanzierungsmodelle (arbeitgeberfinanzierte Weiterbildung, arbeitnehmerfinanzierte WB und mischfinanzierte WB) und berücksichtigen sowohl den Erwerbskontext und die konkreten Tätigkeitsanforderungen als auch die individuelle Kompetenzausstattung und berichtete Lernstrategien, um individuelle Teilnahmekancen zu erklären.

Becker-Mrotzek diskutiert die Beiträge abschließend aus deutschdidaktischer Perspektive.

Donnerstag
12. März

11:15 Uhr

HGC 30

Chair(s)

Nienkemper, Barbara

Heterogenität und Teilhabe der erwachsenen Bevölkerung im Hinblick auf ihre Kompetenzniveaus unter besonderer Berücksichtigung der Weiterbildungsbeteiligung

Bönisch, Markus

Lesekompetenzstufen und niedrige Lesekompetenz – Ein internationaler Vergleich mit einem Fokus auf die Situation in Österreich

Der Beitrag beschreibt wie die PIAAC-Lesekompetenzstufen und das Modul der Reading Components zur Abgrenzung von Personen mit niedriger bzw. sehr niedriger Lesekompetenz verwendet wurden. Er gibt Einblick in das soziodemographische Profil dieser Gruppen in Österreich und in Einflussfaktoren und multivariate Zusammenhänge in Bezug auf niedrige Lesekompetenz.

Auf Basis dieser Analysen verfügen 17,1% der 16- bis 65-Jährigen in Österreich, das entspricht rund einer Million Personen, über nur niedrige Lesekompetenz. Ein kleiner Teil dieser Personen verstehen die deutsche Sprache überhaupt nicht bzw. kann nicht lesen. Großteils können sie jedoch Sätze und leicht verständliche Texte lesen. Längere Texte sind für sie nur schwierig bis gar nicht zu bewältigen. Mit diesem Anteil liegt Österreich im internationalen Vergleich im Durchschnitt. Schränkt man die Betrachtung auf Personen mit sehr niedriger Lesekompetenz ein, d.h. auf Personen, die erhebliche Probleme mit Texten haben und bei denen auch das Verständnis für Satzstrukturen nur in geringem Ausmaß vorhanden ist, so liegt dieser Anteil in Österreich bei 4,3%, das entspricht rund einer viertel Million Personen (240.000). Der vergleichbare internationale Anteil für alle an PIAAC teilnehmenden OECD-Länder liegt bei 4,5%.

Multivariate Analysen (Logistische Regressionen) für Österreich zeigen die hohe Relevanz des formalen Bildungsabschlusses in Bezug auf niedrige Lesekompetenz. Einen weiteren relevanten Einfluss stellt die soziale Herkunft einer Person dar. Dies zeigt sich deutlich bei Betrachtung eines Indikators für das kulturelle Kapital im elterlichen Haushalt (vgl. Bourdieu 1982, OECD, 2011a bzw. auf Basis der PISA-Daten: Wößmann, 2005): Je weniger Bücher im elterlichen Haushalt (zum Zeitpunkt als die befragte Person 16 Jahre alt war) vorhanden waren, desto höher ist das Risiko, von niedriger Lesekompetenz betroffen zu sein. Darüber hinaus hat der Migrationshintergrund einer Person, und hier vor allem die Sprache und etwas weniger stark das Geburtsland, einen deutlichen Einfluss auf die im Rahmen von PIAAC gemessene Lesekompetenz (in Deutsch). Signifikante geschlechtsspezifische Unterschiede sind in Bezug auf niedrige Lesekompetenz in Österreich nicht vorhanden und auch der Wohnort einer Person (Stadt oder Land), hat keinen Einfluss auf das Risiko, von niedriger Lesekompetenz betroffen zu sein. Jedoch ist sehr niedrige Lesekompetenz in Österreich ein eher urbanes Phänomen.

Der Definition von Lesekompetenz bei PIAAC liegt die Annahme zugrunde, dass ein gewisses Maß an Lesekompetenz notwendig ist, um am Gesellschaftsleben teilzunehmen, d.h. niedrige Lesekompetenz kann mit Benachteiligungen im Alltag und Beruf verbunden sein (vgl. OECD 2012 und 2013). Die Analysen in diesem Beitrag zeigen, dass in Österreich zwar ein im internationalen Vergleich relativ hoher Anteil an Personen mit niedriger Lesekompetenz in den Arbeitsmarkt integriert sind, jedoch sind Personen mit höherer Lesekompetenz stärker am Erwerbsleben beteiligt und haben auch einen höheren Stundenverdienst. Darüber hinaus zeigt sich ein signifikanter Zusammenhang zwischen Lesekompetenz und subjektiver Einschätzung der Gesundheit.

Donnerstag
12. März

11:15 Uhr

HGC 30

Chair(s)

Nienkemper, Barbara

Heterogenität und Teilhabe der erwachsenen Bevölkerung im Hinblick auf ihre Kompetenzniveaus unter besonderer Berücksichtigung der Weiterbildungsbeteiligung

Perry, Anja; Rammstedt, Beatrice; Nienkemper, Barbara; Grotlüschen, Anke

Reading Components und niedrigere Kompetenzniveaus - Ein besseres Verständnis von Personen mit niedriger Kompetenz

Die Level-One-Studie (LEO) zeigt, dass 7,5 Millionen erwerbsfähige Erwachsene im Alter von 18-64 Jahren in Deutschland funktionale Analphabeten sind (Grotlüschen & Riekmann, 2012). Weiterhin belegen die Daten des Programme for the International Assessment of Adult Competencies (PIAAC), dass ein Großteil der Personen im erwerbsfähigen Alter (16-65 Jahre) in Deutschland (rund 18%) nur über niedrige Lesekompetenzen verfügt (Zabal et al., 2013). Die Auswirkungen dieser geringen oder fehlenden Kompetenz sind für die Betroffenen vielfältig. Sie reichen von Problemen beim Schreiben und Lesen von Notizen bis hin zur fehlenden Teilhabe an Gesellschaft und Arbeitsmarkt (Döbert & Hubertus, 2000). Die Ergebnisse der LEO-Studie haben gezeigt, dass die Angebote der Grund- und Weiterbildung ihrem kompensatorischen Anspruch noch nicht gerecht werden. So liegt die Beteiligungsquote funktionaler Analphabet/inn/en (28 Prozent) deutlich niedriger als in der erwerbsfähigen Bevölkerung insgesamt (42 Prozent). Die Gruppe der funktionalen Analphabet/inn/en, deren Erstsprache nicht Deutsch ist, sticht dabei hervor, denn sie besucht vorwiegend die bundesweiten Integrationskurse (vgl. Bilger 2012). In unserem Beitrag wollen wir die Fragestellung bearbeiten, wie mit Hilfe der PIAAC-Daten weitere Zielgruppen innerhalb der Gruppe von Personen mit niedriger Lesekompetenzen identifiziert und ihre Lebensumstände beschrieben werden können. Unser Ausgangspunkt sind dabei auch die Studien von Strucker, Yamamoto & Kirsch (2005 und 2007), denen es gelang, die Typiken im unteren Leistungsbereich zumindest nach Lernbehinderung, Zweitsprache oder Einkommenssituation zu unterscheiden.

Die PIAAC- und die LEO-Studie verfolgen unterschiedliche Ansätze zur Messung niedriger Lesekompetenz. Der Lower-Rungs-Ansatz der LEO-Studie nimmt die Kompetenzstufen von Literalitätserhebungen zum Ausgangspunkt und betrachtet die Lese- und Schreibfähigkeit als Leiter mit entsprechenden Leiterstufen. Der Fokus liegt nunmehr auf dem untersten Kompetenzbereich (Level One). Dieser wird in kleinere Leiterstufen (Sprossen), sogenannte Alpha-Levels ausdifferenziert, die enger beieinander liegen (Grotlüschen & Riekmann, 2011). Demgegenüber nimmt der Reading-Components-Ansatz innerhalb der PIAAC-Studie (Sabatini & Bruce, 2009) Theorien des kindlichen Schriftspracherwerbs zum Ausgangspunkt und unterteilt den niedrigsten Kompetenzbereich (unter Stufe 1) in grundlegende Komponenten, die für das Lesen bedeutsam sind. So wurden in der PIAAC Studie die drei Komponenten: Word Recognition, Sentence Comprehension und Passage Fluency getestet.

Dieser Beitrag wird in einem ersten Teil beide Ansätze zur Messung von Lese- bzw. Schreibkompetenz vergleichend darstellen und das Projekt „Reading Components und niedrigere Kompetenzniveaus – ein besseres Verständnis von Personen mit niedriger Kompetenz“ vorstellen. Ziel des Projekts ist es, basierend auf beiden Studien genauere Kenntnisse der betroffenen Zielgruppen sowie ihrer Lebensumstände und Erreichbarkeit abzuleiten. Aus diesen Erkenntnissen sollen Handlungsempfehlungen entwickelt werden, wie Personen mit niedriger Lesekompetenz gezielt gefördert werden können. Hierfür werden die vorhandenen Daten zu den Reading Components und der LEO-Studie zunächst nach Subpopulationen, Variablen der gesellschaftlichen Teilhabe sowie Weiterbildungsvariablen analysiert. In diesem Schritt sollen besonders relevante Merkmale identifiziert werden. Im nachfolgenden Schritt werden, basierend auf diesen Ergebnissen, mittels latenter Klassenanalyse Zielgruppen identifiziert und beschrieben. Da PIAAC eine international vergleichende Studie ist, können diese Analysen auch länderübergreifend verglichen werden und Best-Practice-Beispiele sowie Spezifika des deutschen unteren Kompetenzniveaus herausgearbeitet werden. Eine besondere Schwierigkeit ist hierbei die Berücksichtigung der verschiedenen Sprachfamilien, die die internationale Vergleichbarkeit erschwert.

Im zweiten Teil des Beitrags werden erste Ergebnisse des Projekts präsentiert. Hierzu werden international vergleichend die Anteile richtiger Antworten sowie Dauer der Aufgabenbearbeitung der Reading Components analysiert. Diese Indikatoren werden mit verschiedenen Hintergrundvariablen, wie Herkunft, Bildungsabschluss, Einkommen, Verwendung von Fertigkeiten im Alltag und Beruf und Motivation, in Relation gesetzt (N = 22.163). In einer weiteren Regressionsanalyse je Land wird die Leistung anhand dieser Variablen erklärt, um genauere Zusammenhänge zwischen Hintergrundvariablen und den grundlegenden Lesekomponenten zu analysieren.

Donnerstag
12. März

11:15 Uhr

HGC 30

Chair(s)

Nienkemper, Barbara

Heterogenität und Teilhabe der erwachsenen Bevölkerung im Hinblick auf ihre Kompetenzniveaus unter besonderer Berücksichtigung der Weiterbildungsbeteiligung

Kastner, Monika; Schlögl, Peter

Modellierung einer „Risikogruppe“ und Ergebnisse zu Teilhabe-/Integrationsdimensionen

Hintergrund des Beitrags zum Symposium: Österreich hatte sich an den PIAAC-Vorgängerstudien IALS und ALL (ab 1994) nicht beteiligt, daher bestand bislang Unklarheit über das dort untersuchte Kompetenzniveau der Bevölkerung und über die Adressatengruppe mit niedrigen Kompetenzen im Besonderen. Mit PIAAC 2011/12 stehen für Österreich nun erstmals Daten zu grundlegenden Kompetenzdimensionen zur Verfügung. Für einen Beitrag zum nationalen Zusatzbericht (Statistik Austria 2014) wurde eine Risikogruppe modelliert, die in allen drei gemessenen Testdomänen niedrige Kompetenzwerte erzielt hat. Das Konstrukt der „Risikogruppe“ wird in Analogie zu Basisbildungsangeboten für Erwachsene gewählt. Diese Programme umfassen wahlweise die Grundfertigkeiten „Lesen, Schreiben, Rechnen“ sowie den „Umgang mit technischer Informationsvermittlung“ (vgl. Brugger et al., 1997) oder in Erweiterung noch den Erwerb von Deutsch als Zweitsprache sowie weitere Schlüsselkompetenzen, wie z.B. ‚Lernen lernen‘ (Länder-Bund-ExpertInnengruppe ‚Initiative Erwachsenenbildung‘, 2011).

Für die Analyse der „Risikogruppe“ wurden diejenigen Personen einbezogen, die bei Lesen und Alltagsmathematik Kompetenzstufe 1 und darunter erreicht haben; bei Problemlösen im Kontext neuer Technologien mussten unterschiedliche Gruppen differenziert werden: Personen ohne Computererfahrung, jene, die am Computer-Maus-Test gescheitert waren, und die sogenannten „Computerverweigerer“ wurden zur Risikogruppe gezählt, wenn sie in Lesen und Alltagsmathematik entsprechend niedrige Werte erzielt haben, weil damit Aneignung und Anwendung von IKT-Fertigkeiten sowie Herausbildung von technologiebezogener Problemlösekompetenz erschwert sind. Von jenen Personen, die den Aufgabenteil am Computer bearbeitet haben, wurden diejenigen der Risikogruppe zugerechnet, die hier maximal die Kompetenzstufe unter 1 (weniger als 241 Punkte) und bei Lesen und Alltagsmathematik weniger als 226 Punkte erreicht haben. Die Größenordnung dieser Risikogruppe wurde mit 11,4% der 16- bis 65-Jährigen in Österreich bestimmt, das entspricht rund 640.000 Menschen. Diese Gruppe wurde in Hinblick auf deren Sozialprofil sowie gesellschaftliche Teilhabe-/Integrationsdimensionen untersucht.

Es konnte unter anderem gezeigt werden, dass Ältere in höherem Ausmaß betroffen sind, signifikante Geschlechtsunterschiede sind keine feststellbar. Sowohl in Österreich geborene Personen mit Deutsch als Erstsprache als auch nicht in Österreich geborene Personen mit anderen Erstsprachen sind gleichermaßen betroffen. Bei Bildungshintergrund und Bildungsabschlüssen zeigen sich erwartete Muster, wie beispielsweise die bekannte Vererbung des Bildungsstatus und die Dominanz von niedrigen Bildungsabschlüssen innerhalb der Risikogruppe. Ungeachtet dessen ist über die Hälfte der zur Risikogruppe zählenden Personen in Erwerbsarbeit eingebunden, vielfach in der Berufshauptgruppe der Hilfsarbeitskräfte, jedoch auch auf Arbeitsplätzen mit höherem Anforderungsniveau (mehrheitlich ISCO-Level 2); die Arbeitszufriedenheit der Erwerbstätigen innerhalb der Risikogruppe ist hoch. Weitere Analyseergebnisse liegen vor zur Beteiligung an Weiterbildung sowie zu gesellschaftlichen Teilhabedimensionen wie beispielsweise ehrenamtlichem Engagement, zur subjektiven Einschätzung von demokratiepolitischer Einflussnahme, der eigenen Gesundheit sowie der sozialen und familiären Einbindung.

Beitrag zum Symposium: Basierend auf oben skizzierten Analyseergebnissen – die im Rahmen des Einzelbeitrags erläutert werden – wurden zum einen Evidenzen für bildungspolitische Steuerungsfragen herausgearbeitet. Im Symposium soll nun „im Blick zurück“ untersucht werden, ob und wie diese Hinweise (von u.a. Ministerien, Bildungspraxis) aufgenommen wurden. Zum anderen wurden Forschungsdesiderata für vertiefende und weiterführende Analysen bereitgestellt. Hinsichtlich des Forschungsinteresses generierte diese Analyse nämlich vertiefende Forschungsfragen (und weniger abschließende Antworten). Hier zeigt sich unter anderem, dass solche weiterführenden Forschungen oftmals eines qualitativ-empirischen Forschungszugangs bedürften, um gerade der Heterogenität der Personen bzw. der Zielgruppen mit niedrigen Kompetenzen gerecht zu werden (u.a. Bildungsbedarfe/-bedürfnisse; Erreichbarkeit, Ansprache und Angebotsformate, Weiterbildungsbarrieren).

Donnerstag
12. März

11:15 Uhr

HGC 30

Chair(s)

Nienkemper, Barbara

Heterogenität und Teilhabe der erwachsenen Bevölkerung im Hinblick auf ihre Kompetenzniveaus unter besonderer Berücksichtigung der Weiterbildungsbeteiligung

Kaufmann, Katrin; Widany, Sarah

Weiterbildungsteilnahmechancen in Abhängigkeit von Kompetenzen und Tätigkeitsanforderungen

Theoretischer Hintergrund: Berufliche Weiterbildungsbeteiligung ist das Ergebnis komplexer Zusammenhänge von Angebots- (Gelegenheits-) und Nachfrage- (Nutzungs-)strukturen, die auf mehreren Ebenen wirksam werden: Individuelle Teilnahmechancen werden durch Konstellationen verschiedenster Faktoren auf der Mikroebene (bspw. Geschlecht, Qualifikation), Mesoebene (bspw. Beruf, Betrieb) und Makroebene (bspw. Staat) beeinflusst. Zur Erklärung berufsbezogener Weiterbildungsbeteiligung kommen mehrere theoretische Bezüge in Betracht. In bildungsökonomisch und bildungssoziologisch geprägten Forschungsarbeiten werden insbesondere etablierte Theorien, wie Humankapitaltheoretische Ansätze, Rational-Choice-Theorien, Signaltheorie oder Arbeitsmarktsegmentationstheorien herangezogen (vgl. Büchter 2010; Becker/Hecken 2009). Humankapitaltheoretische Ansätze verweisen u.a. auf die Bedeutung von Investitionsmerkmalen und Renditeerwartungen, die mit berufsbezogener Weiterbildungsteilnahme verbunden sind und erklären damit unterschiedliche Beteiligungsselektivitäten an arbeitgeber- und arbeitnehmerfinanzierter Weiterbildung (Becker 1993). Segmentierungstheoretische Ansätze wiederum stellen verschiedene Teilarbeitsmärkte und damit unterschiedliche Gelegenheitsstrukturen für Weiterbildungsbeteiligung heraus (Lutz/Sengenberger 1974). Im erziehungswissenschaftlichen Kontext existieren mehrere Teilnahmemodelle (bspw. Boeren et al. 2010; Helms-Jørgensen/Warring 2003; Desjardins/Rubenson 2011), deren empirische Überprüfung jedoch vielfach aussteht. Sie sind weniger als abgeschlossener Theorieentwurf sondern vielmehr dahingehend konzeptioniert, einen Analyserahmen zur Verfügung zu stellen, der sowohl empirische Befunde zur Weiterbildungsteilnahme als auch verschiedene Theorieansätze integriert. Dieser integrative Ansatz, der Faktoren der Angebots- und der Nachfrageseite berücksichtigt, hat ein hohes Analysepotential für die Teilnahmeforschung in einem hochgradig heterogenen Bildungsbereich (Widany/Kaufmann 2013). Um selektive Teilnahmestrukturen an berufsbezogener Weiterbildung zu erklären, hat sich zudem eine Differenzierung in Weiterbildungssegmente nach Investitionskriterien bewährt (vgl. Behringer 1999; Schiener 2006; Kaufmann/Widany 2013). Die Auswahl konkreter Merkmale der Angebots- und Nachfrageseite zur Modellierung ist nur bedingt theoriegeleitet möglich, da o.g. Ansätze bezüglich der Operationalisierung zentraler Konstrukte oft wenig gehaltvoll sind. Die Beobachtung unterschiedlicher Gelegenheits- und Nutzungsstrukturen der Weiterbildungsbeteiligung ist somit wesentlich von berücksichtigten (verfügbaren) Erklärungsfaktoren beeinflusst.

Fragestellung: Vor dem Hintergrund der skizzierten theoretischen Annahmen zur Wechselwirkung von Gelegenheits- und Nutzungsstrukturen zur Erklärung berufsbezogener Weiterbildungsbeteiligung wird anknüpfend an den Forschungsstand geprüft, ob bestimmten Merkmalen der Angebotsseite in verschiedenen Weiterbildungssegmenten eine unterschiedliche Bedeutung zukommt. Ungeklärt ist beispielsweise die Frage, ob der Zusammenhang zwischen hohen Bildungsabschlüssen und hohen Beteiligungschancen an (arbeitgeberfinanzierter) Weiterbildung auf eine Signalwirkung der formalen Abschlüsse zurückzuführen ist (Signal-/Filtertheorie) oder ob die individuelle Kompetenzausstattung eine bedeutendere Rolle spielt (Humankapitaltheorie). Obgleich die Ergebnisse der ersten PIAAC-Erhebung einen deutlichen Zusammenhang zwischen formalen Abschlüssen und individueller Kompetenzausstattung nachweisen (vgl. Maehler et al. 2013), ist der jeweils eigenständige Einfluss dieser Merkmale auf Weiterbildungsbeteiligung in verschiedenen Segmenten noch zu spezifizieren, insbesondere, wenn neben der Kompetenzausstattung auch individuelle Lernstrategien Berücksichtigung finden. Der Einfluss dieser individuellen Merkmale ist zudem in Abhängigkeit der Gelegenheitsstrukturen zu prüfen. Auch diesbezüglich liegt der Erkenntnisgewinn in der Berücksichtigung differenzierter Informationen. Neben übergeordneten Merkmalen des Erwerbskontexts, wie bspw. Unternehmensgröße wird der Einfluss konkreter Tätigkeitsanforderungen untersucht. Es wird davon ausgegangen, dass insbesondere arbeitgeberfinanzierte Weiterbildungsbeteiligung stärker von konkreten Tätigkeitsanforderungen beeinflusst ist, während Weiterbildung mit individuellen Investitionsanteilen stärker vom individuellen Kompetenzniveau und Lerneinstellungen beeinflusst ist.

Methode: Die Analysen basieren auf den deutschen Daten der ersten Erhebung des „Programme for the International Assessment of Adult Competencies“ (PIAAC). Die Analysen beziehen sich auf berufsbezogene Weiterbildung; daher wird das Sample auf erwerbstätige, nicht in Ausbildung befindliche Personen eingegrenzt (N=2.928). Zur Differenzierung von Weiterbildungssegmenten werden zeitliche und finanzielle Investitionen von Arbeitgebern und Arbeitnehmern berücksichtigt und zwischen betrieblicher, individueller und mischfinanzierter Weiterbildung unterschieden. Die PIAAC-Daten erlauben, sowohl Informationen über Kompetenzausstattungen sowie konkreter Tätigkeitsanforderungen von Erwerbstätigen zu berücksichtigen (JRA, OECD 2009). Das Zusammenspiel dieser Faktoren sowie weiterer Merkmale der Angebots- und Nachfrageseite in Bezug auf Weiterbildungsteilnahmechancen wird mit Hilfe logistischer Regressionen geprüft (Best/Wolf 2010). Es werden identische Modelle für Beteiligungschancen in den jeweiligen Weiterbildungssegmenten gerechnet und die Ergebnisse zueinander in Bezug gesetzt.

Ergebnisse: Vor dem Hintergrund heterogener Investitionskalküle auf Arbeitgeber- und Arbeitnehmerseite erweisen sich für arbeitgeberfinanzierte Weiterbildung insbesondere Erwerbskontextmerkmale und Tätigkeitsanforderungen als bedeutsam. Teilnahmechancen an Weiterbildung mit individuellen Investitionsanteilen gehen mit einem stärkeren Einfluss individueller Merkmale einher. Unter Kontrolle von Bildungsabschluss und Erwerbskontextmerkmalen zeigen das Kompetenzniveau, berichtete Lernstrategien und Tätigkeitsanforderungen teilweise einen signifikanten Einfluss. Differenziertere Informationen tragen damit zur Erweiterung des Kenntnisstands über Bedingungen beruflicher Weiterbildungsbeteiligung bei.

Donnerstag
12. März

11:15 Uhr

HGC 30

Chair(s)

Nienkemper, Barbara

Heterogenität und Teilhabe der erwachsenen Bevölkerung im Hinblick auf ihre Kompetenzniveaus unter besonderer Berücksichtigung der Weiterbildungsbeteiligung

Wilde, Annett

Experimentelle Studien mit Lehramtsanwärtern zu Aspekten der professionellen Kompetenz

In der empirischen Bildungsforschung sind Studien zur Entwicklung von Kompetenzen von Lehrkräften zahlreich und beschreiben verschiedene mögliche Einflussfaktoren auf individueller und kontextueller Ebene (Terhart, Bennewitz, & Rothland, 2001; Zlatkin-Troitschanskaia, Beck, Sembill, Nickolaus, & Mulder, 2009). Die bisherige Forschung nutzte dabei bisher überwiegend Querschnittsdesigns mit Gruppenvergleichen oder Längsschnittstudien. Diese methodischen Herangehensweisen haben jedoch Grenzen in Bezug auf kausale Rückschlüsse zur Erklärung der Ergebnisse. Dagegen gilt das Experiment als „gold standard“, wenn es darum geht, die Kausalität von Befunden einzuordnen. In diesem Symposium haben wir vier Studien zusammengestellt, die sich bei der Erforschung der professionellen Kompetenz von Lehramtsanwärtern experimenteller Ansätze bedient haben. Dabei fokussieren die Beiträge inhaltlich auf verschiedene Aspekte der professionellen Kompetenz bzw. den kognitiven Prozessen, die der Entwicklung professioneller Kompetenz zugrunde liegen. Durch kontrollierte Lernbedingungen und die experimentelle Manipulation dieser können Bedingungsfaktoren von Lernprozessen aufgezeigt werden.

Der Beitrag von Tröbst et al. untersucht, wie fachdidaktisches Wissen bei angehenden Lehrkräften entsteht. Die Vermittlung von Fachwissen, pädagogischem Wissen und/oder fachdidaktischem Wissen wurde dabei systematisch variiert. Im zweiten Beitrag (Wilde, Seiz, Decker & Kunter) wurde untersucht, wie sich die Bewertung einer Lernsituation als Bedrohung oder Herausforderung auf die Verarbeitungstiefe beim Lernen auswirkt. Der Beitrag von Heitzmann, Fischer und Fischer beschäftigt sich mit der Wirkung von Selbsterklärungsprompts und adaptivem Feedback auf den Erwerb von Diagnosekompetenz durch Lösungsbeispiele mit integrierten Fehlern. Im letzten Beitrag (Markova et al.) geht es darum, welche Rolle Selbstwirksamkeit in Bezug auf Inklusion und wahrgenommene Kompetenz bei der Wahl pädagogischen Förderbedarfs spielen. Dabei wurde zwischen Schülern mit Verhaltensauffälligkeiten und Schülern mit Lernauffälligkeiten unterschieden.

Donnerstag
12. März

11:15 Uhr

HGC 40

Chair(s)

Wilde, Annett
Seiz, Johanna**Experimentelle Studien mit Lehramtsanwärtern zu Aspekten der professionellen Kompetenz**

Tröbst, Steffen; Kleickmann, Thilo; Heinze, Aiso; Bernholt, Andrea; Rink, Roland; Kunter, Mareike

T-KnoX – Eine experimentelle Untersuchung der Bedeutung fachlichen und pädagogischen Wissens für den Erwerb fachdidaktischen Wissens

In Anlehnung an Shulman (1986) wird das professionelle Wissen von Lehrkräften für gewöhnlich in fachliches, pädagogisches und fachdidaktisches Wissen unterteilt. Fachdidaktisches Wissen umfasst dabei jenes Wissen, welches Lehrkräfte benötigen, um Schülerinnen und Schülern fachliche Inhalte zugänglich zu machen (Depaepe, Verschaffel, & Kelchtermans, 2013). Die Bedeutung dieses Wissens für die Unterrichtsqualität und den Lernerfolg von Schülerinnen und Schülern wurde empirisch nachgewiesen (Baumert et al., 2010). Die Befundlage zur Frage, wie angehende Lehrkräfte fachdidaktisches Wissen entwickeln, ist allerdings spärlich. In einer experimentellen Studie untersuchten wir, welche Rolle verfügbares Fachwissen und verfügbares pädagogisches Wissen für die Entwicklung fachdidaktischen Wissens spielen. Drei Annahmen wurden geprüft: (1) Fachwissen und pädagogisches Wissen verschmelzen zu fachdidaktischem Wissen („Amalgam-Hypothese“), (2) fachdidaktisches Wissen entsteht „en passant“ aus Fachwissen („Mitnahme-Hypothese“) und (3) Fachwissen unterstützt die Nutzung von Lerngelegenheiten für fachdidaktisches Wissen („Moderator-Hypothese“).

Die teilnehmenden 100 Lehramtsstudierenden aus Berlin und Brandenburg wurden zufällig einem von fünf Kursen zugeteilt. Die Kurse bestanden aus je zwei Blöcken á vier Stunden. Die „Amalgam-Bedingung“ umfasste einen Block zum Fachwissen und einen Block zum pädagogischen Wissen, die „Mitnahme-Bedingung“ zwei Blöcke zum Fachwissen und die „Moderator-Bedingung“ einen Block zum Fachwissen und einen Block zum fachdidaktischen Wissen. Darüber hinaus erhielt eine schwache Kontrollbedingung ausschließlich Blöcke zum pädagogischen Wissen, während eine starke Kontrollbedingung nur mit Blöcken zum fachdidaktischen Wissen konfrontiert wurde. Thematisch behandelten die fachdidaktischen Blöcke das Unterrichten von Bruchrechnung in der sechsjährigen Grundschule. Die fachlichen und pädagogischen Blöcke wiesen jeweils spezifische inhaltliche Überschneidungen mit den fachdidaktischen Blöcken auf: Die fachlichen Blöcke befassten sich mit Brüchen und Bruchrechnung ohne die Vermittlung dieser Inhalte im Schulunterricht zu thematisieren. Die pädagogischen Blöcke widmeten sich allgemein dem Thema Lernen und Lehren, d.h. sie behandelten Repräsentationsformen, Schülervorstellungen, Lernschwierigkeiten und Lehrstrategien ohne Bezug zur Mathematik herzustellen. Entsprechend gab es keine inhaltliche Überschneidung zwischen fachlichen und pädagogischen Blöcken.

Die zentrale abhängige Variable, das fachdidaktische Wissen, wurde vor, zwischen und unmittelbar nach den beiden Kursblöcken sowie sechs Wochen nach Ende der Kurse erfasst. Der Itempool zur Erfassung des fachdidaktischen Wissens umfasste 41 Items, welche bei starker Verankerung teilrotiert über die Messzeitpunkte hinweg präsentiert wurden. Im Zuge der inferenzstatistischen Auswertung wurden die Antworten der Lehramtsstudierenden einer Analyse mit einem explanatory item response model mit kreuzklassifizierten Zufallseffekten für Personen und Items unterzogen. Es konnten statistisch signifikante Effekte sowohl für die „Amalgam-“ als auch für die „Mitnahme-Bedingung“ nachgewiesen werden. Diese waren jedoch im Vergleich zur starken Kontrollgruppe, welche in beiden Kursblöcken explizit zum fachdidaktischen Wissen unterrichtet worden war, und auch im Vergleich zur „Moderator-Bedingung“ vergleichsweise schwach ausgeprägt. Moderatoreffekte des verfügbaren Fachwissens im Sinne einer besseren Nutzung von Lerngelegenheiten für fachdidaktisches Wissen konnten nicht nachgewiesen werden; in der starken Kontrollgruppe und in der „Mitnahme-Bedingung“ erwarben die teilnehmenden Studierenden in vergleichbarem Ausmaß fachdidaktisches Wissen. In Hinblick auf die interne Validität unserer Studie bestätigten Ratings von Ausschnitten aus Videoaufnahmen der durchgeführten Kurse die Umsetzung der intendierten Kursinhalte. Ebenso wiesen vor, zwischen und nach den beiden Kursblöcken eingesetzte Tests zum fachlichen und pädagogischen Wissen nach, dass die Studierenden in den fachlichen und pädagogischen Blöcken entsprechende Wissensbestände aufbauten.

Obwohl sich unsere Operationalisierung fachlichen und pädagogischen Wissens stark an der Struktur korrespondierenden fachdidaktischen Wissen orientierte, registrierten wir lediglich schwache Amalgamierungs- und Mitnahmeeffekte. Insgesamt sprechen die Befunde vor allem gegen eine bedeutende Rolle des pädagogischen Wissens für den Erwerb fachdidaktischen Wissens und heben das Potential einer direkten Vermittlung fachdidaktischer Inhalte hervor. Nichtsdestotrotz besitzt pädagogisches Wissen eine eigenständige Relevanz für die professionelle Tätigkeit von Lehrkräften und ist nicht bloß Mittel zum Zweck des Aufbaus fachdidaktischen Wissens.

Donnerstag
12. März

11:15 Uhr

HGC 40

Chair(s)

Wilde, Annett
Seiz, Johanna

Experimentelle Studien mit Lehramtsanwärtern zu Aspekten der professionellen Kompetenz

Wilde, Annett; Seiz, Johanna; Decker, Anna-Theresia; Kunter, Mareike

Challenge oder Threat - unter welcher Bedingung lernen angehende Lehrkräfte besser?

Wissen und Überzeugungen bilden zentrale kognitive Aspekte der professionellen Kompetenz von Lehrkräften (Baumert & Kunter, 2011). Lehrerfortbildungen zielen oft auf den Aufbau und die Veränderung von Überzeugungen ab (Lipowsky, 2011). Dabei werden jedoch strukturelle Faktoren stark betont, weniger im Fokus steht hingegen, welche Bedingungen auf individueller Ebene gegeben sein müssen, damit Lernprozesse erfolgreich sind. Das kognitiv-affektive Modell der Überzeugungsveränderung von Gregoire (2003) geht davon aus, dass substanzielle Veränderungen in Lehrerüberzeugungen nur unter bestimmten motivationalen Voraussetzungen möglich sind. Weiterhin trifft es die Annahme, dass eine systematische, intensive Verarbeitung neuer Informationen notwendig ist, um eine langfristige Überzeugungsveränderung zu erzielen. Die Voraussetzung für eine systematische Verarbeitung von Lerninhalten ist wiederum, dass die Situation, in der die neuen Inhalte präsentiert werden, als Herausforderung und nicht als Bedrohung wahrgenommen wird.

Ziel unserer Studie war es, die Annahme zu überprüfen, dass die Wahrnehmung einer Lernsituation als Bedrohung oder Herausforderung zu einer jeweils unterschiedlichen Verarbeitung von Lerninhalten führt. Dazu wurde ein computerbasiertes Experiment durchgeführt, in dem die Bewertung einer Lernsituation als Herausforderung vs. Bedrohung anhand von manipulierten Rückmeldungen systematisch variiert wurde. Die Teilnehmer/innen bearbeiteten einen Wissenstest zum Thema Klassenführung und erhielten danach zufällig eine von drei Rückmeldungen (Bedrohung, Herausforderung, neutral). Anschließend folgte die Lerngelegenheit, ein Text zum Thema Klassenführung. Danach wurden Wissen, Verarbeitungstiefe und Anstrengung beim Lesen sowie motivationale Variablen (Selbstwirksamkeit, Interesse am Thema) erfasst.

In zwei Pilotstudien wurden bisher insgesamt 59 Lehramtsstudierende befragt. Bis zur Tagung im März werden neben den Daten der Pilotierungen zusätzlich die Daten der Hauptstudie (N=90) vorliegen.

Die vorläufigen Ergebnisse zeigen eine Reihe erwarteter Effekte. Wer eine bedrohliche Rückmeldung erhalten hatte, strengte sich beim Lesen am wenigsten an, wendete weniger Zeit zum Lesen des Textes auf und neigte zu weniger tiefer Verarbeitung. Teilnehmer/innen der Bedrohungsbedingung zeigten auch weniger Interesse und erlebten sich als weniger selbstwirksam in Bezug auf Klassenführung. Personen mit einer bedrohlichen Rückmeldung konnten zudem vor allem Wissen reproduzieren, während Teilnehmer/innen mit herausfordernder Rückmeldung Schlüsse aus dem Gelesenen ziehen konnten, was für eine tiefere Verarbeitung spricht. Insgesamt weisen die Ergebnisse darauf hin, dass die Bewertung der Lernsituation einen direkten Einfluss auf die Lernbereitschaft, Verarbeitungstiefe und Lernerfolg hat. Daraus lassen sich Hinweise für die lernförderliche Gestaltung von Lehrerfortbildungen ableiten.

Donnerstag
12. März

11:15 Uhr

HGC 40

Chair(s)

Wilde, Annett
Seiz, Johanna

Experimentelle Studien mit Lehramtsanwärtern zu Aspekten der professionellen Kompetenz

Heitzmann, Nicole; Fischer, Frank; Fischer, Martin

Förderung von Diagnosekompetenz mit Lösungsbeispielen: Effekte von Scaffolding mit Selbsterklärungsprompts

Einen Instruktionsansatz im Klassenzimmer zu implementieren und zu analysieren ist für Lehrkräfte eine entscheidende Kompetenz bei der die Anwendung wissenschaftlichen Wissen und eigener Erfahrung eine Rolle spielt. Die dafür notwendige Diagnosekompetenz ist eine wichtige Kompetenz für Lehrkräfte (Artelt & Gräsel, 2009). Der Erwerb von Diagnosekompetenz ist jedoch ein komplexer Prozess. Um Lehrkräfte bei der Anwendung von wissenschaftlichem Wissen im Klassenzimmer zu unterstützen liegt ein fallbasierter Ansatz nahe. Eine erfolgreich zur Diagnosekompetenzförderung eingesetzte Methode sind Lösungsbeispiele mit integrierten Fehlern (Stark et al. 2011). Eine Möglichkeit um Lerner zu aktiver Verarbeitung der Fehler anzuregen ist die Integration von Selbsterklärungsprompts (Schworm & Renkl, 2007). Um aus den integrierten Fehlern zu lernen ist es notwendig, dass Lerner die Fehler erkennen und verstehen (Schank, 1999). Dazu kann Feedback notwendig sein. Feedback kann Lernern zusätzlich helfen die Hintergründe eines Fehlers sowie Strategien zu dessen Vermeidung zu verstehen (Kolodner, 2006). Ausführliches Feedback kann jedoch im Lernverlauf redundant werden (van Gog, Paas, & van Merriënboer, 2008). Eine Möglichkeit Feedback an die Bedürfnisse des Lerners anzupassen, ist die Nutzung von adaptierbarem Feedback (Leutner, 2002).

In der vorliegenden Studie wird der Frage nachgegangen inwiefern der Erwerb von Diagnosekompetenz durch Selbsterklärungsprompts und adaptierbares Feedback in einer computerbasierten Lernumgebung gefördert werden kann.

Dazu wurde eine experimentelle Studie mit einem 2x2-faktoriellen Design durchgeführt. Es nahmen N = 108 Studierende des Lehramts oder der Pädagogik teil. Die variierten Faktoren waren Selbsterklärungsprompts (mit vs. ohne) und adaptierbares Feedback (mit vs. ohne). Lerner bearbeiteten fehlerhafte Lösungsbeispiele in einer online Lernumgebung. Die Lerner wurden gebeten sich in einen fiktiven Praktikanten in einer Schule hineinzusetzen, der Lerneinheiten nach Problem-based Learning (Hmelo-Silver, 2004) vorbereitet und durchführt. Während der Diagnose macht der fiktive Praktikant Fehler. Lerner in der Bedingung mit Selbsterklärungsprompts wurden anschließend zur Selbsterklärung des Fehlers angeregt. Eine fiktive erfahrende Lehrkraft gab nach jedem Fehler in allen vier Bedingungen Feedback. In der Bedingung mit adaptierbarem Feedback konnten Lerner die Elaboriertheit des Feedbacks an die eigenen Bedürfnisse anpassen. Diagnosekompetenz wurde mittels eines Modells von Stark und Kollegen (2011) operationalisiert. Diagnosekompetenz besteht darin aus konzeptuellem Wissen und aus Handlungswissen. Konzeptuelles Wissen wurde im Vor- und Nachtest mit einem 21-Item umfassenden MC-Test gemessen. Handlungswissen wurde mit 10 Key-Feature Aufgaben und 10 Problemlöseaufgaben erfasst.

Die Ergebnisse zeigten, dass Selbsterklärungsprompts in Kombination mit adaptierbarem Feedback einen positiven Effekt auf konzeptuelles Wissen hatten ($F(1,103) = 5,57$, $p < ,05$, $\text{partial } \eta^2 = ,05$). Lernende die mit Selbsterklärungsprompts und mit adaptierbarem Feedback lernten ($M = 61,96$; $SD = 5,70$) erwarben mehr konzeptuelles Wissen als Lerner ohne Selbsterklärungsprompts und mit adaptierbarem Feedback ($M = 59,19$; $SD = 8,32$, $\text{post-hoc comparison } p < ,05$). Im Gegensatz dazu erwarben Lernende mit Selbsterklärungsprompts ($M = 19,69$; $SD = 8,12$) weniger Handlungswissen als Lernende ohne Selbsterklärungsprompts ($F(1,103) = 12,60$, $p < ,01$, $\text{partial } \eta^2 = ,11$; $M = 26,17$; $SD = 9,41$, $\text{post-hoc comparison } p < ,01$).

Während Prompts oft generell mit positiven Wirkungen assoziiert werden, fehlen bislang systematische Untersuchungen der differenzierten Effekte insbesondere im Kontext des Lerners mit Fehlern. In dieser Studie hatten Selbsterklärungsprompts insgesamt einen unerwarteten negativen Einfluss auf Handlungswissen. Möglicherweise befanden sich die Lerner in einer frühen Phase des kognitiven Fertigkeitserwerb und konzentrierten sich auf das Verständnis von zugrundeliegenden Prinzipien (Renkl & Atkinson, 2010). Erst in einer mittleren Phase des Fertigkeitserwerbs sind Lerner fähig über abstrakte Strategien zur Problemlösung zu reflektieren (Anderson, 1982). Selbsterklärungsprompts und adaptierbares Feedback konnten nur in Kombination konzeptuelles Wissen fördern. Möglicherweise realisierten Lerner mithilfe der Selbsterklärungsprompts ihren Mangel an zugrundeliegenden Konzepten und fokussierten dann anschließend auf das konzeptuelle Wissen in dem bereitgestellten Feedback.

Donnerstag
12. März

11:15 Uhr

HGC 40

Chair(s)

Wilde, Annett
Seiz, JohannaExperimentelle Studien mit Lehramtsanwärtern zu Aspekten der
professionellen Kompetenz

Markova, Maria; Pit-ten Cate, Ineke; Krolak-Schwerdt, Sabine; Glock, Sabine

„Das habe ich vorher noch nie versucht, also bin ich völlig sicher, dass ich es schaffe!“*: Selbstwirksamkeit und wahrgenommene Kompetenz von Lehramtsstudierenden im Umgang mit Schülern mit sonderpädagogischem Förderbedarf

Mit der Ratifizierung der UN-Behindertenrechtskonvention 2009 wurde auch in Deutschland der inklusive Unterricht eingeführt. Kinder mit und ohne sonderpädagogischem Förderbedarf (SPF) werden in einer Klasse gemeinsam unterrichtet. Jedoch fühlen sich viele Lehrkräfte unzureichend vorbereitet, um auf den individuellen Lernbedarf von Kindern in einer derart zusammengesetzten Klassengemeinschaft einzugehen und effektiv zu unterrichten (Blanton, Pugach, Florian, 2011). Lehramtsstudierende und Lehrkräfte stehen inklusivem Unterricht generell positiv gegenüber (Avramidis, Bayliss & Burden, 2000), allerdings empfinden sie den Umgang mit Kindern mit Lernschwierigkeiten oder sozialen-emotionalen Verhaltensweisen von SchülerInnen als schwierig (Killu, 2008; Lambe & Bones, 2006). Inwieweit Lehrkräfte sich den Umgang mit SchülerInnen mit SPF zutrauen hängt u.a. von der Art und Schwere des SPF der Schüler ab (Avramidis & Norwich, 2002). Forschungsergebnisse zeigen, dass die Einstellungen von Lehrkräften gegenüber inklusivem Unterricht von SchülerInnen mit Verhaltens- oder emotionalen Störungen negativer ist als gegenüber der Inklusion von SchülerInnen mit Lernschwierigkeiten (Levins, Bornholt, & Lennon, 2005; Ward, Center, & Bochner, 1994). Ein weiterer Faktor ist die wahrgenommene Kompetenz und Selbstwirksamkeit der Lehrperson (Soodak & Podell, 1993). Dabei stehen Selbstwirksamkeit und Kompetenzerleben im positiven Zusammenhang: Je höher die wahrgenommene Kompetenz, desto eher fühlen sich Personen selbstwirksam (Bandura, 1977). In der vorliegenden Studie wird die Wahl von pädagogischen Maßnahmen gegenüber SchülerInnen mit SPF in Abhängigkeit von Art der SPF, sowie der Selbstwirksamkeit und wahrgenommener Kompetenz von Lehramtsstudierenden experimentell untersucht.

Am Experiment nahmen 50 deutsche Lehramtsstudierende (37 weiblich) mit einem Durchschnittsalter von 23,04 Jahre (SD = 2,94) und einer Lehrerfahrung von 2,19 Monaten, SD = 3,27, Range: 0-14 Monate. Die Lehramtsstudierenden wurden gebeten, sich vorzustellen, dass sie die folgenden zwei Schüler in ihrer Klasse unterrichten. Ein Fall beschrieb einen lernauffälligen Schüler, der andere einen verhaltensauffälligen Schüler. Die Darbietung der Schülerfälle erfolgte in zufälliger Reihenfolge und für jeden Schülerfall satzweise. Die Teilnehmer beurteilten nach jedem Schülerfall, ob es sich um ein Problem mit einer Diagnose handelt (ja/nein) und wurden gebeten, die Wahrscheinlichkeit auf einer 5-Punkt-Likert-Skala („würde ich in keinem Fall tun“ bis „würde ich in jedem Fall tun“) anzugeben, mit der sie 11 vorgegebene pädagogische Maßnahmen vornehmen würden. Das Material der Studie wurde in Anlehnung an Lanfranchi und Jenny (2005) entwickelt. Nach dem Bearbeiten der Schülerfälle wurden die Teilnehmer gebeten, einen Fragebogen zur Erfassung ihrer Selbstwirksamkeit für inklusive Bildung (Sharma, Loreman & Forlin, 2012) sowie einen Fragebogen zur Erfassung der wahrgenommenen Kompetenz (in Anlehnung an Dessemontet, Benoit, & Bless, 2011) auszufüllen.

Es zeigte sich, dass Lehramtsstudierende sich als selbstwirksam und kompetent bezüglich Inklusion wahrnehmen. Der Zusammenhang zwischen Lehrerselbstwirksamkeit und der wahrgenommenen Kompetenz ($r = .39$) war positiv. In 53% der Beurteilungen wurde die beschriebene Schülersymptomatik nicht korrekt identifiziert, dennoch wurden für beide Schülerfälle im Durchschnitt drei pädagogische Maßnahmen mit hoher Anwendungswahrscheinlichkeit gewählt. Es zeigten sich Unterschiede in den gewählten pädagogischen Maßnahmen in Abhängigkeit von der Art des SPF und der Kompetenz. Es zeigte sich jedoch weder ein Effekt der Diagnose noch der Selbstwirksamkeit. Der Schüler mit Verhaltensproblemen wurde häufiger zur externen Beratung und an Schulpsychologen überwiesen, während für den Schüler mit Lernschwierigkeiten öfter eine Klassenwiederholung in Betracht gezogen wurde. Lehramtsstudierende mit relativ hoher wahrgenommener Kompetenz schlugen tendenziell öfter vor, die Eltern zu kontaktieren, besonders für den verhaltensauffälligen Schüler.

Zusammenfassend nahmen sich Lehramtsstudierende zwar als kompetent und selbstwirksam bezüglich inklusiven Unterrichts wahr, sie waren jedoch nicht in der Lage, SPF korrekt zu identifizieren. Die gewählten pädagogischen Maßnahmen variierten abhängig von der Art des SPF. Jedoch konnte kein Zusammenhang zwischen wahrgenommener Kompetenz, Selbstwirksamkeit und der Auswahl pädagogischer Maßnahmen gefunden werden. Die Befunde deuten darauf hin, dass subjektiv empfundene Kompetenz und Selbstwirksamkeit nicht automatisch zu hohen diagnostischen Fähigkeiten führen, die Lehramtsstudierende befähigen, differenzierte Maßnahmen zur Unterstützung von Schülern mit SPF auszuwählen.

Donnerstag
12. März

11:15 Uhr

HGC 40

Chair(s)

Wilde, Annett
Seiz, Johanna

Experimentelle Studien mit Lehramtsanwärtern zu Aspekten der professionellen Kompetenz

Hilkenmeier, Johanna

Gespräche zwischen Eltern und Lehrkräften

Im Zuge des Wandels zum inklusiven Unterricht und eines damit verbundenen Umgangs mit heterogenen Lernvoraussetzungen kann eine verstärkte Kooperation von Elternhaus und Schule ein gewinnbringender Ansatz sein. Das Elternhaus hat unumstritten einen wesentlichen Einfluss auf die Kompetenzentwicklung im Kindesalter (z.B. Bos et al., 2012). Hierbei spielt das elterliche Unterstützungsverhalten eine wichtige Rolle. Prominente Konzeptualisierungen (Epstein, z.B. 1986; Grolnick & Sloviaczek, 1994) verdeutlichen, dass das Konstrukt eine Vielfalt an Aktivitäten umfasst, beispielsweise die Teilnahme am Schulleben, Gremienarbeit oder die Gestaltung der häuslichen Lernumgebung. Je nach Klassenstufe und Schulform sind verschiedene Facetten unterschiedlich bedeutsam (z.B. Hill & Tyson, 2009).

Um elterliches Schulengagement zu unterstützen, sollte die Kommunikation und Kooperation zwischen Elternhaus und Schule optimal gestaltet und gefördert werden. Zwar sind im deutschen Sprachraum vielzählige Ratgeber zum Thema „Elternarbeit“ zu finden, die Anzahl wissenschaftlicher Publikationen ist jedoch vergleichsweise überschaubar (z.B. Wild & Lorenz, 2010). Kommunikation zwischen Elternhaus und Schule in Deutschland beschränkt sich größtenteils auf formelle Kontakte wie Elternabende und Elternsprechtage (z.B. Sacher, 2008). Das oftmals beschriebene Spannungsfeld zwischen den Sozialisationsinstanzen ist zum Teil von gegenseitigen Verantwortungszuschreibungen (z.B. Neuenschwander et al., 2005) und kritischen Einstellungen zu elterlichem Schulengagement geprägt (z.B. Witjes & Zimmermann, 2000). Viele Eltern fühlen sich schlecht informiert und haben Beratungsbedarf (z.B. Krumm, 1996; Sacher, 2008), und auch Lehrkräfte sind damit konfrontiert, sich in ihrer Ausbildung optimal auf Elternarbeit und Elternberatung vorzubereiten (z.B. Landesinstitut für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2001).

Das Symposium nähert sich aus unterschiedlichen Perspektiven und mittels verschiedener methodischer Herangehensweisen den Herausforderungen und Möglichkeiten, die sich durch Gespräche zwischen Lehrkräften und Eltern ergeben. Die Beiträge lassen sich hierbei den Disziplinen Erziehungswissenschaft, Linguistik, Lehrerbildung und Psychologie zuordnen.

Im ersten Beitrag befassen sich Hertel et al. anhand von Daten aus PISA 2012 mit Rahmenbedingungen für die Zusammenarbeit von Schule und Familie, insbesondere im Kontext von Elterngesprächen. Zudem wird der Zusammenhang zwischen elterlichem Schulengagement und kindbezogenen Variablen beleuchtet.

Der Beitrag von Hilkenmeier et al. kombiniert in einem längsschnittlichen Design quantitative mit qualitativen Analysemethoden, um zu prüfen, welche Wirkung Elternsprechtage auf das elterliche Schulengagement haben.

Ergänzend dazu betrachten Bennewitz und Wegner Elterngespräche mit soziologisch-linguistischen und erziehungswissenschaftlichen qualitativen Methoden auf Mikroebene.

Der letzte Beitrag widmet sich der Verbesserung von Kommunikation zwischen Eltern und Lehrern, indem er Möglichkeiten des Kompetenzaufbaus der Lehrer aufzeigt. Aich et al. befassen sich vor einem transaktionsanalytischen Hintergrund mit der Konzeption und Evaluation eines Lehrertrainings zur Förderung von Gesprächskompetenz in Bezug auf die Beratung von Eltern durch Lehrpersonen.

Donnerstag
12. März

11:15 Uhr

HGC 50

Chair(s)

Buhl, Heike M.
Hilkenmeier, Johanna

Hertel, Silke; Jude, Nina; Sälzer, Christine

Die Zusammenarbeit von Lehrkräften und Eltern im Rahmen von Elterngesprächen – Befunde aus PISA 2012

Die Familie ist, neben der Schule, eine zentrale Lernumgebung von Kindern und Jugendlichen. Zahlreiche empirische Befunde weisen darauf hin, dass das schulbezogene Engagement der Eltern in einem positiven Zusammenhang mit der Leistung sowie leistungsrelevanten Verhaltensweisen und Einstellungen steht (Christenson, 1995; Crosnoe, 2009). Schülerinnen und Schuler profitieren sowohl in kognitiver als auch in nicht-kognitiver Hinsicht davon, wenn Eltern in ihrer schulischen Entwicklung involviert sind, beispielsweise indem sie anregende Lernumgebungen zu Hause schaffen, im Kontakt mit den Lehrerinnen und Lehrern bleiben und an Schulaktivitäten teilnehmen (Alexander, Entwisle & Olson, 2007; Christenson, 2004; Hoover-Dempsey & Sandler, 1997; Ma, 1999; Raver, Gershoff & Aber, 2007). Wichtig ist hier allerdings die Unterscheidung zwischen einer förderbezogenen Zusammenarbeit mit der Schule, z.B. im Rahmen von Gesprächen mit Lehrkräften, und einem eher breiten Engagement, z.B. die Mitarbeit bei Schulfesten.

Der Fokus dieses Beitrags liegt auf der Beschreibung der Rahmenbedingungen für die Zusammenarbeit von Schule und Familie sowie der Zusammenhänge von Elternengagement – insbesondere der Teilnahme an Elterngesprächen – und Leistung sowie Interesse in der Domäne Mathematik (Schwerpunkt in PISA 2012). Den theoretischen Rahmen bilden das Angebots-Nutzungs-Modell für die Elternberatung (Hertel et al., 2013) sowie das Modell der elterlichen Beteiligung von Green et al. (2007).

Grundlage für die Analysen sind (1) die Angaben von N=2885 Eltern der 15-Jährigen Schülerinnen und Schüler in PISA 2012 in Deutschland sowie (2) die Daten zu Leistung und Interesse im Bereich Mathematik der Schülerinnen und Schüler. Insgesamt weisen 9.1% der Eltern, von denen ein Elternfragebogen vorliegt, Migrationshintergrund auf. Ausfallanalysen zeigen einen systematischen Ausfall in dieser Gruppe der befragten Eltern.

Wichtige Rahmenbedingungen für die Zusammenarbeit von Schule und Eltern sind ein partizipationsfreundliches Schulklima sowie die Einladung von Lehrkräften, z. B. zu Elternabenden und -gesprächen. Hier zeigt sich, dass die Eltern die Schulqualität sehr positiv und statistisch signifikant höher als in PISA 2009 beurteilen. Allerdings wünschen sie sich regelmäßiger Informationen zu den Fortschritten ihres Kindes. Zudem zeigt sich, dass Gespräche mit den Lehrkräften häufiger durch die Eltern selbst angeregt als von Lehrkräften initiiert werden. Wenn Eltern sich an der Schule ihres Kindes beteiligen, dann am häufigsten im Rahmen von Gesprächen mit Lehrkräften über das Verhalten und die Fortschritte ihres Kindes bzw. über die Entwicklung des Kindes. Etwa zwei Drittel der Eltern geben an, im letzten Schuljahr auf ihre Initiative hin mindestens ein Gespräch mit der Lehrkraft ihres Kindes über dessen Verhalten bzw. Fortschritt geführt zu haben. Andere Formen der Beteiligung, wie beispielsweise Mitarbeit im Unterricht oder das Halten eines Gastvortrags, kommen deutlich seltener vor (Zustimmung jeweils <20%).

Vergleichsweise selten ist auch die mathematische Förderung im Elternhaus. Etwa jedes fünfte Elternteil gibt an, mit dem Kind über die Anwendungsmöglichkeiten von Mathematik zu sprechen bzw. dem Kind bei den Hausaufgaben in Mathematik zu helfen. Das zur Verfügung stellen zusätzlicher Ressourcen – konkret das Kaufen von Mathematikmaterialien – ist ebenfalls sehr selten. Hier stimmt nur jedes 10. Elternteil zu.

Regressionsanalysen zeigen, dass Eltern mit geringerem sozioökonomischen Hintergrund – unter Kontrolle von Bildungsniveau und Migrationshintergrund – seltener an Beratungsgesprächen teilnehmen. Die Teilnahme an Gesprächen mit der Lehrkraft steht, so wie die individuelle Förderung im Elternhaus, in einem negativen Zusammenhang mit der Leistung und dem Interesse des Kindes bezogen auf Mathematik.

Elterngespräche stellen demnach einen wichtigen Zugang und einen Startpunkt für die Zusammenarbeit von Schule und Eltern dar. Die negativen Zusammenhänge der Teilnahme an Elterngesprächen bzw. der Förderung in der Familie und Leistung sowie Interesse für Mathematik lassen sich möglicherweise darauf zurückführen, dass entsprechende Maßnahmen dann einsetzen, wenn Schwierigkeiten bestehen (Remedialeffekt). Da es sich um querschnittliche Analysen handelt, lassen sich keine kausalen Aussagen über die Wirkrichtung treffen. Auf der Grundlage von Daten aus PISA PLUS werden die längsschnittliche Betrachtung und die Ableitung von Wirkungsaussagen möglich.

Donnerstag
12. März

11:15 Uhr

HGC 50

Chair(s)

Buhl, Heike M.
Hilkenmeier, Johanna

Hilkenmeier, Johanna; Wiescholek, Sabrina; Buhl, Heike M.

Was passiert am und nach dem Elternsprechtag? Wahrnehmung von Elterngesprächen und ihre Wirkung auf elterliches Schulengagement.

Unter elterlichem Schulengagement, oder auch Parental Involvement, kann die elterliche Teilhabe am schulischen Alltag des Kindes verstanden werden. Dass verschiedene Facetten dessen einen positiven Einfluss auf die Kindesentwicklung haben, konnte bereits vielfach bestätigt werden (z.B. Henderson & Mapp, 2002; Hill & Tyson, 2009). Was bewegt Eltern jedoch dazu, sich im schulischen Alltag ihres Kindes zu engagieren? Mögliche Prädiktoren elterlichen Schulengagements wurden bereits konzeptualisiert und in Teilen empirisch untersucht (Eccles & Harold, 1993; Grolnick et al., 1997; Hoover-Dempsey & Sandler, 2005). Auch für den deutschen Sprachraum lassen sich hierzu erste Befunde finden (Hilkenmeier, Wiescholek & Buhl, in Vorb.; Yotodying & Wild, 2014). Insgesamt ist festzuhalten, dass in Abhängigkeit von der erfassten Unterstützungsform Faktoren wie die elterliche Selbstwirksamkeit und Rollenwahrnehmung, die Ressourcen der Eltern oder die Einladungen durch das Kind und die Schule eine bedeutsame Rolle spielen.

Im vorliegenden Beitrag wird der Frage nachgegangen, was sich hinter dem Prädiktor der Einladungen durch die Schule verbirgt und wie bzw. ob dieser auf das elterliche Unterstützungsverhalten wirkt. Da sich Kontakte zwischen Elternhaus und Schule im deutschen Sprachraum häufig auf Elternabende und Elternsprechtage beschränken (z.B. Neuenschwander et al., 2005; Sacher, 2008), wurde daher die elterliche Wahrnehmung von Gesprächen am Elternsprechtag sowie ihre Wirkung auf elterliches Unterstützungsverhalten im Bereich Lesen fokussiert. Hierzu wurden 219 Schulkinder der dritten Klasse drei bis vier Wochen vor und nach dem Elternsprechtag befragt. Außerdem wurden die Eltern am Elternsprechtag sowie drei bis vier Wochen danach befragt.

Neben strukturellen Variablen und der Leseleistung des Kindes wurden auf Kindes- und Elternseite Variablen zum elterlichen Unterstützungsverhalten erfasst (z.B. gemeinsame lesebezogene Aktivitäten). Zusätzlich wurden Eltern und Kinder zur wahrgenommenen Veränderung im elterlichen Unterstützungsverhalten befragt. Direkt nach dem Gespräch wurden den Eltern Fragen zur Quantität und Qualität des Elterngesprächs gestellt. Freitext-Angaben zu Gesprächsinhalten wurden inhaltsanalytisch ausgewertet (vgl. Mayring, 2010) und von drei Ratern kodiert (Krippendorff's $\alpha = .70-.81$). Auf Grundlage der Selbstbestimmungstheorie (Decy & Ryan, z.B. 2004) und darauf aufbauenden Arbeiten zur Qualität häuslicher Lernumgebung (Lorenz & Wild, 2007) wurde die wahrgenommene Qualität des Elterngesprächs durch die Variablen Autonomieunterstützung, Strukturgebung und Kontrolle operationalisiert.

Die Ergebnisse bestätigen einen Zusammenhang zwischen wahrgenommener Gesprächsqualität und einer berichteten Veränderung im elterlichen Unterstützungsverhalten. Eine niedrige Leseleistung führt, vermittelt über die wahrgenommene Autonomieunterstützung durch die Lehrkraft, zu einer berichteten Veränderung lesebezogenen Unterstützungsverhaltens. Varianzanalytisch zeigen sich zusätzlich im Rahmen einer signifikanten Interaktion Gruppenunterschiede im Unterstützungsverhalten zu den zwei Messzeitpunkten. Wird zu T1 von einem geringen lesebezogenen Unterstützungsverhalten berichtet, ist zu T2 ein bedeutsamer Zuwachs zu verzeichnen. Wird hingegen zu T1 von einem hohen Ausmaß elterlichen Unterstützungsverhaltens berichtet, ist zu T2 eine Abnahme zu beobachten. Zusätzlich zeigt sich bezüglich der Leseleistung ein signifikanter Gruppenunterschied über die Zeit. Kinder mit einer niedrigeren Leseleistung zu T1 weisen zu T2 einen höheren Zuwachs in der Leseleistung auf als jene mit hoher Leseleistung zu T1.

Zusammenfassend sprechen die Ergebnisse für eine Wirksamkeit von Elternsprechtagen. So scheinen die Gespräche, zumindest kurzfristig, einen motivierenden und in Teilen förderlichen Einfluss auf das elterliche Schulengagement und die Leistung des Kindes zu haben. Die Ergebnisse werden im Lichte aktueller Befunde und (Alltags-)theorien zur Bedeutsamkeit von Elternsprechtagen diskutiert.

Donnerstag
12. März

11:15 Uhr

HGC 50

Chair(s)

Buhl, Heike M.
Hilkenmeier, Johanna

Bennewitz, Hedda; Wegner, Lars

Elternsprechtagsgespräche – Begegnungen zwischen Schule und Elternhaus aus empirischer Perspektive

Elternsprechtagsgespräche sind soziale Ereignisse, an denen Eltern bzw. Familien und Lehrpersonen einander begegnen. Obwohl die Zusammenarbeit von Schule und Elternhaus zu einem ‚Dauerbrenner‘ pädagogischer Diskurse gehört (vgl. Busse & Helsper, 2007), steht die empirische, auf authentischen Daten fußende Analyse dieser konkreten Kommunikationssituation weitgehend aus. Ziel unseres Beitrags ist es, sich dem konstatierten Forschungsdesiderat anzunehmen und die Verknüpfung familiärer und schulischer Interaktionen in der besagten Gesprächssituation mittels eines interdisziplinären Zugangs zum Datenmaterial zu rekonstruieren. Die empirische Datenbasis hierfür liefert das im Rahmen des Dissertationsprojekts "Gespräche zwischen Lehrkräften, Eltern und SchülerInnen an schulischen Elternsprechtagen – eine gattungsanalytische Untersuchung" von Lars Wegner erhobene Korpus von 142 Elternsprechtagsgesprächen. Diese Gespräche entstammen unterschiedlichen Schulformen (Grundschule, Hauptschule, Gesamtschule, Gymnasium) und wurden in verschiedenen Städten Nordrhein-Westfalens erhoben und im Anschluss daran gemäß gesprächsanalytischer Konventionen verschriftlicht, um sie damit der konkreten Untersuchung zugänglich zu machen.

Gemäß der beschriebenen interdisziplinären Ausrichtung bedienen wir uns sowohl des soziologisch-linguistischen wie auch des erziehungswissenschaftlichen Methoden-instrumentariums, dessen Schnittpunkt in ethnomethodologischen bzw. konversationsanalytischen Grundannahmen zu bestimmen ist (vgl. Bergmann, 2000, Deppermann, 2001): Mittels eines sequenzanalytischen Vorgehens geht es uns zunächst um die Untersuchung relevanter lokaler Phänomene, also das Identifizieren und Beschreiben jener sprachlichen Mittel, die von den Beteiligten eingesetzt werden, um die soziale Gesprächswirklichkeit, in der sie sich gerade befinden, zu erzeugen (vgl. Goffman, 1982, 1986). Die daraus resultierenden Befunde werden in eine erziehungswissenschaftliche Perspektive überführt.

Im Rahmen der Analyse fokussieren wir auf verschiedene Aspekte, die – wie der empirische Zugang zum Datenmaterial offenbart – für die soziale Situation "Elternsprechtagsgespräch" konstitutiv sind: So zeigt sich beispielsweise, dass Eltern, LehrerInnen und SchülerInnen in den Gesprächen Verantwortlichkeiten und Zuständigkeiten aushandeln, die die erfolgreiche Realisierung des gemeinsamen, vom Schulgesetz deklarierten Bildungs- und Erziehungsauftrags am Kind betreffen. Diese Aushandlungsprozesse gestalten sich keineswegs immer unproblematisch: Die Beteiligten realisieren vorwurfsähnliche Handlungen (Günthner, 1999), die an das jeweilige Gegenüber gerichtet sind. Statt – wie vom Schulgesetz deklariert – in "partnerschaftlicher Zusammenarbeit" nach konstruktiven Lösungen zu suchen, sind die Interagierenden primär darauf bedacht, ein positives Selbstbild ("gutes Schülerin", "gute Lehrkraft) zu bewahren. Es wird versucht, der anderen Partei die Verantwortung für misslungene Bildungs- und Erziehungsprozesse zuzuschreiben. Statt also über konkrete Förderungsmaßnahmen zu beraten (vgl. Schwarzer & Posse, 2005), die der konstatierten negativen Entwicklung entgegenwirken, wird die "Schuld"-Frage diskutiert. Die Gespräche scheinen gestört und ihre Dauer zu kurz, als dass noch produktive Lösungswege erwogen werden könnten.

Im Rahmen unseres Beitrags kann schließlich gezeigt werden, dass und wie die Konfliktlinien zwischen Schule und Familie in diesem spezifischen Gesprächstyp eine konstruktive, partnerschaftliche Zusammenarbeit behindern.

Donnerstag
12. März

11:15 Uhr

HGC 50

Chair(s)

Buhl, Heike M.
Hilkenmeier, Johanna

Aich, Gernot; Kuboth, Christina; Behr, Michael

Das Gmünder Gesprächsmodell zur Gesprächsführung mit Eltern – Entwicklung und Evaluation eines Lehrertrainings

Hintergrund: Die Gesprächsführung und Beratung von Eltern durch Lehrpersonen wird sowohl aus wissenschaftlicher Sicht als auch von den handelnden Akteuren selbst als ein wichtiger Bestandteil der Lehrerverberufung angesehen (Krumm, 1996, Sacher, 2008). Trotzdem ist die Ausbildung von Lehramtsstudierenden in Beratung und Gesprächsführung mit Eltern zumindest in Baden-Württemberg noch sehr rudimentär (Behr, 2005). Aus diesem Grund wurde in Schwäbisch Gmünd das „Gmünder Gesprächsmodell zur Gesprächsführung mit Eltern“ (Aich, 2011, Aich & Behr, 2010) entwickelt. Theoretischer Hintergrund sind die Transaktionsanalyse und der personenzentrierte Ansatz. Das Training ist in fünf Module gegliedert. Es werden die drei Kernbedingungen nach Rogers – Empathie, bedingungsfreie Wertschätzung und Authentizität – und aus der Transaktionsanalyse die Grundpositionen, die Ich-Zustände, die Kommunikationsregeln und die Spieltheorie nach Berne in integrierter Form vermittelt. Die einzelnen Konzepte sind so vernetzt, dass sie auf das spezifische Beratungssetting der Schule bezogen sind. Ziel ist es, die Kommunikationskompetenzen zu verbessern, die laut dem Bildungsplan 2004 in Baden-Württemberg von Lehrpersonen gefordert werden.

Das Training wurde auf drei verschiedenen Levels durchgeführt und evaluiert: (1) Lehramtsstudierende der Pädagogischen Hochschule Schw. Gmünd (N=151). Hier betrug die Trainingszeit ca. 35 h. (2) Referendarinnen und Referendare des Staatlichen Seminars für Didaktik und Lehrerbildung Schw. Gmünd (N= 48) mit einer Trainingszeit von ca. 42 h. (3) Berufserfahrene Lehrerinnen und Lehrer aus Baden-Württemberg (N=48) mit einer Trainingszeit von ca. 45 h.

Flankierend wurde auf Wunsch Supervision angeboten. Aus diesem Grund differiert die Trainingszeit zwischen den einzelnen Levels.

Fragestellungen: (1) Verbessert sich das Gesprächsverhalten der Teilnehmerinnen und Teilnehmer hinsichtlich der Kernbedingungen nach Rogers – „Empathische Intervention der Lehrperson“, „Authentizität der Lehrperson“ und „Wertschätzende Interventionen der Lehrperson“? (2) Welche subjektiven Ziele setzen sich die Teilnehmerinnen und Teilnehmer und haben sie diese aus ihrer Sicht erreicht?

Methode: Der Fragestellung wird in einem quasiexperimentellen Versuchsdesign nachgegangen. Die Lehramtsstudierenden und die Referendarinnen und Referendare wurden in einem Wartegruppensdesign mit verschiedenen Messzeitpunkten und die berufserfahrenen Lehrerinnen und Lehrer in einem Kontrollgruppensdesign mit zwei bzw. drei Messzeitpunkten untersucht (Prä-Post und teilweise Follow up). Für die Evaluation wurden bereits erprobte Rollenspielszenarien sowie reale Gespräche zwischen Eltern und Lehrkräften eingesetzt, die auf Audiobändern aufgenommen wurden und dann mittels einer Fremdratingprozedur durch zwei unabhängige Rater evaluiert wurden. Als Instrument dienten bewährte Ratingskalen von Pfeiffer (1977) in Anlehnung an Carkhuff (1969a, 1969b; Tausch & Tausch 1990).

Die Trainingsziele der Teilnehmerinnen und Teilnehmer und der Grad der Zielerreichung wurden über die Goal-Attainment-Scale (GAS) erhoben.

Ergebnisse: Hinsichtlich der drei Kernbedingungen nach Rogers zeigen die verschiedenen Trainingsgruppen fast immer einen signifikanten bis hoch signifikanten Zuwachs ihrer Gesprächskompetenz mit mittleren bis hohen Effektstärken.

Die Auswertung der Goal-Attainment-Scale (GAS) zeigt einen hohen subjektiven Zielerreichungsgrad und belegt somit die Bedarfsangemessenheit des Trainings. Die qualitative Auswertung der von den Teilnehmerinnen und Teilnehmern genannten Ziele lassen sich den Themenbereichen Selbstmanagement, Gesprächskultur und Gesprächsführung zuordnen.

Donnerstag
12. März

11:15 Uhr

HGC 50

Jäntsich, Christian; Bosse, Stefanie; Henke, Thorsten; Lambrecht, Jennifer; Spörer, Nadine

Schritt für Schritt zur Inklusion? Lehrereinstellungen und Concerns im Prozess der Implementation inklusiven Unterrichts

Aktuell finden sich nunmehr bundesweit Bestrebungen, dem inklusiven Unterrichten in vielfältiger Hinsicht auf die Spur zu kommen (Preuss-Lausitz, 2014). Den Lehrkräften in den Schulen obliegt dabei die praktische Umsetzung des Inklusionsgedankens im Unterricht (Stellbrink, 2012). Somit geraten in der wissenschaftlichen Auseinandersetzung mit dem inklusiven Unterrichten auch die Lehrenden in den Blick. Um die Akzeptanz von Innovationen (als welche die bildungspolitisch motivierte Umstellung zur Inklusion gesehen werden kann) durch Lehrkräfte zu erfassen, fokussieren sogenannte concerns-theoretische Ansätze auf die Beschreibung kognitiver und affektiver Reaktionen innerhalb eines Implementationsprozesses (Pant, Vock, Pöhlmann, & Köller, 2008). Das Concerns Based Adoption Model (Hall & Hord, 2011) bietet einen konzeptionellen und analytischen Rahmen zur Steuerung schulischer Implementationsvorhaben. Mithilfe des darin enthaltenen diagnostischen Instruments der Stages of Concern (SoC) (Hall, George, & Rutherford, 1977) kann die Perzeption von Innovationen durch Lehrkräfte im zeitlichen Verlauf abgebildet werden. Das SoC-Modell beinhaltet sieben distinkte Stufen: (0) Bewusstsein, (1) Informationsbeschaffung, (2) persönliche Betroffenheit, (3) Aufgabenmanagement, (4) Auswirkung auf Lernende, (5) Kooperationsmöglichkeiten sowie (6) Revision und Optimierung. Der idealtypische Entwicklungsverlauf sieht für einen gelingenden Implementationsprozess eine wellenförmige Verschiebung von der ersten bis zur siebenten Stufe vor.

Aufgrund allein Prozess beschreibender Funktionen der SoC steht dieser Ansatz in der Kritik, theoretisch nur unzureichend ausgearbeitet zu sein und keine Anknüpfungspunkte an bestehende psychologische Konzepte zu bieten (Anderson, 1997; Watzke, 2007). Insofern als dass Lehrereinstellungen wichtige handlungsregulierende zugleich jedoch stabile Persönlichkeitsdispositionen sind, welche die individuelle Wahrnehmung gegenüber bestimmten Objekten beeinflussen, kommt diesen für die Implementation inklusiven Unterrichts eine besondere Bedeutung zu (Kunz, Luder, & Moretti, 2010). Die Bedeutsamkeit individueller Lehrermerkmale für den Unterricht bzw. die Unterrichtsgestaltung wurde überdies bereits verschiedentlich theoretisch eingebettet und modelliert (Baumert & Kunter, 2006; Blömeke, Felbrich, Müller, Kaiser, & Lehmann, 2008; Lipowsky, 2010).

Um aufzuzeigen, an welchem Punkt bzw. inwiefern die Lehrereinstellungen gegenüber der Inklusion im Implementationsprozess verortet werden können, werden im vorliegenden Beitrag die Stufen der SoC im Kontext der inklusiven Grundschule in Zusammenhang mit den Lehrereinstellungen gegenüber dem inklusiven Unterrichten gebracht. Welche entscheidende Stufe – oder vielmehr Hürde – muss nun überwunden sein, um positive Einstellungen gegenüber dem inklusiven Unterrichten auszubilden? Unter der Annahme, dass ein erfolgreicher Implementationsprozess mit einer positiven Einstellung aufseiten der Lehrer korrespondiert, ist davon auszugehen, dass höhere Stufen der SoC mithin positive Lehrereinstellungen zeitigen. Folgende Hypothese wird daraus abgeleitet:

Je höher die Stufe innerhalb der SoC, desto positiver ist der Zusammenhang zur Lehrereinstellung gegenüber dem inklusiven Unterrichten.

Grundlage dieser Studie bildet die erste Lehrkräftebefragungen von rund 900 Lehrerinnen und Lehrern im Rahmen des Pilotprojektes „Inklusive Grundschule“ des Landes Brandenburg, das seit dem Schuljahr 2012/13 von der Universität Potsdam wissenschaftlich begleitet wird (Bosse & Spörer, 2014a). Zur Messung der SoC wurde eine deutschsprachige Adaption des Originalinstrumentes (George, Hall, & Stiegelbauer, 2006) herangezogen. Die Einstellungen zur Inklusion werden über Teilskalen des KIESEL-Fragebogens (Bosse & Spörer, 2014b) abgebildet.

Erste regressionsanalytische Befunde ($R^2 = .37$) weisen darauf hin, dass mitnichten ein ansteigender Zusammenhang von Stufe zu Stufe innerhalb der SoC und den Lehrereinstellungen gegenüber dem inklusiven Unterrichten zu beobachten ist. Vielmehr sticht die Stufe des „Aufgabenmanagements“ durch einen hohen negativen Zusammenhang zu den Einstellungen hervor ($\beta = -.38$), wohingegen von den davor stehenden Stufen einzig die Stufe „Information“ einen geringen signifikanten Zusammenhang zu diesen aufweist ($\beta = -.12$). Für die Stufen „Auswirkung auf Lernende“ und „Kooperationsmöglichkeit“ sind ansteigend leichte positive Effekte vorzufinden ($\beta = .11$ bzw. $\beta = .16$). Weiterführende Multilevel-Analysen sollen den Zusammenhang zwischen Lehrereinstellungen zur Inklusion und den SoC konkretisieren.

Donnerstag
12. März

11:15 Uhr

GC 03/142

Chair(s)

Jäntsich, Christian

Opalinski, Saskia

Berufsbezogene Selbstwirksamkeitserwartungen und inklusionsbezogene Einstellungen von Lehrkräften – binationale Untersuchung in Finnland und Deutschland

Mit Ratifizierung der UN-Behindertenrechtskonvention hat sich Deutschland zur Sicherstellung eines inklusiven Bildungssystems verpflichtet. Im Schuljahr 2011/2012 wurden jedoch nur 28,2% der Schüler_innen mit sonderpädagogischem Förderbedarf inklusiv beschult (KMK 2014). Bei der Umsetzung eines inklusiven Bildungssystems nach Art. 24 der UN-Behindertenrechtskonvention müssen Haltungen, Einstellungen und Kenntnisse der Lehrkräfte als zentrales Moment inklusiver Schulentwicklung betrachtet werden (Baumert & Kunter, 2006; Avramidis & Norwich, 2002). Dazu vorliegende nationale und internationale Forschungsbefunde sind uneindeutig bis gegensätzlich: Allgemeine Einstellungen zur schulischen Inklusion variieren von positiv bis negativ. Sie korrelieren nicht eindeutig mit Geschlecht, Alter oder Qualifikation der Lehrkräfte (de Boer et al., 2011; Hwang & Evans, 2011; Avramidis & Norwich, 2002). Es konnte jedoch gezeigt werden, dass berufsbezogene Selbstwirksamkeitserwartungen der Lehrkräfte mit deren Einstellungen zu schulischer Inklusion und der Unterstützung von Schulreformbewegungen zusammenhängen (Savolainen et al., 2012; Tschannen-Moran et al., 1998; Bender et al., 1995). Für Deutschland und insbesondere die neuen Bundesländer fehlen allerdings aktuelle Befunde zu berufsbezogenen Selbstwirksamkeitserwartungen bzw. inklusionsbezogenen Einstellungen und ihrer Beziehung. Dies ist bedauerlich, da vor diesem Hintergrund auch keine international vergleichenden Analysen durchgeführt werden können.

In dem hier vorgestellten Forschungsprojekt soll daher zum einen der Frage nachgegangen werden, welche berufsbezogenen Selbstwirksamkeitserwartungen und inklusionsbezogenen Einstellungen Thüringer Lehrkräfte verschiedener Schularten aufweisen. Zum anderen werden Vergleiche zu Finnland - einem in Bezug auf Inklusion aber auch Schülerleistungen und Chancengerechtigkeit führenden Land – gezogen. Es wird erwartet, dass sich die Unterschiede im jeweiligen Entwicklungsstand der Bildungssysteme in den inklusionsbezogenen Einstellungen und berufsbezogenen Selbstwirksamkeitserwartungen der Lehrkräfte widerspiegeln. Vor dem Hintergrund internationaler Befunde wird außerdem erwartet, dass sich in den beiden Lehrerstichproben Zusammenhänge zwischen den Selbstwirksamkeitserwartungen und den inklusionsbezogenen Einstellungen nachweisen lassen.

Im Schuljahr 2010/2011 wurden Thüringer Lehrkräfte aller Schularten (N=362) und finnische Grundschullehrkräfte (N=112) befragt. Die berufsbezogene Selbstwirksamkeit wurde mittels der Kurzversion (16 items) der Teacher Efficacy Scale (Gibson & Dembo, 1984), die Einstellung zur schulischen Inklusion mit der Moberg-Scale (22 items) (Moberg & Savolainen, 2003; Moberg, 1997) erhoben.

Die bisher erfolgten Analysen zeigen, dass die Thüringer Lehrkräfte eine eher kritische Einstellung zu schulischer Inklusion aufweisen, während finnische Lehrkräfte eher positiv eingestellt sind. Die Lehrkräfte beider Gruppen schätzen sich vergleichbar als (moderat) selbstwirksam ein. Der Zusammenhang von berufsbezogener Selbstwirksamkeit und Einstellung zur Inklusion fällt für die Thüringer Lehrkräfte allerdings signifikant höher aus als für die Finnischen. In weiteren Auswertungen wird geprüft, wodurch die Unterschiede erklärt werden. Dazu werden Resultate multivariater Analysen (Strukturgleichungsmodelle, Typenanalysen), die zum Zeitpunkt der Beitragseinreichung noch ausstehen, berichtet.

Aus den Ergebnissen können Hinweise auf die Notwendigkeit einer veränderten Aus- und Weiterbildung sowie den Implementationsprozess schulischer Inklusion begleitender und unterstützender Maßnahmen abgeleitet werden.

Donnerstag
12. März

11:15 Uhr

GC 03/142

Chair(s)
Jäntschi, Christian

Görel, Gamze; Hellmich, Frank

Prädiktoren für Einstellungen von Lehrerinnen und Lehrern zur Inklusion in der Grundschule

Von der Europäischen Agentur für Entwicklungen in der sonderpädagogischen Förderung (2012, S. 8f.) wird die Wertschätzung der Vielfalt der Lernenden als ein wichtiges Profil von Lehrkräften für den inklusiven Unterricht erachtet. Hierzu gehören geeignete Einstellungen zur inklusiven Bildung, die als die Bereitschaft verstanden werden, auf Aspekte des persönlichen Umfelds durch Zuspruch oder Ablehnung wertend zu reagieren (vgl. Eagly & Chaiken, 1993). Einstellungen von Lehrkräften zum gemeinsamen Unterricht von Kindern mit und ohne sonderpädagogischem/n Förderbedarf gelten als wichtige Prädiktoren für erfolgreiche Umsetzungen von Lehr-Lernprozessen. Vor diesem Hintergrund stellt sich aktuell im Zusammenhang mit der Realisierung inklusiver Schulen die Frage, durch welche Faktoren Einstellungen zur Inklusion von Lehrkräften beeinflusst werden und auf welche Weise günstige Einstellungen zur Inklusion in Schule und Unterricht im Rahmen der Aus-, Fort- und Weiterbildung von (angehenden) Lehrkräften entwickelt werden können. Bislang vorliegende Studien geben Hinweise darauf, dass günstige Einstellungen von Lehrkräften aus Erfahrungen aus dem gemeinsamen Unterricht von Kindern mit und ohne sonderpädagogischem/n Förderbedarf resultieren. Auch der Teilnahme an Fort- und Weiterbildungen zu den Themen Inklusion, Integration und sonderpädagogische Förderung kommt eine wichtige Bedeutung für die Entwicklung positiver Einstellungen gegenüber der Inklusion in Schule und Unterricht bei Lehrerinnen und Lehrern zu (vgl. Avramidis & Kalyva, 2007; Villa, Thousand, Meyers & Nevin, 1996). Weitere Befunde verweisen auf einen bedeutsamen Zusammenhang zwischen Einstellungen und persönlichen Ressourcen wie Selbstwirksamkeitsüberzeugungen: Lehrkräfte mit hohen Selbstwirksamkeitsüberzeugungen in Bezug auf die Gestaltung inklusiven Unterrichts zeigten in Untersuchungen positivere Einstellungen zur Inklusion als Lehrkräfte, die sich als weniger selbstwirksam wahrnahmen (vgl. z.B. Savolainen, Engelbrecht, Nel & Malinen, 2012).

Ungeklärt ist gegenwärtig die Frage, welche Auswirkungen Einstellungen zur Inklusion auf die Motivation von Lehrkräften und ihre Vorstellungen in Hinblick auf die eigene inklusive Unterrichts- und Schulentwicklung haben.

Im Rahmen der vorliegenden Studie wurde vor diesem Hintergrund anhand einer Stichprobe von N=201 Grundschullehrkräften untersucht, ob und inwiefern Einstellung und Motivation bezüglich der Gestaltung inklusiven Unterrichts in der Grundschule durch Vorerfahrungen und Auffassungen von Inklusion sowie über persönliche Ressourcen wie Selbstwirksamkeitsüberzeugungen erklärt werden können. Die Erhebung der Daten erfolgte dabei anhand eines Fragebogens.

Die Befunde aus einem Strukturgleichungsmodell zeigen, dass Lehrkräfte mit Erfahrungen aus dem gemeinsamen Unterricht, geeigneten Auffassungen von Inklusion sowie hohen Selbstwirksamkeitsüberzeugungen signifikant positivere Einstellungen bezüglich der Gestaltung inklusiven Unterrichts haben als Lehrkräfte, die bislang nicht über diese Voraussetzungen verfügen. Dabei stellt die Variable 'Selbstwirksamkeitsüberzeugung in Bezug auf die Gestaltung inklusiven Unterrichts' den größten Erklärungsfaktor für die Vorhersage von Einstellungen zur Inklusion bei den befragten Lehrkräften dar.

Weiter wird deutlich, dass Effekte der Einstellungen zur Inklusion auf die Motivation, sich mit inklusionspädagogischen Themen auseinanderzusetzen, durch die Selbstwirksamkeitsüberzeugungen in Bezug auf die Durchführung inklusiven Unterrichts moderiert werden. Vorstellungen bezüglich einer gelungenen Gestaltung inklusiven Unterrichts in der Grundschule lassen sich über das Ausmaß an intrinsischer Motivation und den zugrunde liegenden Auffassungen von Inklusion bei den befragten Grundschullehrkräften erklären.

Die vorliegenden Befunde geben zusammenfassend Hinweise auf die Bedeutung von Erfahrungen aus dem gemeinsamen Unterricht, geeigneten Auffassungen von Inklusion sowie Selbstwirksamkeitsüberzeugungen für die Entwicklung positiver Einstellungen zur Inklusion bei Grundschullehrkräften. Sie unterstreichen darüber hinaus aber auch die Bedeutung motivationaler Faktoren für die Entwicklung von Vorstellungen über die Gestaltung des eigenen inklusiven Unterrichts in der Grundschule.

Donnerstag
12. März

11:15 Uhr

GC 03/142

Hövel, Dennis C.; Hennemann, Thomas; Fitting-Dahlmann, Klaus; Casale, Gino; Hillenbrand, Clemens; Wilbert, Jürgen; Grosche, Michael

Wie beurteilen Lehrkräfte die Teilnahme an einer landesweiten Qualifizierungsmaßnahme zum Thema Inklusion? - Erste Ergebnisse der Evaluation der Ausbildung von Moderatorinnen und Moderatoren zur Fortbildungsreihe „Auf dem Weg zur inklusiven Schule“ in NRW

Einleitung: Die inklusive Beschulung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf stellt Lehrerinnen und Lehrer vor neue Herausforderungen und verstärkt die Komplexität eines bereits hoch komplexen Berufs (Forlin, Au & Chong, 2008). Die mit Inklusion einhergehenden Veränderungen im Bildungssystem wecken Besorgnisse bei Lehrkräften, die vor allem den Umgang mit Verhaltensproblemen sowie mangelnde Fertigkeiten im Umgang mit Heterogenität betreffen (Forlin et al., 2008). In diesem Zusammenhang werden ein Gefühl mangelnder Kompetenz, fehlende Fort- und Weiterbildungsangebote sowie ein unzureichendes Fachwissen über Behinderungen als „Killer-Faktoren“ schulischer Inklusion betrachtet (Lindsay, 2007). Um Lehrerinnen und Lehrer auf die Anforderungen eines solchen Systemwandels vorzubereiten, haben sich Lehrerfortbildungen als eine effektive Möglichkeit erwiesen (Nougaret, Scruggs & Mastropieri, 2005).

Die landesweite Qualifizierungsmaßnahme „Auf dem Weg zur inklusiven Schule“ in Nordrhein-Westfalen greift genau diese Aspekte auf. Im Rahmen eines „indirect service delivery model“ (Hagermoser, Sanetti & Kratochwill, 2009) werden Multiplikatorinnen und Multiplikatoren zum Thema Inklusion fortgebildet. Die ausgebildeten Lehrkräfte bieten dann schulinterne Fortbildungen als prozessbezogene Unterstützung der Schulen bei der Umsetzung der Inklusion in der jeweiligen Schulform an. Die Qualifizierung gliedert sich in 13 Module zu den Themen Inklusion als Auftrag, Schulentwicklung, Beratung (I+II), Diagnostik (I+II), Classroom Management, Prävention von Verhaltensstörungen (I+II), Lernstörungen (I+II), Sprachstörungen (I+II) und Schulabsentismus. Es werden jeweils Tandems aus einem Regel- und einem Sonderschullehrer gebildet, die gemeinsam die schulinternen Fortbildungen anbieten.

Fragestellung: Der vorliegende Beitrag geht der Frage nach, wie die teilnehmenden Lehrkräfte die einzelnen Module der Qualifizierungsmaßnahme beurteilen. Es werden Unterschiede zwischen der Bewertung der einzelnen Module, zwischen den Bewertungen von Lehrkräften aus verschiedenen Schulformen (Haupt-, Real-, Gesamt-, Grund-, Förderschule, Gymnasium, Berufskolleg), zwischen den Bewertungen von Lehrerinnen und Lehrern mit und ohne Erfahrung im Gemeinsamen Unterricht (GU) untersucht.

Methode: Von November 2011 bis Oktober 2014 wurden 300 Lehrerinnen und Lehrer in den 13 Modulen an jeweils zwei aufeinander folgenden Fortbildungstagen von Mitarbeitern der Universitäten Köln, Oldenburg und Würzburg qualifiziert. Nach jedem Fortbildungsmodul wurden die Teilnehmerinnen und Teilnehmer zur Beurteilung der Veranstaltung analog zu Schulnoten (1=sehr gut – 6=ungenügend) befragt. Die Frage nach der Gesamtbeurteilung der Veranstaltungen steht im Fokus des vorliegenden Beitrags. Die Datenanalysen zur Prüfung der Unterschiedshypothesen wurden mit einfaktoriellen Varianzanalysen durchgeführt.

Ergebnisse: Insgesamt liegen 2513 Gesamteinschätzungen zu den Modulen vor, die die Maßnahme über alle Module hinweg mit gut ($m=2,12$) bewertet. Fast dreiviertel aller Beurteilungen (71,2%) liegen im guten bis sehr guten Bereich. Die Ergebnisse der ANOVA zeigen signifikante Unterschiede zwischen den Beurteilungen der einzelnen Module ($df=13$, $F=41,807$, $p=.000$), wobei die Module „Beratung II“ ($m=1,76$) und „Prävention von Verhaltensstörungen II“ ($m=1,65$) am besten benotet werden. Die niedrigste Benotung erfolgte für das Modul „Lernförderung II“ ($m=2,70$). Weiterhin konstatieren die Ergebnisse keine signifikanten Unterschiede zwischen den Lehrkräften aus verschiedenen Schulformen sowie keine nennenswerten Unterschiede zwischen den Lehrkräften mit und ohne GU-Erfahrung.

Donnerstag
12. März

11:15 Uhr

GC 03/142

Chair(s)
Jäntsich, Christian

Ludwig, Tobias; Priemer, Burkhard

Real vs. virtuell: Welchen Einfluss hat das Medium auf Argumentationsprozesse beim Experimentieren?

Motivation und Zielsetzung: Simulationen werden als effiziente Methode genutzt, um das Verhalten von Lernenden beim Experimentieren zu untersuchen (z. B. Gößling, 2010; Künsting et al., 2008; de Jong & van Joolingen, 1998). Bisher ist u. E. jedoch nur unzureichend untersucht, ob es möglich ist, Ergebnisse, die in einer virtuellen „Experimentierumgebung beobachtet werden, didaktisch sinnvoll auf den realen naturwissenschaftlichen Unterricht, insbesondere den Physikunterricht, übertragen zu können“ (Künsting et al., 2008, S. 2). Die Zulässigkeit dieser Übertragung ist zu prüfen, da die experimentellen Vorgehensweisen und die epistemologischen Voraussetzungen in beiden Lernsettings sehr unterschiedlich sein können, z. B. hinsichtlich der Datengenerierung, des Zeitaufwands, der Dimensionalität des Aufbaus, der Vereinfachung bzw. Komplexität des Handlings und der Manipulation, sowie hinsichtlich des Umgangs mit Fehlern (z.B. Olympiou & Zacharia, 2012).

Bisherige Ergebnisse bzgl. der Unterschiede beim Arbeiten mit realen und virtuellen Experimenten zeigen divergierende Befunde (Finkelstein et al., 2005; Schreiber et al., 2009; Smetana & Bell, 2012; Triona & Klahr, 2003). Auffällig ist hierbei die Auswahl der abhängigen Variablen. Der überwiegende Teil dieser Arbeiten fokussiert eher auf Lernergebnisse (z. B. Wissenserwerb, Effizienz) als auf zugrunde liegende kognitive Prozesse. Diese Arbeit untersucht daher Argumentationsprozesse beim nicht-hypothesenkonformen Experimentieren in realen und virtuellen Settings. Auf der Basis von Vorarbeiten zur Kategorisierung dieser Argumentationen (Ludwig & Priemer, 2012, 2013) sowie deren quantitativer Erfassung (Ludwig & Priemer, 2014) stellt der Vortrag die Ergebnisse einer Vergleichsstudie vor.

Fragestellung (Auszug): Welchen Einfluss hat der Typ der Experimentierumgebung – real vs. virtuell – auf die Stärke der Verwendung bestimmter Argumente?

Methode: Die aufgeführten Forschungsfragen wurden in einem randomisiert-experimentellen Design mit 934 Probanden (14.3 Jahre) untersucht. Nach der Erhebung der persönlichen Merkmale wurden die Probanden aufgefordert, zunächst eine Hypothese zu einem physikalischen Zusammenhang aufzustellen, um diese im Anschluss im realen bzw. virtuellen Experiment zu testen. Danach wurden die angeführten Argumentationen zum Beibehalten bzw. Verwerfen der eingangs aufgestellten Hypothese fragebogenbasiert erfasst. Dieser Test erfasst die folgenden vier Argumentkategorien: Intuition, Expertenwissen, Messunsicherheiten und Evidenz (Ludwig & Priemer, 2014).

Ergebnisse: Von 934 Probanden haben 815 Probanden vor dem Experiment eine falsche Hypothese aufgestellt. Nur diese Fälle wurden in weiteren Analysen berücksichtigt. Die Unterschiedshypothesen wurden mit Hilfe eines Multigruppen-Strukturgleichungsmodells getestet ($\chi^2(1220) = 1916.1$, CFI = .91, RMSEA = .04). Im Hinblick auf die Forschungsfrage konnte gezeigt werden, dass der Typ der Experimentierumgebung Einfluss auf zwei Argumentkategorien nimmt: Probanden, die mit dem virtuellen Experiment arbeiten, argumentieren weniger mit den Kategorien Intuition ($d = -0.27$, $p < .05$) sowie Messunsicherheiten ($d = -0.52$, $p < .05$). In den Kategorien Expertenwissen und Evidenz zeigen sich keine Unterschiede zwischen den Gruppen.

Diskussion: Bezüglich der Gruppenunterschiede in der Verwendung bestimmter Argumente zeigt sich lediglich ein bedeutsamer Effekt: Die deutlich geringere Ausprägung der Kategorie Messunsicherheiten bei den Probanden, die mit dem virtuellen Experiment gearbeitet haben, ist erwartungsgemäß, da Messunsicherheiten in Simulationen nicht implementiert sind. In den drei anderen Argumentationskategorien zeigten sich hingegen keine Unterschiede. Dies ist aus Sicht der o.g. Studien ein dienliches Ergebnis, da dadurch die Legitimität der Folgerungen gegeben ist. Aus didaktischer Sicht ist dieses Ergebnis allerdings problematisch, da Schüler den im virtuellen Experiment erhobenen Daten und Beobachtungen offenbar die gleiche Aussagekraft zuschreiben wie Realexperimenten, obwohl Simulationen u. E. nur ein unzutreffendes Bild des Experimentierens zeigen: Die Manipulation einer Programmoberfläche zum Entdecken einer bereits integrierten analytischen Modellierung eines physikalischen Problems lässt eine Beobachtung des eigentlichen Naturphänomens nicht zu.

Donnerstag
12. März

11:15 Uhr

GC 03/149

Chair(s)

Priemer, Burkhard

Motyka, Marc; Lipowsky, Frank

Das Lernpotenzial digitaler Spiele in der politischen Bildung am Beispiel von Food Force

Theoretischer Hintergrund: Digitale Lernspiele dürften geeignet sein, um der Methodenarmut des Politikunterrichts (Janssen, 2013) entgegenzuwirken, da sie prinzipiell die Anforderungen an lernförderlichen Unterricht erfüllen: Sie können zur Entstehung intrinsischer Motivation beitragen (Ryan, Rigby & Przybylski, 2006), die kognitive Architektur des Menschen effizient nutzen (Mayer, 2014) sowie durch die Übernahme einer anderen Rolle einen Perspektivwechsel ermöglichen (Peng, Lee & Heeter, 2010). Die Auswahl an geeigneten Computerspielen ist für das Fach Politik besonders groß (Macleod, Hamish, Heywood, Heywood & Littleton, 2004). Bisher existieren jedoch nur wenige empirische Befunde zu den Auswirkungen digitaler Lernspiele in der politischen Bildung (Motyka, 2012; Scholz, 2007). Metaanalytische Untersuchungen kommen zu dem Schluss, dass das digitale, spielbasierte Lernen (Prensky, 2007) ebenso effektiv sein kann wie traditionelle Unterrichtsmethoden (Ke, 2008; Randel, Morris, Wetzel & Whitehill, 1992; Vogel et al., 2006; Wouters, van Nimwegen, van Oostendorp & van der Spek, 2013). Ein großer Teil der bisherigen Befunde bezieht sich allerdings auf das Lernen in mathematisch-naturwissenschaftlichen Domänen. Eine Übertragbarkeit auf sozialwissenschaftliche Fächer ist nicht notwendigerweise gegeben (Westelinck, Valcke, Craene & Kirschner, 2005). Weiterhin gibt es kaum gesicherte Erkenntnisse zum Einfluss individueller Schülermerkmale wie dem privaten Computerspielkonsum (Heeter, Lee, Magerko & Medler, 2011). Speziell für die politische Bildung ist zu berücksichtigen, dass die politische Einstellung von Schülern bei der Verwendung eines Computerspiels stärker beeinflusst werden könnte als beim Lesen eines traditionellen Textes (Bogost, 2007; Motyka & Künsting, 2013).

Hypothesen: (1) Das Lernen mit Computerspiel ist im Fach Politik mindestens ebenso lernwirksam wie das Lernen mit einem vergleichbaren Papiertext. (2) Schüler, die mit einem Computerspiel lernen, zeigen eine höhere Motivation als Schüler, die mit einem traditionellen Papiertext lernen. (3) Die politische Einstellung der Schüler entwickelt sich unterschiedlich in Abhängigkeit des eingesetzten Lernmediums.

Methode: Zur Überprüfung der Hypothesen wurde ein multikriteriales 2x2-faktorielles Vergleichsgruppen-Design mit Pre, Post und Follow-Up-Test eingesetzt. N = 179 Gymnasialschüler der Jahrgangsstufe 9 wurden randomisiert den Bedingungen PC-Spiel (das Computerlernspiel Food Force der Vereinten Nationen zum Thema Nahrungsmittelhilfe; WFP, 2014) und Lehrbuchtext (papierbasierte Transkription des PC-Spiels) zugewiesen. Die Probanden wurden nach der Auseinandersetzung mit dem Lernmaterial zusätzlich zwei Formen der Nachbesprechung auf (Gruppe vs. Einzel) zugewiesen. Als abhängige Variablen wurden unter anderem die Motivation, der Wissenserwerb sowie die politische Einstellung der Schüler untersucht. Individuelle Schülermerkmale wie etwa die private Computerspielnutzung wurden erhoben, um ATI-Effekte untersuchen zu können. Bei der Messung der abhängigen Variablen kamen bestehende Instrumente wie der Fragebogen zur aktuellen Motivation von Rheinberg, Vollmeyer und Burns (2001), ein selbstentwickelter Wissenstest mit zwölf Fragen zu den vermittelten Lerninhalten sowie ein selbstentwickeltes kontextspezifisches, semantisches Differential (Heidenreich, 1999) zur Nahrungsmittelhilfe zum Einsatz. Die verwendeten Skalen zur Motivation und dem Wissenserwerb weisen eine hinreichende Reliabilität auf (Cronbachs $\alpha \geq .70$).

Ergebnisse: Varianzanalytisch konnte dem verwendeten Computerspiel zum dritten Messzeitpunkt eine ebenso hohe Lernförderlichkeit attestiert werden wie dem Lehrbuchtext ($F(1,175) = .33$, n.s.). Die Schüler der PC-Gruppe waren sowohl zum ersten ($F(1,171) = 7.68$, $p < .01$) als auch zum zweiten Messzeitpunkt ($F(1,173) = 6.41$, $p < .05$) etwas motivierter als die der Textgruppe. Die politische Einstellung der Probanden entwickelte sich nicht unterschiedlich in Abhängigkeit der Vergleichsgruppe ($F(1,134.32) = 1.59$, n.s.). Auch wenn die praktische Bedeutsamkeit der Befunde eher gering ausfällt, stellt sich zumindest für das untersuchte Computerspiel Food Force heraus, dass Computerspiele auch in sozialwissenschaftlichen Domänen den Unterricht bereichern und eine ernsthafte Alternative zu bisherigen Unterrichtsmethoden darstellen können. In zukünftigen Forschungsbemühungen sollten die Auswirkungen komplexerer Computerspiele mit politikwissenschaftlich relevanten Inhalten (z. B. SimCity oder Civilization) empirisch untersucht werden.

Donnerstag
12. März

11:15 Uhr

GC 03/149

Chair(s)

Priemer, Burkhard

Gerth, Sabrina; Dolk, Thomas; Fließer, Michael; Nottbusch, Guido; Fischer, Martin H.; Klassert, Annegret; Festman, Julia
Handschrift auf dem Tablet oder auf Papier – macht es einen Unterschied? Eine graphomotorische Studie mit Vorschülern, Zweitklässlern und Erwachsenen

Schreiben – damit meinen wir heute immer häufiger Tippen auf Computer oder Smartphone. Trotzdem ist das Schreiben mit einem Stift in der Hand immer noch eine grundlegende Fähigkeit, die von Grundschulern ab Schuleintritt erworben und angewendet wird. Schüler in der Grundstufe verbringen beispielsweise bis zu 2/3 ihrer Zeit mit dem Schreiben auf dem Papier [1]. Und das zu Recht, denn vor allem die komplexen graphomotorischen Aspekte des Schreibens mit dem Stift im Vergleich zum Tippen fördern maßgeblich das Schreiben-Lernen wie auch das Lernen im Allgemeinen [2-4].

Schreiben ist ein äußerst komplexer Prozess. Auf unterschiedlichen Verarbeitungsebenen interagieren grundlegende Prozesse (graphomotorische Planung und Kontrolle) mit komplexen, konzeptuellen Prozessen (Ideenbildung, Phonemanalyse, Phonem-Graphem-Übersetzung) [5]. Gemeinsam stehen diese Prozesse im ständigen Wettbewerb um begrenzte Arbeitsgedächtniskapazitäten [6]. Da während des Schriftspracherwerbs besonders die graphomotorischen Aspekte des Schreibens die Arbeitsgedächtniskapazität beanspruchen, sind weniger Ressourcen für komplexe Prozesse verfügbar [7]. Kommt es aufgrund von Schwierigkeiten mit grundlegenden Prozessen des Schreibens zu dauerhaften Kapazitätsengpässen bei komplexen Prozessen [8-11], kann sich dies negativ auf den Schulerfolg auswirken [12,13].

Bisherige Studien untersuchten meist das Produkt des Schreibens auf dem Papier in Hinblick auf Qualität und Quantität [13-15]. Der Einsatz von Tablets ermöglicht es zusätzlich den Prozess des Schreibens anhand von Maßen wie Schreib- und Pausenzeit, Schreibgeschwindigkeit und Anzahl von Geschwindigkeitsumkehrungen während des Schreibens (NIVs) genauer zu analysieren [16-19]. Diese Schreibprozessmaße gelten als Indikator für den Grad der Automatisierung des Schreibens [19,20] und bieten Lehrern, Therapeuten und Wissenschaftlern neben dem Schreibprodukt wichtige Informationen, um Schüler beim Schreiben(-lernen) gezielter fördern zu können. Vor dem Hintergrund heterogener Lehr- und Lernvoraussetzungen von Schülern in inklusiven Schulen erscheint dies heute wichtiger denn je [21].

Bisher fehlen Studien, die systematisch prüfen, inwieweit Schreiben mit einem Stift auf dem Tablet und dem Papier denselben Schreibprozess abbilden. In diesem Projekt untersuchen wir, inwieweit sich verschiedene Grade des automatisierten Schreibens zwischen den beiden Medien (Tablet vs. Papier auf einem Tablet) mit Hilfe von Schreibprozessmaßen dokumentieren lassen. Hierzu nahmen drei Gruppen mit unterschiedlichen Niveaus des Schriftspracherwerbs (29 Vorschüler, 30 Zweitklässler und 30 Erwachsene) an zwei Erhebungsterminen teil (Dauer jeweils 25-35min). In jeder Gruppe begann die Hälfte der Probanden mit dem Tablet und führte die gleichen Aufgaben auf dem Papier aus. Auf beiden Medien wurden drei verschiedene Aufgabentypen gestellt: (a) graphomotorische Vorläuferbewegungen, (b) visuomotorische Aufgaben und (c) Schreiben der Phrase „Sonne und Wellen“ (nur Zweitklässler und Erwachsene).

Die Ergebnisse zeigen für alle erfassten Schreibprozessmaße und Aufgabentypen ein aufsteigendes Stufenmuster des Schriftspracherwerbs. Vorschulkinder schrieben im Vergleich zu Zweitklässlern ebenso wie Zweitklässler im Vergleich zu Erwachsenen langsamer (Schreibzeit und –geschwindigkeit), machten mehr Pausen und zeigten einen geringeren Automatisierungsgrad in der motorischen Ausführung (mehr NIVs).

Der Vergleich Tablet-versus-Papier ergab für alle Gruppen und Aufgaben eine schnellere Schreibgeschwindigkeit auf dem Tablet. Die Analyse weiterer Schreibprozessmaße zeigte, dass Vorschüler für alle Aufgaben auf dem Papier mehr Zeit benötigten, mehr Pausen, Schriftzüge und NIVs machten. Ebenso hatten Zweitklässler längere Schreibzeiten und mehr NIVs für die Vorläuferbewegungen auf dem Papier. Interessanterweise zeigten Zweitklässler und Erwachsene jedoch für die Schreibaufgabe längere Schreib- und Pausenzeiten sowie mehr NIVs auf dem Tablet.

Zusammenfassend lässt sich festhalten, dass der Schreibprozess anhand der erhobenen Schreibprozessmaße in allen drei Altersgruppen abgebildet werden konnte. Diese Maße liefern detailliertere Informationen für Erzieher, Lehrer, Therapeuten und Wissenschaftler als dies das Schreibprodukt auf dem Papier allein zu bieten vermag. Während graphomotorische Vorläuferbewegungen und visuomotorische Aufgaben schneller und leichter auf dem Tablet auszuführen sind, scheint ein flüssiges, automatisiertes Schreiben von Wörtern auf dem Tablet schwerer zu sein. Neben dem Einfluss heterogener Schreibprozesse auf das Schreibprodukt werden Vor- und Nachteile der Nutzung von (a) Prozess- im Vergleich zu Produktmaßen und (b) Tablets im Vergleich zu Papier für den Schriftspracherwerb diskutiert.

Donnerstag
12. März

11:15 Uhr

GC 03/149

Chair(s)
Priemer, Burkhard

Antosch-Bardohn, Jana; Stegmann, Karsten

Insights beim Nicht-Lernen - Eine explorative ESM Studie zu Inkubations- und Insight-Effekten im Lernprozess

Experimentelle Studien im Bereich der Lernforschung sind gewöhnlich auf aktive Lernprozesse fokussiert. Im Regelfall kann davon ausgegangen werden, dass Inhalte ohne ein weiteres bewusstes Auseinandersetzen im Laufe der Zeit vergessen werden (Anderson, 2000). Daher werden Phasen, in denen sich nicht mit dem Lernstoff befasst wird, meistens vernachlässigt. Es gibt jedoch Indizien, dass unbewusstes Denken zu einem Wissenszuwachs führen kann (Lewicki et al., 1992). Eine empirisch untersuchte Phase unbewussten Denkens ist die Inkubationsphase im kreativen Problemlösungsprozess, in der man sich bewusst mit anderen Inhalten beschäftigt (Dodds et al., 1999). Inkubationsstudien, die ausschließlich im Kontext von Problemlöseprozessen durchgeführt werden, liefern uneinheitliche Ergebnisse darauf, welcher kognitive Zustand Inkubationseffekte fördert. So weisen einige Studien nach, dass Aufgaben mit geringer kognitiver Anstrengung Inkubationseffekte fördern (Browne & Cruse, 1988; Sio & Ormerod, 2009), andere Studienergebnisse belegen das Gegenteil (Segal, 2004) oder können keinerlei Unterschiede feststellen (Olton & Johnson, 1976; Smith & Blankenship, 1991). Indikatoren für eine Inkubation sind nicht-intentionale Erinnerungen, die spontan in das Bewusstsein treten (Mandler, 1994) und das Auftreten von Insights, also das plötzliche Gewinnen einer Einsicht (Finke et al., 1992). Treten Inkubationseffekte und Insights auch in Lernumgebungen auf, in denen keine kreative Problemlösung gefordert ist?

Um Inkubations- und Insighteffekte im Rahmen von Vorlesungen zu untersuchen, wurde eine explorative Studie mit folgenden Forschungsfragen durchgeführt: (1) Wie oft erinnern sich Studierende nicht-intentional an den Lernstoff einer Vorlesung? Wie oft erlangen Studierende Insights bezüglich des Lernstoffs? (2) Zu welchen Zeitpunkten erinnern Studierende den Lernstoff nach einer Vorlesung nicht-intentional und zu welchen Zeitpunkten treten Insights auf? (3) Bei der Ausübung welcher Tätigkeiten treten nicht-intentionale Erinnerungen und Insights auf?

Um die Forschungsfragen zu beantworten wurde die Experience-Sampling Methode ESM gewählt (Csikszentmihalyi & Reed, 1987). 89 (N=89) Studierende der LMU München erhielten im Anschluss an eine Vorlesung eine Woche lang zu randomisiert generierten Zeitpunkten 15 mal Fragebögen. Die Teilnehmer beantworteten, ob und wann sie in den letzten drei Stunden den Lerninhalt erinnerten, ob sie Insights erfahren hatten und welche Aktivität sie in diesen Momenten mit welcher kognitiven Anstrengung (hoch, mittel oder niedrig) getätigt hatten.

(1) Zu der Frage, wie häufig die Studierenden den Lernstoff erinnerten, gaben von insgesamt 912 Fragebögen 132 nicht-intentionale Erinnerungen an den Lerninhalt (14,5%) an. Insgesamt wurden 18 Insight-Erfahrungen berichtet, was extrapoliert auf die Gesamtstichprobe bedeutet, dass 49% aller Studierenden im Laufe der Untersuchungswoche zumindest einen Insight zu dem Vorlesungsthema gewonnen haben.

(2) Betrachtet man die Zeitpunkte der Erinnerungen, so nahmen die Erinnerungen im Verlauf der untersuchten Woche exponentiell ab, die Anzahl an Insights blieb täglich ungefähr gleich. Die Analyse der Tageszeiten zeigt ein leichtes Abnehmen von Erinnerungen am Abend.

(3) Der Vergleich von Aktivitäten in Erinnerungs- und Insight-Momenten mit sämtlichen Aktivitäten, die im Laufe der Woche unternommen wurden zeigt, dass Alltagsroutinen bei Erinnerungs- und Insight-Momenten signifikant seltener auftraten: $\chi^2(2) = 39,500$; $p < .000$. Zudem gaben 70,6% der Personen, die einen Insight berichteten, zu keinem Zeitpunkt an gelernt zu haben. Das Auftreten von Insights scheint also nicht in Beziehung zu aktiven Lernphasen zu stehen. In Erinnerungs- und Insightmomenten waren die Untersuchungsteilnehmer meistens im Zustand mittlerer kognitiver Anstrengung (Erinnerung: 48,35%, Insight: 61,11%). Allerdings traten Insights signifikant häufiger mit einem hohen kognitiven Level auf: $\chi^2(2) = 9,953$; $p < .007$.

Die Ergebnisse zeigen, dass Inkubationseffekte auch in Lernumgebungen auftreten, in denen Lernende weniger aktiv und problemlösend agieren, wobei kognitiv anspruchsvollere Aktivitäten Inkubationseffekte in Nicht-Lernphasen zu fördern scheinen. Phasen zwischen aktiven Lernphasen sollten daher weiter untersucht werden um zu ergründen, welche Rolle das unbewusste Denken in der Wissensverarbeitung spielt. Im nächsten Schritt könnte untersucht werden, inwiefern eine Inkubationsphase den Wissenszuwachs fördert um daraufhin Rückschlüsse auf das Planen von Lernphasen und Unterrichtseinheiten zu ziehen.

Donnerstag
12. März

11:15 Uhr

GC 03/149

Chair(s)

Priemer, Burkhard

Beege, Barbara; Antosch-Bardohn, Jana

Die richtige Methode für den passenden Zweck – Die intuitive und zielorientierte Systematisierung von Lehr-/Lernmethoden des Münchner Methodenkastens

In der Hochschullehre versuchen Lehrende immer häufiger den traditionellen frontalen Lehrstil aufzubrechen und nach konstruktivistischen bzw. studierendenorientierten Ansätzen zu lehren (Krapp & Weidenmann, 2006). Damit stehen nicht mehr die Auswahl und Quantität der Inhalte im Fokus, sondern welches Wissen, welche Fähigkeiten und Einstellungen die Lernenden am Ende der Lehrveranstaltung erworben haben (Metz-Göckel et al., 2012; Stelzer-Rothe & Brinker, 2008). Der Einsatz von interaktiven Lehr-/Lernmethoden unterstützt den studierendenorientierten Lernprozess. Allerdings ist es für die Lehrenden oft schwierig aus der Vielzahl an Methoden eine inhaltlich passende Methode auszuwählen.

Um einen schnellen Zugriff auf passende Lehr-/Lernmethoden zu gewährleisten, ohne sich intensiv mit didaktischen Fragen auseinandersetzen zu müssen, wurde die Methodensystematisierung „Münchner Methodenkasten“ entwickelt. Die zwei zentralen Dimensionen des Methodenkastens sind „Lernziele“, „Sozialformen“ und „Gruppengröße“. Die Dimension „Lernziele“ besteht aus den Kategorien „Ausrichten“, „Vorwissen aktivieren“, „Informieren“, „Verarbeiten“ und „Auswerten“, angelehnt an das kompetenz- und lernendenorientierte Phasenschema AVIVA (Städeli et al., 2010). Eine zweifarbige Kodierung zeigt an, welche Methode für welche Gruppengröße geeignet ist. Mit diesem einfachen nachvollziehbaren Aufbau ist der Münchner Methodenkasten komplexitätsreduzierend, intuitiv zu nutzen und ohne großes Vorwissen anzuwenden, neue Methoden sind in trennscharfen Kategorien einzuordnen (Meyer et al., 2014). Damit hebt sich der Münchner Methodenkasten von anderen Systematisierungen ab (vgl. z.B. Adl-Amini, 1993; Baumgartner, 2011; Berendt et al., 2006; Siebert, 2008; Winteler, 2011).

Um die Anwendbarkeit des Münchner Methodenkastens zu prüfen, wurden in einer explorativen Studie folgende Forschungsfragen gestellt: (1) Nach welchen Kriterien wählen Lehrende interaktive Lehr-/Lernmethoden aus? (2) Finden Lehrende den Aufbau des Münchner Methodenkastens intuitiv verständlich? (3) Trägt die Methodensystematisierung dazu bei, dass Lehrende in der Praxis häufiger Lehr-/Lernmethoden anwenden? Durch einen online bereitgestellten Fragebogen wurden 75 Lehrende zum Münchner Methodenkasten befragt. Alle Befragten besuchten im Vorfeld ein hochschuldidaktisches Seminar, indem der Methodenkasten vorgestellt wurde. Die Rücklaufquote beträgt 42,7% (n=32; weiblich=13, männlich=19). Die Stichprobe repräsentiert 62,5% Sozial- und Geisteswissenschaftler und 31,3% Naturwissenschaftler. 46,9% der Befragten gaben an wenig Wissen über Lehr-/Lernmethoden zu besitzen, 53,1 % berichteten von viel Wissen über den Einsatz von Methoden.

(1) Die Analyse, nach welchen Kriterien Lehrende Methoden aussuchen (Mehrfachantwort möglich) ergibt, dass in 98,8% der Fälle Lehrende Methoden nach Zielen, 98,8% nach Gruppengröße, 87,5% nach Sozialformen, 87% nach der zur Verfügung stehenden Zeit, 84,0% nach dem AVIVA Schema und 72,0% nach Inhalten auswählen. (2) Betrachtet man den Aufbau, so finden 87,6% den Münchner Methodenkasten leicht erlernbar und 75% bestätigen seine intuitive Anwendbarkeit, wobei es keinen signifikanten Unterschied zwischen Lehrenden mit viel Wissen zu Lehrenden mit wenig Vorwissen zum Thema Methoden gibt. 91,6% aller Befragten benutzen den Münchner Methodenkasten in ihrer Lehrpraxis. (3) Die Untersuchung der Fragestellung, ob die Kenntnis über den Münchner Methodenkasten den Einsatz von Methoden in der Praxis fördert zeigt, dass 68,7% der Befragten häufiger Methoden in ihren Lehrveranstaltungen einsetzen, wobei es keinen signifikanten Unterschied zwischen Lehrenden der Sozial- und Geistes- gegenüber den Naturwissenschaften gibt.

Die Ergebnisse zeigen, dass zentrale Auswahlkriterien von Lehr-/Lernmethoden für Lehrende die Lernziele, die Gruppengröße und die Sozialformen sind und nicht wie anzunehmen wäre nur die Lerninhalte. Alle drei Kategorien werden in der Münchner Methodensystematisierung repräsentiert, womit sie fächerübergreifend einsetzbar ist. Die Studie kann bestätigen, dass der Münchner Methodenkasten leicht erlernbar, intuitiv verständlich ist und in unterschiedlichsten Fachgebieten einsetzbar ist, unabhängig davon ob Lehrende kein bzw. viel Methodenwissen haben. Außerdem führt die Kenntnis über den Methodenkasten zu einem vermehrten Einsatz interaktiver Methoden. Damit trägt der Münchner Methodenkasten dazu bei, dass der „Shift from Teaching to Learning“ in der Praxis gelebt werden kann. In einem nächsten Schritt soll anhand einer größeren Stichprobe untersucht werden, inwieweit sich die Einstellung der Lehrenden in Bezug auf interaktive Lehr-/Lernmethoden durch die Nutzung des Münchner Methodenkastens verändert.

Donnerstag
12. März

11:15 Uhr

GC 03/149

Chair(s)

Priemer, Burkhard

3

Becker, Rolf

Heterogenität als Problem oder als Potenzial?

Dem Bildungssystem werden Aufgaben wie Qualifizierung – Förderung und Forderung von Talenten – und Selektion nach Leistung – d.h. Gewährung angemessener Lerngelegenheiten – und Allokation – Zuordnung von Individuen auf Positionen nach erworbener Bildung – zugeschrieben. Neben der formalen Chancengleichheit sind hierbei auch die Grundsätze von „Chancengerechtigkeit“ zu erfüllen. Die Fakten sind andere: Bildungs- und Kompetenzarmut auf der einen Seite und soziale Exklusivität der Chancen von Bildungschancen und für den Erwerb von Bildungsabschlüssen. Offensichtlich gelingt es dem deutschen Bildungssystem nicht, vorhandene Bildungspotenziale in den nachwachsenden Jahrgängen effizient zu nutzen, Talente in ausreichender Menge zu fördern und für ausreichende Bildungsergebnisse zu sorgen. Die Vielfalt und Heterogenität von Begabungen der Kinder und Jugendlichen bei der Einschulung wird offensichtlich zu wenig genutzt und scheinen – in der Regel gekoppelt an deren soziale Herkunft, Geschlecht und Migrationshintergrund – im Schul- und Ausbildungsbereich verschwendet zu werden. Bei der Selektion und Allokation werden Talente häufig falsch diagnostiziert und aussichtsreiche Talente von nützlichen Lerngelegenheiten und weiterführenden Bildungswegen abgelenkt. Nicht selten ist hierbei ambivalent, ob hierbei leistungsfremde Kriterien die Förderung und Forderung von Bildungspotenzialen in der Schulbildung, Berufsausbildung, Hochschulbildung und Weiterbildung dominieren.

Es stellt sich die Frage, was sich eine Gesellschaft an Talenten und Begabungen verschenkt durch ein ineffizientes und ineffektives Bildungssystem? Wie viele Qualifikationen gehen dem Arbeitsmarkt sowie für die gesellschaftliche Entwicklung verloren? Warum lassen sich – trotz Bildungsexpansion und Bemühungen um Bildungsreformen – Bildungserfolge, Bildungserwerb und die Struktur der Bildungsergebnisse anhand sozialer Herkunft, Geschlecht und Migrationshintergrund vorhersagen? Wie müsste ein Bildungssystem aussehen, welche die Vielfalt erkennt und nutzt und die Heterogenität von Begabungen eher als Potenzial ansieht statt als Problem behandelt? Was müsste ein Bildungssystem in Zukunft für die Sozial- und Systemintegration leisten? Inwieweit müssen Stratifizierung und Segmentierung des Bildungssystems reduziert werden, um Bildungspotenziale optimal ausschöpfen zu können oder um Voraussetzungen zu schaffen, dass Bildungschancen wahrgenommen werden können ohne dass neue Prozesse sozialer Schließung in Gang gesetzt werden? Welche gesellschaftlichen Voraussetzungen außerhalb des Bildungssystems müssen geschaffen werden, damit im und durch das Bildungssystem soziale Heterogenität als Potenzial erkannt und genutzt wird, damit Individuen Lebenschancen nicht vorenthalten werden und damit die Allgemeinheit daraus Nutzen zieht? Welche Beiträge kann eine sozialwissenschaftliche Bildungsforschung in der nächsten Zukunft für diese Problemstellung leisten?

Donnerstag
12. März

14:00 Uhr

HGB 10

Chair(s)

Drewek, Peter

Betz, Anica; Rothstein, Björn

Authentische linguistische Wissenschaftsvermittlung im Schülerlabor zur Förderung des Interesses am Grammatikunterricht

Schülerlabore wurden als interesselördernde, außerschulische Lernorte primär im MINT-Bereich geschaffen, um SchülerInnen Lernen in authentischen Settings zu ermöglichen (vgl. Engeln 2004). Das an die Ruhr-Universität Bochum angeschlossene Alfred-Krupp Schülerlabor (AKS) zeichnet sich insbesondere durch sein interdisziplinäres Angebot aus, welches nicht nur die Naturwissenschaften, sondern auch ein geisteswissenschaftliche Angebot umfasst. Die fächerübergreifende Konzeption des AKS spiegelt sich auch in den heterogenen Zugriffsweisen auf das Lernen im Schülerlabor des vorzustellenden Promotionskollegs wider. Dabei liegt der Forschungsschwerpunkt auf der Authentizität in der Wissenschaftsvermittlung im Schülerlabor.

Ausgangspunkt für das hier vorgestellte und im Kolleg verortete sprachdidaktische Forschungsvorhaben ist das Desinteresse, auf das die Beschäftigung mit der Grammatik des Deutschen bei vielen SchülerInnen stößt (z.B. Bremerich-Vos 1981). Um diesem Mangel an Interesse entgegenzuwirken entwickelte Menzel (u.a. 2008) die Grammatikwerkstatt als wissenschaftspropädeutische und schüleraktivierende Methode für den muttersprachlichen Grammatikunterricht. Deren Wirkung auf die SchülerInnen ist allerdings dadurch beschränkt, dass es sich um einen rein methodischen Ansatz handelt, der andere Aspekte der Wissenschaftsvermittlung wie z.B. den Ort und die Person unberücksichtigt lässt. Dagegen bieten Schülerlabore Wissenschaftsvermittlung in einer ganzheitlichen und vor allem authentischen Lernumgebung.

Das Konzept der Authentizität ist in vielen Fachdidaktiken (z.B. Englisch- oder Physikdidaktik, z.B. Amor 2002, Kuhn 2010) untersucht worden. In der Fremdsprachendidaktik beispielsweise werden Aspekte wie ‚Authentizität des Textes‘, ‚Authentizität als Bezug auf die Lebenswirklichkeit‘ und ‚Authentizität der Person‘ (z.B. Siebolt 1997) diskutiert. Im MINT-Bereich wurde außerdem in Fragebogenstudien nachgewiesen, dass die Authentizität von Schülerlaboren einen wesentlichen Einfluss auf den von den SchülerInnen wahrgenommenen Interessantheitsgrad der Lernumgebung hat (z.B. Engeln 2004, Pawek 2009). In der germanistischen Sprachdidaktik allerdings fehlt die Forschung zur Rolle von Authentizität im Unterricht sowie in der Wissenschaftsvermittlung fast vollständig.

Diesem Desiderat folgend wird im vorliegenden Forschungsvorhaben der Frage nachgegangen, inwiefern sich eine authentische Lernumgebung auf das Interesse der SchülerInnen an der expliziten Beschäftigung mit der Grammatik auswirkt. Ziel ist dabei, am Beispiel des Ruhrdeutschen als Gegenstand linguistischer Forschung die unterrichtliche Rolle von Authentizität aus sprachdidaktischer Perspektive näher zu untersuchen und daraus mögliche Implikationen für einen Grammatikunterricht abzuleiten, der das Interesse an linguistischen Themen fördert.

Zur Operationalisierung des Konzeptes der Authentizität werden verschiedene Teilbereiche, die zu einem authentischen Unterricht beitragen können, zusammengetragen und einer authentischen linguistischen Wissenschaftsvermittlung zum Ruhrdeutschen entsprechend modelliert: Ort (Schülerlabor), Akteur (Forscher), Material (Korpus der gesprochenen Sprache des Ruhrgebiets), Innovation (Erkenntnisgewinn und Beantwortung einer Forschungsfrage) und Methode (linguistische Verfahrenstechniken, mit denen sich die SchülerInnen das Material erschließen). Der Schwerpunkt soll auf die Überprüfung des Einflusses des Lernortes und des Akteurs gelegt werden.

Dazu ist eine Interventionsstudie mit sechs Gymnasialklassen der Jahrgangsstufe acht in einem Kontrollgruppendesign vorgesehen, die von einer quantitativen Fragebogenerhebung im pre-, post-, follow-up-Design begleitet wird. Jede Klasse wird dabei in vier unterschiedliche Gruppen eingeteilt und nimmt jeweils an einem 90-minütigen Projekt zum Ruhrdeutschen als einem Forschungsschwerpunkt der Bochumer Linguistik teil. Im Rahmen dieser vier Grammatikprojekte werden unterschiedliche Grade von Authentizität in ihrer Wirkung auf das Interesse der SchülerInnen untersucht. Um zu überprüfen, inwieweit sich der jeweilige Workshop auf das Schülerinteresse an der expliziten Beschäftigung mit der Grammatik des Deutschen auswirkt, werden die Faktoren Ort und Akteur unter Einbezug von drei Kontrollgruppen variiert. Die Inhalte, Methoden und das Material, an dem die SchülerInnen arbeiten, sind in allen Gruppen konstant.

Zur Messung des Interesses der SchülerInnen wurde auf der Basis der pädagogisch-psychologischen Person-Gegenstands-Theorie des Interesses (z.B. Krapp 1998, Krapp 1999) und u.a. in Anlehnung an den quantitativ ausgerichteten Fragebogen zur Erhebung des Interesses an den Naturwissenschaften von Pawek (2009) ein eigener Fragebogen erstellt. Auf der Grundlage der so dokumentierten Entwicklung des Interesses sollen im Vergleich der vier untersuchten Projektgruppen Rückschlüsse auf die Rolle der Authentizität und die Möglichkeit der Interessensteigerung durch linguistische Schülerlabore gezogen werden.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

Mierwald, Marcel; Brauch, Nicola

Authentizität als Chance der Vermittlung von Geschichtswissenschaft im Schülerlabor

Schülerlabore wurden als interessefördernde, außerschulische Lernorte primär im MINT-Bereich geschaffen, um SchülerInnen das Lernen in authentischen Settings zu ermöglichen (vgl. Engeln 2004). Das Alfred-Krupp Schülerlabor (AKS) der Ruhr-Universität Bochum zeichnet sich durch sein interdisziplinäres Angebot aus, welches nicht nur die Naturwissenschaften, sondern auch die Geisteswissenschaften umfasst. Die fächerübergreifende Konzeption des AKS spiegelt sich ebenfalls in den heterogenen Zugriffsweisen auf das Lernen im Schülerlabor des vorzustellenden Promotionskollegs wider. In diesem liegt der Forschungsschwerpunkt auf der Authentizität in der Wissenschaftsvermittlung im Schülerlabor.

Eine zentrale Aufgabe von Schülerlaboren wird darin gesehen, den Lernenden einen authentischen Einblick in die jeweilige Wissenschaftsdisziplin zu vermitteln, indem diese in eigens geschaffenen Lern-Lern-Situationen an Forschungsfragen sowie Denk- und Arbeitsweisen einer Wissenschaft herangeführt werden (vgl. Glowinski 2007; Pawak 2009; Engeln 2004). Das dem Kolleg angeschlossene Dissertationsprojekt untersucht erstmals die Auswirkungen von unterschiedlichen Arrangements in der Vermittlung von Geschichtswissenschaft innerhalb einer authentischen Lernumgebung im Schülerlabor. Für die Geschichtswissenschaft ist hierbei domänenspezifisch, dass die SchülerInnen authentisches Quellenmaterial nach Forschungsfragen des Faches sichten, interpretieren und unterschiedliche Deutungen der Vergangenheit in Form von argumentativen Erzählungen entwickeln (z.B. van Drie & van Boxel 2008). In diesen Erzählungen wird die sogenannte „Argumentative historische Narrativität“ sichtbar, die als Ausdruck fachspezifischer Argumentationsfähigkeit verstanden werden kann. Für die Wissenschaftsvermittlung im Schülerlabor stellt besonders die Vermittlung von Kenntnissen und Fähigkeiten zum Argumentieren einen geeigneten Anknüpfungspunkt dar, weil die SchülerInnen dadurch lernen, wie in einer Wissenschaft Erkenntnisse generiert werden (vgl. u.a. Fleischer 2013). Die für die Domäne Geschichte zentrale Fähigkeit des historischen Argumentierens ist dabei stark von den epistemologischen Überzeugungen der SchülerInnen zur Natur der Geschichtswissenschaft abhängig (vgl. Havekes et al. 2012; Kuhn et al. 1994).

Im Promotionsprojekt werden Möglichkeiten zur Förderung von epistemologischen Überzeugungen zu Geschichte und zur Anbahnung von Argumentationsfähigkeit untersucht. Im Fokus steht die empirische Überprüfung folgender Fragestellung: Wie wirken sich unterschiedliche mediale Repräsentationen zeitgenössischer Quellen zur Verfolgung und Ermordung der Juden als Teil einer Intervention im Schülerlabor auf die Ausprägung der Kompetenz „Argumentative historische Narrativität“ aus?

Für die Durchführung der Untersuchung wird ein standardisiertes Lehr-/Lernangebots entwickelt, in dem wissenschaftliche Authentizität in der historischen Fragestellung, dem Quellenmaterial und der argumentativen Erkenntnisweise als Bedingungen zur Förderung der „Argumentativen historische Narrativität“ modelliert werden. Die Art der Quellenrepräsentation wird als unabhängige Variable variiert, so dass die SchülerInnen je Intervention mit Quellen in Textform, Audioform oder einer gemischten Zusammenstellung arbeiten. Als abhängige Variable fungieren die „Argumentative historische Narrativität“ und die epistemologischen Überzeugungen der Lernenden.

Die Datenerhebung erfolgt zum einen über ein Pre-Posttest-Design. Mit diesem werden die epistemologischen Beliefs zur Natur der Geschichtswissenschaft, das faktuale Wissen zum historischen Inhalt, das prozedurale Wissen über das fachliche Handeln sowie das Interessen und die Motivation zur Beschäftigung mit Geschichte zu Beginn und am Ende der Intervention gemessen. Zum anderen liegen am Ende jedes Angebots im Schülerlabor argumentative Essays der Lernenden als Antwortnarrativ zur inhaltlichen Fragestellung vor. Das Sample besteht aus 25 SchülerInnen der Sekundarstufe II Gymnasium NRW pro Angebot. Dies ergibt eine Stichprobe von N= 50 je Format der Quellenrepräsentation und eine Gesamtstichprobe von N = 150.

Für die Erhebung der epistemologischen Überzeugungen wird im Pre-/Posttest das Instrument von Urhahne et. al. (2008) adaptiert. Die Überprüfung des faktualen und prozeduralen Wissens basiert auf Items aus einem Kompetenztest für das Fach Geschichte (HiTCH-Projekt). Interesse und Motivation werden mittels adaptierter Messinstrumenten aus bereits existierenden Studien zum Lernen in Schülerlaboren erhoben (u.a. Engeln 2004; Brandt 2005; Pawak 2009). Die inhaltsanalytische Auswertung der argumentativen Texte erfolgt über ein Kodiersystem, das deren formale und inhaltliche Gestalt erfasst (vgl. Toulmin 1975; Riemeier et al. 2012; Hodel et al. 2013) und es erlaubt, diese der Qualität nach unterschiedlichen Niveaustufen zuzuordnen.

Das Poster veranschaulicht die Analyse der argumentativen SchülerInnen-Essays beispielhaft.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

Vanderbeke, Marie; Wilden, Eva

Affordances für authentisches fremdsprachliches Handeln in der Wissenschaftsvermittlung bilingualer Schülerlabore

Schülerlabore wurden als interessefördernde, außerschulische Lernorte primär im MINT-Bereich geschaffen, um SchülerInnen das Lernen in authentischen Settings zu ermöglichen (vgl. Engeln 2004). Das an die Ruhr-Universität Bochum angeschlossene Alfred-Krupp-Schülerlabor (AKS) zeichnet sich durch sein interdisziplinäres Programm aus, welches nicht nur die Naturwissenschaften, sondern auch ein geisteswissenschaftliches Angebot umfasst. Die fächerübergreifende Konzeption des AKS spiegelt sich auch in den heterogenen Zugriffsweisen auf das Lernen im Schülerlabor des vorzustellenden Promotionskollegs wider. Dabei liegt der Forschungsschwerpunkt auf der Authentizität in der Wissenschaftsvermittlung im Schülerlabor.

Das im Kolleg verortete Promotionsprojekt der englischen Fremdsprachendidaktik untersucht, ob bilinguale Schülerlabore (BSLab) SchülerInnen affordances zum fremdsprachlichen Handeln bieten und wie diese von den SchülerInnen wahrgenommen werden. Während die empirische Erforschung des bilingualen Sachfachunterrichts seit der Jahrtausendwende intensiv vorangetrieben wurde (vgl. z.B. Heine 2013), blieb die Untersuchung der Sprachhandlungen in BSLabs bislang aus.

Das in dieser Studie untersuchte Konzept der affordances wurde von Gibson 1979 wie folgt definiert: "The affordance of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (Gibson 1979: 127). Affordances sind demnach als Handlungsoptionen (Greeno 1994) zu verstehen, die der Umwelt zwar inhärent sind, aber nur bedeutsam werden, wenn das Subjekt mit ihnen in Interaktion tritt. Diesem Konzept folgend sollte auch das BSLab affordances aufweisen, die das fremdsprachliche Handeln von SchülerInnen bedingen. Bereits durchgeführte Studien in monolingualen Schülerlaboren charakterisierten die authentische Lernumgebung Universität, das Lehrpersonal sowie die genuin naturwissenschaftlichen Arbeitsweisen als wichtige Variablen des Schülerlabors. In Fragebogenerhebungen wurde die Authentizität des Schülerlabors als prominenter Faktor für den Grad an Interessantheit der Lernumgebung in der Schülerwahrnehmung herausgearbeitet (vgl. z.B. Glowinski 2007; Engeln 2004; Pawek 2009). Eine solche Erhebung steht für BSLabs noch aus. Hierbei soll neben dem bereits genannten sachfach- auch auf den fremdsprachendidaktischen Diskurs zu Authentizität Bezug genommen werden, in dem seit der Entwicklung hin zum Communicative Language Teaching in den 1980er Jahren die Debatte um kontextualisierte, genuine Materialien und authentische Sprechansätze zu einem zentralen Anliegen avanciert ist (vgl. Weskamp 1999).

Vor diesem Hintergrund wird in einem qualitativen Design ein molekularbiologisches BSLab auf folgende Forschungsfragen hin untersucht: (1) Bietet das BSLab SchülerInnen affordances zum fremdsprachlichen Handeln auf Englisch? (2) Welche fremdsprachlichen affordances werden im BSLab geboten und wie werden diese von SchülerInnen wahrgenommen? (3) Nehmen die SchülerInnen ihr sprachliches Handeln im BSLab als authentisch wahr im Hinblick auf die Lernerauthentizität und die authentische Lernumgebung?

Die Studie untersucht dabei die Verwendung von Cognitive Discourse Functions, Bedeutungsaushandlungen und den Erwerb von Cognitive Academic Language Proficiencies.

Die Datenerhebung wird mit einer fragebogengestützten Erhebung der Hintergrunddaten der SchülerInnen eingeleitet. Es ist davon auszugehen, dass Schülerwahrnehmungen über Sprachhandlungen und Authentizitätsempfindungen stark von ihrem Sprachstand und ihren wahrgenommenen sprachlichen und sachfachlichen Kompetenz beeinflusst wird (Glowinski 2007: 135ff.). Dementsprechend werden u.a. die letzten Zeugnisnoten und das Selbstkonzept der SchülerInnen in den Fächern Englisch und Biologie erfragt. Während ihrer Laborarbeit werden SchülerInnen mittels teilnehmender Beobachtung untersucht und videografiert. Im Anschluss an den Besuch im Schülerlabor diskutiert eine Gruppe von ca. sechs SchülerInnen in Fokusgruppen ihre Eindrücke. Gruppenteilnehmer werden auf der Basis der erhobenen Hintergrunddaten ausgewählt und die Diskussionen mit Audiogeräten aufgezeichnet. Insgesamt werden an der Hauptuntersuchung vier Klassen der gymnasialen Oberstufe teilnehmen. Somit werden ca. 80 SchülerInnen videografiert; 24 nehmen zusätzlich an den Fokusgruppen teil. Die Auswertung der Videografie findet mittels Interaktionsanalyse (vgl. Jordan und Henderson 1995), die Analyse der Fokusgruppen mit der dokumentarischen Methode nach Bohnsack statt (vgl. z.B. Bohnsack 2013).

Neben dem theoretischen Hintergrund, soll auf dem Poster das Forschungsdesign skizziert und die Auswertung der Datenerhebungen dargestellt werden. Da zum Zeitpunkt der Posterpräsentation die Datenerhebung abgeschlossen sein wird, können vorläufige Ergebnisse aus den ersten Auswertungsschritten präsentiert werden.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

Eckes, Alexander; Wilde, Matthias

Novelty Space in einer Mitmachausstellung am außerschulischen Lernort Universität

Theoretischer Hintergrund: Außerschulische Lernorte aktivieren Besucher (Gudjons, 2001), führen zu einer erhöhten Bereitschaft im Umgang mit Arbeitsmaterialien (Winkler & Scherer, 2000) und fördern erkundendes Verhalten (Falk & Dierking, 2000). Besuche an außerschulischen Lernorten werden u.a. durch die Neuheit der Umgebung (Falk & Adelman, 2003) beeinflusst. Dieser „Novelty Space“, wird nicht nur durch den Bekanntheitsgrad des Standortes (Martin, Falk & Balling, 1981, Falk, Martin & Balling, 1978), sondern auch durch das Vorwissen relevanter Konzepte und Vorerfahrungen mit außerschulischen Lernorten im Allgemeinen (Orion, Hofstein & Mazor, 1986) bestimmt. Nach Falk et al. (1978) kann die Neuheit eines außerschulischen Lernortes zu Desorientierung führen. Die Bereitstellung strukturierender Maßnahmen erhöht Klarheit und Qualität an Informationen, die benötigt werden, um vorgenommene Aufgaben erfolgreich erfüllen zu können (Connell & Wellborn, 1991; Skinner & Belmont, 1993). Zusätzliche Struktur könnte die Qualität des Besuches durch Reduktion der Neuheit erhöhen (Falk & Dierking, 2000) und den Wissenserwerb fördern (Falk, 1983; Falk & Adelman, 2003). Gerade Schülerinnen und Schülern (SuS) mit großem Novelty Space soll so ermöglicht werden, an außerschulischen Lernorten besser zu arbeiten.

Forscherfrage: Kann die Bereitstellung zusätzlicher Struktur auch SuS mit großem Novelty Space ermöglichen, Arbeitsmaterialien an außerschulische Lernorte effektiver zu nutzen und den Wissenserwerb erhöhen?

Methode: 224 SuS der sechsten und siebten Jahrgangsstufe im Alter von 11,7(±0,6) Jahren besuchten eine Mitmachausstellung rund um das Thema Bewegung an der Universität für den Zeitraum von vier Stunden. Die Ausstellung unterteilte sich in drei Räume, wobei jeder Raum einem anderen Themenschwerpunkt zugeordnet war. Diese quasi-experimentelle Untersuchung nutzte ein Prä-Post-Test Design. Die Experimentalgruppe bestand aus 112 SuS (58% Jungen), die Kontrollgruppe aus 111 SuS (54% Jungen). Die Kontrollgruppe erhielt ein Basisniveau an Struktur, wohingegen die Experimentalgruppe zusätzliche Struktur in Form einer Vorbereitung auf den Besuch, strukturunterstützender Anweisungen und Erklärungen zu den Arbeitsmaterialien erhielt. In beiden Gruppen arbeiteten SuS in Kleingruppen mit Hilfe von Arbeitsblättern an einer Reihe von Arbeitsstationen. Bei den Lehrern handelte es sich um Biologiestudenten fortgeschrittener Semester. Alle Lehrer verhielten sich eher kontrollierend. Sie zeigten Präsenz und gaben Zeitfenster für die Bearbeitung der Stationen an. Die drei Faktoren des Novelty Space wurden durch binäre Items erhoben. Als Testinstrumente wurden offene Wissensfragen (oW) während der Ausstellung und geschlossene Wissensfragen (gW) nach der Ausstellung erhoben. Als Implementationskontrolle für die bereitgestellte Struktur kam ein in der Arbeitsgruppe entwickelter Fragebogen zum Einsatz (Connell & Wellborn, 1991; McKenzie, 2000; Jang, Reeve & Deci, 2010, Falk & Storksdieck, 2005).

Ergebnisse: Für Vorerfahrung, Bekanntheitsgrad und Vorwissen wurden jeweils Gruppen durch Aufteilen am Median gebildet. Es wurde jeweils eine zweifaktorielle ANOVA mit einem der drei Faktoren und dem Maß an Strukturvorgabe berechnet.

SuS mit weniger Vorerfahrungen profitieren bei gW von der Bereitstellung von Zusatzstruktur mehr als SuS mit höherer Vorerfahrung. Haupteffekte für Strukturvorgabe ($F(1,209)=53,27, p<0,001, \eta^2=.20$) und Vorerfahrung ($F(1,209)=3,89, p<0,05, \eta^2=.02$) sind signifikant. Sie profitieren ebenfalls bei oW. Haupteffekt für Strukturvorgabe ($F(1,209)=10,05, p<0,01, \eta^2=.05$) und Vorerfahrung ($F(1,209)=4,73, p<0,05, \eta^2=.02$) sind ebenfalls signifikant.

Die Interaktion ist für gW und oW jeweils nicht signifikant.

SuS, die noch nicht an der Universität Bielefeld waren, ziehen nur im Hinblick auf oW einen größeren Vorteil von zusätzlicher Strukturvorgabe als SuS, die bereits dort waren. Haupteffekt der Strukturvorgabe ($F(1,203)=11,10, p<0,01, \eta^2=.05$) und Interaktion ($F(1,203)=5,33, p<0,05, \eta^2=.03$) sind signifikant.

Vorwissen zeigt sich als bedeutender Faktor. Im Falle der gW sind Haupteffekte für Strukturvorgabe ($F(1,212)=53,74, p<0,001, \eta^2=.20$) und Vorwissen ($F(1,212)=7,26, p<0,01, \eta^2=.03$) sowie die Interaktion ($F(1,212)=7,48, p<0,01, \eta^2=.03$) signifikant.

Für oW zeigen sich ebenfalls signifikante Haupteffekte für Strukturvorgabe ($F(1,212)=10,24, p<0,01, \eta^2=.05$) und Vorwissen ($F(1,212)=51,91, p<0,001, \eta^2=.20$). Die Interaktion ist ebenfalls signifikant ($F(1,212)=22,91, p<0,001, \eta^2=.10$).

Fazit: SuS, denen Vorwissen und Vorerfahrungen fehlten, wurde durch die Bereitstellung zusätzlicher Struktur ermöglicht, die Arbeitsmaterialien vergleichbar gut zu bearbeiten und einen ähnlichen Wissenserwerb zu erzielen.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

Maier, Janosch; Spielmann, Jonathan; Baumann, Nadine; Franz, Simone; Polzin, Sarah

Prezi in der Schule: Wird damit mehr und besser organisiertes Wissen erworben als mit PowerPoint?

In der Hochschule sind computerbasierte Präsentationen in Vorlesungen die Regel (Levasseur & Sawyer 2006), und auch im Schulalltag werden PowerPoint-Präsentationen immer häufiger eingesetzt. Neben PowerPoint, welches Folien in einer linearen Folge abspielt, gibt es auch neuere Programme, die Lerninhalte auf andere Weise darzustellen erlauben. Das Präsentationsprogramm Prezi soll durch die graphische Repräsentation der Struktur der zu vermittelten Informationen den Wissenserwerb fördern (Perron & Stearns, 2011). Das Programm bietet dazu die Möglichkeit zur freien Anordnung von Informationsinhalten ähnlich der bei Graphical Organizers (Stull & Mayer 2007), Mind- und Concept-Maps. Studien zur Wirksamkeit von Prezi in Vorlesungen und Unterrichtsvorträgen fehlen bislang jedoch. Nach der kognitiven Theorie des multimedialen Lernens (CTML) kann erwartet werden, dass diese Darstellungsform den Lernenden dabei hilft, die aufgenommenen Informationen zu organisieren (Mayer, 2009). Von einer besseren Wissensorganisation ist auch ein höherer Erwerb von Detailwissen zu erwarten.

Fragestellungen: Daraus ergeben sich folgende Fragestellungen: (1) Führt eine Prezi-Präsentation zu einer besseren Wissensorganisation eines Themas als eine PowerPoint-Präsentation? (2) Führt eine Prezi-Präsentation zu umfangreicherem Detailwissen als eine PowerPoint-Präsentation? (3) Geht eine bessere Wissensorganisation mit umfangreicherem Detailwissen einher?

Vor dem Hintergrund der CTML wurden positive Effekte von Prezi auf die Wissensorganisation und auf das Detailwissen unmittelbar nach der Lernphase und zwei Wochen später sowie ein positiver Zusammenhang zwischen der Wissensorganisation und dem Detailwissen erwartet.

Methode.

Stichprobe: An der Studie nahmen 57 Schüler (23 männliche, 34 weibliche) aus zwei elften Klassen eines Wirtschaftsgymnasiums teil. Diese Stichprobengröße ist zum Nachweis großer Effekte, wie sie bei nicht zu komplexem Material in früheren Untersuchungen beobachtet wurden (Stull & Mayer 2007) bei üblichem Signifikanzniveau und einer angestrebten Teststärke von 80% angemessen. Das durchschnittliche Alter der Stichprobe betrug 16.93 Jahre (SD = .75).

Design: In einem Pre-Post-Kontrollgruppenplan wurde innerhalb der beiden Klassen der Faktor „Präsentationsform“ randomisiert manipuliert.

Material und Manipulation der unabhängigen Variable: Zum Ausschluss von Effekten des Vortragenden wurden mit PowerPoint bzw. Prezi zwei inhaltsgleiche Präsentationen zum Thema „Preis- und Produktpolitik“ erstellt und mit derselben Audiospur unterlegt und dauerten 26 Minuten 34 Sekunden.

Abhängige Variablen: Das Vorwissen zum Thema wurde mit 23 Multiple-Choice-Fragen gemessen. Zur Überprüfung der Wissensorganisation sollten die Teilnehmer Begriffe in teilweise ausgefüllte Baumdiagramme mit der Struktur des Themengebietes einsetzen. Das Detailwissen wurde unmittelbar im Anschluss an die Präsentation sowie erneut nach zwei Wochen mit demselben Test gemessen, der 19 Multiple-Choice-Fragen aus dem Test zum Vorwissen sowie zwei offenen Fragen enthielt, die anhand eines Codierschemas mit je 4 Codierdimensionen ausgewertet wurden.

Ergebnisse: Unterschiede zwischen den beiden Gruppen mit unterschiedlichen Präsentationsformen konnten (1) weder für die Wissensorganisation, $F(1, 52) = 0.00$; $p = .95$, $\eta_p^2 = 0.00$, (2) noch für das Detailwissen im unmittelbaren Posttest, $F(1, 49) = 0.43$; $p = .52$, $\eta_p^2 = 0.01$, oder im verzögerten Posttest festgestellt werden, $F(1, 45) = 1.95$; $p = .17$, $\eta_p^2 = 0.04$. Allerdings (3) korrelierte die Wissensorganisation statistisch signifikant mit dem Detailwissen im unmittelbaren, $r = .43$, $p = .00$, sowie im verzögerten Posttest, $r = .42$, $p = .00$.

Diskussion: Entgegen unseren Hypothesen konnten keine positiven Effekte von Prezi auf die Wissensorganisation sowie das Detailwissen festgestellt werden. Zu berücksichtigen ist dabei jedoch, dass die in der vorliegenden Studie eingesetzten Instrumente, insbesondere für die Wissensorganisation, möglicherweise keine ausreichende Sensitivität aufwiesen. Die statistisch signifikante mittlere bis große Korrelation zwischen Wissensorganisation und Detailwissen steht jedoch im Einklang mit der theoretisch begründeten Annahme, dass der Erwerb von Detailwissen mit durch eine günstige Wissensorganisation gestützt werden könnte. In dem Ausmaß, indem es durch zukünftige Forschungsbemühungen gelingt, die Wissensorganisation durch Prezi oder vergleichbare Medien zu fördern, könnte sich dies auch förderlich auf den Erwerb von Detailwissen auswirken. Welche Formen der Präsentation dazu geeignet sind, bleibt jedoch weiterhin eine offene Frage.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

3

Berner, Nicole; Theurer, Caroline; Rogh, Wida; Lipowsky, Frank

Wirkung kultureller Bildung auf die Kreativität im fünften Schuljahr – Vorstellung des Forschungsprojekts KuBiK-5

Theoretischer Hintergrund und Forschungsfrage: Kreativität kann als die Fähigkeit umschrieben werden, neuartige Ideen zu entwickeln und dabei zielgerichtet vorzugehen, um zu einem originellen Produkt zu gelangen (Drevdahl, 1956). Als Problemlösekompetenz beeinflusst Kreativität auch schulisches Lernen und wird daher als leistungsrelevantes Merkmal anerkannt, das es in der Schule zu fördern gilt (z. B. Holm-Hadulla, 2004; Serve, 2000). Obwohl Kreativitätsförderung fest in den Lehrplänen der Bundesländer verankert ist, existiert national wie international besonders zur Frage, wie Kreativität systematisch und gezielt gefördert werden kann, ein erhebliches Forschungsdesiderat. In der aktuellen interdisziplinären Kreativitätsforschung besteht Konsens darüber, dass Kreativitätsentwicklung von Kontextfaktoren abhängig ist und demnach gerade (schulische wie außerschulische) kulturelle Bildung eine relevante Einflussgröße darstellen dürfte (z. B. Amabile, 1996; Berner, 2013; Ebert, 1973; Preiser, 1976; Preiser & Bucholz, 2004; Urban, 2004). Obwohl der kulturellen Bildung eine hohe Bedeutung für die Persönlichkeits- und Kreativitätsentwicklung zugeschrieben wird (Kultusministerkonferenz, 2013), wird deren Wirkung auf die Kreativität(sentwicklung) bislang wissenschaftlich nur wenig untersucht. Hier setzt das Projekt KuBiK-5 an, mit dem die Wirkung Kultureller Bildung auf die Kreativität innerhalb des fünften Schuljahres untersucht wird.

Methode und erwartete Ergebnisse: Das durch die Mercator-Stiftung geförderte Forschungsprojekt ist als quantitative Längsschnittstudie mit zwei Messzeitpunkten im Schuljahr 2015/2016 angelegt. Konkret soll der Frage nachgegangen werden, inwieweit die Nutzung von schulischen und außerschulischen kulturellen Bildungsangeboten die Kreativitätsentwicklung von Schülern des fünften Schuljahres erklärt. Schülerkreativität wird über domänenspezifische und domänenunspezifische Kreativitätsmaße sowie die divergente Denkfähigkeit erfasst. Mit unterschiedlichen Testverfahren sowie Fragebögen werden zusätzlich Schüler- und Elternvariablen erfasst, um theoretisch denkbare Wirkmechanismen überprüfen zu können. Die kulturellen Bildungsangebote werden mittels Fragebögen bei Schülern und Eltern erhoben. Darüber hinaus werden die Intensität sowie die individuell erlebte Qualität der in Anspruch genommenen schulischen wie außerschulischen Bildungsangebote erfragt. Aus diesen Informationen wird - in Anlehnung an das Vorgehen innerhalb des Projekts AMseL (Schneider & Seither-Preiser, 2013) und an die Studie von Burgart (1961) - ein Index kultureller Teilhabe gebildet, der Angaben zur Dauer und zur Intensität der kulturellen Beschäftigung zusammenfasst und die subjektiv erlebte Qualität dieser Angebote berücksichtigt. Es wird erwartet, dass die Kreativität durch die in Anspruch genommenen kulturellen Bildungsangebote positiv beeinflusst wird, dass aber die Qualität und die erlebte Intensität solcher Bildungsangebote dabei von entscheidender Bedeutung sind. Die Studie unterscheidet sich von bisherigen Studien dadurch, dass der Effekt kultureller Bildungsangebote auf die Entwicklung von Schülerkreativität unter Berücksichtigung von primären sowie mediierenden individuellen, häuslichen sowie schulischen und außerschulischen Kontextmerkmalen untersucht werden kann.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

Bellenberg, Gabriele; im Brahm, Grit; Reintjes, Christian

Zur Nutzung gymnasialer Lernzeit im neuen G9-Bildungsgang in Nordrhein-Westfalen: Befunde und Perspektiven aus der Begleitforschung zum Modellversuch „Abitur nach 12 oder 13 Jahren“

Die Dauer der gymnasialen Schulbesuchszeit bis zum Abitur spielt in gegenwärtigen Bildungsdebatten eine prominente Rolle: Nach der entsprechenden KMK-Regelung kann die Hochschulzugangsberechtigung über das Abitur entweder nach acht oder nach neun Jahren erworben werden. Um die Qualität des schnell eingeführten Regelwegs von G8 zu erhöhen, haben Bildungsverwaltungen in den Ländern in den Jahren nach dessen Umsetzung zahlreiche Nachkorrekturen vorgenommen und entstandene Probleme bearbeitet. In der jüngeren Vergangenheit wird um die Vergleichbarkeit und die Güte beider Wege wieder heftig gestritten: Während die neuen Bundesländer nie einen neunjährigen Gymnasialen Bildungsgang eingeführt hatten, haben fast alle übrigen Länder zwischen dem Schuljahr 2001/02 und 2008/09 die Schulzeit am Gymnasium generell auf 8 Jahre (G8) verkürzt. Dabei haben sich zwei Modelle etabliert (6 plus 2 bzw. 5 plus 3). Beide Modelle sind nicht widerspruchsfrei umzusetzen, da sie in Konflikt mit anderen schulischen Vorgaben stehen. Seit 2011/12 hat in einigen Bundesländern ein Rollback eingesetzt, im Rahmen dessen G9 entweder flächendeckend wiedereingeführt wurde und G8 ersetzt oder als Wahloption in unterschiedlichen Varianten bzw. im Rahmen eines Modellversuchs ermöglicht wurde.

Vor dem Hintergrund dieser Situation widmet sich der Beitrag dem Thema der gymnasialen Schulzeit auf Grundlage der wissenschaftlichen Begleitforschung des nordrhein-westfälischen Modellversuchs „Abitur nach 12 oder 13 Jahren“ und geht der Frage nach, wie die Modellgymnasien mit der zusätzlichen nominellen Unterrichtszeit umgehen.

Dabei vergleicht der Beitrag zum einen die aus Dokumentenanalysen der Stundentafeln und schulinternen Curricula gewonnenen Daten beider Bildungsgänge – G8 und G9neu. Zum anderen dienen Ergebnisse von Befragungen der unterschiedlichen am Modellversuch beteiligten Akteure mittels Fragebögen (Eltern der neuen Jg.-St. 5 und Schülerinnen und Schüler der jetzigen Jg. 7 [G9neu] und 8 [G8] im Längsschnitt aus den vergangenen drei Schuljahren) sowie Interviews (Schulleitungen, Lehrerinnen und Lehrer) als Grundlage für den Beitrag.

Aus den Dokumentenanalysen geht hervor, dass berechnet für die Sekundarstufe I Schülerinnen und Schülern in G9neu am Ende der Sekundarstufe I knapp 1.000 Unterrichtsstunden mehr Zeit für schulisches Lernen zur Verfügung haben als ihre Mitschüler im Bildungsgang G8. Demnach führt G9neu bis zur 9. Jahrgangsstufe im Vergleich zu G8 zu einer Verminderung des wöchentlichen Stundenvolumens um zwei bis drei Unterrichtsstunden gegenüber G8 und erst in Jahrgangsstufe 10 zu einem deutlichen zusätzlichen Stundengewinn. In der Orientierungsphase gibt es nur einen sehr geringen zeitlichen Unterschied. Zudem zeigen die Dokumentenanalysen eine Restrukturierung der Verteilung der wöchentlichen Unterrichtsstunden innerhalb der Stundentafel in G9neu. So profitieren alle Fächer durch das zusätzliche 10. Schuljahr, jedoch nicht in gleichem Stundenumfang und zu gleicher Zeit. In der Orientierungsstufe haben alle Schulen die Wochenstundenzahl in Kernfächer erhöht. Die Stundenzahl der sog. Nebenfächer hingegen steigt proportional weniger an. So weisen einige Fächergruppe in der Orientierungsstufe keine Veränderung zu G8 auf, andere hingegen – hier ist vor allem der musisch-künstlerische Bereich zu nennen – verlieren zunächst zwei Wochenstunden, erhalten diese jedoch im Laufe der Mittelstufe größtenteils zurück und überholen dann in ihrer wöchentlichen Gesamtstundenanzahl in der Sekundarstufe I den Wert der Fächergruppen im G8-Bildungsgang um zwei Wochenstunden.

Bezogen auf die Unterrichtsgestaltung und deren Inhalte deuten die Dokumentenanalysen darauf hin, die Erwartungen und Einschätzungen der Lehrkräfte sowie ihrer Schulleitungen zu bestätigen, die zusätzliche Zeit für verstärktes Üben, Vertiefen, Wiederholen oder auch für die Hinzunahme neuer Themen auf- und hierbei eher die Möglichkeit zu haben, offene bzw. schülerorientierte Unterrichtsformen anzuwenden. Dies äußere sich in einer entspannteren und stärkenorientierteren Lernatmosphäre und erhöhe den Blick auf die individuelle Förderung von SuS. Zudem gebe die Mehr-Zeit auch mehr Möglichkeiten zu einer Stärkung der Beziehung zwischen Lehrkräften und ihren SuS sowohl im Unterricht als auch außerhalb. Gleichzeitig ermögliche es ihnen in ihrer Persönlichkeitsentwicklung voranzuschreiten und für bestimmte Themen und Aufgaben eher die dafür notwendige Reife (Reflexions-/Abstraktions-/ Urteilsvermögen) mitzubringen.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

Schafer, Yves; Baeriswyl, Franz

Erfolg in der Berufsausbildung. Faktoren des Ausbildungserfolges bei Absolventen/-innen der dualen kaufmännischen Berufslehre in der Schweiz.

Im vorgeschlagenen Beitrag werden relevante Faktoren des Ausbildungserfolgs mit unterschiedlichem Status (Input-, Prozess-, Output- und Outcomefaktoren) und auf unterschiedlichen Ebenen (Mikro-, Meso, Makroebene) anhand eines theoretisch ausgearbeiteten Rahmenmodells systematisch verortet und empirisch geprüft.

Der Ausbildungserfolg in der Berufslehre hat einen entscheidenden Einfluss auf die unmittelbare berufliche Anschlusslösung. So beeinflusst die Qualifikation am Ende der Berufslehre den erzielten Einstiegslohn positiv (vgl. Wydra-Somaggio et al., 2010). Zudem wurde für tiefere Lehrabschlussnoten ein erhöhtes Risiko inadäquat beschäftigt (vgl. Bertschy, Cattaneo & Wolter, 2008, S.2) oder arbeitslos zu werden (vgl. Müller & Schweri, 2009, S.212) nachgewiesen.

Gelingt der direkte Übergang von der Berufslehre in die Anschlusslösung nicht problemlos, erhöht das die Wahrscheinlichkeit auch zukünftig nicht, oder nicht adäquat beschäftigt zu werden. Schöngen und Westhoff (1992) aber auch Bender, Haas und Klose (2000) zeigen, dass ein stabiler Erwerbsverlauf in den ersten Jahren den weiteren Karriereverlauf nachhaltig prägt (vgl. Dorau, 2010, S.1). Nach Büchel und Pollman-Schult (2003) schaffen es nur bis zu 50% der unterwertig Beschäftigten nach 6 Jahren wieder eine ausbildungsadäquate Anstellung zu kriegen (vgl. Winkelmann, 2006, S.89). Zusätzlich konnte gezeigt werden, dass die beruflichen Einstiegsschwierigkeiten nach der Berufslehre längerfristige Auswirkungen auf neue Jobmöglichkeiten (Heckman & Borjas, 1980; Margolis et al., 2004) oder den Lohn (Gray, 2000) haben (vgl. Bertschy, Cattaneo & Wolter, 2011, S.221).

Aus den dargestellten Forschungsergebnissen geht hervor, wie wichtig der Ausbildungserfolg für die weitere Berufslaufbahn ist. Bisher konnte gezeigt werden, dass der Ausbildungserfolg am besten mit den durchschnittlichen Schulabschlussnoten prognostiziert wird (vgl. Baron-Boldt, Schuler & Funke, 1988). Eignungsdiagnostische Studien finden zusätzlich allgemeine Intelligenz, fachspezifisches Vorwissen, mathematische Kenntnisse, Lesekompetenz und Gewissenhaftigkeit als einflussreich für den Ausbildungserfolg (vgl. Abele & Nickolaus, 2009). Allerdings gibt es kaum Forschungsergebnisse zu den Faktoren, die den Ausbildungserfolg in der dualen kaufmännischen Berufslehre in der Schweiz beeinflussen.

Für den Ausbildungserfolg wird eine objektive und eine subjektive Komponente unterschieden. Dazu wird unter dem objektiven Ausbildungserfolg die erreichte Abschlussnote und unter dem subjektiven Ausbildungserfolg die Zufriedenheit mit dem Beruf verstanden. In Bezug auf den Berufserfolg zeigen metaanalytische Untersuchungen einen mittleren Zusammenhang zwischen Kriterien des objektiven und des subjektiven Berufserfolgs (vgl. Dette, Abele & Renner, 2004; Ng et al., 2005). Ähnliches wird für den Ausbildungserfolg erwartet, weshalb die Suche nach unterschiedlichen Prädiktoren sinnvoll erscheint.

Das Ziel des Beitrags besteht in der Prädiktion des objektiven und subjektiven Ausbildungserfolgs durch die Identifizierung von relevanten Faktoren. Daraus ergibt sich die folgende Fragestellung: Welche Faktoren prognostizieren den objektiven, welche den subjektiven Ausbildungserfolg in der dualen kaufmännischen Berufslehre in der Schweiz?

Im dualen System setzt sich die Qualifikation aus einer Note des berufsfachschulischen und einer des ausbildungsbetrieblichen Teils zusammen, welche je die Hälfte der Abschlussnote ausmachen. Diese lernortspezifischen Noten werden getrennt betrachtet, da die beiden Lernorte unterschiedliche Anforderungen stellen und unterschiedliche didaktische Konzeptionen verfolgen (vgl. Sembill & Seifried, 2010, S. 158).

Die Stichprobe besteht aus 320 Lernenden aus vier Berufsfachschulen in 18 unterschiedlichen Klassen. Alle Daten wurden vor dem abschliessenden Qualifikationsverfahren anhand von Fragebogen im Rahmen des Berufsfachschulunterrichts im Klassenverbund erhoben. Erfasst wurden unterschiedliche Persönlichkeitsmerkmale, psychologische Skalen und Einschätzungen auf der Mikroebene, sowie unterschiedliche Indikatoren auf der Mesoebene, die aufgrund von anderen Forschungsergebnissen und von theoretischen Annahmen als einflussreich für die Vorhersage des objektiven respektive subjektiven Ausbildungserfolgs erwartet wurden.

Zur Beantwortung der Fragestellung werden die längsschnittlichen Daten anhand von (mehrebenen) Regressionsmodellen in MPlus ausgewertet.

Erste Berechnungen zeigen, dass vor allem Faktoren auf der Mikroebene (Persönlichkeit und Schullaufbahn) den Ausbildungserfolg erklären. Es wurden aber jeweils unterschiedliche Faktoren zur Prognose des lernortspezifischen objektiven, und des subjektiven Ausbildungserfolgs gefunden, wobei der objektive und der subjektive Ausbildungserfolg nicht korreliert. Darüber hinaus kann gezeigt werden, dass die Klasse in der Berufsfachschule als Referenzrahmen den objektiven Ausbildungserfolg mitbestimmt.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

3

Tork, Jennifer

Vermeidung von geringer Bildung durch Resilienz

Geringe Bildung ist in der heutigen Gesellschaft ein Ausschlusskriterium für den Erhalt einer Ausbildung oder dem allgemeinen Einstieg in die Berufswelt. Ursache von geringer Bildung bzw. Bildungsarmut ist häufig die soziale Herkunft (Becker 2007; Büchner 2008). Dennoch lassen sich immer wieder Ausnahmen finden, wo Jugendlichen trotz benachteiligenden Bedingungen der Herkunft sich positiv in Bildung entwickeln (Schoon 2006). Es stellt sich die Frage „Wie kann es sein, dass manche Kinder oder Jugendliche trotz bestehender Bildungsbenachteiligung aufgrund existierender negativer Entwicklungsbedingungen, die unter anderem durch die soziale Herkunft bestimmt sind, sich positiv in Bildung entwickeln?“. Die Beantwortung der Fragestellung kann zur Vermeidung der Entstehung geringer Bildung beitragen, die in der eingangs genannten Form gesellschaftlich inakzeptable ist (Hackett et al. 2001; Solga 2009).

Es gibt verschiedene Erklärungsansätze in der soziologischen Bildungsforschung, die Ursachen und Mechanismen von Bildungserfolg untersuchen. Der unterschiedliche Bildungserfolg nach sozialer Herkunft wird von den Ansätzen der rationalen Handlungstheorie oder auch anderen ökonomischen Theorien häufig durch die staturerhaltende Bildungsentscheidung der Eltern, die durch den Status selbst resultieren, erklärt (Breen/Goldthorpe 1997). Sekundäre Herkunftseffekte spielen dabei eine Rolle. Es werden dabei Unterschiede jedoch lediglich zwischen Gruppen untersucht, nicht innerhalb von Gruppen. Zur Beantwortung der Fragestellung und der Vermeidung von geringer Bildung, die häufiger bei Jugendlichen aus bildungsfernen oder einkommensschwachen Familien auftritt, ist daher das Augenmerk auf den Bildungserfolg jener zu richten, die sich nicht in Status oder Klassen unterscheiden. Klassische bildungssoziologische Erklärungsansätze berücksichtigen diesen Aspekt eher selten. Deshalb wird hier als ergänzender und alternativer Erklärungsansatz das Konstrukt der Resilienz verwendet. Das Konstrukt stammt ursprünglich aus den Gesundheitswissenschaften (Antonowsky 1997) und ist in der deutschsprachigen soziologischen Bildungsungleichheitsforschung bisher kaum verwendet worden (Boeckenhoff et al. 2012). In der angelsächsischen Bildungsforschung wurde das Konstrukt der Resilienz hingegen häufiger in Bezug auf die Erklärung der Bildungsentwicklung angewendet (Schoon 2006; Johnsons 1997). Resilienz beschreibt eine Art psychische Widerstandsfähigkeit gegenüber biologischen, psychologischen und psychosozialen Entwicklungsrisiken (Wustmann 2005). Dies bedeutet, dass einschneidende Krisen unter Rückgriff auf persönliche und sozial vermittelte Ressourcen überwunden werden, sodass die Bildungsentwicklung positiv verläuft (Welter-Enderlin 2008: 13; Masten/Coatworth 1998). Bei den Ressourcen handelt es sich um sogenannte protektive Faktoren, die innerhalb der Familie (Beziehung zu Eltern), in Form von Persönlichkeitseigenschaften (Kontrollüberzeugung) oder Merkmalen des sozialen Umfeldes außerhalb der Familie (Relevanz der Peergroup) bestehen (Garmezy 1985). Sie wirken den Entwicklungsrisiken entgegen. Es ist anzunehmen, dass diese auch der Bildungsentwicklung förderlich sind bei bestehenden Entwicklungsrisiken.

Die Datenbasis bildet das deutsche Sozio-Oekonomischen Panel (SOEP; Wagner et al. 2008). Informationen der Jugendlichen können mit Informationen der Eltern verknüpft werden. Dadurch kann ein direkter Zusammenhang zur sozialen Herkunft hergestellt werden. Als benachteiligendes Entwicklungsrisiko wird das Erleben von Armutserfahrung angenommen. Durch die Befragung der Jugendlichen, anhand des Jugendfragebogens, ist es möglich Auskunft über die Beziehung der Jugendlichen zu ihren Eltern sowie die Relevanz der Peergroup oder Lehrer als Bezugspersonen außerhalb des familialen Umfeldes zu bekommen. Zudem können auch Persönlichkeitseigenschaften der Jugendlichen berücksichtigt werden. Zur Überprüfung der Annahmen, dass soziale und personale Faktoren protektiv gegenüber dem negativen Einfluss der sozialen Herkunft auf Bildung wirken, werden logistische Regressionsmodelle gerechnet.

Der Einfluss der Beziehung zu den Eltern wirkt sich auf die Bildungsentwicklung positiv aus, wenn keine Armutserfahrungen vorliegen. Die Relevanz der Peergroup sowie der Lehrer weisen einen protektiven Effekt auf. Ähnliche Ergebnisse finden sich beim Einfluss der Persönlichkeitseigenschaften. Die meisten protektiven Faktoren wirken auch unabhängig vom Auftreten von Entwicklungsrisiken. Das Augenmerk zur Vermeidung von geringer Bildung sollte auf der Förderung der protektiven Faktoren liegen, welche den negativen Einfluss der sozialen Herkunft innerhalb der betroffenen Gruppen unterbinden können.

Donnerstag
12. März

15:00 Uhr

Foyer GB - Ost

3

Gabriel, Lena

Interkulturelles Lernen im und durch Sport

Theoretischer Hintergrund: Der Sport stellt eine zentrale Funktion in unserer Gesellschaft dar. Neben gesellschaftlichen Wandlungsprozessen sind auch im Sport Veränderungen zu beobachten. Ein erweitertes Sportverständnis überwiegt, das neben dem Wettkampfgedanken, den Erholungs-, Freizeit-, Kommunikations- und Integrationsgedanken impliziert (vgl. Artikel Apuz, 16-19/2011, S. 3 ff.). Das Dissertationsprojekt „Die Bedeutung von Integration und Sport - interkulturelles Lernen im und durch Sport“ richtet seinen Blick auf den Sport in seiner integrierenden/kulturellen Dimension. Integration und Sport in Relation zueinander schaffen Aussagen wie „Sport verbindet“ oder „Sport spricht alle Sprachen“ und verdeutlichen somit die Chancen und die Potentiale, die im Sport gesehen werden. Ferner stellt der Bereich des Sports (formelle sowie informelle Lernprozesse im Sport) einen wichtigen Bildungspartner dar und darf sich aus diesem Grund der Bildungsarmut/Bildungsbenachteiligung in Deutschland nicht verschließen. Der Integrationsgedanke und die Vorstellung interkultureller Lernprozesse sind längst im Sport sowie den damit verbundenen Institutionen (Sportverein, Schule) angekommen. „Interkulturelles Lernen“ als Prozess ist längst in den Lehrplänen verankert. Das gemeinsame Ziel umfasst eine gelungene Integration von Kindern und Jugendlichen mit Migrationshintergrund und eine damit verbundene Reduzierung der Bildungsbenachteiligung.

Der Forschungsgedanke dieser Arbeit analysiert zum einen die Beziehung zwischen Sport und Integration, zum anderen wird die Umsetzung interkulturellen Lernens in den Bereichen „Bewegung-, Spiel- und Sport“ sowie im Bereich „Bewegung und Tanz“ in der offenen Ganztagschule im Primarbereich betrachtet und ausgewertet. Im Fokus steht eine Modulkonzeption zum Thema „Interkulturelles Lernen im Sport“ (kontinuierliche Sportangebote in der offenen Ganztagschule), welche sich zurzeit in einer Erprobungsphase an Essener Grundschulen befindet. Die wissenschaftliche Begleitung der Module dient der Überprüfung der Effekte in Bezug auf die kontinuierlichen Sportangebote.

Fragestellung: Es wird sich an folgender Forschungsfrage (und dessen implizierten Unterfragen) orientiert: Lassen sich in konzeptorientierten „Bewegungs-, Spiel- und Sportangeboten“ sowie in Angeboten zum Bereich „Bewegung und Tanz“ Ansätze interkulturellen Lernens umsetzen? Wie muss ein solcher Lernprozess sportpädagogisch/sportdidaktisch aufbereitet werden? Und inwiefern spielt die Bedeutung von interkulturellem Lernen aus unterschiedlichen Perspektiven eine Rolle?

In Anlehnung an Degele & Winker (2010) wird eine Mehrebenenanalyse angestrebt, die eine kontextspezifische Untersuchung impliziert. Die analytische Vorgehensweise fokussiert drei Ebenen, die in einer Wechselwirkung zueinander stehen: Strukturen (z.B. Schulstruktur, soziales Milieu), symbolische Repräsentationen (z.B. Zuschreibungen aufgrund des Migrationshintergrunds) und Identitäten (Individualität jedes einzelnen Individuums).

Gerade in Bezug auf die Forschungsfrage (interkulturelles Lernen im und durch Sport) scheint eine Mehrebenenanalyse sinnvoll, da der Prozess des „interkulturellen Lernens“ von unterschiedlichen Ebenen (Schule, Familie, Peer-Group, Individuum selbst) beeinflusst wird. Darüber hinaus wird die mehrperspektivische Vorgehensweise (Gruppendiskussionen Kinder, Interviews Übungsleiter, Interviews OGS Mitarbeiter) die kontextspezifische Untersuchung unterstützen.

Methode: Die Datenerhebung erfolgt durch ein qualitatives Forschungsdesign, da sich dieses durch die Konkretisierung der Forschungsfrage und der damit verbundenen kontextspezifischen Untersuchung besonders gut eignet. Darüber hinaus wird eine offene Herangehensweise im untersuchten Feld fokussiert:

Die Absicht der qualitativen Forschungsmethode besteht darin, soziale Wirklichkeit von innen, d.h. aus der subjektiven Sichtweise der Beteiligten, zu beschreiben. Ziel ist es ein besseres Verständnis über bestimmte Abläufe, Strukturmerkmale und Deutungsmuster in unserer sozialen Wirklichkeit zu bekommen (Flick, 2000, S.14).

Beobachtungsbögen der Sportstunden (teilnehmende Beobachtungen), Videoaufnahmen einzelner Modulstunden, Interviews mit den Übungsleitern/Sportstudenten, Interviews mit den OGS-Verantwortlichen und Gruppendiskussionen mit den Kindern werden das Datenmaterial repräsentieren.

Ergebnisse: Für die Auswertung des Datenmaterials wird die Idee der qualitativen Inhaltsanalyse verfolgt. Die Vorgehensweise im Prozess der Auswertung beinhaltet eine Mischform der Kategorienbildung. Die Hauptkategorien werden aus der Bezugstheorie zum interkulturellen Lernen entwickelt und im Leitfadenterview berücksichtigt. Nach Kuckarts (2014) wird dieser Schritt als grobe Kategorisierung bezeichnet. Die induktive Bildung von Subkategorien wird direkt am Material entwickelt (vgl. Kuckarts, 2014, S. 72 ff.). Die Ergebnispräsentation wird ein Teil der deduktiven und induktiven Kategorien vorstellen und beschreiben.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

Türkyilmaz, Aytüre; Morgenroth, Stefanie; Böttner, Miriam; Bühler-Niederberger, Doris; Gräsel, Cornelia

Kinder als erfolgreiche Akteure – Ein quantitativ-induktiver Ansatz zur Heterogenität

Theoretischer Hintergrund: Forschungsansätze der Soziologie der Kindheit und der Pädagogischen Psychologie postulieren, dass das Kind als selbstorientierter und selbstständiger Akteur zu konzipieren sei (Bühler-Niederberger, 2011; Dickhäuser, 2006). Sozialisationsinflüsse, die stets sozial strukturiert sind und als mehr oder weniger begünstigend für Bildungschancen zu betrachten sind (vgl. Baumert, Watermann & Schümer, 2003; Bos, Tarelli, Bremerich-Voß & Schwippert, 2012), werden also durch das Kind zunächst adaptiert und „übersetzt“ und produzieren (erst) in dieser Weise tatsächlichen Bildungserfolg. In diesem Projekt wird – teils als Ausgangspunkt, teils aber auch induktiv in der Aufarbeitung der Ergebnisse – ein Interaktionsmodell zur Erklärung von unterschiedlichem Bildungserfolg entworfen und empirisch zu belegen versucht. Dieses stellt das Kind als eine reflektierende und (re-)agierende Instanz ins Zentrum von Interaktionen zwischen allen Beteiligten (Eltern-Kind-Lehrkräfte). Diese Leistung des Kindes nähern wir uns mit den Begriffen von „Selbst“ und „Selbstprozessen“ (Gräsel, Bühler-Niederberger & König, 2011; Bühler-Niederberger & Türkyilmaz, 2014).

Fragestellung: Im Einklang mit den Annahmen der kindheitssoziologischen und pädagogisch psychologischen Forschungsansätze wurde in der vorliegenden Studie untersucht, welche Wirkungspfade zwischen dem sozio-ökonomischen Kapital, dem elterlichen Sozialisationsverhalten, dem akademischen Selbstkonzept sowie verschiedenen weiteren Aspekten von Selbstprozessen und den Schülerleistungen bestehen. Das Modell folgt einem quantitativ-induktiven Ansatz, bei dem angenommen wurde, dass die reflektierende Selbstinstanz des Kindes in Interaktion mit Schule einen stärkeren Effekt auf die Schulleistungen zeigt als die Kultivierung über verschiedene Erziehungspraktiken.

Methode: 801 Grundschul Kinder der dritten Klasse (433 Mädchen, 368 Jungen; Mage = 8.79; SD = 0.57) füllten Fragebögen zu den Schulnoten der Fächer Mathematik, Deutsch und Sachunterricht, dem akademischen Selbstkonzept (Dickhäuser, 2006), dem sozio-ökonomischen Kapital (Bos, Bensen, Kummer, Lintorf & Frey, 2009) sowie dem praktizierten Eltern-Kind-Verhältnis aus Sicht des Kindes (ebd.) aus. Die Stichprobe umfasst Kinder aus 31 Grundschulen in Wuppertal, Solingen und Remscheid, die zufällig ausgewählt wurden. Anhand eines Strukturgleichungsmodells wurden die postulierten Zusammenhangsmuster der Variablen untersucht. Für 529 Kinder dieses Samples liegen auch bereits Daten aus der zweiten Klasse vor, auf die zurückgegriffen werden kann.

Ergebnisse: Wie erwartet, zeigte sich ein deutlicher Wirkungspfad zwischen dem akademischen Selbstkonzept und der Schulleistung, der (das zeigen die Berechnungen) in beide Wirkrichtungen gedeutet werden kann. Ein weiterer deutlicher Wirkungspfad zeigt sich zwischen dem sozio-ökonomischen Kapital und den Schulleistungen. Die Wirkungskraft des indirekten Effektes ausgehend vom sozio-ökonomischen Kapital über die elterlichen Erziehungspraktiken und das akademische Selbstkonzept als auch des indirekten Effektes ausgehend vom sozio-ökonomischen Kapital über das akademische Selbstkonzept auf die Schulleistungen, sind wiederum so schwach, dass sie vernachlässigt werden können. Die Ergebnisse legen nahe, dass primär dem akademischen Selbstkonzept und in zweiter Linie dem sozio-ökonomischen Kapital eine hohe Bedeutung für die Schulleistung zugesprochen werden kann. Im Weiteren testen wir in Pfadmodellen unterschiedliche Variablenkombinationen auf der Suche nach der Bedeutung von Selbstprozessen in schulischen Interaktionen. Implikationen für die weitere Forschung zur Bedeutung von Interaktionsprozessen zwischen Eltern, Schule und Schülern für Schulleistungen als auch die praktische Anwendbarkeit der Ergebnisse werden diskutiert.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

Stahl, Juliane F.

(Sozial)Räumliche Verteilung von Kita-Qualität: Hängt Zugang zu guter frühkindlicher Bildung und Betreuung vom Wohnort ab?

Der Besuch von Kindertageseinrichtungen (Kitas) wird häufig als eine Möglichkeit gesehen, systematische Unterschiede in den Fähigkeiten von Kindern je nach sozioökonomischer Herkunft zu reduzieren. Dabei ist die Qualität der Betreuung für die kindliche Entwicklung von entscheidender Bedeutung (Hasselhorn & Kuger 2014; Sylva et al. 2013; Anders et al. 2012; NICHD 2006; Shonkoff & Phillips 2000). Bisherige deutsche Studien deuten jedoch auf eine insgesamt nur mittelmäßige Kita-Qualität mit großen Unterschieden zwischen Einrichtungen hin (Kuger & Kluczniok 2008; Tietze et al. 2013). Es ist anzunehmen, dass eine der Ursachen für die ausgeprägte Heterogenität in der dezentralen Organisation der öffentlichen Kindertagesbetreuung liegt. US-amerikanische Studien zeigen außerdem positive Zusammenhänge zwischen der Qualität einer Kita und dem sozioökonomischen Hintergrund von Kindern, die diese Kita besuchen (Crosnoe et al. 2010; Phillips et al. 1994; Pianta et al. 2002).

Im vorliegenden Beitrag wird der Frage nachgegangen, ob der Zugang zu guter Kita-Qualität und damit frühkindliche Bildungschancen gleich verteilt sind oder systematisch zwischen sozialstrukturell unterschiedlichen Nachbarschaften sowie zwischen Regionen variieren. Die Untersuchung stützt sich auf Daten der Kindergartenkohorte des Nationalen Bildungspanels (NEPS), um zunächst deskriptiv die Verteilung von strukturellen Qualitätsindikatoren (d.h. Gruppengröße, Betreuungsschlüssel, Gruppenzusammensetzung, Qualifikation und Umfang der Weiterbildung von pädagogischen Fachkräften) anhand eines repräsentativen Datensatzes von Kitas in Deutschland abzubilden. Diese strukturellen Indikatoren gelten als notwendige Voraussetzungen für gute Prozessqualität und optimale kindliche Entwicklung. Anschließend werden ausgehend vom organisationssoziologischen Neo-Institutionalismus, der die Einbettung von z.B. Bildungsinstitutionen in sogenannte organisationale Felder mit heterogenen Umweltvorgaben hervorhebt (Arum 2000), Hypothesen zur (sozial)räumlichen Verteilung von Kita-Qualität abgeleitet und getestet. Dieser theoretische Ansatz lässt sich gut auf Kindertageseinrichtungen anwenden, da er auf der Annahme basiert, dass Organisationsstrukturen nicht so sehr aus Effizienzbestrebungen resultieren. Sie entstehen vielmehr durch das Vorhandensein von ‚institutionalisierten Regeln‘, denen Organisationen ausgesetzt sind und an die sie sich anpassen müssen, um die eigene Legitimation und damit den Fortbestand zu sichern - ein Prozess, der Organisationen zunehmend isomorph werden lässt (Meyer & Rowan 1977). Im Kontext der öffentlichen Kindertagesbetreuung in Deutschland beziehen sich solche Regeln beispielsweise auf gesetzliche Vorgaben zur Strukturqualität (z. B. Gruppengröße, Betreuungsschlüssel), auf öffentliche Mittel, die für die Kindertagesbetreuung freigegeben werden, oder auch auf lokale Wettbewerbsbedingungen. Mithilfe von Mehrebenenmodellen soll überprüft werden, ob sich Kitas in Deutschland bezüglich Strukturqualität innerhalb einer Region stärker ähneln als zwischen Regionen. Anschließend wird untersucht, ob ein Zusammenhang zwischen sozioökonomischen Merkmalen der Nachbarschaft einer Kita und deren Qualitätsmerkmalen besteht, und Zugang zu guter frühkindlicher Bildung somit ungleich über sozioökonomische Gruppen verteilt ist. Zudem wird analysiert, ob systematische regionale Unterschiede in der Kita-Qualität teilweise auf regional differierende öffentliche Ausgaben der Kinder- und Jugendhilfe sowie gesetzlich festgelegte Mindeststandards struktureller Betreuungsqualität zurückzuführen sind.

Erste Auswertungen bestätigen, dass große Variationen in Gruppengröße, Betreuungsschlüssel und sozialer Komposition von Gruppen bestehen. Dagegen existieren deutlich geringere Unterschiede hinsichtlich Aus- und Weiterbildung des pädagogischen Personals. Es ergeben sich zudem einige Hinweise darauf, dass sich nachteilige Ausprägungen auf verschiedenen Qualitätsindikatoren kumulieren. So geht etwa eine schlechtere Qualifikation des Betreuungspersonals in Westdeutschland mit größeren Gruppengrößen und in Ostdeutschland mit einem höheren Anteil an Kindern mit Migrationshintergrund pro Gruppe einher. Teilweise erhebliche Diskrepanzen in der mittleren Strukturqualität zwischen Ost- und Westdeutschland lassen des Weiteren darauf schließen, dass der Zugang zu hoher Kita-Qualität vom Wohnort einer Familie abhängt. Entsprechend können substantielle Abweichungen sowohl im Hinblick auf bundeslandspezifische Qualitätsvorgaben als auch landkreisspezifische öffentliche Ausgaben je betreutem Kind beobachtet werden.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

Fäsche, Anika; Sturmhöfel, Nicole; Arndt, Petra A.

Zusammenhänge mathematischer und sprachlicher Basiskompetenzen im letzten Kindergartenjahr mit sozioemotionalen Schulerfahrungen der Kinder nach dem Übergang in die Grundschule

Der Übergang von der Kindertagesstätte in die Grundschule erfordert von Kindern eine Anpassungsleistung sowohl an eine veränderte Lernumwelt als auch an eine neue soziale Gruppe. Der Übertritt gilt als gelungen, wenn sich die Kinder wohl fühlen und sie die schulischen Lernmöglichkeiten ihren Fähigkeiten entsprechend nutzen können. Während der Zusammenhang zwischen sprachlichen und mathematischen Fertigkeiten am Ende der Kindergartenzeit und den entsprechenden schulischen Lernerfolgen naheliegend ist (z. B. Krajewski & Schneider, 2006), sind Auswirkungen sprach- und mathematikbezogener Basiskompetenzen auf die soziale und emotionale Anpassung weniger offensichtlich. Vielmehr werden für den sozialen und emotionalen Bereich häufig familiäre (Belastungs-) Faktoren sowie Persönlichkeitseigenschaften des Kindes diskutiert (z.B. Beelmann, 2006; Faust et al., 2012).

Das Ziel der vorliegenden Untersuchung war eine Prüfung möglicher Zusammenhänge zwischen mathematischen und sprachlichen Basiskompetenzen im letzten Kindergartenjahr und den sozialen und emotionalen Erfahrungen der Kinder in der ersten Klasse. Der Untersuchung im Rahmen der „Wissenschaftlichen Begleitung des „Bildungshaus 3 – 10“, gefördert mit Mitteln des BMBF und des ESF der EU, liegen neben Daten von Kindern, die zwischen regulär kooperierenden Einrichtungen wechseln, auch solche von Kindern vor, deren Einrichtungen im baden-württembergischen Modellprojekt „Bildungshaus 3-10“ sehr intensiv kooperieren. Das „Bildungshaus 3 – 10“ bietet regelmäßige, institutions- und jahrgangsübergreifende Bildungsarrangements, die es bereits Kindergartenkindern ermöglichen mit Schule und schulischen Inhalten in Kontakt zu kommen.

Teilnehmer waren N=2121 Kinder (51% Mädchen) aus drei Kohorten (Alter M=6.04 Jahre, SD=.38, 59% aus Bildungshauseinrichtungen, 41% aus Vergleichseinrichtungen). Sie wurden am Ende des ersten Kindergartenjahres hinsichtlich ihrer Zahlenbegriffsentwicklung (OTZ; van Luit et al., 2001) und Sprachentwicklung (SETK 3-5; Grimm et al., 2010) untersucht. Ein halbes Jahr nach der Einschulung beantworteten die Kinder einen Fragebogen zur Erfassung ihrer emotionalen und sozialen Schulerfahrungen (FEES 1-2, Rauer & Schuck, 2004). Hiervon wurden im betrachteten Modell die Subskalen Selbstkonzept der Schulfähigkeit, Schuleinstellung, Anstrengungsbereitschaft, Lernfreude und Gefühl des Angenommenseins durch die Lehrkraft einbezogen. In den Analysen wurde neben Alter, Geschlecht, Migrationshintergrund und Kohortenzugehörigkeit des Kindes, auch für Bildungsniveau und Erwerbstätigkeit der Mutter kontrolliert.

Die ersten Ergebnisse multipler Regressionen im Mehrebenenendesign mit Bayes-Schätzung zeigten, dass am Ende der ersten Klasse Kinder aus Einrichtungen mit Bildungshaus-Kooperation von einem höheren Selbstkonzept der Schulfähigkeit ($\beta=.08$, BCI .001 .147), jedoch von einer geringeren Anstrengungsbereitschaft ($\beta=-.09$, BCI -.165 -.021), als Kinder ohne Bildungshauserfahrung berichteten. Darüber hinaus zeigten die Analysen, dass die Zahlenbegriffsentwicklung am Ende des Kindergartenjahres positiv mit allen Subskalen des FEES zusammenhing ($\beta=.28$, BCI .180 .375 für Selbstkonzept; $\beta=.10$, BCI .003 .203 für Schuleinstellung; $\beta=.22$, BCI .126 .328 für Anstrengungsbereitschaft; $\beta=.18$, BCI .079 .292 für Lernfreude; $\beta=.15$, BCI .047 .242 für Gefühl des Angenommenseins). Für die Subskala Lernfreude wurde dieser Zusammenhang durch die Bildungshauserfahrung moderiert ($\beta=.10$, BCI .004, .214). Das Niveau der Sprachentwicklung am Ende des Kindergartenjahres hing zudem positiv mit dem Selbstkonzept der Schulfähigkeit ($\beta=.10$, BCI .011, .184) sowie dem Gefühl des Angenommenseins durch die Lehrkraft ($\beta=.10$, BCI .013, .182) zusammen. Hier zeigten sich keinerlei Moderationseffekte durch die Bildungshauserfahrung der Kinder.

Die vorliegende Arbeit betont die Bedeutung mathematischer und sprachlicher Basiskompetenzen für die emotionale und soziale Bewältigung des Übergangs in die Grundschule. Insbesondere die Fertigkeiten in Mathematik hingen konsistent mit den sozioemotionalen Schulerfahrungen der Kinder zusammen. Dabei zeigte sich, dass ein Angebot an jahrgangs- und institutionsübergreifenden Lern- und Spielangeboten vor und nach dem Eintritt in die Schule Einfluss auf die Schulerfahrungen als solche wie auch den Zusammenhang dieser mit mathematischen und sprachlichen Basiskompetenzen ausüben kann. Implikationen dieser Ergebnisse werden sowohl für die zukünftige Forschung als auch für die praktische Anwendung zur Unterstützung der Kinder bei der Bewältigung des Übergangs vom Kindergarten in die Grundschule diskutiert. Darüber hinaus sollen zusätzliche Analysen hinsichtlich möglicher Moderatoreffekte durch familiäre Hintergrundvariablen einbezogen werden, um familiär bedingte Heterogenität besser berücksichtigen zu können.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

3

Sappok, Christopher; Fay, Johanna

Die Erfassung von prosodischem Lesen bei Grundschulkindern über Ratingprozeduren und akustische Messungen

Der geplante Beitrag befasst sich mit dem prosodischen Lesen, das neben Genauigkeit und Geschwindigkeit einen Aspekt von Leseflüssigkeit darstellt (Rosebrock/Nix 2008). Dieser Aspekt wird in den letzten Jahren auch als Prädiktor für die Entfaltung von Bildungserfolg insgesamt erforscht (NAEP 2005). Viele Fragen, z. B. die nach den genauen Zusammenhängen zwischen Prosodie und Verstehen, sind dabei noch nicht beantwortet. Auch theoretische Fundierung und methodologische Vorgehensweisen entwickeln sich noch. Klärungsbedarf lässt sich anhand der Gegenüberstellung von angemessener und ansprechender Prosodie verdeutlichen. In der Deutschdidaktik wird ein eher enges Konzept von Leseflüssigkeit verfolgt, bei dem Prosodie hauptsächlich als syntaxgemäße Gliederung eine Rolle spielt. In der internationalen Leseforschung hingegen wird mittlerweile auch ein weiteres Konzept von Prosodie verfolgt, das alle Aspekte von expressivem, d. h. sinngestaltend-ansprechendem Vorlesen berücksichtigt (Kuhn/Schwanenflugel/Meisinger 2010). Bei dieser Gegenüberstellung tut sich ein Konflikt auf: Die Konzeption einer angemessenen Prosodie droht in didaktischen Kontexten darin zu resultieren, dass ein Text von allen auf dieselbe Weise vorgelesen wird. Die Konzeption von sinngestaltend-ansprechender Prosodie hingegen kann die Verwirklichung von individuellem Stil bedeuten und darin resultieren, dass ein Text höchst heterogen vorgelesen wird (Sappok/Arnold 2012 für erwachsene Sprecher).

Um entsprechende Förderkonzepte für prosodisches Lesen fundiert entwickeln zu können, bedarf es einer sorgfältigen theoretischen und methodologischen Erfassung des Phänomens. Dies ist jedoch nicht unproblematisch: Im Gegensatz zu anderen Aspekten von Leseflüssigkeit kann prosodisches Lesen nicht einfach quantitativ gemessen werden. Bislang versucht man prosodisches Lesen z. B. über Ratingprozeduren zu erfassen, in denen ganze Vorträge durch Experten über multidimensionale Ratingskalen beurteilt werden. Ein Hauptproblem ist dabei mangelnde Interrater-Reliabilität (Kuhn et al. 2010). In unserem Beitrag stellen wir dem einen anderen Zugang gegenüber: In einem Perceptionsexperiment fokussieren wir bei Aufnahmen von 12 Drittklässlern ausschließlich ein einziges prosodisch zu meisterndes Phänomen, nämlich die Deutlichkeit, mit der bei durch eine Person vorgelesenem Dialog der Wechsel der Sprechenden Figur prosodisch markiert wird. Vorlesestimulus war ein hauptsächlich dialogischer Text, bei dem die Figuren durch kleine Icons kenntlich gemacht wurden. Wer gerade „spricht“, musste vom Vorlesenden also rein stimmlich vermittelt werden.

Die 12 Kinder wurden aus einem Korpus von N = 48 ausgewählt. Nach den in NAEP (2005) herangezogenen Maßen repräsentieren sie das oberste Stratum der Gesamtstichprobe. Der Textvortrag jedes Kindes wurde in wiederum 12 Einzelsequenzen zerlegt, die je einen Sprecherwechsel inklusive der flankierenden Repliken enthalten. Die resultierenden 144 Einzelsequenzen wurden insgesamt 96 Ratern vorgespielt, wobei ein Rater in einer computerbasierten experimentellen Sitzung 12 Einzelsequenzen in drei randomisierten Durchgängen zu beurteilen hatte. Die Instruktion dabei lautete: Wie deutlich ist ein Sprecherwechsel zu hören?. Die Antwortmöglichkeit bestand in einer vierstufigen Skala von sehr undeutlich bis sehr deutlich.

Dieses Setting erlaubte die Ermittlung von Intra-Rater-Reliabilität (Gwet 2008) und Inter-Rater-Reliabilität (Hallgren 2012) unter Verwendung der Softwareumgebung R (R Core Team 2013). Zusätzlich wurden in allen Aufnahmen mittels der Software Praat (Boersma & Weenink 2014) verschiedene akustische Korrelate von prosodischen Mitteln zur Markierung von Sprecherwechsel gemessen (z. B. Pausendauer). Die Ergebnisse zeigen signifikante Unterschiede bei den beurteilten Kindern, was aufgrund der Vorauswahl bedeutet, dass sich die Methode zur Feindifferenzierung im oberen Leistungsbereich bewährt hat. Die Reliabilitätswerte sind hoch bis sehr hoch und weisen darauf hin, dass vergleichbare Ergebnisse auch mit bedeutend geringeren Raterzahlen erzielt werden können. Die Korrelation der Ratingmittelwerte je Kind mit den Mittelwerten akustischer Korrelate ist sehr hoch und zeigt Wege auf, wie eine weitgehend automatische Messung von Deutlichkeit bei der Markierung von Sprecherwechsel umgesetzt werden kann. In Planung sind weitere Experimente zur Klärung der Frage, inwiefern der Umgang mit Sprecherwechsel durch Leser als repräsentativ für allgemeinere Aspekte von prosodischem Lesen bzw. Leseflüssigkeit allgemein gelten kann.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

3

Decristan, Jasmin; Rauch, Dominique; Richter, Katja; Schastak, Martin

Förderung der Deutsch-Lesekompetenz bilingualer Grundschüler durch Peer-Learning: Zur Bedeutung des Sprachhintergrunds der Peers sowie der Sprache der Peer-Kommunikation (BiPeer)

Hintergrund: Insbesondere Schülerinnen und Schüler, deren Familien aus der Türkei migriert sind, weisen einen erhöhten Förderbedarf hinsichtlich ihrer Deutsch-Lesekompetenz auf (Haag, Böhme & Stanat, 2012). Das Projekt BiPeer erforscht die optimale Förderung der Deutsch-Lesekompetenz von türkisch-deutsch bilingualen Grundschulkindern durch Peer-Learning.

Allgemein lässt sich beobachten, dass bilingual aufwachsende Kinder in Deutschland kommunizieren im familiären Umfeld und in ihrer Peer-Group in ihren beiden Sprachen (vgl. Reich, 2009). Hierbei befinden sie sich in einem „bilingualen Sprachmodus“ (Grosjean, 2001), in dem es zu intensiven Formen des Code-Switching, dem schnellen Wechsel zwischen den Sprachen, und dem Code-Mixing, der Mischung von Elementen beider Sprachen, kommt. Ausgehend von Cummins (2000) Interdependenzhypothese, die Transfereffekte zwischen akademisch sprachlichen Fähigkeiten der Erst- und Zweitsprache postuliert, kann die Erstsprache von bilingual aufwachsenden Kindern eine Ressource zum Erwerb bildungssprachlicher Fähigkeiten in der Zweitsprache bzw. weiteren Erstsprache darstellen. Im schulischen Kontext interagieren und lernen bilingual aufwachsende Kinder jedoch in der Regel monolingual (deutsch). Hierbei greifen bilingual aufwachsende Kinder in der Regel nicht auf ihre Erstsprache und ihre mehrsprachigen Kommunikationsstrategien als Ressource zum Wissenserwerb zurück. Für die häufig geäußerte Forderung, diese Ressourcen von bilingual aufwachsenden Kindern fruchtbar in den schulischen Unterricht zu integrieren (vgl. Rehbein, 2010), gibt es allerdings bisher wenig konkrete und empirisch überprüfte Vorschläge. Ein vielversprechender Ansatz stellt in diesem Zusammenhang die Methode Peer-Learning dar. Hierbei lernen zwei Kinder gemeinsam in Tandems und nach vorgegebenen Interaktionsregeln. Diese Methode hat sich für basale schulische Fertigkeiten wie Lesekompetenz und mathematische Fähigkeiten bei einer heterogenen Schülerschaft im Hinblick auf kognitive, soziale und motivationale Aspekte als äußerst wirksam erwiesen (Adesope, Lavin, Thompson & Ungerleider, 2010; McElvany & Schneider, 2009; Munser-Kiefer & Kirschhock, 2012; Slavin & Cheung, 2005). Bisherige Forschung fokussiert als Merkmale zur Bildung der Lern tandems insbesondere leistungs- oder altershomogene bzw. -heterogene Zusammensetzungen (vgl. z. B. Robinson, Schofield & Steers-Wentzell, 2005), während die mit Sprache verbundene Merkmale für die Tandem-Bildung bislang kaum Beachtung gefunden haben.

Forschungsfragen: Das Projekt „Förderung der Deutsch-Lesekompetenz bilingualer Grundschüler durch Peer-Learning (BiPeer)“ (Start: 08/2014, Förderung: BMBF) versucht einen Beitrag zur Schließung dieser für die Praxis relevanten Forschungslücke zu leisten, indem die Bedeutung des Sprachhintergrunds sowie der Sprache der Peer-Kommunikation für die Förderung der Deutsch-Lesekompetenz durch peer-gestützte Lernarrangements untersucht werden soll.

Hierbei wird zum einen überprüft, wie die Tandemzusammensetzung abhängig vom Sprachhintergrund der Peers (monolingual-bilingual vs. bilingual-bilingual) zu gestalten ist. Zum anderen wird untersucht, ob in Tandemzusammensetzungen mit zwei bilingual aufwachsenden Kindern ein monolingualer Sprachmodus in Deutsch oder ein bilingualer Sprachmodus zur Aufgabenbearbeitung zu favorisieren ist.

Methode: Zur Untersuchung dieser Fragestellung wird eine 10-wöchige Intervention zur Förderung der Lesekompetenz von 240 monolingual deutschen und bilingual deutsch-türkischen Dritt- und Viertklässlern durchgeführt. Innerhalb der Interventionsstudie soll ein allgemeiner Interventionseffekt (Lesen vs. Mathematik) von einem Effekt der Paarbildung (monolingual-bilingual vs. bilingual-bilingual) und einem Effekt der Sprache der Aufgabenbearbeitung im Peer-Tandem (monolingual Deutsch vs. bilingual Deutsch/Türkisch) unterschieden werden. Das beschriebene Vorhaben stützt sich auf das Leseförderprogramm „Peer Assisted Learning Strategies“ (PALS) (Fuchs, Fuchs, Simmons & Mathes, 2008) als strukturierte und evaluierte Methode (vgl. z. B. Robinson et al., 2005; Rohrbeck et al., 2003) und variiert experimentell die Peer-Konstellation mit Blick auf den Sprachhintergrund der Schülerinnen und Schüler (monolingual-bilingual vs. bilingual-bilingual) sowie – im Fall des bilingualen Sprachhintergrunds beider Peers – die im Peer-Tandem gesprochene Sprache (monolingual Deutsch vs. bilingual Deutsch/Türkisch). Die Wirksamkeit des Lesekompetenztrainings mittels Peer-Learning wird durch den mittleren Lesekompetenzzuwachs der Tandems aus den Experimentalgruppen und den Kontrollgruppen verglichen. Die Kontrollgruppen sind als Treatment-Kontrollgruppen konzipiert und erhalten eine Intervention im Bereich der Mathematik. Eine Pilotierung der Trainings (N=60) ist im Frühjahr 2015 vorgesehen. Im angedachten Posterbeitrag sollen die Projektkonzeption vorgestellt und diskutiert sowie aus den laufenden Pilotierungen der Trainings berichtet werden.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

Stark, Linda

Heterogenität in Vorlesegesprächen als Chance für schriftsprachliches Lernen

Sprachliche Fähigkeiten sind unumstritten eine zentrale Voraussetzung für die Teilhabe an schulischem Unterricht. Die Basis sprachlicher Fähigkeiten wird bereits in den ersten Lebensjahren eines Kindes – vornehmlich innerhalb seiner Familie – gelegt. Ist eine solche Basis bei der Einschulung nur unzureichend vorhanden, hat dies Auswirkungen auf den weiterführenden Bildungsweg (vgl. Gogolin 2014). Insofern können Familien durch vorschulische Sprachförderung ihrer Kinder einen Beitrag dazu leisten, den Weg für späteren Bildungserfolg zu ebnen. Ein gängiges Verfahren zur Erzeugung sprachlicher Kompetenz im Vorschulalter stellt das Vorlesen von Bilderbüchern dar. Sowohl in Familien als auch in Institutionen wird dabei auf Studienergebnisse gebaut, die dem Vorlesen von Bilderbüchern positive Auswirkungen auf den Spracherwerb, insbesondere den Wortschatz, bescheinigen (vgl. z.B. Sénéchal et al. 2008)). Die entsprechenden Studien basieren zumeist auf einem Pre-Post-Design, indem sie die Sprachstände der Kinder vor und nach einer bestimmten Vorleseintervention erheben und diese Pre-Post-Daten miteinander oder mit denjenigen einer Kontrollgruppe ohne entsprechende Intervention vergleichen. Ein solches Design ist darauf angelegt, mögliche Effekte bestimmter, kontrollierter Vorleseinteraktionen auf unterschiedliche kindliche Sprachfertigkeiten nachzuweisen und zu beziffern. Nahezu unberücksichtigt bleibt bei einem solchen methodischen Vorgehen jedoch die Frage danach, WIE das Vorlesen diese Effekte auf die kindliche Sprachentwicklung erzeugt. Während also auf Grundlage bisheriger Forschungsergebnisse unstrittig ist, dass das Vorlesen von Bilderbüchern als spezifischer Input im Spracherwerb fungiert (vgl. Meibauer 2011: 9), ist nur wenig darüber bekannt, was diesen spezifischen Input als solchen auszeichnet. Daher zielte die vorliegende Untersuchung darauf ab, diejenigen sprachlichen Eigenschaften des Vorlesens zu ergründen, die zu den positiven Effekten auf die kindliche Sprachentwicklung führen. Insbesondere wurden dabei Lernmöglichkeiten zum Erwerb des schriftsprachlichen Präteritumgebrauchs in den Blick genommen, der ein zentrales Kriterium zur Beurteilung narrativer und damit auch bildungssprachlicher Kompetenzen darstellt.

Dazu wurden insgesamt 40 Vorlesegespräche von 20 Elternteil-Kind-Paaren zu je einem texthaltigen und einem textlosen Bilderbuch audiographiert. Die transkribierten Vorlesegespräche wurden daran anschließend auf Anpassungen an die kindlichen Aufgaben im Erwerb des Präteritumgebrauchs hin untersucht, die – so die Annahme – die Grundlage eines spezifischen Inputs bilden. Diese Anpassungen, in denen sich die kindlichen Erwerbsaufgaben spiegeln, wurden zunächst quantitativ mittels eines deduktiv entwickelten, tempustheoretischen Kategorienrasters in Hinblick auf den Gebrauch des Präteritums unter unterschiedlichen kommunikationssituativen Bedingungen erfasst. Dabei zeigte sich, dass sich Vorlesegespräche aus aufeinanderfolgenden Abschnitten zusammensetzen, die sich sowohl hinsichtlich der zugrundeliegenden kommunikationssituativen Bedingungen als auch hinsichtlich des Gebrauchs der miteinander konkurrierenden Vergangenheitstempora Präteritum und Perfekt unterscheiden. Diese stark heterogene Zusammensetzung der einzelnen Vorlesegespräche führte zu einer qualitativen Betrachtung der Übergänge, die zwischen den unterschiedlichen Abschnitten im Vorlesegespräch vom Vorleser geschaffen werden. Ein induktives, auf gesprächsanalytischen Prinzipien beruhendes Verfahren stellte heraus, dass diese Übergänge im Vorlesegespräch vom Vorleser überwiegend verbal markiert werden. Durch diese Markierungen kann die sprachliche und kommunikationssituative Heterogenität innerhalb von Vorlesegesprächen für das zuhörende Kind transparent gemacht und mit Sinn gefüllt werden: Dadurch dass an den Übergängen von dem einen zum anderen Bedingungsgefüge verbale Markierungen eingesetzt werden, die den Beginn oder das Ende eines Abschnitts und damit die Veränderung der kommunikationssituativen Bedingungen signalisieren, wird eine Grundlage geschaffen, die Kinder zur Verknüpfung von kommunikationssituativer Bedingung einerseits und Tempusgebrauch andererseits nutzen können. Insofern liegen gerade in der heterogenen Zusammensetzung von Vorlesegesprächen Chancen begründet, die schriftsprachliche Lernmöglichkeiten im Allgemeinen eröffnen und Rückschlüsse auf die sprachsystematische Regularität des narrativen Präteritumgebrauchs im Speziellen ermöglichen.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

3

Wiescholek, Sabrina; Hilkenmeier, Johanna; Buhl, Heike M.

LIFE – Lesen in Familie erleben: Erste Evaluationsergebnisse eines Family Literacy Programms

Durch viele Studien ist mittlerweile bestätigt, dass neben der Schule auch das Elternhaus einen wesentlichen Einfluss auf die Lesekompetenzentwicklung hat (Hurrelmann, 2004). Die wachsende Heterogenität im inklusiven Unterricht an Schulen lenkt den Fokus auch auf die Förderung der Kooperation zwischen den beiden Instanzen Schule und Elternhaus, da positive Zusammenhänge zwischen einer engen Zusammenarbeit zwischen Schule und Elternhaus und dem Leistungszuwachs von Schulkindern bestehen (Barnard, 2004; Epstein & Dauber, 1991).

In Bezug auf schulbezogene Aktivitäten sprechen Befunde dafür, dass gerade Eltern mit niedrigem SES vergleichsweise selten schulische Angebote wahrnehmen (Hertel et al., 2013; Schwanenberg et al., 2013). Family Literacy Programme bieten einen erfolgsversprechenden Ansatz, bildungsferne Eltern, Eltern mit Migrationshintergrund sowie Eltern mit niedrigem SES in das Schulleben einzubinden. Zentrale Ziele dieser Programme sind die Förderung der Zusammenarbeit von Schule und Elternhaus, die Förderung der Sprachfähigkeiten und Kompetenzen der Eltern und die Unterstützung des Schriftspracherwerbs der Kinder (Brooks, 1996). Im angelsächsischen Sprachraum existiert mittlerweile eine Fülle von verschiedensten Programmen. International angelegte Metaanalysen zeigen kleine, aber signifikant positive Effekte der Programme auf den Leseerwerbsprozess (Mol et al., 2008; Sénéchal & Young, 2008; van Steensel et al., 2011). In Deutschland ist das Konzept von Family Literacy allenfalls im Elementarbereich verbreitet. Programme für Grund- bzw. weiterführende Schulen sind rar. Neben Programmen wie HIPPY und Rucksack, die insbesondere den frühen Spracherwerb fokussieren, gibt es nur ein größeres Family Literacy-Programm, das den oben genannten Zielen von Brooks (1996) gerecht wird: FLY (Family Literacy) aus Hamburg. Obwohl in der Praxis ausgesprochen erfolgreich, fehlt den Programmen größtenteils eine theoretische Fundierung und darauf aufbauend eine summative Evaluation. Besonders über die Effektivität von Grundschulprogrammen und die zugrundeliegenden Wirkprozesse ist wenig bekannt (Stanat & Felbrich, 2013; Wild & Lorenz, 2010).

Ziel dieser Studie ist die Evaluation des Family Literacy Programms LIFE – Lesen in Familie erleben. LIFE ist ein Förderprogramm, das Schüler/-innen und deren Eltern gleich zu Anfang ihrer Schullaufbahn im Schriftspracherwerb unterstützt. Besonderes Augenmerk liegt dabei auf der Förderung von Lesemotivation. Mit niedrigrschwellig konzipiertem Material und handlungs- und produktionsorientierter Ausrichtung der Veranstaltungen soll einer möglichst heterogenen Gruppe von Eltern und Kindern ermöglicht werden, in die Welt des Lesens einzutauchen. In sechs Veranstaltungen während des erstens Schuljahres erleben Kinder und Eltern Leselernsituationen. Implizite Lerngelegenheiten zu Hause sollen dadurch bewusst gemacht und die Motivation der Eltern für die Unterstützung ihrer Kinder geweckt werden.

Auf der Basis des Erwartungs-Wert-Modells (Eccles, 1998) ist zu prüfen, welchen Einfluss das LIFE-Projekt bei Eltern auf motivationale Aspekte hinsichtlich des Engagements in der Schule, der Unterstützung ihrer Kinder zu Hause und des eigenen Leseverhaltens hat.

Im ersten Jahrgang wurden in einem quasiexperimentellen Design 46 Eltern-Kind-Dyaden, die sich in Interventions- und Kontrollgruppe aufteilen, am Anfang und Ende der ersten Klasse mittels eines standardisierten Fragebogens befragt. Der Fragebogen enthielt Aspekte wie die Einstellung der Eltern zum Lesen, die Selbstwirksamkeitserwartung der Eltern in Bezug auf die Unterstützung ihrer Kinder, die Anwendung von Lesestrategien, Lesemotivation sowie die Häufigkeit von gemeinsamen Leseaktivitäten zu Hause.

Im Gegensatz zur Kontrollgruppe konnte bei den Eltern der Interventionsgruppe eine positivere Veränderung der elterlichen Einstellung zum Lesen sowie der Selbstwirksamkeitserwartung im Hinblick auf die Leseunterstützung gefunden werden. Auch hinsichtlich der Wahrnehmung von zeitlichen Ressourcen für das Engagement in der Schule wiesen die Eltern der Interventionsgruppe eine positivere Veränderung auf. Des Weiteren wendeten Eltern der Interventionsgruppe zum zweiten Messzeitpunkt häufiger Strategien zum tieferen Verständnis eines Textes an, nachdem sie eine Geschichte mit ihrem Kind gelesen hatten.

Die methodischen Einschränkungen, v.a. durch die kleine Stichprobe, werden im Hinblick auf die Interpretierbarkeit der Ergebnisse ebenso diskutiert wie die praktischen Implikationen hinsichtlich der Förderung von Zusammenarbeit zwischen Elternhaus und Schule.

Donnerstag
12. März

15:00 Uhr

Foyer GB - West

Schmidt, David; Thillmann, Hubertina; Stebner, Ferdinand; Wirth, Joachim
Diagnostische Fähigkeiten von Lehrkräften: Die Validierung eines theoretischen Modells

Einleitung: Ausgangspunkt des Gesamtprojektes ist zum einen die Bedeutsamkeit der diagnostischen Fähigkeiten von Lehrkräften für alle an der professionellen Lehrerbildung beteiligten Expertengruppen (Helmke & Schrader et al. 2008; Helmke, Hosenfeld & Schrader 2004; Schrader 2010; McElvany et al., 2009; Lorenz & Artelt, 2009; Horstkemper, 2006; Langenfeldt, 2006; Bos, Helmke, Leutner & Wirth, 2011), zum anderen die sehr breite Forschungslage zum Themenkomplex der diagnostischen Fähigkeiten von Lehrkräften, die bislang unsystematisch und weder einheitlich noch eindeutig ist (Brünken, 2009; Artelt, 2009; Jürgens, 2012; Schrader, 2009). Statt eines einheitlichen theoretischen Modells, scheinen vielmehr verschiedene subjektive Theorien eine Grundlage der unterschiedlichen Forschungs- und Definitionsansätze zu bilden (Dann, 1983; Bergold & Breuer, 1987; Groeben, Wahl, Schlee & Scheele, 1988).

Ziele des Projektes: Vor diesem Hintergrund war es das Ziel des Gesamtprojektes, bestehend aus einer Delphi-Studie mit zwei Befragungsrunden, die subjektiven Theorien der unterschiedlichen an der professionellen Lehrerbildung beteiligten Expertengruppen empirisch zu untersuchen und gegenüberzustellen. Aufbauend auf einer Analyse der Unterschiede und Gemeinsamkeiten sollte zudem untersucht werden, ob ein gemeinsamer Definitionskern identifiziert werden kann.

In der ersten explorativen Befragungsrunde stand die empirische Klärung der Teilaspekte des Themenkomplexes der diagnostischen Fähigkeiten von Lehrkräften im Mittelpunkt. Dies geschah mit Hilfe von Mind Maps und der Auszählung der zwanzig meist genannten Begriffe je Expertengruppe. Anschließend wurden die zwanzig meist genannten Begriffe und ihre Zusammenhänge je Expertengruppe in einem vorläufigen theoretischen Modell dargestellt. Ziel der nun präsentierten zweiten Befragungsrunde ist es, dieses vorläufige theoretische Modell der Ergebnisse der ersten Befragungsrunde zu validieren. Zudem soll über die zweite Befragungsrunde weiterer Aufschluss über die Struktur im Sinne der Strukturbildung der Inhalte der unterschiedlichen subjektiven Theorien zu diagnostischen Fähigkeiten von Lehrkräften erlangt werden.

Stichprobe: Die Stichproben der ersten (N=323) und zweiten (N=30) Befragungsrunde wurden aus den unterschiedlichen an der professionellen Lehrerbildung beteiligten Expertengruppen gezogen, d.h. aus der Gruppe der Dozenten, die an der universitären Lehrerbildung beteiligt sind, der Fachleiter, der Referendare und der Lehrkräfte.

Methode: In der hier präsentierten zweiten Befragungsrunde der Delphi-Studie wurden mittels einer Stichprobe aus den unterschiedlichen Expertengruppen im Sinne der Strukturlegetechnik (Bonato, 1990; Scheele & Groeben, 1988) Concept Maps aus den zwanzig meist genannten Begriffen der ersten Befragungsrunde generiert.

Anhand dieser Concept Maps wurden hierbei die mentalen Repräsentationen zum Thema diagnostische Fähigkeiten von Lehrkräften erneut erfasst und unter Verwendung der Visualisierungssoftware Akovia (Pirnay-Dummer & Ifenthaler, 2010) für jede Expertengruppe agglomeriert und visualisiert. Neben den deskriptiven Ergebnissen der zweiten Befragungsrunde, die in Form von agglomerierten und visualisierten Akovia-Netzen der einzelnen Expertengruppen vorliegen, werden Ergebnisse der empirischen Validierung eines vorläufigen theoretischen Modells zu diagnostischen Fähigkeiten von Lehrkräften mittels unterschiedlicher statistischer Methoden präsentiert.

Ergebnisse: Zum einen kann mit Hilfe der in der zweiten Befragungsrunde gewonnenen Daten gezeigt werden, dass das vorab definierte theoretische Modell mit seinen Schwerpunkten nicht ausreichend zu verifizieren ist. Zum anderen kann anhand der agglomerierten und visualisierten Akovia-Netze der einzelnen Expertengruppen ein weiterer Beitrag zur empirischen Klärung der Teilaspekte des Themenkomplexes der diagnostischen Fähigkeiten von Lehrkräften und deren mentalen Repräsentationen geleistet werden.

Diskussion & Fazit: Als Fazit kann festgehalten werden, dass eine empirische Validierung eines vorab aufgestellten theoretischen Modells zu diagnostischen Fähigkeiten von Lehrkräften unter Verwendung unterschiedlicher statistischer Methoden vorerst nicht gelungen ist. In Zukunft ist daher anzuraten, andere theoretische Modelle zu generieren und weitere Analysemethoden einzubeziehen. Des Weiteren kann festgehalten werden, dass die Expertengruppen gruppenintern weitaus homogener in ihren subjektiven Theorien bezüglich diagnostischer Fähigkeiten von Lehrkräften sind als gruppenextern. Es ist daher fraglich, ob eine zusammenfassende Definition sinnvoll ist oder eine Existenz mehrerer parallel existierender Definitionen anzustreben ist.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

Corvacho del Toro, Irene

Förderdiagnostische Kompetenzen als Grundlage eines adaptiven Rechtschreibunterrichts

Der Begriff adaptiv beschreibt die optimale Nutzung des Entwicklungspotenzials jeden Schülers im Klassenkontext. Die genaue Einschätzung der Fähigkeiten des Schülers sowie die Antizipation seines nächsten Lernschrittes sind für die Gestaltung adaptiver Lernangebote ausschlaggebend (Karing et al. 2011). Adaptivität von Rechtschreibunterricht kann aus fachdidaktischer Perspektive als Resultat einer optimalen Aufbereitung von Lernmaterialien unter Berücksichtigung von individuellen Ausgangslagen definiert werden. Im Bereich der Rechtschreibdidaktik gibt es erste Hinweise für die Effektivität von Einzelförderungen, die auf dem Fehlerprofil des Schülers basieren und einen schriftsystematischen lernförderlichen Aufbau des Lernmaterials (Corvacho del Toro 2014) sowie konstruktivistische Lehr- und Lernprinzipien berücksichtigen (Kochan 1986, vgl. Hardy et al. 2011; Hattie & Timperley 2007, Hattie 2012).

Im Rahmen der Interventionsstudie „Effektiver Rechtschreibunterricht“ (effRU) sollen Lehrkräfte eine Fortbildung erhalten, bei der sie zunächst die Anwendung eines diagnostischen Instruments zur Erhebung der Rechtschreibleistung erlernen. Darüber hinaus sollen die Lehrkräfte sich mit einer schriftsystematischen Aufbereitung von Lernmaterial sowie mit Merkmalen lernförderlicher Rückmeldungen auseinandersetzen. Ziel der Studie ist, die Übertragung (Bromme & Haag 2008, McCutchen & Berninger 1999) dieses Fachwissens und fachdidaktischen (diagnostischen) Wissens auf den Rechtschreibunterricht zu begleiten und dessen Wirksamkeit für die Gestaltung eines adaptiven lernförderlichen Rechtschreibunterrichts zu überprüfen. Die Forschungsfragen lauten: (1) Gelingt es den Lehrkräften nach der Fortbildung und mithilfe des Coachings, das angeeignete Wissen in ihren Unterricht erfolgreich zu integrieren? (2) Verbessert sich die Rechtschreibleistung der Schüler nach dieser Intervention statistisch signifikant und ist diese Veränderung praktisch bedeutsam?

Im März 2014, nach einer von der Studienautorin geleiteten vierstündigen Fortbildung zur individuellen Rechtschreibförderung, wurde mit zwei Lehrkräften eine weitergehende Kooperation vereinbart. Im Rahmen einer Pilotstudie erhält eine Klasse (Experimentalgruppe, EG, N = 11) ab Oktober 2014 eine Förderung über 20 Sitzungen in Form eines Lehrercoachings zum adaptiven Rechtschreibunterricht. Die andere Klasse dient zunächst als Kontrollgruppe (KG, N = 11) und erhält die Förderung nachträglich.

Im Juni 2014 (Ende des Schuljahres, Normierungszeitraum) wurde die Rechtschreibleistung der Experimental- und der Kontrollgruppe sowohl standardisiert (HSP, May 2010) als auch nach einem halbstandardisierten Schreibanlass (OLFA 3-9, Thomé & Thomé 2014) erhoben (Prä-Messung). Für die Experimentalgruppe wurden nach der OLFA die orthographischen Bereiche identifiziert, die während der Förderung (ab November 2014) behandelt werden sollen. Die Projektmitarbeiterinnen (Studienautorin und studentische Hilfskraft) haben zu jedem Thema exemplarisch Arbeitsblätter erstellt, die die postulierten Richtlinien zur schriftsystematischen und lernförderlichen Instruktion und Auseinandersetzung mit der deutschen Orthographie beachten. Nach den 20 Sitzungen wird die Rechtschreibleistung standardisiert und im freien Texten wieder erhoben (Post-Messung). Ein eigens konstruiertes Diktat zu den behandelten orthographischen Phänomenen sowie ein Lückentext zu den grammatischen Bereichen werden zusätzlich eingesetzt, um spezifische Effekte der Förderung genauer eingrenzen zu können. Alle Testergebnisse werden mit den Ergebnissen der Kontrollklasse verglichen werden (Prä-Post). Es werden Unterrichtstagebücher zur Dokumentation der Unterrichtsstunden geführt. Darüber hinaus werden die Materialien und die Schüleraufgaben dokumentiert. Ein Leitfadenterview mit der Lehrkraft soll zum Projektabschluss die Evaluation der Maßnahme aus Lehrerseite festhalten.

Im Oktober 2014 wird die Lehrkraft über das Fehlerprofil der Klasse, die Rangordnung nach der orthographischen Leistung und Besonderheiten in den Fehlerprofilen informiert. Ebenfalls werden die Arbeitsblätter präsentiert und die schriftsystematischen und lernförderlichen Richtlinien eingehend besprochen. Die Lehrkraft kann und soll während der Maßnahme die Studienautorin konsultieren und Unterstützung bei der Klärung von Fragen erhalten. Ein Coaching-Verhältnis ist angestrebt. Hospitationen sind für die ersten Sitzungen vorgesehen. Von dieser Pilotstudie erwarten wir wichtige Erkenntnisse für die Planung und Durchführung einer nachfolgenden vertiefenden Untersuchung an einer größeren Stichprobe.

Das Poster soll die Ergebnisse der Pilotphase darstellen und die Anlage der Folgestudie präsentieren. Die Folgestudie „Effektiver Rechtschreibunterricht“ (effRU) soll Fragen der Lehrerprofessionalisierung aufwerfen und die Wirksamkeit der postulierten Inhalte unter einem stringenten Forschungsdesign statistisch sichern.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

Kempka, Franziska; Sauer, Susanne; Mulder, Regina H.,

Entwicklung eines Messinstruments zur Erfassung pädagogischer Professionalität von Lehrenden - FALKO Teilprojekt Allgemeine Pädagogik

Nach Shulman (1987) lassen sich drei Wissenskategorien für Lehrende unterscheiden: (a) Fachwissen, (b) fachdidaktisches Wissen und (c) pädagogisches Wissen. Während es, vor allem im mathematisch-naturwissenschaftlichen Bereich, einige Studien zu Fachwissen, fachdidaktischem Wissen und deren Beziehung zueinander gibt (Kunter et al., 2011), sind bislang in der internationalen Forschung nur wenige Messinstrumente, welche das pädagogische Wissen von Lehrenden erfassen, vorhanden (z.B. Voss & Kunter, 2011). Diese fokussieren sich auf die Hauptaufgabe eines Lehrenden, dem Unterricht, und beinhalten die Unterrichtsplanung und -gestaltung. Lehrende sind jedoch weit mehr als nur Unterrichtsgestalter und Wissensvermittler (Terhart, 2007). Nach einer Literaturrecherche (u.a. Dubs, 2009; Cochran-Smith, 2005; Terhart, 2007) wurde in diesem Projekt ein Professionalitätsmodell für Lehrende erstellt, welches das professionelle Wissen und Handeln Lehrender als (1) Unterrichtsgestalter, (2) Entwickler, (3) Coach, (4) Wissenschaftler und (5) Lebenslang Lernender umfasst (Mulder, Messmann & Gruber, 2009). Ziel dieses Projekts war die Entwicklung eines Messinstruments, das, mit besonderer Berücksichtigung des genannten Professionalitätsmodells, die pädagogische Professionalität von Lehrenden umfassend misst.

Hieran anschließend ergibt sich folgende Fragestellung: Wie lässt sich pädagogische Professionalität Lehrender aufbauend auf einem mehrdimensionalen Professionalitätsmodell angemessen erfassen?

Method: Ausgangspunkt für die Entwicklung des Messinstruments war das o.g. Professionalitätsmodell von Lehrenden. Ziel war es hierbei das professionelle Handeln im Kontext der fünf angesprochenen Rollen zu erfassen. Um das professionelle Handeln über alle Rollen hinweg im Handlungskontext Lehrender zu verankern, wurden vier pädagogisch relevante Themen- bzw. Handlungsfelder ausgewählt. Diese waren der (1) Umgang mit migrationsbedingter Heterogenität sowie mit (2) Leistungsheterogenität, der (3) kompetente Umgang mit Medien und die (4) Evaluation von Lernprozessen. Für die Entwicklung eines Messinstruments ergab sich daraus die Maßgabe, professionelles Handeln im Kontext von 5 Rollen und 4 Handlungskontexten abzubilden.

Um dies umzusetzen, wurde auf die Vignettenteknik zurückgegriffen, bei der Vignetten im Sinne kurzer Beschreibungen realitätsnaher Arbeitssituationen genutzt werden, um die Beantwortung einer offenen Frage kontextuell zu verankern (Finch, 1987; Jeffries & Maeder, 2004; Mulder, in press). Die Antworten spiegeln demzufolge das Handlungswissen der Lehrenden wider.

Die Entwicklung der Vignetten wurde anhand dreier Studien, einer (1) Entwicklungsstudie, einer (2) Interviewstudie und einer (3) Pilotierungsstudie durchgeführt.

Im Rahmen der Entwicklungsstudie wurden 46 Vignetten konstruiert. Die Vignetten wurden N=17 Lehrenden verschiedener allgemeinbildender Schulen zur Beantwortung vorgelegt. Anschließend wurden die Vignetten gemeinsam hinsichtlich Augenschein-, Inhalts- und Konstruktvalidität besprochen. Kriterien hierbei waren u.a., (1) ob die Fragestellung klar ist; (2) ob die beschriebenen Situationen tatsächlich im Arbeitsalltag eines Lehrenden vorkommen und (3) ob die Vignetten einer der fünf Lehrerrollen zugeordnet werden kann. Im Anschluss hieran wurden die Vignetten überarbeitet und modifiziert. Zusätzlich wurde pro Vignette ein Kodierschema entwickelt.

Die Ziele der danach folgenden Interviewstudie waren die Sicherung der Inhalts- und Konstruktvalidität sowie die erneute Überprüfung, ob die Vignetten realistische Arbeitsalltagssituationen von Lehrenden darstellen. Hierfür wurden die optimierten sowie neu entwickelte Vignetten N=22 Lehrenden verschiedener allgemeinbildender Schulen vorgelegt und wiederum gemeinsam bezüglich der o.g. Kriterien besprochen. Zusätzlich wurden die Kodierschemata weiterentwickelt, indem neue Antwortalternativen aufgenommen wurden.

Das Ziel der Pilotierungsstudie war dann die quantitative Sicherung der Konstruktvalidität und die Fertigstellung eines Testheftes. An dieser Studie nahmen N=250 Lehrende verschiedener allgemeinbildender sowie beruflicher Schulen teil. Die qualitativen Daten wurden von mehreren Ratern quantifiziert. In diesem Zuge wurde auch das in den vorherigen Studien entwickelte Kategoriensystem präzisiert und das Kodierschema optimiert. Anhand der erzielten Interraterreliabilitäten wurde schließlich eine finale Auswahl an Vignetten für das Testheft vorgenommen. Die Interraterreliabilitäten liegen im Bereich von Krippendorffs' $\alpha = .67 - .98$. Vignetten, bei denen das Minimum von Krippendorffs' $\alpha = .67$ nicht erreicht wurde, wurden aus dem Testheft entfernt.

Ausblick: Im Vortrag werden das Professionalitätsmodell, das methodische Vorgehen zur Entwicklung des Messinstruments sowie die Ergebnisse unserer Studie näher erläutert.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

Göhring, Anja; Schödl, Anja

Fachspezifische Lehrerkompetenzen (FALKO) – Teilprojekt Physik

Das fachspezifische Professionswissen, welches durch das jeweilige Schulfach den Handlungsrahmen einer Lehrkraft vorzugeben scheint, ist von zentraler Bedeutung (Baumert & Kunter, 2006; Kunter & Pohlmann, 2009; Loewenberg Ball, Theule Lubinski & Spangler Mewborn, 2001). Es lässt sich weiter in die Dimensionen Fachwissen (CK: Content Knowledge) und Fachdidaktisches Wissen (PCK: Paedagogical Content Knowledge) unterteilen (Shulman, 1986; Baumert & Kunter, 2011).

Bei der Konstruktion des Testinstruments FALKO-Physik erfolgte hinsichtlich der Erfassung des CK eine Orientierung am Niveau des „vertieften Schulwissens“ bzw. des „vertieften Hintergrundwissens zu Themengebieten des Fachlehrplans“, wie es sich auch bei Fachwissenstests anderer Forschergruppen (vgl. z. B. MT21-, TEDS-M-, COACTIV-Studie, Forschungsverbund ProfiLe-P) wiederfinden lässt. Während bisherige Tests (vgl. z. B. Riese & Reinhold, 2009) primär das CK in der Mechanik erheben, werden bei FALKO-Physik darüber hinaus weitere für die Sekundarstufe I relevante Themengebiete (Elektrizitätslehre, Optik, Wärmelehre) im Test abgebildet. Beim PCK werden – in Anlehnung an die COACTIV-Studie – die Subfacetten „Wissen über Instruktion“ (Erklären und Repräsentieren von Sachverhalten) und „Wissen über Schülerkognition“ (Kennen typischer Schülerfehler und Lernschwierigkeiten) unterschieden. Als weitere Teildimension wurde das „Wissen über Messen und Experimentieren“ erfasst. Die einzelnen Aufgaben des Tests zum PCK wurden so konzipiert, dass sie sowohl das deklarative und prozedurale, als auch das konditionale Wissen (Paris, Lipson & Wixson, 1983) erfassen. Da Letztere situations- und ablauforientiert organisiert sind (sog. Handlungsschemata) (Blömeke et al., 2008), wurden zahlreiche Testaufgaben (Papier-und-Bleistift-Test) so formuliert, dass sie für den Physikunterricht typische Unterrichtssituationen beschreiben.

Mit dem Testinstrument FALKO-Physik soll das fachspezifische Professionswissen von Lehrkräften erfasst und die folgenden Forschungsfragen beantwortet werden: (1) Lässt sich das fachspezifische Professionswissen (PCK, CK) von Physiklehrkräften reliabel und valide messen? (2) Wie hängen PCK und CK im Fach Physik zusammen? (3) Wie unterscheiden sich Physiklehrkräfte an Haupt-/Mittelschulen und Realschulen in Bezug auf die beiden Wissensdimensionen? (4) Wie unterscheiden sich Lehrkräfte an Haupt-/Mittelschulen, die Physik als Hauptfach studiert haben, von denen, die Physik als Nebenfach studiert haben oder ohne ein Physikstudium unterrichten, hinsichtlich der beiden Wissensbereiche?

Eine erste Fassung des Testhefts (36 Items CK, 17 Items PCK) wurde im Januar 2014 mit sechs Fachdidaktikern verschiedener Universitäten sowie sechs erfahrenen Seminarlehrkräften (Realschule) und zwei PCB-Multiplikatoren (Haupt-/Mittelschule) prä-pilotiert. Die Probanden wurden während der Testheftbearbeitung teilweise videografiert und zum sog. Lauten Denken aufgefordert. Durch die Prä-Pilotierung konnten Items modifiziert und reduziert sowie Distraktoren zur Schließung von offenen Items und Einschätzungen zur Augenscheinvalidität der Aufgaben gewonnen werden. Das modifizierte Testheft (19 Items CK, davon 12 offen; 14 Items PCK, alle offen) wurde von Februar bis Mai 2014 an einer gemischten Stichprobe (N = 75) pilotiert. Parallel hierzu wurden ein Kodiermanual erstellt und Rater zur Auswertung der Testhefte geschult.

Die Pilotstudie ergab hinsichtlich der Augenscheinvalidität (vierstufige Skala, 1 = trifft nicht zu, 4 = trifft genau zu), dass sowohl die PCK-Aufgaben als auch die CK-Aufgaben als eindeutig gestellt eingeschätzt werden (Subskalen PCK: $3,56 \leq M \leq 3,66$; CK: $M = 3,67$) und aus Sicht der Lehrkräfte berufsrelevantes Wissen beinhalten (Subskalen PCK: $3,40 \leq M \leq 3,53$; CK: $M = 3,26$). Dass die erhobenen Inhalte im Lehramtsstudium/Referendariat berücksichtigt werden sollten, fand ebenfalls Zustimmung (Subskalen PCK: $3,37 \leq M \leq 3,53$; CK: $M = 3,31$). Erste Auswertungen bzgl. der Testleistung zeigen, dass für die Hauptstudie weder Decken- noch Bodeneffekte zu erwarten sind. Die Skalenreliabilitäten (Cronbachs Alpha) liegen bei .69 (PCK) und .88 (CK). Die Interraterreliabilitäten (Spearman's Rho) bzgl. der PCK-Items sind zufriedenstellend bis gut (Instruktion: .72 - .91, Schülerkognition: .71 - .86, Messen und Experimentieren: .86 - .96), die der CK-Items liegen zwischen .61 und 1.0.

Darüber hinaus sollen auf der Tagung Zusammenhänge zwischen PCK und CK sowie Ergebnisse nach Subgruppen differenziert präsentiert werden. Eine Konstruktvalidierung des FALKO-Physik-Tests an Kontrastpopulationen wird abschließend vorgenommen.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

3

Goreth, Sebastian; Geißel, Bernd

Erfassung fachdidaktischer Lehrkompetenz im technikbezogenen Unterricht der Sekundarstufe 1. - Instrumentenkonstruktion und Validierungsprozess

Theoretischer Hintergrund: In der aktuellen Forschungsdiskussion stellen die Bereiche Fachwissen (CK), pädagogisches Wissen (PPK) und fachdidaktisches Wissen (PCK) die Kernbereiche des Professionswissens von Lehrer(innen) dar. Das Poster zeigt in diesem Zusammenhang das für den Fachbereich Technikdidaktik entwickelte, theoretisch abgeleitete Kompetenzstrukturmodell und die darauf basierende Instrumentenentwicklung zur Erfassung fachdidaktischer Kompetenzen angehender Lehrkräfte des Sekundarbereichs I. Dabei wird ein an internationalen Standards der Professionalisierungsforschung ausgerichtetes video- und textvignettengestütztes Testverfahren vorgestellt und der Validierungsprozess offengelegt.

Forschungsstand: Baumert und Kunter (2006) legten ein Modell professioneller Handlungskompetenz vor, in dem professionelles Wissen als grundlegender Bestandteil für die Profession des Lehrerberufs aufgefasst wird (Baumert/Kunter 2006; Vogelsang/Reinhold 2013). Nach dem von Shulman (1987) postulierten Konzept werden verschiedene Formen professioneller Fähigkeiten unterschieden. In der aktuellen Forschungsdiskussion über professionelle Lehrkompetenzen haben drei der Facetten eine übergeordnete Bedeutung gewonnen: content knowledge (CK), general pedagogical knowledge (PPK) sowie pedagogical content knowledge (PCK) (Shulman 1986, 1987), die vielfach aufgegriffen wurden.

Das Forschungsprojekt COACTIV konnte für die Domäne des Mathematikunterrichts zeigen, dass Fachwissen die Grundlage bildet, um lernwirksames Unterrichten stattfinden lassen zu können (Kunter et al. 2011). Tepner et al. (2012) untersuchen im Projekt ProWiN das Professionswissen von Lehrkräften in den Naturwissenschaften. Unterteilt in den Dimensionen Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen wird eine Vergleichbarkeit der Fachdisziplinen Biologie, Chemie und Physik angestrebt (Tepner et al. 2012; Tepner & Dollny 2014). Erstmals für den Lernbereich Technik wurde an der Technischen Universität Eindhoven, für Lehrpersonen im Primarbereich, ein erster Ansatz zur Definition des fachdidaktischen Wissens (PCK) (vgl. Rohaan/Taconis/Jochems 2012) entwickelt. Als bedeutsam stellten sich dort die Facetten: (1) Wissen über Fehlvorstellungen und Vorverständnisse der Schüler(innen); (2) Wissen über Art und Zweck Technischer Bildung; (3) Wissen über pädagogische Konzepte und Lehrstrategien des Technikunterrichts heraus. Diese empirisch-quantitativen ausgerichteten Projekte sowie Publikationen aus der Technikdidaktik und benachbarter Disziplin der Naturwissenschaftsdidaktik orientierten für die Entwicklung eines theoretischen Arbeitsmodells und bieten gute Anschlussmöglichkeiten für die Forschungsfragen des Projekts EKoL.

Fragestellung: Das Teilprojekt geht unter vergleichender Perspektive der Frage nach, ob fachwissenschaftliche wie fachdidaktische Studieninhalte aus dem Lehramtsstudium der Naturwissenschaften den Erwerb fachdidaktischer Kompetenzen im Bereich Technik positiv beeinflussen. Hierzu wird dieser Kompetenzausschnitt Studierender mit affiner Studienfachkombination (etwa Physik und Technik oder Chemie und Technik) mit nicht-affinen Studienfachkombinationen (etwa Theologie und Technik) verglichen. Zur Prüfung der Hypothese muss ein Instrument zur Erfassung fachdidaktischer Kompetenzen neu entwickelt werden.

Methode: Um fachdidaktische Kompetenzen zu erfassen und eine Hypothesenprüfung vornehmen zu können, wurde ein an internationalen Standards der Professionalisierungsforschung (s.o.) ausgerichtetes, video- und textvignettengestütztes Testverfahren (Bernholt/Parchmann/Commons 2009; Oser/Curcio/Düggeli 2007; Lindmeier/Heinze/Reiss 2013; Kersting 2008) mit dem Fokus auf die fachdidaktische Kompetenz (PCK) entwickelt. Diese wurden mit Fachleitern an Staatlichen Seminaren für Didaktik und Lehrerbildung in BW durch eine horizontale Expertenbefragung hinsichtlich der fachdidaktischen Relevanz sowie der Eindeutigkeit und Alltagsnähe der Testvignetten validiert. In einem zweiten Validierungsschritt wird der Gütemaßstab der Testitems ermittelt. Dazu dient ein mehrstufiges Expertendelphi (bestehend aus Fachdidaktikern, Fachleitern und Fachberatern). Dieser Schritt dient der Erstellung einer Expertennorm zur Identifizierung fachdidaktisch adäquater Handlungsweisen und Bewertungen zu den in den Vignetten dargestellten (kritischen) Unterrichtssituationen. Die Urteile der Studierenden werden dann über einen Paarvergleich mit den Urteilen der Expertennorm verglichen.

Ergebnisse: Neben der Präsentation des Kompetenzstrukturmodells sollen Ergebnisse des quantitativ angelegten Expertendelphis und darauf basierend Auszüge des Instrumentariums (Unterrichtsvignetten) vorgelegt und diskutiert werden. Die Datenerhebung im Delphi erfolgt bis Ende Dezember 2014, so dass bis zur Posterpräsentation im Frühjahr 2015 die Ergebnisse aufbereitet vorliegen.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

Janík, Tomáš; Minaříkova, Eva; Janík, Miroslav

Video Clubs: Professionelle Wahrnehmung von EaF-Lehrer/innen vorher und nachher

Zum Einsatz von Videos zur Förderung professionellen Wissens, Wahrnehmens und Handelns der Lehramtsstudierenden bzw. Lehrer/innen sind heutzutage zahlreiche Studien vorhanden. Es werden unterschiedliche Formen der Intervention in der Lehrer(Innen)ausbildung bzw. Fort- und Weiterbildung eingesetzt, oft in den Fächern Mathematik und Naturwissenschaften (z. B. Interventionsprogramme: problem-solving cycle, video clubs, oder dialogic video cycle). EaF/EFL- Lehrer/innen (Englisch als Fremdsprache, English as a foreign language) stehen in den aktuellen Forschungsaktivitäten nicht im Vordergrund. Aus diesem Grund zielt unsere Studie auf diese Berufsgruppe. Die Studie wird im Rahmen eines Forschungsprojektes zur Untersuchung der professionellen Wahrnehmung von EaF-Lehrer/innen (GA13-21961S) realisiert. Der vorliegende Beitrag fokussiert auf den Einsatz von Video Clubs (vgl. van Es & Sherin, 2006) zur Förderung der professionellen Wahrnehmung. Unsere Video Clubs gehen von den van Es und Sherin durchgeführten Video Clubs aus, d.h. kleine Gruppen von EaF-Lehrer/innen treffen sich regelmäßig, um Videosequenzen aus ihrem eigenen Unterricht zu diskutieren. Zusätzlich wurde noch ein Element aus dem Programm problem-solving cycle implementiert (z.B. Koellner et al., 2007): Die Lehrer/innen entwerfen gemeinsam eine Aufgabe (lesson design), deren Realisierung im Unterricht auf Video angenommen und im Rahmen des Video Clubs diskutiert wird. Das Ziel unserer Studie ist es, die Veränderungen in der professionellen Wahrnehmung der EaF- Lehrer/innen nach dem Video Club zu erfassen. An unserer Studie haben 11 an der Sekundarstufe I. in der Tschechischen Republik praktizierende Lehrer/innen teilgenommen. Sie wurden in drei Gruppen aufgeteilt, in denen sie innerhalb von einem akademischen Jahr im Rahmen der Video Clubs zusammengearbeitet haben. Die Daten wurden mittels videostimulierten Interviews vor und nach den Video Clubs gesammelt. Die Lehrer/innen wurden darum gebeten, eine fremde Videosequenz zu kommentieren (es wurde kein „scaffolding“ angeboten), um von den Lehrer/innen wahrgenommene Inhalte zu erfassen. Die Daten wurden aus zwei Perspektiven analysiert: (a) welche Themen werden von den Lehrer/innen eingeführt und wie werden diese Themen erarbeitet (d.h. selektive Aufmerksamkeit, selective attention), und (b) welche kognitive Prozesse sind in ihren Kommentaren ersichtlich (d.h. Indikatoren des knowledge-based reasoning). Die entworfenen, theoriegeleiteten Kategoriensysteme dienten zur Erfassung der selektiven Aufmerksamkeit (Kategorien: Lehrer/in, Schüler/innen, Unterrichtsprozesse, Inhalt, Kontext) und der knowledge-based reasoning (Kategorien: Beschreiben, Interpretation, Erklärung und Vorhersagen, Evaluation, Alteration). Die anschließende Analyse bezieht nicht nur Häufigkeit der Codes in Kategorien ein, sondern auch den Zusammenhang zwischen der selektiven Aufmerksamkeit und der knowledge-based reasoning. Da die Kategorien der selektiven Aufmerksamkeit relativ umfassend entworfen wurden, erfolgte zusätzlich eine offene Codierung der Interviewsegmente. Somit konnte der Fokussierungsgrad und die Differenzierung der selektiven Aufmerksamkeit von Lehrer/innen innerhalb der jeweiligen Kategorie erfasst werden. Zurzeit stehen nur die vorläufigen Ergebnisse zur Verfügung, aber zur GEBF-Tagung können die Endergebnisse präsentiert werden. Unsere Analyse möchte einen Beitrag zur Erforschung leisten, wie Video Clubs die professionelle Entwicklung von EaF-Lehrer/innen sowie den Charakter ihrer professionellen Wahrnehmung beeinflussen.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

3

Austermann, Nora; McElvany, Nele

Wie beeinflussen strukturelle und personalen Faktoren der Lehrkräfte den Einsatz neuer Rahmenlehrpläne im schulischen Unterricht?

Lehrpläne zählen – neben Studentafel und Prüfungsbestimmungen – zu den wichtigsten staatlich autorisierten Vorgaben für schulischen Unterricht und sind als entscheidende Instrumente zur politisch-staatlichen Steuerung schulischer Lernprozesse zu verstehen (Vollstädt, 2003). Inwiefern Lehrpläne ihre intendierte Wirkkraft auch entfalten, also durch die Nutzung der Lehrkräfte ihren Weg in den schulischen Unterricht finden, ist jedoch noch nicht hinlänglich geklärt (Wacker, Rohlfis & Kramer, 2013). Bisherige Forschungsergebnisse weisen darauf hin, dass Lehrpläne die staatlich intendierte Steuerungsfunktion nur bedingt erfüllen, da diese den Lehrkräften zwar häufig bekannt zu sein scheinen, allerdings nur selten zur tatsächlichen Unterrichtsplanung herangezogen werden (Vollstädt, Tillmann, Höhmann & Tebrügge, 1999; Pant, Vock, Pöhlmann & Köller, 2008).

Die Hinwendung zu einem stärker am ‚Output‘ orientierten Bildungssystem hat ihren Kern in der Konzeption neuer Lehrpläne. So wurden in Deutschland nationale, kompetenzorientierte Bildungsstandards eingeführt, welche unter anderem den ‚Output‘ in Form von Kompetenzen, Qualifikationen, Wissensstrukturen und Haltungen bei Schülern messbar machen sollten (Klieme et al., 2003). Ob und wie ‚Neues‘ seinen Weg in die Praxis findet, kann von vier Bedingungsfaktoren abhängen (Gräsel, 2010): (1) den Merkmalen der Innovation selbst, (2) den Merkmalen der beteiligten Lehrkräfte, (3) den Merkmalen der einzelnen Schulen und (4) den Merkmalen der Unterstützungssysteme.

Dieser Beitrag beschäftigt sich mit den Fragen, inwiefern die Nutzungshäufigkeit neuer Rahmenlehrpläne durch Faktoren auf personaler Ebene (Erschöpfungs- und Erfolgserleben, Innovationsbereitschaft) einerseits und durch strukturelle Bewertungsaspekte der Lehrkräfte bezüglich der Rahmenlehrpläne (Zufriedenheit mit formulierten Kompetenzerwartungen und Inhaltskontexten, wahrgenommener Mehrwert) andererseits erklärt werden kann. Zudem ist von Interesse, welche Prädiktoren sich hierbei in einem Gesamtmodell als bedeutsam erweisen.

Zur Beantwortung dieser Fragen wurden Daten des Projekts „Evaluation der Rahmenpläne der DG Belgien“ genutzt, welche im Frühjahr 2014 anhand eines Fragebogens erhoben wurden. Hierbei wurden 223 Lehrkräfte der Deutschsprachigen Gemeinschaft in Belgien befragt (Nmännlich = 47; Nweiblich = 176; MAlter = 40,9 Jahre [SD = 11,2]).

Zur Erfassung der Nutzungshäufigkeit wurde eine Skala mit neun Items verwendet, welche sich auf verschiedene Aspekte von möglichen Einsatzanlässen für Rahmenlehrpläne beziehen. Als personale Merkmale wurden die Einstellung zu Innovationen, das Erschöpfungserleben und das Erfolgserleben erfragt. Zur Erfassung der strukturellen Merkmale der Rahmenpläne wurden Skalen zum wahrgenommenen Mehrwert, zur Zufriedenheit mit den formulierten Kompetenzerwartungen und Inhaltskontexten eingesetzt. Alle Reliabilitäten waren mindestens zufriedenstellend.

Zur Erklärung der Nutzungshäufigkeit der Rahmenpläne wurden Modelle, von denen eines den Einfluss der personalen Merkmale, eines den Einfluss der strukturellen Merkmale und ein weiteres beide Einflussbereiche berücksichtigt, spezifiziert und mittels einer multiplen Regressionsanalyse berechnet.

Die Befunde zeigen, dass die Rahmenlehrpläne von den befragten Lehrkräften insgesamt nicht sehr häufig genutzt werden. Für beide Prädiktorbereiche, personale sowie strukturelle, ließen sich jedoch Prädiktoren identifizieren, welche die Nutzungshäufigkeit statistisch signifikant vorhersagten. Im ersten Modell war die Einstellung zu Innovationen ein signifikanter Prädiktor der Nutzungshäufigkeit ($\beta = .39, p < 0.01$), während sich im zweiten Modell der wahrgenommene Mehrwert ($\beta = .21, p < 0.01$) und die Zufriedenheit mit den formulierten Kompetenzerwartungen ($\beta = .27, p < 0.05$) als signifikante Aspekte erwiesen. Im Gesamtmodell, welches 34 % der Varianz erklärt, sind die Werte für die Einstellung zu Innovationen ($\beta = .18, p < 0.05$) und den wahrgenommen Mehrwert der Rahmenpläne ($\beta = .17; p < 0.05$) als statistisch signifikant und die Aspekte Zufriedenheit mit den Kompetenzerwartungen ($\beta = .18, p < 0.10$) und den Inhaltskontexten ($\beta = .18, p < 0.10$) als tendenziell signifikant abzusichern.

Lehrplänen wird in Literatur und Forschungsarbeiten eine hohe Wichtigkeit für den schulischen Unterricht zugesprochen, jedoch ist noch nicht hinreichend geklärt, wie die Nutzung dieser durch die Lehrkräfte begünstigt werden kann. Dieser Beitrag bietet Hinweise darauf, dass bestimmte personale Merkmale, wie die Einstellung zu Innovationen, und strukturelle Merkmale, wie beispielsweise der wahrgenommene Mehrwert der Rahmenpläne, die Nutzungshäufigkeit positiv vorhersagen können.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

Stubben, Sina A.; Häusler, Janina; Jurik, Verena; Seidel, Tina

Heterogenität im Zusammenspiel von Schülerlernvoraussetzungen methodisch erfassen: Schüler- und Lehrerperspektive über das achte Schuljahr

Theoretischer Hintergrund: Schüler sind heterogen bezüglich ihrer Voraussetzungen (z.B. Kognitive Fähigkeiten, Vorwissen, Interesse, Selbstkonzept), die wichtig für erfolgreiches Lernen sind (Snow, Corno & Jackson, 1996). Um Schülerinnen und Schülern individuelle Unterstützung anzubieten, müssen Lehrer diese Schülerlernvoraussetzungen korrekt diagnostizieren (Baumert & Kunter, 2006). Bisher wurden sowohl aus Schüler- als auch aus Lehrerperspektive kognitive und motivational-affektive Schülermerkmale meist nur getrennt voneinander untersucht (Seidel, 2011). Daher stellt sich die Frage nach dem Zusammenspiel der unterschiedlichen Voraussetzungen aus Lehrer- und Schülersicht.

Methodisch gibt es mehrere Ansätze, die Heterogenität in dem Zusammenspiel unterschiedlicher Schülerlernvoraussetzungen zu erfassen. Das methodische Vorgehen ist gerade aufgrund der unterschiedlichen Erfassung von Lehrer- und Schülerangaben entscheidend. Viele Studien untersuchen zunächst die Korrelationen der unterschiedlichen Merkmale, etwa Interesse und Leistung (Krapp, Schiefele, & Schreyer, 1993). Zudem hat sich die latente Klassenanalyse für die Gruppierung von Schülern in Profile mit einer unterschiedlichen Anordnung einzelner Merkmale etabliert (cf. Yang, Shaftel, Flasnapp, & Poggio, 2005). Eine weitere Möglichkeit das Zusammenspiel unterschiedlicher Variablen zu untersuchen ist die Konfigurationsfrequenzanalyse (von Eye, Mair, & Mun, 2010). Ein inhaltlicher Vergleich der beiden unterschiedlichen Perspektiven von Lehrern und Schülern über ein Schuljahr kann so in diesem Beitrag durch einen methodischen Vergleich erweitert werden.

Fragestellungen: Betrachtet wurden folgende Fragestellungen: 1a) Wie korrelieren die Schülermerkmale (allgemeine kognitive Fähigkeiten, Vorwissen, Interesse und Selbstkonzept) nach der Einschätzung durch die Lehrkräfte und nach den tatsächlichen Schülerangaben? 1b) Zeigen sich zu Beginn und Ende des Schuljahres, in den unterschiedlichen Fächern (Deutsch & Mathematik), und bezogen auf das Schüलगeschlecht Unterschiede in den Korrelationen? 2a) Welche Gruppierung unterschiedlicher Schülercluster zeigt sich durch latente Klassenanalyse und Konfigurationsfrequenz nach der Einschätzung durch die Lehrkräfte und nach den tatsächlichen Schülerangaben? 2b) Zeigen sich in der Analyse durch die unterschiedlichen Verfahren bzw. die unterschiedlichen Kodierung der Variablen als stetige bzw. diskrete Unterschiede in den Ergebnissen der Gruppierung? Welche?

Methode:

Stichprobe: Der vorliegende Beitrag basiert auf Daten der Studie „Interaction“. Die Stichprobe umfasst N = 20 Gymnasialklassen der 8. Jahrgangsstufe mit 20 Lehrkräften und insgesamt 472 Schülerinnen und Schülern (3 Mädchenklassen, 53% Mädchen in koedukativen Klassen) aus Bayern.

Design: Die Schülerinnen und Schüler wurden bezüglich ihrer allgemeinen kognitiven Fähigkeiten (KFT) zu Beginn des Schuljahres, zu ihrem Vorwissen und Interesse sowie zu ihrem Selbstkonzept (PISA) bezogen auf Deutsch und Mathematik zu Beginn und zum Ende des Schuljahres getestet und befragt. Gleichzeitig wurden die Lehrkräfte nach ihrer Einschätzung der jeweiligen individuellen Schülevoraussetzung auf einer dreistufigen Skala befragt.

Statistische Analysen: Korrelationen mit SPSS Version 22, Latente Klassenanalysen mit diskreten und kontinuierlichen Variablen mit Mplus, Konfigurationsfrequenzanalyse mit R.

Ergebnisse: In der Untersuchung des Zusammenspiels der Schülermerkmale, deuten die Korrelationen zwischen den untersuchten Schülermerkmalen an, dass die Lehrerperspektive von den gemessenen Schülerangaben abweicht. Lehrer sehen kognitive und motivational-affektive Fähigkeiten enger verknüpft, als es die Schülerfragebogendaten anzeigen. Während Lehrer über das Schuljahr und die unterschiedlichen Fächer sowie das Geschlecht hinweg ähnliche Zusammenhänge sehen, zeigen sich auf Schülerenebene Unterschiede. Die anschließenden näheren Untersuchungen der Schüler- und Lehrerdaten durch latente Klassenanalysen bestätigen dieses Bild. Während Lehrer vor allem homogene Profile wahrnehmen, zeigt eine Gruppierung der Schülerdaten ein differenzierteres Bild. Diesen Trend zeigen auch die Konfigurationsfrequenzanalysen. Im Vergleich der beiden Verfahren erlaubt die Konfigurationsfrequenzanalyse einen detaillierteren Blick auf bestimmte Schülergruppen. Die Kodierung der Variablen verändert die Schülergruppierung und sollte daher bei der latenten Klassenanalyse immer berücksichtigt werden. Insgesamt zeigten die Analysen, erstens, dass es keine methodisch einzig richtige Erfassung von Heterogenität in Schülerdaten gibt. Zweitens wurde gezeigt, dass dennoch über die Verfahren hinweg deutliche Unterschiede in der Erfassung dieser Heterogenität zwischen Lehrereinschätzungen und Schülerangaben bestehen. Dies impliziert sowohl für die Forschung als auch für die Praxis, dass eine differenziertere Betrachtung der Heterogenität von Schülerlernvoraussetzungen nötig ist.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

3

Görtner, Lara; Hüber, Tanja; Käser, Udo; Röhr-Sendlmeier, Una Maria

Kompetenzerhalt und Kompetenzförderung älterer Arbeitnehmer: erste Evaluationsergebnisse zum Projekt LiA (Lernen im Arbeitsalltag)

In der modernen Arbeitswelt wird die Situation von Arbeitnehmern, die etwa 50 Jahre alt oder älter sind, durch verschiedene Aspekte charakterisiert. Je nachdem, welchen Ausbildungs- und Berufsweg sie eingeschlagen haben, handelt es sich einerseits um Arbeitnehmer, die schon seit möglicherweise mehr als 20 Jahren berufstätig sind. Andererseits liegen Verrentung oder Pensionierung noch ca. 10 bis 15 Jahre in ihrer persönlichen Zukunft. Damit ist das Ausscheiden aus dem Arbeitsleben noch kein unmittelbar aktuelles Thema, wenn auch die Arbeitszeit, welche noch in der Zukunft liegt, schon deutlich kürzer ist als das bisherige Berufsleben der Arbeitnehmer dieser Altersgruppe. Aufgrund ihrer langjährigen Erfahrung haben solche Arbeitnehmer oft auch hohe Expertise in ihrem Arbeitsfeld erreicht, die es ihnen langfristig erlaubt eine hohe Arbeitsproduktivität zu erhalten (Börsch-Supan, Düzgün & Weiss, 2007). Ihre Aufstiegsmöglichkeiten sind jedoch in der Regel nur noch gering. Außerdem werden alltägliche Anforderungen häufig nur noch als Routine wahrgenommen und vielfach sind sich ältere Arbeitnehmer gar nicht bewusst, über welche vielfältigen Kompetenzen sie verfügen, da ihnen die eigenen automatisierten Fähigkeiten selbstverständlich geworden sind (Plath, 2002). Schließlich bleibt es für ältere Arbeitnehmer trotz möglicherweise hoher Expertise wichtig sich fort- und weiterzubilden: Einerseits aktualisiert und ermöglicht Weiterbildung eine innere Ordnung als Ausdruck von Bildung, die angesichts einer sich schnell wandelnden äußeren Welt und eines sich auch im Alter noch entwickelnden Selbst unverzichtbar ist (Röhr-Sendlmeier & Käser, 1999, Staudinger & Heidemeier, 2009). Andererseits kann Weiterbildung motorischen und kognitiven Abbauprozessen entgegenwirken, wie sie im Prozess des Alterns zwangsläufig auftreten (Jaeggi, Buschkuhl, Jonides & Perrig, 2008).

Dies macht Fort- und Weiterbildungsangebote erforderlich, in denen ältere Arbeitnehmer nicht bloß über neue technologische Entwicklungen informiert und in deren Handhabung eingeführt werden. Solche Maßnahmen greifen zu kurz, da sie lediglich die zeitnahe Optimierung der beruflichen Fähigkeiten von Mitarbeitern zum Ziel haben. Vielmehr sind ganzheitliche Konzepte notwendig, welche die berufliche Selbstwirksamkeit der Teilnehmer stärken, den Umgang mit beruflichen Belastungen verbessern sowie eine motorische, kognitive und metakognitive Förderung umfassen (Zimmermann, 2009). Angesichts der hohen Heterogenität der Zielgruppe stellt die Konzeption solcher Weiterbildungsmaßnahmen eine besonders schwierige Aufgabe dar, da in einem solchen Training individuelle Besonderheiten der Teilnehmer besondere Berücksichtigung finden müssen – die Teilnehmer müssen dort abgeholt werden, wo sie stehen –, sich zugleich aber interindividuelle Unterschiede im Entwicklungsverlauf zunehmend vergrößern (Berk, 2004).

Vor diesem Hintergrund werden erste Ergebnisse zur Wirksamkeit des Konzepts LiA („Lernen im Arbeitsalltag“) vorgestellt. Hierbei handelt es sich um ein laufendes Forschungsprojekt, in dem eine Weiterbildungsmaßnahme für ältere Arbeitnehmer entwickelt wurde und in Kooperation mit Arbeitgebern aus der Region Köln / Bonn umgesetzt wird. Die Maßnahme besteht aus einem modular aufgebauten Training und beinhaltet ein berufsbiografisches Kompetenzmodul, welches in Abhängigkeit von den Tätigkeitsfeldern der Teilnehmer spezifisch angepasst wird, und vier bereichsspezifischen Modulen, die ein psychomotorisches Training, ein Training zum Stressmanagement, ein kognitives und ein metakognitives Training umfassen. Die einzelnen Module sind eng miteinander verzahnt und greifen in den insgesamt 15 wöchentlichen Trainingseinheiten systematisch ineinander. Sie beinhalten in allen Bereichen neben psychoedukativer Aufklärung umfangreiche Praxiselemente, die mit den Teilnehmern eingeübt werden.

Vorgestellt werden erste Evaluationsergebnisse auf Basis eines Prä-Post-Vergleichs für eine Stichprobe von N = 30 Teilnehmern der bislang ersten Erhebungswelle. Bislang zeigt sich eine hohe Wirksamkeit von LiA: Der Nutzen des Trainings wird von den Arbeitgebern positiv eingeschätzt, es wird von den Teilnehmern sehr gut angenommen und erfährt hohe Akzeptanz. Des Weiteren zeigt sich eine Verbesserung der Teilnehmer in Selbstberichten zu ihrem Kompetenzerleben und Stressmanagement ebenso wie in Tests für ihre kognitiven und metakognitiven Leistungen. Vor diesem Hintergrund wird einerseits aufgezeigt, wie die Konzeption von LiA weiter verbessert und seine Wirksamkeit weiter untersucht werden soll. Andererseits werden Perspektiven diskutiert, wie Maßnahmen wie LiA in der Arbeitswelt etabliert und implementiert werden können.

Donnerstag
12. März

15:00 Uhr

Foyer GC - Ost

Behrmann, Lars; Souvignier, Elmar

Brauchen Jungen männliche Lehrkräfte, um bessere Leistungen in der Schule zu erbringen?

Theoretischer Hintergrund: Bereits seit Jahrzehnten sind Mädchen in vielen schulischen Belangen erfolgreicher als Jungen: Sie erzielen bessere Noten (Aktionsrat Bildung, 2009), besuchen häufiger das Gymnasium und schließen ihre Schullaufbahn 30% häufiger als Jungen mit dem Abitur ab (StBA, 2009). Darüber hinaus werden Jungen bereits in der Grundschule bei gleichen Kompetenzen schlechter benotet als Mädchen (Bos et al., 2005). Als Erklärung für dieses Befundmuster wird häufig die „Feminisierungstheorie“ aufgegriffen, nach der es zum einen in der Schule an männlichen Vorbildern für eine optimale Entwicklung der Jungen fehle (vgl. Budde, 2008), zum anderen die Überzahl an weiblichen Lehrkräften dafür Sorge, dass der Unterricht flächendeckend eher den Bedürfnissen von Mädchen angepasst werde (vgl. Diefenbach & Klein, 2002). In der Konsequenz dieser Annahmen wird deshalb häufig eine Erhöhung des Männeranteils im Lehrerberuf gefordert (vgl. Helbig, 2010). Bislang wurde jedoch kaum überprüft, ob sich die so angenommenen Vorteile von männlichen Lehrkräften auch tatsächlich positiv auf die Benotung und die Kompetenzen von Jungen auswirken. Ebenso ist weitgehend unklar, welche Auswirkungen das Geschlecht der Lehrkraft auf Noten und Leistungen von Mädchen hat.

Fragestellungen: Folgende drei Hypothesen sollen überprüft werden: (1) Im Mittel werden Jungen von Deutschlehrern besser als von Deutschlehrerinnen und Mädchen von Deutschlehrerinnen besser als von Deutschlehrern beurteilt. (2) Es wird angenommen, dass sich Jungen in Klassen von Deutschlehrern und Mädchen in Klassen von Deutschlehrerinnen in ihrer Leistungsentwicklung optimaler entfalten können. Demnach sollte sich sowohl eine Interaktion zwischen dem Geschlecht der Lehrkraft und dem der Schulkinder abbilden lassen hinsichtlich der Entwicklung a. des Leseverständnisses als auch b. der Leseflüssigkeit der Schulkinder. (3) Da angenommen wird, dass Lehrer die „typisch männlichen“ Verhaltensweisen von Jungen im Vergleich zu Lehrerinnen eher als "normal" empfinden, wird davon ausgegangen, dass männliche im Vergleich zu weiblichen Lehrkräften Jungen im Mittel als weniger verhaltensauffällig einschätzen.

Methode: Die Untersuchung wurde mit 83 Sekundarstufenlehrkräften (16 männlich, 67 weiblich) und ihren 1960 Schüler/innen (1054 männlich, 906 weiblich, Durchschnittsalter = 11.6 Jahre) der 5.-7. Klasse durchgeführt. Sowohl Leseverständnis (erhoben mit dem „Frankfurter Leseverständnistest“, FLVT) als auch Leseflüssigkeit (erhoben mit dem „Salzburger Lesescreening“, SLS) wurden jeweils zu Beginn des Schuljahres als auch zum Ende des ersten Schulhalbjahres erhoben. Zur Erfassung der Schulnoten wurden die Kinder gebeten, ihre Deutschnote aus dem Halbjahreszeugnis anzugeben. Die Lehrerurteile über externalisierende Verhaltensauffälligkeiten der einzelnen Schüler/innen wurden ebenfalls zum Halbjahresende über die SNAP IV-ADHS-Rating-Skala erhoben.

Ergebnisse: Alle drei Hypothesen wurden mittels Intercepts-and-Slopes-as-Outcomes-Mehrebenenanalysen überprüft.

Hypothese 1. Unter Berücksichtigung der FLVT- und SLS-Ausgangsleistungen zum Beginn des Schuljahres zeigte sich keine signifikante Interaktion zwischen Lehrer- und Schülergeschlecht im Hinblick auf die Benotung ($p = .186$).

Hypothese 2a: Bezüglich der Entwicklung des Leseverständnisses der Schüler/innen zeigte sich keine Interaktion zwischen Lehrer- und Schülergeschlecht ($p = .899$).

Hypothese 2b: Im Hinblick auf die Entwicklung der Leseflüssigkeit konnte keine Lehrer*Geschlecht x Schüler*Geschlecht-Interaktion nachgewiesen werden ($p = .276$).

Hypothese 3. Es zeigte sich, dass männliche Lehrkräfte tendenziell Jungen als mehr und Mädchen als weniger verhaltensauffällig einschätzten als dies von weiblichen Lehrkräften der Fall war ($p = .091$).

Insgesamt konnte keiner der angenommenen Vorteile für Jungen durch männliche Lehrkräfte bestätigt werden: Von männlichen und von weiblichen Lehrkräften unterrichtete Schüler/innen unterscheiden sich weder hinsichtlich ihrer objektiven Leistungsentwicklung noch hinsichtlich der Beurteilungen durch ihre Lehrkräfte. Aus den Befunden lässt sich also ableiten, dass eine pauschale Erhöhung des Anteils männlicher Lehrkräfte nicht notwendigerweise zur Leistungsverbesserung von Schülern führen dürfte. Die „Feminisierungstheorie“ konnte demnach nicht bestätigt werden.

Donnerstag
12. März

15:00 Uhr

Foyer GC - West

Schmidt, Jennifer; Öztürk, Halit; van Ophuysen, Stefanie

Gender und Ethnie im Kontext erwachsenenpädagogischer Professionalität. Welchen Anforderungen begegnen Kursleitende in Integrationskursen und wie gehen sie mit diesen um?

Seit dem Jahr 2005 werden in Deutschland Integrationskurse über das Bundesamt für Migration und Flüchtlinge angeboten. Diese richten sich an Ausländerinnen und Ausländer und werden je nach Aufenthaltsstatus verpflichtend, mit einem gesetzlichen Anspruch oder freiwillig besucht. Ziel der Integrationskurse ist die erfolgreiche Vermittlung der Sprache, der Rechtsordnung, der Kultur sowie der Geschichte Deutschlands, um das politische und gesellschaftliche Ziel der Integration zu fördern.

Da sich Integrationskurse neben den Jugendintegrationskursen hauptsächlich an Erwachsene richten, ist es eine Aufgabe der Erwachsenenbildung auf die Heterogenität, die durch unterschiedliche Kulturen und Herkunftsländer sowie durch unterschiedliche Geschlechterrollen geprägt ist, zu reagieren, um für alle Teilnehmenden gleiche Lernmöglichkeiten zu schaffen.

Durch die Zielsetzung der Integrationskurse wird die kulturelle Identität eines jeden Teilnehmenden in den Mittelpunkt gestellt, sodass eine besondere Lernsituation entsteht. Das Geschlecht der Teilnehmenden stellt eine weitere Heterogenitätsdimension dar. Das soziale Geschlecht (in Abgrenzung zum biologischen Geschlecht, im Englischen „sex“ vs. „gender“) wird nach dem Konzept „Doing Gender“ nach West und Zimmerman (1987) als soziale Kategorie beschrieben, in der den dichotomen Geschlechtern bestimmte Verhaltensweisen und Stereotype zugeordnet werden. Diese werden in alltäglicher Interaktion immer wieder reproduziert. Innerhalb verschiedener Kulturen liegen unterschiedliche Geschlechterrollen vor. Dadurch ergibt sich in Integrationskursen insofern eine besondere Situation, als nicht nur Ethnie und Gender als separate Kategorien Interaktionen beeinflussen, sondern insbesondere auch ihre Interdependenz bedeutsam wird.

Innerhalb der Diversitätsdebatte wird postuliert, dass Menschen aus unterschiedlichen Kulturen unterschiedliche Ressourcen mit sich bringen, die positiv betrachtet werden (vgl. z.B. Sickendiek, 2007). Entsprechend wird vermutet, dass sich durch die Heterogenität bereichernde und anregende Lernmöglichkeiten ergeben. Studien aus dem schulischen Bereich deuten darauf hin, dass Lehrkräfte die Ethnie für das Lernen innerhalb einer Gruppe eher als Bereicherung ansehen, der Geschlechterzugehörigkeit jedoch weniger Relevanz beimessen (vgl. z.B. PLAZ-Forum, 2012). Gleichzeitig ist es unumstritten, dass die Potentiale, die die Heterogenität mit sich bringt, nicht immer genutzt werden können. Es entstehen immer wieder auch Schwierigkeiten, die eine besondere Herausforderung an die Kursleitenden und ihre Professionalität stellen. Doch welche Sicht haben Lehrende aus Integrationskursen auf diese Thematik? Inwiefern erleben sie die Heterogenität hinsichtlich Gender und Ethnie als Potential, inwiefern als Problem? Welche Strategien verwenden sie, um diese Potentiale zu nutzen bzw. Probleme zu vermeiden oder zu lösen?

Zur Beantwortung dieser Fragen wurden fünf Kursleiterinnen in leitfadengestützten Interviews mit narrativem Charakter zu ihren persönlichen Erfahrungen aus konkreten Integrationskursen befragt. Dabei sollten sie ihre Erzählungen auf Situationen fokussieren, in denen ihnen der Umgang mit der Heterogenität hinsichtlich Gender und Ethnie besonders gut oder eher schlecht gelungen ist.

Die Interviews wurden mittels der qualitativen Inhaltsanalyse nach Mayring (2002) sowie einer induktiven Kategorienbildung (vgl. z.B. Diekmann, 2009) ausgewertet. In der Auswertung wurde zum einen nach Kategorisierung der Kodiereinheiten hinsichtlich der thematisierten Heterogenitätsdimension (Gender, Ethnie, Interdependenz, Sonstiges) für jede Einheit die Bewertungsrichtung (problematisch, neutral, bereichernd) bestimmt. Die resultierenden Häufigkeitsdaten und Ratings wurden deskriptiv statistisch ausgewertet. Zum anderen wurden aus dem Material heraus mithilfe der induktiven Kategorienbildung Handlungsstrategien, die den Umgang der Kursleitenden mit der Heterogenität betrachten, herausgestellt.

Die kategoriale Analyse macht deutlich, dass die Ethnie bewusster wahrgenommen und in der Lernsituation stärker durch die Kursleitenden berücksichtigt wurde als das soziale Geschlecht. Es zeigte sich, dass die Kursleitenden weitere Heterogenitätsdimensionen (z.B. den Leistungsstand) als relevant erleben. Insgesamt nehmen die Kursleiterinnen die Heterogenität in ihren Kursen überwiegend als potentielle Bereicherung denn als Problem wahr. Es lässt sich somit eher eine positive Bewertungsrichtung der Heterogenität feststellen.

Die Implikationen der Ergebnisse für die Gestaltung von Integrationskursen und die Schulung/Vorbereitung professionellen Handelns von Kursleitenden werden diskutiert. Forschungsdesiderata (z.B. Entwicklung eines standardisierten Fragebogens zur breiteren Überprüfung und Replikation der Befunde) werden aufgezeigt.

Donnerstag
12. März

15:00 Uhr

Foyer GC - West

Mozhova, Anastasia

Gendereffekt in der wissenschaftlichen Produktivität: Welche Mechanismen stehen dahinter?

Theoretischer Hintergrund: Wissenschaftliche Produktivität stellt den Grundstein einer akademischen Karriere, umso wichtiger ist es, dass nur die Qualität der wissenschaftlichen Arbeit und nicht die persönlichen Merkmale, wie zum Beispiel Geschlecht, einen Einfluss darauf haben. In vielen Untersuchungen konnte allerdings gezeigt werden, dass Frauen im Durchschnitt weniger Artikel publizieren als ihre männlichen Kollegen (Long; Fox 1995; Xie; Shauman 1998; Zuckerman 1991). Verschiedene Mechanismen sind für diesen Gendereffekt verantwortlich: So konnten zahlreiche Studien zeigen, dass weibliche Studierende eine geringere Selbstwirksamkeit aufweisen, als ihre männlichen Kommilitonen (z.B. Berweger/ Keller 2005, Spies/Schute 1999). Das Konzept von Selbstwirksamkeit beschreibt die Selbsteinschätzung, eine Handlung bis zu ihrem Erfolg ausführen zu können (Bandura 1982). Ausgehend von den internalisierten Rollenstereotypen, also bestimmten Erwartungen, die an Mitglieder einer Gruppe (Männer oder Frauen) gestellt werden, können auch unterschiedliche Selbstwirksamkeitserwartungen ausgebildet werden (Cialdini; Trost 1998). Diese Erwartungen drücken schon oft beobachtete Verhaltensmuster der Geschlechter aus, wie zum Beispiel die Rolle des Ernährers oder die fürsorgliche Rolle der Mutter (Eagly et al., 2000). Angewendet auf die hier untersuchte wissenschaftliche Produktivität können Geschlechterstereotype dazu beitragen, dass Frauen seltener als Erst-Autor einer wissenschaftlichen Publikation fungieren. Dies kann aus unterschiedlichen Gründen passieren: Das „Sich nicht trauen“ überhaupt eine Publikation einzureichen ist das Naheliegende. Aber auch eine Beschäftigung mit Routineaufgaben im Labor, die eher als Fürsorgetätigkeiten definiert werden und weniger mit wissenschaftlichem Fortschritt assoziiert sind, kann den Frauen zum Verhängnis werden: obwohl nicht weniger wichtig, münden diese Aufgaben seltener in einer anerkannten Publikation (Müller 2013).

Fragestellung, Methode: In der hier vorliegenden Forschungsarbeit wird es der Frage nachgegangen, welche Mechanismen hinter dem Geschlechtereffekt in der wissenschaftlichen Produktivität sich verbergen. Die untersuchte Stichprobe umfasst 402 Promovenden aus dem Bereich Life Science, die zwischen April 2013 und April 2014 an den Hochschulen in Bayern, Sachsen und Nordrhein-Westfalen ihren Dokortitel bekommen haben. Die Anzahl der publizierten Artikel während der Promotionsphase (als Erst-Autor/in) ist die interessierende abhängige Variable. Mit Hilfe der negativen binomialen Regression wird der Effekt der folgenden unabhängigen Variablen berechnet: wissenschaftliche Identität, Partner im gleichen Haushalt, Kinder, Wunsch nach einer wissenschaftlichen Karriere, akademische Selbstwirksamkeit, Zufriedenheit mit der Betreuung, Zusammenarbeit in einer Gruppe, Arbeit am Lehrstuhl. Es werden vier hierarchische Modelle gerechnet, in jedem der Modelle sind die Spezifikationen des Vorgängermodells enthalten.

Ergebnisse: Die deskriptiven Ergebnisse bestätigen den Geschlechtereffekt: während der Promotionszeit haben Frauen nur 1,16 Artikel publiziert, Männer dagegen 1,87. Wenn man allerdings weitere Variablen in das Modell aufnimmt, wird der Gendereffekt geringer, im letzten Modell, durch die Aufnahme der Selbstwirksamkeitsvariable, verschwindet er sogar. Neben der Selbstwirksamkeit hat auch höhere wissenschaftliche Identität und Arbeit an einem Lehrstuhl einen signifikanten positiven Effekt auf die Produktivität.

Schlussfolgerung: Wie sind diese Ergebnisse zu interpretieren? Frauen publizieren im Schnitt weniger Artikel als ihre männlichen Kollegen. Dieser Effekt kann durch geringere akademische Selbstwirksamkeit und wissenschaftliche Identität erklärt werden. Obwohl diese Ergebnisse in vielen anderen Bereichen bestätigt werden konnten, ist es doch erstaunlich, dass die Geschlechterstereotype auch in der Promotionszeit einen so deutlichen Effekt zu haben scheinen. Diese Ergebnisse sind ein wichtiger Schritt, um nachzuvollziehen, warum sich (noch) so wenige Frauen für eine wissenschaftliche Karriere entscheiden und vielleicht eine Hilfe, um diesen Gender Gap zu überwinden.

Donnerstag
12. März

15:00 Uhr

Foyer GC - West

Preuße, Daja; Mater, Olga; Zlatkin-Troitschanskaia, Olga; Stump, Martin; Förster, Manuel

Einfluss schulischer Strukturen und transformationaler Führung der Schulleitung auf die Evidenzorientierung von Lehrkräften

Das Konzept evidenzbasierter Steuerung gewinnt im Bildungssystem verstärkt Beachtung. Hierbei wird der Politik sowie der pädagogischen Praxis mittels verschiedener Evaluationsinstrumente (wie Leistungstests, Schulinspektionen) Wissen als Handlungsgrundlage zur Verfügung gestellt. Eine förderliche Wirkung der neuen outputorientierten Steuerungsinstrumente kann für die Schule jedoch nur dann angenommen werden, wenn das dabei erzeugte Wissen in die Entwicklung und Realisierung unterrichtlicher und schulischer Aktivitäten einfließt.

Die (inter)nationale Forschungslandschaft liefert bisher nur wenige Befunde, welche schulischen Bestimmungsgrößen die Evidenzorientierung von Lehrkräften als eine Bedingung für das Nutzen dieses Wissen beeinflussen können. Die Studien beziehen sich zumeist auf strukturelle Bedingungen wie Kommunikations-, Kooperations- und Partizipationsstrukturen, die fördernd oder hemmend für die aktive und erfolgreiche Mitwirkung von Lehrkräften an Schul- und Unterrichtsentwicklung sind (Schildkamp & Kuiper, 2010; Dederich, 2012). Vorliegende Untersuchungen fokussieren einzelne dieser strukturellen Faktoren und den Zusammenhang mit evidenzbasierten Prozessen. Die Ergebnisse weisen auf den positiven Einfluss von Informationsbereitstellung, Austausch über Evidenzen, interner Kooperation etc. hin (Pant & Thiel, 2012; Wayman et al., 2012; Maier et al., 2012; Breiter & Karbautzki, 2012).

Allerdings zeigt der aktuelle Forschungsstand, dass die positive Wirkung nicht notwendigerweise auf die formalen Strukturbedingungen zurückzuführen ist. Die Wahrnehmung dieser Strukturen auf Seiten der Lehrkräfte erweist sich als relevanter Einflussfaktor (Lingkost & Meister, 2011). So zeigen Huppert und Abs (2008) für Schulevaluationen, dass die Wahrnehmung der Strukturen einen stärkeren Einfluss auf die Durchführung und Umsetzung hat als die tatsächlichen implementierten Strukturen. Zudem weist der internationale Forschungsstand auf eine zentrale Rolle der Schulleitung und v.a. des Führungsstils der Schulleitung (insb. transformationaler Führung) für Schul- und Unterrichtsentwicklung hin (Leithwood & Sun, 2012), auch wenn bislang noch keine eindeutige Evidenz zu Effekten seitens der transformationalen Führung auf die Evidenzorientierung der Lehrkräfte vorliegt (Levin & Datnow, 2012).

Eine Betrachtung der verschiedenen schulischen Faktoren sowie des Schulleitungshandelns und ihres Einflusses auf die Evidenzorientierung der Lehrkräfte steht bisher aus. Der Vortrag setzt an diesem Forschungsdefizit an und untersucht die zwei nachfolgenden Hypothesen:

H1: Die positiv wahrgenommenen schulischen Strukturen haben einen positiven Einfluss auf die Evidenzorientierung der Lehrkräfte.

H2: Ein transformationaler Führungsstil der Schulleitung hat einen positiven Einfluss auf die Evidenzorientierung der Lehrkräfte.

Zur Analyse wird auf Daten aus dem Projekt „Evidenzbasiertes Handeln im schulischen Mehrebenensystem“ (EviS) zurückgegriffen. Im Jahr 2011 und 2012 wurden insgesamt 2.922 Lehrkräfte und 297 Schulleitungen verschiedener Schulformen im Bundesland Rheinland-Pfalz mittels standardisiertem Papierfragebogen befragt. Für die folgenden Analysen wird auf einen Teildatensatz zurückgegriffen, bei dem alle hier betrachteten AVs und UVs erfasst wurden: 1.334 Lehrkräfte und 250 Schulleitungen von insgesamt 119 Schulen.

Die Evidenzorientierung der Lehrkräfte wurde anhand der Dimensionen Externe (EE) und Interne Evidenzorientierung (IE) mittels einer Skala von Stumm, Mohr und Dormann (2010) erhoben. Als Schulstrukturen wurden Informationsbereitstellung und Kommunikation, interne sowie externe Kooperation und Partizipation mit auf den Schulkontext adaptierten Skalen von Stumm et al. (2010) erfasst. Zur Erhebung des transformationalen Führungsstils wurde die deutsche Übersetzung der Skala aus dem Multifactor Leadership Questionnaire (MLQ; Felfe & Gohli, 2002) eingesetzt. Die Befragung erfolgte mit 5-stufigen Likertskalen (1-stimme überhaupt nicht zu, 5-stimme voll zu). Die Faktorreliabilitäten liegen zwischen .69 und .88 (s. Zlatkin-Troitschanskaia et al., 2014; Stump et al., 2014).

Um die verschiedenen Einflussfaktoren auf der Schulleitungs- und Lehrerebene berücksichtigen zu können werden die Hypothesen mit Mehrebenen-Regressionsanalysen in MPlus 6.0 geprüft. Die Ergebnisse unterstützen die Annahme des Einflusses wahrgenommener Schulstrukturen auf die Evidenzorientierung der Lehrkräfte mit einer Varianzaufklärung der abhängigen Variablen IE in Höhe von 46% und EE in Höhe von 32%. Im Mehrebenenmodell konnte ebenso der Einfluss des transformationalen Führungsstils der Schulleitung mit einer Varianzaufklärung in Höhe von 16% auf Schulleitungsebene für die abhängige Variable EE festgestellt werden. Für die Interne Evidenzorientierung der Lehrkräfte kann die Annahme des Einflusses der Schulleitung nicht bestätigt werden.

Donnerstag
12. März

15:30 Uhr

GA 03/49

Chair(s)
Mater, Olga

Semper, Ina; Hermstein, Björn; Manitius, Veronika; Berkemeyer, Nils; Bos, Wilfried

Zur Heterogenität regionaler Schulsysteme in Deutschland - Ausprägungen und Einflussfaktoren

Theoretischer Hintergrund und Fragestellung: Mit dem Aufkommen von Large-Scale-Studien und dem damit einhergehenden gesteigerten Bedeutungszuwachs der empirischen Schul- und Bildungsforschung wurde zugleich ein klassischer Forschungsstrang der Bildungsforschung, regionale Disparitäten im Schul- und Bildungswesen (Geipel 1965; Peisert 1967), reaktiviert. In der jüngeren Vergangenheit haben verschiedene Studien (z.B. Sixt 2010; Terpoorten 2014) instruktiv aufgezeigt, wie das weiterentwickelte forschungsmethodische Repertoire sowie neuartige Datensätze für die Beobachtung und Analyse geografischer Bildungsungleichheiten und – benachteiligungen genutzt werden können. Zudem wird im Bereich der Ursachenklärung divergierender Bildungschancen verstärkt der Einfluss diverser kontextueller Einbettungen von Bildungssystemen untersucht (z.B. Becker & Schulze 2013).

In Referenz zu diesen „Paradigmen“ der empirischen Bildungsforschung sollen mit dem Beitrag Forschungsergebnisse zu strukturellen, prozess- und outputbezogenen Differenzen der deutschen Schulsysteme aus dem Projekt „Chancenspiegel“ (Berkemeyer et. al., im Ersch.) auf Ebene der Kreise und kreisfreien Städte vergleichend vorgestellt und diskutiert werden. Neben einer konzentrierten Einführung in Dimensionen und Ausmaße der Heterogenität von regionalen Schulsystemen sind vier Einzelfragestellungen leitend, die allesamt auf die übergeordnete Frage nach den strukturellen und kontextuellen Bedingungen schulsystemischer Unterschiede verweisen: (a) Inwiefern moderieren divergierende strukturelle Angebotslagen regionaler Schulsysteme im Sekundarbereich das Nutzungsverhalten von Schülern? (b) Wie hängen interregional disparate Angebotslagen mit dem schulwahlrelevanten Faktor der Schulweglänge zusammen? (c) Sind Korrespondenzen zwischen spezifischen sozialräumlichen Kontextbedingungen und den Angebotsstrukturen regionaler Schulsysteme sowie ihrer systembezogenen Leistungen empirisch beobachtbar? (d) Auf welchen Regulierungsebenen von Schulsystemen (Kreis/Stadt und Land) sind Einflüsse für regionale Disparitäten festzustellen?

Der Beitrag schließt nach einer kritischen Reflexion der vorgenommenen Analysen und verwendeten Datensätze sodann mit Überlegungen zu Forschungsperspektiven, wobei insbesondere auf theoretische Ansätze aus der Raumsoziologie (Löw 2001), der Sozialgeographie (Werlen 2013) zurückgegriffen wird. Hierbei spielt auch Frage nach dem spezifischen (normativen und funktionalen) Wert der Heterogenität regionaler Schulsysteme eine Rolle.

Methodisches Vorgehen: Für die Beantwortung jeder Fragestellung werden unterschiedliche Daten der amtlichen Statistiken der Bundesländer genutzt und statistische Auswertungsverfahren angewendet.

Für a) und b) werden auf Basis schul- und wohnortbezogener Daten Pendelbewegungen von Schülern dreier regionaler Schulsysteme deskriptiv nachgezeichnet sowie Einzugsgebiete und Schulweglängen berechnet.

Im Rahmen der Analysen in Bezug auf Frage c) werden mithilfe des Verfahrens der Latent Class Analysis (LCA) sozialräumliche Profilierungen regionaler Schulsysteme ermittelt und mit schulstatistischen Maßzahlen (Übergangsquoten auf das Gymnasium, Klassenwiederholerquote und Absolventen- und Abgängerquoten) in Zusammenhang gebracht.

Der Frage d) wird mit mehrebenenanalytischen Auswertungsverfahren nachgegangen, wobei die gefundenen Varianzen entlang der Indikatoren Klassenwiederholerquote und Quote der Abgänger ohne Hauptschulabschluss sowie der Hochschulreifequoten aus allgemeinbildenden und beruflichen Schulen aufgezeigt werden.

Ergebnisse: Die Ergebnisse der Analysen zu den Fragen a) und b) zeigen, dass aufgrund stark divergierender Schulangebotslagen regionale Schulsysteme als „Einpendlersysteme“ oder „Auspendlersysteme“ zu klassifizieren sind. Zudem werden schulartbezogene Einzugsbereiche und Schulweglängen sichtbar. Dies kann aus Sicht der Chancengerechtigkeit kritisch diskutiert werden, da die Erreichbarkeit von Schulangeboten schichtspezifisch unterschiedlich wahrgenommen wird (Clausen 2006).

Hinsichtlich c) lässt sich sagen, dass auf Basis sozialstatistischer Daten und dem Verfahren der Latent Class Analysis Profilbildungen möglich sind und dass sich diesbezügliche Unterschiede hinsichtlich der Schulangebote und schulstatistischer Kennwerte in Form von Korrespondenzen nachzeichnen lassen. Die Resultate sind Hinweise darauf, dass bestimmte sozialstrukturelle und schulsystemische Aspekte sowie Ausprägungen in Kennwerten, die Chancengerechtigkeit beschreiben, gemeinsam auftreten.

Die Ergebnisse der Mehrebenenanalysen (Frage d) zum Einfluss der beiden Ebenen Land und Kreis/kreisfreie Stadt auf die schulstatistischen Kennwerte zeigen, dass die Länderebene relevant für die Ergebnisse der Regionen ist. Es macht also hinsichtlich der Werteausprägungen durchaus einen Unterschied, welchem Bundesland eine Region zugehörig ist. Daher sollte für zukünftig angestellte Regionalanalysen die Ebene des Landes als Einflussgröße nicht vernachlässigt werden.

Donnerstag
12. März

15:30 Uhr

GA 03/49

Chair(s)
Mater, Olga

3

Tillmann, Katja; Willems, Ariane S.; Rollett, Wolfram

Typische Formen der multiprofessionellen Kooperation an Ganztagsgrundschulen: Eine Latent-Class-Analyse zu den Einschätzungen des weiteren pädagogisch tätigen Personals.

Der Ausbau von Ganztagsgrundschulen in Deutschland hat die Schullandschaft in den letzten Jahren nachhaltig verändert. Die Anforderungen an die moderne Ganztagsgrundschule sind dabei vielfältig und umfassen neben dem Abbau von Bildungsbenachteiligung auch eine individualisierte Förderung von SchülerInnen sowie die Entwicklung einer differenzierten und erweiterten Lernkultur (Aktionsrat Bildung, 2013; Holtappels 2009). Um SchülerInnen mit unterschiedlichen Lernvoraussetzungen optimal unterstützen zu können, bedarf es u.a. einer organisatorischen und inhaltlich-konzeptionellen Verknüpfung der Lerngelegenheiten im Unterricht sowie in den außerunterrichtlichen Angeboten, die im schulischen Alltag von den vor Ort tätigen Personen – den Lehrkräften und den Personen des weiteren pädagogisch tätigen Personals (wptP) – hergestellt werden muss. Die erfolgreiche Zusammenarbeit beider Akteursgruppen in Form einer multiprofessionellen Kooperation stellt eine zentrale Gelingensbedingung erfolgreicher Ganztagsgrundschularbeit dar (Maykus 2009; Rollett & Tillmann 2009; Speck 2010; Steiner & Tillmann 2011). Analog zu den Befunden der Lehrkooperation (Gräsel et al. 2006, Steinert et al. 2006) zeigt sich allerdings auch, dass die multiprofessionelle Kooperation voraussetzungsvoll ist und umso seltener stattfindet, je anspruchsvoller die Kooperationsformen sind (Beher et al., 2007; Rollett & Tillmann 2009).

Die bislang vorliegenden Studien zur multiprofessionellen Kooperation fokussieren zumeist auf einzelne Aspekte der Zusammenarbeit (Intensität, Qualität, Entlastung) um arbeiten Zusammenhänge zu individuellen Merkmalen (Qualifikation und Beschäftigungsumfang) oder Merkmalen der schulischen Prozessqualität heraus (Bergmann & Rollett 2008; Fussangel et al. 2010). Aufbauend hierauf wird in dem vorliegenden Beitrag die Forschungsperspektive erweitert, indem zeitgleich unterschiedliche Teilaspekte der multiprofessionellen Zusammenarbeit in den Blick genommen werden, um typische, individuelle ‚Kooperationsprofile‘ aus Sicht des wptP zu identifizieren. Dabei gehen wir folgenden Fragen nach:

(1) Lassen sich Personen des wptP anhand ihrer Angaben zur multiprofessionellen Zusammenarbeit zu unterschiedlichen ‚Kooperationstypen‘ gruppieren und wie lassen sich diese personenbezogenen Profile beschreiben? (2) Unterscheiden sich diese ‚Kooperationstypen‘ hinsichtlich individueller Hintergrundmerkmale (u.a. Qualifikation und Beschäftigungssituation)? (3) Treten die einzelnen ‚Kooperationstypen‘ an bestimmten Schulen gehäuft auf? Durch welche ganztagestypischen Struktur- und Prozessqualitätsfaktoren (Willems & Becker, in Druck) sind diese Schulen ggfs. gekennzeichnet? Grundlage der Analysen sind die Daten von n=672 Personen des wptP an 69 Ganztagsgrundschulen aus der aktuellen Erhebung der Längsschnittstudie StEG-P (2013/2014). Mittels Latent-Class-Analysen (Muthén & Muthén 2008) werden die Personen anhand ihrer Einschätzungen zu acht unterschiedlichen Aspekten der multiprofessionellen Kooperation in Kooperationstypen eingeteilt. Die Skalen erfassen die folgenden Konstrukte: Stellenwert, Intensität und Kollegialität innerhalb des wptP sowie Intensität, Qualität, Unterrichtsbezug, Schülerbezug und erlebte Entlastung durch die multiprofessionelle Kooperation. Alle Skalen weisen gute Reliabilitäten auf (Alpha .79-.94). Gleichzeitig wurden die Personen differenziert zu ihrem beruflichen und qualifikatorischen Hintergrund, der Erfahrung mit der Arbeit an ganztägig organisierten Schulen sowie zu unterschiedlichen Aspekten von Schulqualität befragt.

Unsere Befunde weisen auf fünf charakteristische Kooperationstypen hin, die sich im Niveau der untersuchten Kooperationsaspekte und in der Struktur der Profile unterscheiden: Zwei Kooperationstypen zeigen unterdurchschnittliche, zwei weitere überdurchschnittliche Profile. Die Personen im ersten Kooperationstyp zeichnen sich durch eine geringe Intensität der Zusammenarbeit sowohl innerhalb des wptP als auch mit den Lehrkräften aus und schätzen dabei auch den Stellenwert von Kooperation vergleichsweise niedrig ein. Im Gegensatz dazu haben Personen des zweiten Kooperationstyps auffällig niedrige Werte hinsichtlich der verschiedenen Aspekte der multiprofessionellen Zusammenarbeit, obwohl sie diese Form der Kooperation als persönlich wichtig erachten. Es lassen sich auch Typen identifizieren, die sowohl innerhalb der Gruppe des Personals als auch mit den Lehrkräften intensiv kooperieren. Besonders auffällig ist hier der große Unterschied in der Ausprägung der unterrichtsbezogenen multiprofessionellen Kooperation. Es zeigt sich, dass in den Kooperationstypen mit unterdurchschnittlichem Kooperationsprofil anteilig weniger und in den Klassen mit überdurchschnittlichem Kooperationsprofil anteilig mehr Personen mit pädagogischer Qualifikation, hoher zeitlicher Einbindung und hauptamtlicher Beschäftigung sind. Im Vortrag werden weitere vertiefende Analysen zur Frage des Zusammenhangs zwischen Klassenzugehörigkeit und schulorganisatorischen und schulkulturellen Rahmenbedingungen vorgestellt und Implikationen für die weitere Forschung und Praxis diskutiert.

Donnerstag
12. März

15:30 Uhr

GA 03/49

Chair(s)
Mater, Olga

Yilmaz, Aylin; Fischer, Claudia

Merkmale wirkungsvollen Schulleitungshandelns im Programm SINUS an Grundschulen

Hintergrund: Aktuelle Studien machen die große Bedeutung der Lehrkräfte für den Unterricht deutlich (Hattie, 2009; Lipowsky, 2010; Lipowsky, im Druck), die im Zuge von veränderten Bedingungen im Schulsystem ihr Wissen und Können erweitern müssen. Der Schulleitung kommt die Aufgabe zu, die Weiterentwicklung der Lehrpersonen zu unterstützen. Empirische Studien lassen erkennen, dass Schulleitungshandeln direkt auf das Wissen und Können, die Motivation und Arbeitsbedingungen der Lehrkräfte wirkt und auch deren Unterrichtspraxis beeinflussen kann (Bonsen et al., 2002; Leithwood & Levin, 2005). In den Verantwortungsbereich der Schulleitung fallen die Entwicklung und Sicherung der Qualität, das Personal- und Ressourcenmanagement sowie die Unterrichtsentwicklung (Harazd et al., 2008). Letztere bildet den Dreh- und Angelpunkt der Schulentwicklung. Scheerens et al. (2003) ermittelten, dass Schulleitungen besonders wirkungsvoll handeln, wenn sie die Lehrkräfte bei der Verbesserung ihres Unterrichts unterstützen. Die von ihnen identifizierten Merkmale wirkungsvollen Schulleitungshandelns wurden für Deutschland bisher noch kaum überprüft. Fischer und Rieck (2013) berichteten Daten aus einer Einmalmessung von 332 Schulleitungen von Grundschulen aus zehn Bundesländern, deren Schulen an einem mehrjährigen Unterrichtsentwicklungsprogramm beteiligt waren. Sie konnten deskriptiv zeigen, dass bei den befragten Schulleitungen Merkmale wirkungsvollen Schulleitungshandelns – teilweise in hoher Ausprägung – vorhanden waren. Der vorliegende Beitrag nutzt Vergleichsdaten aus einer drei Jahre später durchgeführten Erhebung und betrachtet die Entwicklung der Merkmale im Zeitverlauf.

Forschungsfragen: (1) Welche Merkmale wirkungsvollen Schulleitungshandelns finden sich in Grundschulen, die an einem mehrjährigen Unterrichtsentwicklungsprogramm beteiligt sind? (2) Wie entwickeln sich diese Merkmale im Zeitverlauf über mehrere Messzeitpunkte?

Methoden: SINUS an Grundschulen stellt ein Programm zur Professionalisierung von Lehrpersonen im Kontext schulischer und unterrichtlicher Entwicklungsprozesse dar und wurde bundesweit von 2009 bis 2013 mit 5500 Lehrpersonen an 870 Schulen durchgeführt sowie mit mehreren Studien wissenschaftlich begleitet (Fischer et al., 2013). Für den vorliegenden Beitrag wurden Daten aus zwei Gesamtbefragungen aller Schulleitungen genutzt und mit Daten aus zwei Gesamtbefragungen aller Lehrkräfte verknüpft. An den Befragungen beteiligten sich 332 Schulleitungen und 1662 Lehrpersonen im Jahr 2010 und 432 Schulleitungen und 1937 Lehrkräfte im Jahr 2013. Zur Beantwortung der Forschungsfragen wurden im Rahmen der Datenauswertung quantitative mit qualitativen Verfahren kombiniert. Neben der Auswertung der Schulleitungsfragebögen wurden Items der Lehrkräftebefragung hinzugezogen, um einen mehrperspektivischen Einblick in das Schulleitungshandeln zu erhalten.

Ergebnisse: Zum jetzigen Zeitpunkt sind lediglich Daten aus der Erhebung 2010 ausgewertet. Sie lassen erkennen, dass die befragten Leitungen von Grundschulen die Beteiligung an einem innovativen Programm in ein Gesamtkonzept der Schule zur Unterrichtsentwicklung einbanden. Sie lenkten die Aufmerksamkeit in der Schule insbesondere auf die Entwicklung des Unterrichts. Darüber hinaus waren sie mit ihrer fachlichen und unterrichtlichen Expertise anerkannt, die sie einerseits benötigten, um Lehrpersonen zu beraten und zu unterstützen und andererseits nutzten, um eine Verbesserung des Fachunterrichts zu initiieren. Die Ergebnisse zeigen ferner, dass ein Großteil der Schulleitungen die Weiterbildung der Lehrpersonen wirksam unterstützte. Im Zuge der Professionalisierung der Lehrpersonen förderten sie u.a. die kollegiale Kooperation an ihrer Schule und setzten dafür besondere Mittel zur Unterstützung der Zusammenarbeit ein. Die Überprüfung der eingesetzten Maßnahmen an der Schule wurde von einem Großteil der Schulleitungen im Rahmen einer Fremd- und/oder Selbstevaluation durchgeführt. In geringerem Umfang nutzten sie auch Feedback aus Evaluationsergebnissen zu den Leistungen der Lernenden für die Unterrichts- und Schulentwicklung. Die Befragung der Lehrpersonen bestätigte diese Befunde. Der Konferenzbeitrag vergleicht die Ergebnisse der ersten Erhebungen von 2010 mit den Befunden aus den Befragungen von 2013. Ein solcher Vergleich soll zeigen, welche Veränderungen und Entwicklungsschwerpunkte sich im Vergleich der beiden Messzeitpunkte ausmachen lassen.

Donnerstag
12. März

15:30 Uhr

GA 03/49

Chair(s)
Mater, Olga

3

Pietsch, Marcus

Der Einfluss transformationaler und instruktionaler Führung auf die Unterrichtsgestaltung von Lehrkräften

Theoretischer Hintergrund: In ihrer Studie zum Einfluss von Schulleitungen auf Schülerleistungen konnten Heck und Hallinger (1998) zeigen, dass die Wirkung in der Regel indirekt erzeugt wird. Neuere Studien systematisieren diese Befundlage und zeigen, dass wirksame Schulleitungen ihre Führungsschwerpunkte grundsätzlich auf vier Bereiche legen (vgl. Leithwood & Jantzi, 2008, Scheerens et al., 2012) : 1) den Schulbeteiligten Wege und Ziele vorgeben, 2) Mitarbeiterinnen und Mitarbeiter (weiter) entwickeln, 3) die Schule (neu) gestalten und 4) das Lernen und Lehren an der Schule aktiv steuern. Diesbezüglich stellt sich dann die Frage, welches Konzept von Führung, welcher Führungsstil hilfreich ist, um auf diesen Wegen konkret wirksam zu werden. Die Schuleffektivitätsforschung unterscheidet hier zwei Führungskonzepte voneinander: instruktionale und transformationale Führung (vgl. Hallinger, 2003). Instruktionale Führung erfolgt vor allem aufgaben- und produktorientiert, zielt auf die Optimierung vorhandener Strukturen und Prozesse und entfaltet eine Wirkung erster Ordnung. Transformationale Führung hingegen erfolgt in erster Linie mitarbeiterorientiert, zielt auf die nachhaltige Veränderung der schulischen Lern- und Arbeitskultur ab und entfaltet eine Wirkung zweiter Ordnung. Erstaunlicherweise untersuchen bislang jedoch nur wenige Studien den simultanen Einfluss dieser Führungsstile auf die vermittelnden Faktoren und die Schülerleistung (vgl. Shatzer et al., 2013) und selbst wenn, werden in der Regel keine standardisierten Instrumente zur Messung der Führungskonstrukte eingesetzt und vermittelnde Faktoren in den Analysen nicht ausreichend berücksichtigt (vgl. Scheerens et al., 2012). Diejenigen Studien jedoch, die ein solches Design nutzen, demonstrieren, dass Schulleitungen, die integrativ instruktional-transformationale führen, einen besonders hohen Einfluss auf die Lernentwicklung der Schülerinnen und Schüler an ihren Schulen ausüben (vgl. Marks & Printy, 2003).

Fragestellung: Leithwood et al. (2002, 2008) haben ein Modell vorgelegt, das unterstellt, dass wirksames Schulleitungshandeln vor allem das Ziel hat, den Unterricht an der Schule zu optimieren und zu verbessern. Und dies, so die Modellannahme, geschieht vor allem, indem Schulleitungen die Motivation und Identifikation von Lehrkräften mit ihrer Schule, deren (subjektiv erlebte) Arbeitsbedingungen sowie deren pädagogische Fähigkeiten und ihr Wissen um wirksames Lernen und Lehren beeinflussen. Die vorliegende Studie zielt darauf a) zu klären, in wie weit das vorgeschlagene Pfadmodell empirisch haltbar ist und b) zu evaluieren, in wie weit sich, jenseits eines g-Faktors, Effekte transformationaler und instruktionaler Führung auf die Gestaltung des Unterrichts durch Lehrkräfte nachweisen lassen.

Methode: Im Rahmen der Studie wurden in Vollerhebungen Daten von Lehrkräften an 50 Hamburger Schulen mithilfe von Fragebögen erfasst. Insgesamt liegen Angaben von N=1479 Lehrerinnen und Lehrern vor. Transformationale Führung wurde mithilfe einer Kurzform des Multifaceted Leadership Questionnaire (MLQ, vgl. Bass & Avolio, 1995), instruktionale Führung mithilfe von Skalen aus dem Teaching and Learning International Survey (TALIS, vgl. OECD, 2009) gemessen. Die Daten wurden mithilfe von Strukturgleichungsmodellen in MPLUS analysiert. Hierbei wurden Vorschläge von Anderson und Gerbig (1988) aufgegriffen und eine sequentielle Analyseprozedur genutzt: In einem ersten Schritt wurden die Messmodelle der latenten Variablen spezifiziert und in einem in einem zweiten, saturierten Modell fixiert. In diesem Zusammenhang wurde Führung als Bi-Faktor-Modell, bestehend aus einem General- sowie den zwei spezifischen Führungsfaktoren spezifiziert. Entsprechend ist es möglich den Einfluss der spezifischen Komponenten jenseits des Generalfaktors zu untersuchen (vgl. Chen, 2012). In einem zweiten Schritt wurde das theoriebasierte (composite) Modell inspiziert. Dieses Vorgehen erlaubte es den globalen Gesamtfit des Modells in eine Struktur- und eine Messmodellkomponente zu unterteilen (vgl. O'Boyle & Williams, 2011). Mithilfe verschiedener Angaben aus beiden Berechnungen ließ sich so der isolierte RMSEA-P des Pfadmodells ermitteln.

Ergebnisse: Das theoretische Modell hat einen sehr guten Fit (TLI=0.918, RMSEA=0.031, $\chi^2=5395$ bei $df=2298$), ebenso das Pfadmodell (RMSEA-P=0.014). Direkte Effekte auf die Unterrichtsgestaltung lassen sich nur für die instruktionale Führung nachweisen. Unter Kontrolle des Generalfaktors übertreffen die Effekte instruktionaler diejenigen transformationaler Führung auf vermittelnde Faktoren bei Weitem, wobei jedoch beide Führungsstile unabhängig voneinander einen Einfluss ausüben.

Donnerstag
12. März

15:30 Uhr

GA 03/49

Chair(s)
Mater, Olga

3

Bruckmaier, Georg; Krauss, Stefan; Blum, Werner; Leiss, Dominik

COACTIV-Video: Fachdidaktische Kompetenzen von Mathematiklehrkräften

Ein wichtiges Forschungsziel der empirischen Bildungswissenschaft ist die Erfassung fachdidaktischen Wissens von Lehrkräften. Im vorliegenden Beitrag wird mit „COACTIV-Video“ ein computergestütztes Instrument vorgestellt, mit dessen Hilfe eine handlungsnahe Erfassung des fachdidaktischen Wissens im Bereich der Mathematik umgesetzt wurde.

Eine Reihe von Forschungsprojekten haben für verschiedene Unterrichtsfächer, vor allem jedoch für Mathematik und die Naturwissenschaften, bereits gezeigt, dass videobasierte Instrumente aussagekräftige Informationen über den Umgang von Lehrkräften mit den Anforderungen des Fachunterrichts liefern können (u.a. Klieme et al., 2006; Klieme & Reusser, 2003; Lindmeier, 2013; Seidel et al., 2006). Mit videogestützten Instrumenten kann die Kritik aufgegriffen werden, dass Papier-und-Bleistift-Tests des fachdidaktischen Wissens nur verhältnismäßig geringe ökologische Validität aufweisen, da sie viel kontextuelle Information zum Unterrichtsgeschehen vernachlässigen (Blömeke, Kaiser & Lehmann, 2010; Buchholtz, Kaiser & Blömeke, 2014; Krauss et al., 2011).

Der vorliegende Beitrag entstand im Rahmen des interdisziplinären Forschungsprojekts COACTIV (vgl. Kunter et al., 2011). In diesem Projekt werden Genese, Struktur und Handlungsrelevanz von verschiedenen Facetten der professionellen Kompetenz von Mathematiklehrkräften untersucht. Das Projekt war konzeptuell und organisatorisch an die PISA-Studie 2003/2004 angedockt und nutzt sowohl Schüler- als auch Lehrerdaten aus PISA. Das zugrundeliegende Wirkungsmodell geht davon aus, dass professionelle Lehrerkompetenzen einen qualitativ hochwertigen Unterricht bewirken und so zu Lernfortschritten der Schülerinnen und Schüler beitragen (Baumert et al., 2011).

Mithilfe eines computergestützten Instruments wurde gemessen, inwiefern Lehrkräfte in der Lage sind, den Unterricht in „didaktisch kritischen“ Situationen adäquat weiterzugestalten. Den Lehrkräften wurden dazu drei kurze Videovignetten zu Unterrichtsszenen (verschiedener mathematischer Inhalte der 9. Jahrgangsstufe) gezeigt, an deren Ende sie um didaktisch geeignete Unterrichtsfortführungen gebeten wurden. Die Antworten der Lehrkräfte wurden von zwei Ratern auf insgesamt fünf Dimensionen („Schülerorientierung“, „methodische Orientierung“, „Verständnisorientierung“, „fachliche Präzision“ und „Ergreifen der didaktischen Chance“) jeweils dreistufig kodiert. Aus den Ratings der Lehrerantworten zu den drei Videos wurde ein Summenwert als Maß für die „situative Unterrichtskompetenz“ der jeweiligen Lehrkraft gebildet.

Mit dem dargestellten Instrument wurde unter anderem die Struktur situativer Unterrichtskompetenz untersucht (Modellpassung). Weiterhin wurden Schulformunterschiede und Zusammenhänge mit Drittvariablen analysiert (Validierung). Schließlich wurde der Einfluss situativer Unterrichtskompetenz auf die Unterrichtsqualität untersucht (prädiktive Validität).

Zur Beantwortung der Fragestellungen stand eine Stichprobe von 280 Mathematiklehrkräften verschiedener Sekundarschulformen (HS, RS, Gym) zur Verfügung.

In einer konfirmatorischen Faktorenanalyse bestätigten sich die beiden theoretisch postulierten Subfacetten „methodische Kompetenz“ und „fachspezifische Kompetenz“. Hinsichtlich der betrachteten Schulformen der beteiligten Lehrkräfte zeigte sich ein Anstieg der situativen Unterrichtskompetenz von den Hauptschul- über die Realschul- hin zu den Gymnasiallehrkräften, wobei die Schulformunterschiede bei den fachspezifischen Dimensionen besonders ausgeprägt waren. Wie erwartet bestanden signifikante Zusammenhänge von situativer Unterrichtskompetenz sowohl mit fachdidaktischem Wissen als auch mit Fachwissen, wobei die Korrelationen beider Wissensarten mit der fachspezifischen Komponente situativer Unterrichtskompetenz deutlich höher ausfielen als mit der methodischen Facette. Auch die Zusammenhänge mit lerntheoretischen Überzeugungen wiesen in die erwartete Richtung (signifikante positive Korrelationen mit konstruktivistischen Überzeugungen).

Schließlich wurde untersucht, ob die gemessene Kompetenz ein erklärungsstärkerer Prädiktor der Unterrichtsqualität ist. Dabei stellte sich heraus, dass die fachspezifische, nicht jedoch die methodische Facette situativer Unterrichtskompetenz das Ausmaß kognitiver Aktivierung signifikant vorhersagt. Mit den beiden Dimensionen der Unterrichtsqualität „Klassenführung“ und „konstruktive Unterstützung“ ergaben sich für beide Kompetenzfacetten keine signifikanten Zusammenhänge.

Donnerstag
12. März

15:30 Uhr

GA 03/142

Chair(s)

Bruckmaier, Georg

Sunder, Cornelia; Todorova, Maria; Möller, Kornelia

Lohnt sich der Einsatz von Videos zur Schulung der professionellen Wahrnehmung von Unterricht? – Eine Untersuchung zur Ausbildung von Studierenden des Faches Sachunterricht

Die professionelle Unterrichtswahrnehmung beinhaltet das Erkennen und das theoriegeleitete Interpretieren lernrelevanter Unterrichtsereignisse (Sherin, 2007). Sie gilt als zentral für das Handeln im Unterricht sowie für die Unterrichtsreflexion (Bromme & Haag, 2008; Hammerness, Darling-Hammond & Shulman, 2002) und weist positive Zusammenhänge zum Handeln der Lehrkraft (Hamre et al., 2012) und zum Lernen der Schüler (Kersting et al., 2012, Roth et al., 2011) auf. Konzeptuell kann die professionelle Wahrnehmung als ein wissensbasierter Prozess aufgefasst werden (Stürmer, Könings & Seidel, 2013), der sich auf verschiedene Dimensionen des Unterrichts richten kann. Dabei kann zwischen einem allgemeinpädagogischen Fokus und einem bereichsspezifischen Fokus der Wahrnehmung unterschieden werden (Blomberg, Stürmer & Seidel, 2011).

Zur Förderung der professionellen Wahrnehmung werden in der Lehreraus- und -weiterbildung häufig Videos eingesetzt (Blomberg et al., 2013). So konnten erste Studien zeigen, dass die professionelle Wahrnehmung von Studierenden durch Videoanalysen gefördert werden kann; sowohl im allgemeinpädagogischen Bereich (z.B. Gold, Förster & Holodyski, 2013) als auch bereichsspezifisch (z.B. der Mathematik: Santagata & Guarino, 2011). Neben diesen empirischen Ergebnissen wird dem Einsatz von Videos auch von Studierenden eine positive Bedeutung beigemessen: Sie bewerten Unterrichtsanalysen auf Basis von Videos als nützlicher und präziser als Analysen, die rein auf Erinnerungen basieren (Rosaen et al., 2008). Im spezifischen Bereich des naturwissenschaftlichen Grundschulunterrichts fehlen bisher kontrollierte Studien, in denen der Nutzen von Videos zur Förderung der professionellen Wahrnehmung untersucht wird. Dies ist Gegenstand des vorliegenden Beitrags. Im Projekt „VIU: Early Science“ wird aus fachdidaktischer Sicht die Dimension „Lernunterstützung im naturwissenschaftlichen Grundschulunterricht“ betrachtet. Unter Lernunterstützung werden Maßnahmen der Lehrperson verstanden, mit denen sie die Lernprozesse der Schülerinnen und Schüler anregt und fördert. Konkret wird untersucht, inwiefern die professionelle Wahrnehmung lernunterstützender Maßnahmen im naturwissenschaftlichen Grundschulunterricht durch videobasierte Lehr-Lern-Umgebungen im Bachelor-Studium gefördert werden kann. Hierzu wird ein Seminar, in dem fallbezogene Unterrichtsanalysen mit Videos durchgeführt werden (Gruppe mit Video, n=32), mit einem Seminar verglichen, in dem Unterricht fallbezogenen, aber ohne Videos – anhand von Texten – analysiert wird (Gruppe ohne Video, n=36). Im Rahmen beider Seminare lernen die Studierenden dieselben theoretische Inhalte zu lernunterstützenden Maßnahmen und analysieren den Einsatz dieser Maßnahmen an Unterrichtsbeispielen zum Thema „Schwimmen und Sinken“. Die Erfassung der professionellen Wahrnehmung erfolgte mit einem standardisierten Videotest (Wolters et al., eing.), der Videos zu den Unterrichtsthemen „Schwimmen und Sinken“ und „Aggregatzustände“ beinhaltet. Neben der professionellen Wahrnehmung wurden das Fachwissen und das fachdidaktische Wissen der Studierenden zu den in den Videos unterrichteten Themen erhoben. Alle oben genannten Variablen wurden dabei im prä-post-Design erfasst. Zusätzlich wurde gemessen, ob sich die Studierenden der beiden Interventionsgruppen nach dem Seminar in ihren motivationalen Orientierungen in Bezug auf das Erkennen bzw. Einsetzen lernunterstützender Maßnahmen im naturwissenschaftlichen Sachunterricht und in Bezug auf ihre Seminarbewertung unterscheiden. Die Studie wird mit zwei Kohorten im Wintersemester 13/14 sowie 14/15 durchgeführt (N= ca. 150 Studierende).

Erste varianzanalytische Berechnungen auf Basis der Daten der 1. Kohorte zeigen, dass es keine Unterschiede zwischen den Gruppen im Lernzuwachs in der professionellen Wahrnehmung von lernunterstützenden Maßnahmen allgemein gibt ($F(1,65) = 0.943$, $p = .335$, $\eta_p^2 = .014$). Unterschiede zwischen den Gruppen zeigen sich aber bei Betrachtung des Lernzuwachses der professionellen Wahrnehmung in Bezug auf die einzelnen Themen: So unterscheidet sich die „Gruppe mit Video“ in ihrem Lernzuwachs der professionellen Wahrnehmung bei dem im Seminar instruierten Thema „Schwimmen und Sinken“ positiv von der „Gruppe ohne Video“ ($F(1,65) = 4.393$, $p = .04$, $\eta_p^2 = .063$), während es beim nicht behandelten Thema „Aggregatzustände“ keine Gruppenunterschiede im Lernzuwachs gibt ($F(1,65) = .000$, $p = .994$, $\eta_p^2 = .000$). Dies lässt folgern, dass die professionelle Wahrnehmung lernunterstützender Maßnahmen möglicherweise eine hohe themenbezogene Bereichsspezifität aufweist.

Im Vortrag werden das Interventionskonzept sowie weiterführende Ergebnisse berichtet und diskutiert.

Donnerstag
12. März

15:30 Uhr

GA 03/142

Chair(s)

Bruckmaier, Georg

Etzel, Julian; Nagy, Gabriel

Die Bedeutung der Big Five Persönlichkeitstraits und der wahrgenommenen Person-Umwelt-Kongruenz für die Vorhersage des Studienerfolgs

Die seit längerer Zeit stetig steigende Zahl von Studienanfängern und vorzeitigen Studienabbrechern an deutschen Hochschulen führt zu einer verstärkten Untersuchung der Bedingungsfaktoren des Studienerfolgs. Neben kognitiven Faktoren rücken dabei motivationale und Persönlichkeitsfaktoren immer stärker in den Vordergrund. Als prominentes Rahmenmodell der menschlichen Persönlichkeit haben sich die Big Five Persönlichkeitstraits – Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit – als zuverlässige Prädiktoren für Erfolgsmarker in Studium und Beruf herausgestellt (z.B. Judge, Heller & Mount, 2002, Poropat, 2009). In den vergangenen Jahren wurden Facetten der subjektiv wahrgenommenen Person-Umwelt-Kongruenz (PUK) als leistungsstarke proximale Prädiktoren des beruflichen und akademischen Erfolgs identifiziert (z.B., Cable & DeRue, 2002; Li, Yao, Chen & Wang, 2013; Schmitt, Oswald, Friede, Imus & Merritt, 2008). Li et al. (2013) führten ein 3-Faktoren Modell der wahrgenommenen PUK von Studierenden ein, in dem zwischen wahrgenommener Interessen-, Bedürfnis- und Fähigkeitenkongruenz unterschieden wird. Die Autoren erbrachten Belege, dass die verschiedenen PUK-Typen besonders stark mit jeweils unterschiedlichen Studienerfolgsmarkern zusammenhängen: Die Interessenkongruenz mit der Abbruchsintention, die Bedürfniskongruenz mit der Studienzufriedenheit und die Fähigkeitenkongruenz mit den Studienleistungen. Bislang fehlt es jedoch an einer Replikation der Befunde im deutschsprachigen Raum sowie an einer simultanen Betrachtung der subjektiven PUK-Typen und der Big Five Persönlichkeitsfaktoren.

Gegenstand des vorliegenden Beitrags ist die Untersuchung der Bedeutung der Big Five Persönlichkeitsfaktoren sowie der subjektiv wahrgenommenen PUK für die Vorhersage verschiedener Aspekte Studienerfolgs (Studienzufriedenheit, Studienabbruchsintentionen und Studienleistung) an einer Stichprobe deutscher Studenten. Die vorliegende Studie untersuchte dazu (1) die Struktur der wahrgenommenen PUK, (2) die Zusammenhänge zwischen verschiedenen PUK-Typen, den Big Five und den Studienerfolgsmarkern, (3) die Vorhersagekraft der PUK-Typen und der Big Five Persönlichkeitsfaktoren. Die Datengrundlage bildete eine Online-Erhebung aus dem Jahr 2013. Die Analysestichprobe bestand aus $N = 326$ Studenten (Durchschnittsalter 22.43 (SD = 3.28), 68.1% weiblich) aus einer Vielzahl verschiedener Studiengänge.

Die Big Five wurden durch das Big Five Inventory (BFI; Rammstedt, 1999) erfasst. Wahrgenommene PUK wurde zum einen mittels der „academic fit“ Skala von Schmitt et al. (2008), zum anderen mit zwei modifizierten Skalen von Cable & DeRue (2002) erhoben. Zur Messung Studienzufriedenheit wurde eine Skala von Westermann, Heise, Spies und Trautwein (1996), eingesetzt. Die Studienabbruchsintentionen und die subjektiv eingeschätzte Studienleistung wurden mittels zwei Skalen aus der Arbeit von Nagy (2006) erfasst.

Die Ergebnisse einer Sequenz von konfirmatorischen Faktorenanalysen unterstützten das 3 Faktoren Modell der PUK. Es zeigten sich insgesamt mäßige latente Zusammenhänge zwischen den PUK-Typen und den Big Five. Lediglich Gewissenhaftigkeit korrelierte signifikant positiv mit wahrgenommener Interessen- ($r = .27$) und Fähigkeitenkongruenz ($r = .24$). Offenheit für Erfahrungen war schwach mit Interessenkongruenz assoziiert ($r = .16$). Von den Big Five Faktoren erwies sich vor allem der Faktor Gewissenhaftigkeit als konsistent mit allen Erfolgsmarkern korreliert und alle drei PUK-Typen waren substantiell mit allen Aspekten des Studienerfolgs assoziiert. Auf Basis der latenten Korrelationsmatrix wurden relative weight Analysen (Johnson, 2000) durchgeführt, welche die Interessenkongruenz als besten Prädiktor für Studienzufriedenheit ($\eta^2 \approx .37$) und Abbruchsintention ($\eta^2 \approx .22$) auswiesen. Die Fähigkeitenkongruenz erwies sich als bester Prädiktor für selbsteingeschätzte Studienleistung ($\eta^2 \approx .21$). Schließlich zeigte sich mit Hilfe hierarchischer Regressionsmodelle, dass sich die Vorhersage aller drei Erfolgsmarker mit Hilfe von wahrgenommener PUK durch die Hinzunahme von Persönlichkeitsfaktoren nicht weiter verbessern ließ.

Die ermittelten Befunde dokumentieren, dass die untersuchten PUK-Typen als proximale Prädiktoren des Studienerfolgs verstanden werden können, während Big Five Persönlichkeitstraits eher distale Erfolgsprediktoren darstellen, deren Effekte möglicherweise durch die PUK mediiert werden.

Donnerstag
12. März

15:30 Uhr

GA 03/142

Chair(s)

Bruckmaier, Georg

Kaub, Kathrin; Karbach, Julia; Spinath, Frank M.; Brünken, Roland

Interessen und Anforderungen im Lehrberuf – Vergleiche zwischen Studierenden, Referendaren und erfahrenen Lehrpersonen

Die Passung angehender Lehrer zum Anforderungsprofil des Lehrberufs wird als eine zentrale Voraussetzung für berufliches Engagement, Berufszufriedenheit und die Verweildauer im Job angesehen (Kaub et al., 2014; Kunter, Kleickmann, Klusmann, & Richter, 2011; Nye, Su, Rounds, & Drasgow, 2012; Swanson, 2012). Die meisten Forschungsarbeiten im Bereich beruflicher Interessenorientierungen bei Lehrern beziehen sich auf das Berufs-Interessen-Modell von John Holland (Kaub et al., 2012; Klusmann et al., 2009; Mayr, 2011; Swanson, 2012). Nach Holland (1997) lassen sich berufliche Interessen und Umwelten anhand von sechs Dimensionen beschreiben: praktisch-technisch (realistic), intellektuell-forschend (investigative), künstlerisch-sprachlich (artistic), sozial (social), unternehmerisch (enterprising) und konventionell (conventional). Vorgehende Arbeiten (Gottfredson & Holland, 1996; Kaub et al., 2014; Mayr, 2011; Swanson, 2012) konnten zeigen, dass sich Lehramtsstudierende insbesondere durch hohe Interessen in sozialen und führungsbezogenen Tätigkeitsbereichen auszeichnen und – je nach gewähltem Fachschwerpunkt – unterschiedlich stark ausgeprägte künstlerisch-sprachliche (Lehramt für Geistes- und Sprachwissenschaften) oder intellektuell forschende Interessen (Lehramt für Mathe und Naturwissenschaften) zeigen. Weniger Klarheit besteht im Bereich der Analyse von spezifischen Tätigkeits- und Anforderungsfeldern für das Lehramt. Kaub et al. (2014) fanden auf der Grundlage von Expertenbefragungen (Dozenten- und Professoren aus unterschiedlichen Fachbereichen der universitären Lehramtsausbildung), dass die drei wichtigsten Anforderungsfelder für den ersten, akademischen Teil der Lehramtsausbildung insbesondere durch intellektuell-forschende, soziale und konventionelle (ordnend-verwaltende) Tätigkeitsbereiche beschrieben werden können. Eine weitere Frage im Bereich der Interessenforschung bei Lehrern beschäftigt sich mit Entwicklungs- und Veränderungsprozessen (Abel, 2011; Hanfstingl, 2008; Mayr, 2011; vgl. auch Tracey & Sodano, 2008). Obwohl weitestgehend davon ausgegangen werden kann, dass berufliche Interessen ab der späten Adoleszenz stabil bleiben (für einen Überblick siehe Tracy & Sodano, 2008), gibt es dennoch empirische Hinweise auf substantielle Veränderungsprozesse (Low & Rounds, 2007; Swanson, 1999; Swanson & Hanson, 1988). In Anlehnung an stärker prozessorientierte, dynamische Entwicklungsmodelle wird davon ausgegangen, dass Interessen nicht notwendigerweise stabile Persönlichkeitsmerkmale sind, sondern sich durch akademische und berufliche Ziele und Erfahrungen auch entscheidend verändern und anpassen können (Hanfstingl, 2008; Hirschi, 2010; Roberts, Caspi, & Moffitt, 2003).

Das Ziel des vorliegenden Forschungsvorhabens ist (1) die Analyse fachspezifischer Tätigkeits- und Anforderungsprofile für die Berufspraxis von Lehrern und (2) die Analyse von Interessenorientierungen und Passungsmaßen von Lehramtsstudierenden (N=149) im Vergleich zu Referendaren (N=45) und Lehrern im Beruf (N=72). Die Daten sind Teil des BMBF-Projekts SioS-L (Studie zu individuellen und organisationalen Einflüssen auf den Studienerfolg in der Lehrerbildung). Die befragten Lehrpersonen verfügen über langjährige Berufserfahrung und sind aktuell im Rahmen der Referendarsausbildung als Seminar- oder Fachleiter tätig. Die Befragung der Seminar- und Fachleiter zum Anforderungsprofil im Lehrberuf erfolgte mit dem Umwelttest von Stoll & Spinath (2013; leichte Adaption des UST-R nach Bergmann & Eder, 2005; Cronbach's Alpha = .76 – .86). Zur Erfassung der beruflichen Interessenorientierungen wurde der Allgemeine-Interessen-Struktur Test (AIST-R; Bergmann & Eder, 2005) eingesetzt (Cronbach's Alpha = .82 – .88).

Die Ergebnisse zeigen, dass die erfahrenen Lehrpersonen die wichtigsten Tätigkeiten und Anforderungen des Lehrberufs hauptsächlich im sozialen Bereich verorten. Damit verändert sich der Hauptschwerpunkt der Anforderungsbereiche zwischen Universität und Lehrberuf. Gleichzeitig zeigen die Daten, dass der intellektuell-forschende Bereich für die praktische Arbeit als Lehrperson ebenfalls als essentiell eingeschätzt wird. Zusätzlich ergeben sich in den drei höchsten Expertenratings Fachspezifika für künstlerisch-sprachliche und konventionelle Anforderungen. Die Gruppenvergleiche zeigen Unterschiede in den Interessenorientierungen und Passungsmaßen zwischen Lehramtsanwärtern, die am Anfang ihrer beruflichen Karriere stehen (Lehramtsstudierende im 6./8. Semester und Referendare) und Lehrpersonen mit langjähriger Berufserfahrung. Die Ergebnisse sind ein erster Hinweis auf interessenbezogene Veränderungs- und Anpassungsprozesse im Zuge beruflicher Praxiserfahrungen – insbesondere bei mathematisch-naturwissenschaftlichen Lehrern.

Die Nutzung der Ergebnisse für Studien- und Berufsberatungszwecke wird vor dem Hintergrund unterschiedlicher akademischer und beruflicher Anforderungsschwerpunkte sowie vor dem Hintergrund von Selektions- und Veränderungsprozessen kritisch diskutiert.

Donnerstag
12. März

15:30 Uhr

GA 03/142

Chair(s)

Bruckmaier, Georg

Rieser, Svenja; Naumann, Alexander; Decristan, Jasmin; Fauth, Benjamin; Klieme, Eckhard; Büttner, Gerhard

Kompensation mangelnden Vorwissens durch Instruktion metakognitiver Strategien?

Theoretischer Hintergrund: Das Vorwissen von Lernenden hat sich in verschiedenen Studien immer wieder als wichtige Voraussetzung für den Lernerfolg erwiesen (z.B. Pintrich & Zusho, 2002; Winne 1995). Dieser Zusammenhang kann zumindest teilweise durch Unterschiede im Einsatz metakognitiver Strategien erklärt werden. Metakognitive Strategien sind mental repräsentierte Handlungssequenzen um Lernprozesse zu planen, zu überwachen und zu evaluieren (Friedrich & Mandl, 1992). Da metakognitive Strategien inhaltspezifisch angepasst werden müssen, ist ihr Einsatz durch Lernende mit Vorwissen in einem Themenbereich meist effizienter. Sie haben die nötigen Anpassungen bereits vorgenommen und können sich daher besser auf die Inhalte fokussieren (Veenman, Hout-Wolters & Afflerbach, 2006). Eine potentielle Möglichkeit, den aus niedrigem Vorwissen entstehenden Nachteil zumindest teilweise auszugleichen, besteht in der Instruktion metakognitiver Strategien durch die Lehrperson (z.B. Boekaerts, 1997). Bei der Instruktion metakognitiver Strategien bezieht die Lehrperson die Schülerinnen und Schüler aktiv ein, vermittelt ihnen Wissen über Strategien, ihre Einsatzmöglichkeiten sowie ihren Nutzen und unterstützt die Schülerinnen und Schülern beim praktischen Einsatz der Strategien (z.B. Pressley & Harris, 2006).

Fragestellung und Hypothese: Während das Zusammenspiel von Strategieeinsatz, Vorwissen und Lernerfolg empirisch gut belegt ist, mangelt es noch an Belegen für einen kompensatorischen Effekt der Instruktion metakognitiver Strategien bei niedrigem Vorwissen. Daher soll untersucht werden, ob die Instruktion metakognitiver Strategien den positiven Zusammenhang zwischen Vorwissen und Lernerfolg moderiert. Wir erwarten eine negative Moderation des Zusammenhangs durch die Instruktion metakognitiver Strategien, d.h. einen zumindest teilweisen Ausgleich eines durch niedriges Vorwissen entstehenden Nachteils für den Lernerfolg.

Methode: Zur Überprüfung der Hypothese wurden das Vorwissen und der Lernerfolg von 573 Drittklässlern (49% weiblich, MAlter: 8.8 Jahre (SD = 0.50)) aus 36 Klassen vor bzw. nach einer vorstrukturierten Unterrichtseinheit zum Schwimmen und Sinken getestet. Die Tests wurden aus früheren Studien adaptiert (Hardy, Jonen, Möller & Stern, 2006; Kleickmann et al., 2010) und um selbstentwickelte Items ergänzt. Der Pretest umfasste 16 Items (EAP-Reliabilität = .52), der Posttest 13 Items (EAP-Reliabilität = .70). Zusätzlich wurde die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler mittels eines zwölf Items umfassenden Tests vor der Unterrichtseinheit erfasst, der aus TIMSS 2007 adaptiert wurde (Martin, Mullis & Foy, 2008; EAP-Reliabilität = .70).

Zur Erfassung der Instruktion metakognitiver Strategien durch die Lehrperson wurde jeweils eine Unterrichtsstunde (ca. 90min) pro Klasse videographiert und mittels eines hochinferenten Ratings ausgewertet. Das Rating umfasste drei Items (Unterstützung bei der Planung, der Überwachung und der Evaluation des Lernprozesses), die auf einer vierstufigen Skala von zwei unabhängigen geschulten Ratern eingeschätzt wurden. Die Übereinstimmung zwischen den Ratern war zufriedenstellend (> .76%). Der Mittelwert dieser drei Items wurde als Indikator für die Strategieinstruktion der Lehrperson genutzt.

Zur Überprüfung unserer Hypothese wurden Mehrebenen-Regressionsanalysen verwendet (Raudenbush & Bryk, 2010). Auf individueller Ebene wurde das Vorwissen (am Gruppenmittelwert zentriert) als Prädiktor für den Lernerfolg eingesetzt und für die naturwissenschaftliche Kompetenz (am Stichprobenmittelwert zentriert) kontrolliert. Auf Klassenebene wurde die Instruktion metakognitiver Strategien als Prädiktor für den Lernerfolg sowie für den Zusammenhang zwischen Vorwissen und Lernerfolg eingesetzt. Alle Analysen wurden in MPlus 7 unter Verwendung von MLR Schätzern durchgeführt (Muthén & Muthén, 1998-2011).

Ergebnisse: Wie erwartet wurde ein positiver Zusammenhang zwischen dem Vorwissen und dem Lernerfolg identifiziert ($\beta = .24$, $SE = .04$, $p < .001$). Die Strategieinstruktion und der Lernerfolg hingen nicht bedeutsam zusammen. Übereinstimmend mit unserer Hypothese wurde der Zusammenhang zwischen Vorwissen und Lernerfolg negativ durch die Instruktion metakognitiver Strategien moderiert ($\beta = -.08$, $SE = 0.03$, $p < .01$). Das bedeutet, dass in Klassen mit einer intensiveren Instruktion metakognitiver Strategien das Vorwissen weniger bedeutsam für den Lernerfolg war, als in Klassen mit weniger intensiver Strategieinstruktion. Folglich scheinen Lehrpersonen durch eine Intensivierung der Instruktion metakognitiver Strategien besonders Schülerinnen und Schülern mit einer ungünstigen Lernausgangslage unterstützen zu können.

Donnerstag
12. März

15:30 Uhr

GA 03/149

Chair(s)

Rieser, Svenja

Roelle, Julian; Nowitzki, Christine; Berthold, Kirsten

Förderung von Lernstrategien selbstregulierten Lernens in Lernprotokollen: Die Sequenz von kognitiven und metakognitiven Prompts spielt eine Rolle

Theoretischer Hintergrund: Das Schreiben von Lernprotokollen im Anschluss an Schulunterricht ist eine vielversprechende Methode, um ein tiefes Verständnis der Unterrichtsinhalte zu erreichen. Insbesondere ist das Schreiben von Lernprotokollen dann effektiv, wenn Lernende durch Prompts zu hochqualitativen kognitiven (z. B. Organisations- und Elaborationsstrategien) und metakognitiven (z. B. Monitoring) Lernstrategien angeregt werden (z. B. Berthold, Nückles & Renkl, 2007). Darüber hinaus stellen Nückles, Hübner und Renkl (2009) in ihrer Selbstregulationsperspektive des Schreibens von Lernprotokollen heraus, dass kognitive und metakognitive Strategien in Lernprotokollen in einem zyklischen Prozess miteinander interagieren sollten. Zwei stillschweigende und bislang nicht überprüfte Annahmen dieses zyklisch-interaktiven Modells selbstregulierten Lernens in Lernprotokollen sind die Kognition-fördert-Metakognition-Annahme und die Metakognition-fördert-Kognition-Annahme. Die Kognition-fördert-Metakognition-Annahme besagt, dass die vorhergehende Ausführung von kognitiven Lernstrategien die Qualität anschließender metakognitiver Lernstrategien erhöhen sollte, da das Organisieren und Vertiefen zentraler Inhalte (kognitive Lernstrategien) den Lernenden helfen könnte, Verständnisschwierigkeiten zu erkennen und somit die Qualität des Monitoring (metakognitive Lernstrategie) zu steigern. Die Metakognition-fördert-Kognition-Annahme wiederum besagt, dass die vorhergehende Ausführung von metakognitiven Lernstrategien (z. B. Monitoring) die Qualität anschließender kognitiver Strategien erhöhen könnte, da metakognitive Lernstrategien dazu führen könnten, dass Lernende eine erhöhte Notwendigkeit sehen, Lerninhalte zu organisieren und zu elaborieren (kognitive Lernstrategien).

Fragestellungen: Vor diesem Hintergrund war es das Ziel der vorliegenden Studie, sowohl die Kognition-fördert-Metakognition als auch die Metakognition-fördert-Kognition-Annahme experimentell zu überprüfen. Die zentralen Forschungsfragen waren: (1) Zeigen Lernende, die durch Prompts zunächst zu kognitiven und erst danach zu metakognitiven Lernstrategien angeregt werden, metakognitive Lernstrategien von höherer Qualität als Lernende, die zunächst zu metakognitiven Lernstrategien angeregt werden? (2) Zeigen Lernende, die durch Prompts zunächst zu metakognitiven und erst danach zu kognitiven Lernstrategien angeregt werden, kognitive Lernstrategien von höherer Qualität als Lernende, die zunächst zu kognitiven Lernstrategien angeregt werden?

Neben diesen Forschungsfragen in Bezug auf die Qualität der gezeigten Lernstrategien sollten zudem Effekte auf den Lernerfolg untersucht werden.

Methode: Die Forschungsfragen wurden untersucht im Rahmen einer experimentellen Feldstudie, in der N = 46 ZehntklässlerInnen nach drei aufeinanderfolgenden Deutschstunden jeweils ein Lernprotokoll verfassten. In ihren Lernprotokollen erhielten die Lernenden entweder Prompts in einer Organisation-Elaboration-Metakognition-Reihenfolge oder in einer Metakognition-Organisation-Elaboration-Reihenfolge. Die Zuteilung zu den beiden Reihenfolgen erfolgte randomisiert. In der vierten Deutschstunde bearbeiteten alle Lernenden einen Wissenstest in Bezug auf die Inhalte der drei vorhergegangenen Deutschstunden (drei Lernende verpassten den Nachtest aufgrund von Krankheit).

Ergebnisse und Diskussion: Als zentrale Ergebnisse fanden wir, dass Lernende, die Prompts in einer Metakognition-Organisation-Elaboration-Reihenfolge bearbeiteten, entsprechend der Metakognition-fördert-Kognition-Annahme kognitive Strategien von höherer Qualität zeigten als Lernende, die Prompts in einer Organisation-Elaboration-Metakognition-Reihenfolge erhielten, $t(41) = 2.45, p = .019, d = 0.76$ (mittlerer bis großer Effekt). Entgegen der Kognition-fördert-Metakognition-Annahme zeigte sich jedoch, dass Lernende, die zunächst kognitive Prompts bearbeiteten, auch bezüglich der Qualität ihrer metakognitiven Strategien unterlegen waren, $t(41) = 2.32, p = .025, d = 0.73$ (mittlerer bis großer Effekt). Dieses Ergebnismuster spiegelte sich auch im Lernerfolg wider. So fanden wir, dass Lernende, die in ihren Lernprotokollen zunächst metakognitive Prompts bearbeitet hatten, höhere Leistungen im Nachtest erzielten als Lernende, die zunächst kognitive Prompts bearbeitet hatten, $t(41) = 2.08, p = .044, d = 0.65$ (mittlerer Effekt).

Diese Befunde erweitern die Befundlage zur Förderung von Lernstrategien selbstregulierten Lernens in Lernprotokollen um folgende Punkte: (1) Die Metakognition-fördert-Kognition-Annahme unterstützend fanden wir, dass Lernende, die zunächst metakognitive Prompts bearbeiteten, anschließend kognitive Strategien von höherer Qualität zeigten als Lernende, die kognitive Prompts zu Beginn ihrer Lernprotokolle bearbeiteten. (2) Entgegen der Kognition-fördert-Metakognition-Annahme zeigten Lernende, die zunächst kognitive Prompts bearbeiteten, metakognitive Strategien von geringerer Qualität als Lernende, die zunächst zu metakognitiven Lernstrategien angeregt wurden. (3) Einhergehend mit den Befunden in Bezug auf die Lernstrategiequalität erreichten Lernende, die zunächst metakognitive Prompts bearbeitet hatten, einen höheren Lernerfolg als Lernende, die zunächst kognitive Prompts bearbeiteten.

Donnerstag
12. März

15:30 Uhr

GA 03/149

Chair(s)
Rieser, Svenja

Buhl, Monika; Kuhn, Hans Peter; Mello, Zena R.; Worrell, Frank C.

Zusammenhänge zwischen der Entwicklung von Zeitperspektive und Selbstreguliertem Lernen in der Sekundarstufe

Zeitperspektiven beschreiben, wie Individuen über die Vergangenheit, die Gegenwart und die Zukunft denken und wie sie diese bewerten. Sie beziehen sowohl kognitive Repräsentationen als auch motivationale Aspekte mit ein (Zimbardo & Boyd, 1999). Aufgrund zahlreicher zukunftsrelevanter Entwicklungsprozesse kann hierfür das Jugendalter als bedeutsame Lebensphase angesehen werden (Mello & Worrell, 2014), auch wenn sich die Forschung zu Zeitperspektiven zumeist auf Erwachsene bezieht und vergleichsweise wenige Studien Entwicklungsprozesse bei Heranwachsenden thematisieren. Entsprechende Forschungsarbeiten - häufig mit einem Fokus auf die Zukunft - stellen jedoch Zusammenhänge von Zeitperspektiven mit diversen Aspekten des Bildungskontexts (Phalet et al., 2004; Peetsma & van der Veen, 2001) her.

Ergebnisse der eigenen Arbeitsgruppe zeigen, dass verschiedene Profile, die spezifische Bewertungen der Vergangenheit, der Gegenwart und der Zukunft verbinden, dazu geeignet sind, individuelle Zeitperspektiven abzubilden, die gleichfalls mit Schulzufriedenheit und Schulleistungen in Zusammenhang stehen (Buhl & Lindner, 2009; Andretta et al., 2014). Die geplante Präsentation fokussiert in einem weiteren Schritt darauf Zeitperspektive mit Lernerfahrungen in der Schule in Verbindung zu bringen. Zentrale Fragestellung ist, inwieweit Aspekte des Selbstregulierten Lernens einen Einfluss auf Zeitperspektiven und deren Veränderung haben. In einem ersten Schritt werden hierzu die Stabilität und Veränderung von Zeitperspektivenprofilen über einen Zeitraum von zwei Jahren untersucht. In einem zweiten Schritt wird der wechselseitige Einfluss dieser Profile mit verschiedenen Dimensionen des Selbstregulierten Lernens analysiert.

Als Datengrundlage werden die ersten drei Messzeitpunkte der BiO Studie (Bildungsverläufe an der Offenen Schule Kassel Waldau) herangezogen, eine Multikohorten-Längsschnittstudie, in der zu jedem Messzeitpunkt etwa 600 Schülerinnen und Schüler (51% weiblich) der 5. bis 10. Jahrgangsstufe mit einem standardisierten Fragebogen befragt werden. Für die drei Befragungen umfassenden Längsschnitt liegen Daten von über 450 Heranwachsenden vor. Zeitperspektiven wurden über die deutsche Version (Worrell et al., 2013) der Time Attitude Scale (ATI-TA, Mello & Worrell, 2007) erfasst, die aus sechs Subskalen besteht (Past, Present, Future jeweils in positiver und negativer Ausprägung, $.73 < \alpha < .89$). Der Bereich des Selbstregulierten Lernens wird durch verschiedene - v.a. durch die PISA-Studien etablierte - Instrumente erfragt, die von stärker kognitiven Bereichen (Kontrollerwartung, fachbezogene Selbstkonzepte), über Aspekte der Handlungssteuerung (Ausdauer, Anstrengung) bis hin zu verschiedenen Lernstrategien (Wiederholung, Erarbeitung, Kontrolle) reichen ($.75 < \alpha < .87$) (Baumert et al., 2000).

Über latente Profil Analysen wurden aus den ATI-TA Subskalen, fünf verschiedene Zeitperspektive Profile identifiziert: Positives, Balanced, Negatives, Ambivalent und Optimists. Während die ersten vier Profile vergleichbare Ausprägungen für die drei Zeitdimensionen aufweisen (z.B. Positives auf den positiven Dimensionen über dem Skalenmittelwert, auf den negativen darunter) beschreibt das Profil Optimists die pädagogisch und entwicklungspsychologisch interessante Kombination einer kritischen Evaluation der Vergangenheit bei einem gleichzeitig überdurchschnittlich positiven Blick in die Zukunft. Die Stabilität über zwei Jahre ist für das Profil Positives (60%) am höchsten, für das Profil Ambivalent am Niedrigsten (37%). Die verschiedenen Profile weisen Unterschiede im Bereich des Selbstregulierten Lernens auf. Während die kognitiven Aspekte wie Kontrollerwartung und fachliche Selbstkonzepte in den Profilen Positives und Optimists signifikant über denen im Profil Negatives liegen, zeigt sich hinsichtlich der Handlungssteuerung die Abgrenzung des Profils Positives gegenüber den Profilen Negatives und Optimists. Ausdauer und Anstrengung der Positives sind im Mittel signifikant höher. Die Bereiche der Lernstrategien weisen keine signifikanten Unterschiede auf.

Insgesamt zeigen die Ergebnisse, dass sich Zeitperspektiven Heranwachsender über verschiedene Profile beschreiben lassen, die gleichsam in Zusammenhang mit Aspekten des Selbstregulierten Lernens stehen. Während Unterschiede der Profile Positives und Negatives in die erwartete Richtung gehen, finden sich die Schülerinnen und Schüler des Profils Optimists hinsichtlich der kognitiven Aspekte des SRL wie Kontrollüberzeugungen und fachbezogene Selbstkonzepte in der Gruppe der Positives, hinsichtlich der Bereiche der Handlungssteuerung in der Gruppe der Negatives. Die Diskussion soll darauf fokussieren, welche Veränderungen Zeitperspektiven im Jugendalter erfahren und welche Rolle schulische Erfahrungen hierbei einnehmen können.

Donnerstag
12. März

15:30 Uhr

GA 03/149

Chair(s)

Rieser, Svenja

Karlen, Yves

Differenzen beim selbstregulierten Lernen von Schülerinnen und Schüler in der gymnasialen Oberstufe – eine Typologie

Die Fähigkeit zum selbstregulierten Lernen (SRL) lässt sich grob als Prozess beschreiben, in welchem die Schülerinnen und Schüler verschiedene Selbststeuerungs- und Regulationsmassnahmen ergreifen, die auf kognitiver, metakognitiver und motivationaler Komponente ansetzen. Schülerinnen und Schüler besitzen über unterschiedliche Ausprägungen bezüglich der Fähigkeit zum SRL. Zur Systematisierung von interindividuellen Unterschieden bieten sich Typologien an. Während einige Studien auf mehrere erfolgreiche Typen von Schülerinnen und Schüler beim SRL hinweisen (Barnard-Brak, Lan, & Paton, 2010), zeigen andere Studien auf (Artelt et al., 2004), dass vorwiegend diejenigen Typen am erfolgreichsten sind, die sich durch eine hohe Ausprägung bei allen Komponenten des SRL charakterisieren lassen. Übergreifend scheinen besonders die Typen erfolgreich zu sein, die in der Lage sind, flexibel und angemessen auf unterschiedliche Herausforderungen zu reagieren.

Studien zu Typologien bezüglich des SRL von Schülerinnen und Schülern auf der gymnasialen Oberstufe sind bisher kaum vorhanden. In dieser Studie soll deshalb untersucht werden, inwiefern sich auf der Basis von interindividuellen Unterschieden beim SRL Typologien bilden lassen, inwieweit Zusammenhänge zwischen den ermittelten Typen und den schulischen Leistungen der Schülerinnen und Schülern im Fach Deutsch vorhanden sind, sowie welche intraindividuellen Veränderungen in der Typenklassifikation sich über ein Schuljahr hinweg ergeben.

Eine repräsentative Stichprobe von Schülerinnen und Schüler aus der 10. und 11. Klassenstufe aus mehreren Gymnasien wurde befragt. Zu Beginn des Schuljahres (t1) nahmen N = 1'272 Schülerinnen und Schüler teil, für den zweiten Erhebungszeitpunkt am Ende des Schuljahres (t2) liegen Daten von N = 1'126 Schülerinnen und Schüler vor. Auf der Basis der beschriebenen Stichprobe, konnten die Datensätze von N = 879 Schülerinnen und Schülern zu beiden Messzeitpunkten zugeordnet werden. Diese Längsschnittstichprobe umfasst 62.3% Schülerinnen und 38.7% Schüler mit einem Durchschnittsalter von 16.4 Jahren (SD = 0.97). Die Online-Erhebungen wurden im Zeitraum von neun Monaten jeweils während zwei Schulstunden im Regelunterricht unter Aufsicht einer Lehrperson durchgeführt.

Mit dem eingesetzten metakognitiven Wissenstest wurde das Wissen zu Regulationsstrategien beim Verfassen von grösseren und komplexen Arbeiten und Projekten erfasst (Maag Merki et al. 2013). In sieben verschiedenen Lernszenarien, die einen vollständigen Selbstregulationsprozess umfassen, mussten die Lernenden die aufgelisteten Lernstrategien hinsichtlich ihrer Nützlichkeit zur Lösung der Aufgabenanforderungen einschätzen. Zur Erfassung der selbstberichteten Nutzung von kognitiven und metakognitiven Lernstrategien sowie von mehreren motivationalen Komponenten des SRL (z.B. Selbstkonzept, Selbstwirksamkeit, Leistungsmotivation, intrinsische Motivation) wurde ein standardisierter Fragebogen eingesetzt. Die schulische Leistung wurde über die Schulhabjahresnote im Fach Deutsch erfasst.

Anhand von latenten Clusteranalysen konnten zu beiden Messzeitpunkten vier verschiedene Typologien von Schülerinnen und Schülern identifiziert werden, die sich signifikant voneinander unterscheiden. Zu beiden Messzeitpunkten wiesen die Schülerinnen und Schüler aus der Gruppe mit den höchsten Ausprägungen bei allen Komponenten des SRL über beide Schulhalbjahre im Vergleich zu den anderen Typen die jeweils besten Noten im Fach Deutsch auf. Kontrastierend erzielten die Schülerinnen und Schüler aus der Gruppe mit sehr tiefen Werten bei den motivationalen Komponenten im Durchschnitt jeweils die tiefsten Schulnoten im Fach Deutsch. Vom ersten zum zweiten Messzeitpunkt veränderte sich die Zuordnung zu einem der vier Typen für insgesamt fast die Hälfte der Schülerinnen und Schüler. Die Ergebnisse dieses Artikels liefern folglich einen ersten Hinweis darauf, dass Typologien bis zu einem gewissen Grad veränderbare Handlungsmuster und Verhaltenspräferenzen darstellen und das SRL nicht eine unveränderbare Begabung darstellt. Diese Ergebnisse werden vor dem Hintergrund möglicher Implikationen für die Praxis sowie zukünftigen Analysen diskutiert.

Donnerstag
12. März

15:30 Uhr

GA 03/149

Chair(s)

Rieser, Svenja

Puddu, Sandra

Scaffolding bei der Einführung Forschenden Lernens

Theoretischer Hintergrund: Die Einführung Forschenden Lernens wird von der European Commission (2007) empfohlen, um mit der Unterschiedlichkeit von Schülerinnen und Schülern in einer Klasse umzugehen. Als Ziele für das Forschende Lernen nennen Abrams et al. (2008): „Learning about inquiry“, „Learning to inquire“ und „using inquiry to learn science content“. Für die Einführung Forschenden Lernens wird ein stufenweises Vorgehen empfohlen (Lederman, 2007), z.B. anhand der Level von Inquiry (Blanchard et al., 2010). Bei Level 0 ist die Lehrperson für die Organisation des gesamten Prozesses, der grob in Fragestellung, Methodenwahl und Interpretation der Ergebnisse gegliedert werden kann, zuständig. Die SchülerInnen übernehmen bei jedem Level mehr Verantwortung bis sie bei Level 3 diese für den gesamten Prozess übernehmen. Mit steigender Offenheit (höherem Level) ändern sich die Rollen und Verantwortlichkeiten von Lehrerinnen und Lehrern sowie von Schülerinnen und Schülern (van der Valk & de Jong, 2009). Scaffolding, also die Unterstützung und Begleitung beim forschenden Lernen, ist für den Erfolg entscheidend und eine große Herausforderung für Lehrpersonen, insbesondere in einer heterogenen Klasse (Blanchard et al., 2010). Beim Scaffolding können verschiedene Strategien angewendet werden (vgl. Duckworth, 2009; Furtak, 2008; van Zee & Minstrell, 1997).

Fragestellung: In einer explorativen Fallstudie wurde erforscht, wie Forschendes Lernen in einer Klasse mit hoher Diversität eingeführt und begleitet werden kann. Die Fragestellung lautete: „Wie begleitet die Lehrperson Schülerinnen und Schüler beim Forschenden Lernen unter Berücksichtigung der Diversität?“ Dazu wurde der Laborunterricht einer Chemielehrperson während eines Schuljahres begleitet. Der Laborunterricht fand alle drei Wochen für 150 Minuten statt. Die ausgewählte Klasse (9. Schulstufe, 31 Schülerinnen und Schüler) einer städtischen Handelsakademie (9.-13. Schulstufe mit wirtschaftlichem Schwerpunkt) zeichnete sich durch ihre hohe Diversität aus, wie z.B. 14 Muttersprachen und 3;1 Jahre Altersunterschied. Von Interesse waren die Maßnahmen, durch die die Lehrperson, bei der sukzessiven Implementation von Forschendem Lernen, die Schülerinnen und Schüler begleitet und unterstützt.

Methode: Um das Scaffolding analysieren zu können, wurden Video- und Audioaufnahmen gemacht und passive teilnehmende Beobachtung durchgeführt. Außerdem wurden schriftliche Aufgabenstellungen sowie ausgefüllte Laborprotokolle herangezogen. Als Analyseverfahren wurde die induktive Zusammenfassung nach Mayring (2008) gewählt, um ein möglichst vollständiges Bild der verwendeten Unterstützungsmaßnahmen zu zeichnen.

Ergebnisse: Während der Analyse bildeten sich vier Hauptkategorien des Scaffoldings heraus. Diese sind „Sprache“, „Forschendes Lernen durchführen“, „Lernen über Forschendes Lernen“ und „Inhalte lernen“. Interessant ist, dass sich diese Kategorien (ausgenommen Sprache) mit den Zielen decken, die Abrams et al. (2008) formuliert haben. Im Vortrag soll auf die Kategorie Sprache fokussiert werden. Bei den Scaffoldingmaßnahmen, die zur Kategorie Sprache zusammengefasst wurden, ergaben sich drei Unterkategorien: „Aufgabenblätter“, „Durchführung“ und „Protokoll“. Die Aufgabenblätter zeichnen sich häufig durch Visualisierungen, klare Strukturierung und einfache Sprache aus. Bei der Durchführung achtete die Lehrperson darauf wenig geläufige Wörter wie „spröde“ oder „Apparatur“ zu erklären. Unklarheiten bei der Verwendung von Wörtern wie „schwer“ und „schwierig“, die in einigen Protokollen synonym verwendet wurden („Mit Wasser wird die Masse schwieriger als, ohne Wasser.“ P20111010), wurden im Unterricht aufgegriffen und geklärt. Die Unterscheidung von „Schmelzen“ und „Lösen“ war für die Lehrperson so wichtig, dass hierzu eine komplette Laboreinheit geplant und umgesetzt wurde (vgl. Puddu & Koliander, 2013). Schriftliche Aufgabenstellungen wurden oft mündlich erklärt und zusätzlich von SchülerInnen zusammengefasst, wobei hier die Kombination von Wort und Gestik besonders wichtig erscheint. Bei den Protokollen unterstützte die Lehrperson die Schülerinnen und Schüler durch Beispielsätze und -tabellen sowie durch Leitfragen. Die Lehrperson ließ sich auch Sätze aus dem Protokoll vorlesen und machte Verbesserungsvorschläge, wodurch die SchülerInnen ihre Formulierungen optimieren konnten und die Aussprache geübt wurde.

Ausgehend von dieser Einzelfallstudie zum Scaffolding wird die Passung mit den individuellen Bedürfnissen der SchülerInnen diskutiert. Daraus werden Empfehlungen abgeleitet für die Implementierung von Forschendem Lernen in Klassen mit hoher Diversität.

Donnerstag
12. März

15:30 Uhr

GA 03/149

Chair(s)

Rieser, Svenja

3

Gast, Anna; Stapp, Christiane; Herppich, Stephanie; Wittwer, Jörg; Nückles, Matthias

Ein Training mit Finesse für mehr Akkuratessse – Interaktion und Diagnosegenauigkeit von Lehramtsstudierenden im instruktionalen Dialog

Vor allem Dialogsituationen erlauben es Lehrpersonen Wissen von Schülerinnen und Schüler (SuS) formativ zu diagnostizieren. Diese Art der Diagnose ist für das Unterrichten und Lernen höchst relevant, weil es im Prozess („on the fly“) Schülerdenken sichtbar machen kann (Shavelson et al., 2008). Indem die Lehrperson das Wissen der Lernenden fortlaufend diagnostiziert, kann sie ihr Modell über das Wissen der Lernenden konkretisieren (Nickerson, 1999) und ihren Unterricht anpassen (Adaptives Unterrichten; Beck et al., 2008). Befunde weisen jedoch darauf hin, dass Lehrpersonen im Dialog mit SuS das Ziel, Schülergedanken sichtbar zu machen, nicht verfolgen. So zeigt sich häufig ein Ungleichgewicht im Dialog zwischen Lehrenden und Lernenden. Lehrpersonen haben zum Beispiel einen höheren Sprechanteil als ihre SuS (z.B. Chi, Siler, Jeong, Yamauchi & Hausmann, 2001) und unterschätzen ihren eigenen Sprechanteil systematisch (Helmke et al., 2008). Dennoch wurde bislang selten versucht, diese formative Diagnosekompetenz zu fördern (für eine Ausnahme siehe Herppich, Wittwer, Nückles, & Renkl, 2014).

Vor diesem Hintergrund testeten wir zwei Hypothesen:

Erstens, erwarteten wir, dass Lehramtsstudierende, die ein computerbasiertes Training mit konzeptuellen und strategischen Elementen erhalten, eine stärker schülerzentrierte Kommunikation anwenden, also den Lernenden mehr Möglichkeiten geben, sich zu äußern und weniger Erklärungen geben, als Lehramtsstudierende die ein Kontrolltraining erhalten.

Zweitens, erwarten wir, dass eine stärkere Schülerzentrierung mit einer höheren Diagnosegenauigkeit einhergeht.

An unserer Studie nahmen 70 Personen (35 Dyaden) teil. Lehramtsstudierende des Fachs Biologie (Lehramt für Gymnasien) wurden randomisiert einer Trainingsgruppe (n = 18) und einer Kontrollgruppe (n = 17) zugeordnet. Die Trainingsgruppe erhielt ein computerbasiertes Training zu den Inhalten „Relevanz formativer Diagnose im instruktionalen Lehrer-Schüler-Gespräch“ und „Interaktionsstrategien im instruktionalen Lehrer-Schüler-Gespräch“, das aus vier Bausteinen bestand: (1) Vermittlung von Inhalten, (2) Anwendung und Üben des erworbenen Wissens (u.a. mit Videoclips), (3) Prüfung des Wissens (Übungsfragen), (4) Transfer in den Alltag durch einen „Take-Home-Übungszettel“. Es wurde konzeptuelles Wissen, also theoretischer Input und Informationen zu dem Thema, und strategisches Wissen, also die Anwendung der Inhalte, vermittelt. Die Kontrollgruppe erhielt auch ein Training, welches jedoch lediglich Informationen zu den Themen als Fließtext auf PowerPoint-Folien beinhaltete (nur konzeptuelles Wissen). Drei Tage nach dem Training fand ein Lehrer-Schüler-Dialog zum Thema Meiose statt. Die Lernenden waren Erstsemester unterschiedlicher Fächer (kein Biologie) der Universität Freiburg. Vor dem Dialog wurde das Vorwissen der Lehramtsstudierenden und ihrer Lernenden durch einen Multiple-Choice-Test und einer Zeichnung zu Meioseprozessen erfasst. Nach dem Dialog, der durch einen, in kurze Abschnitte geteilten, Lehrbuchtext sequenziert war, wurde das Wissen der Lernenden wiederum erfasst. Um Diagnosegenauigkeit zu messen, wurden Lehramtsstudierende gebeten, die Fragen aus Sicht ihrer Lernenden zu beantworten. Die Dialoge wurden videographiert und kodiert.

Alle Hypothesen wurden einseitig getestet. Hinsichtlich ihres Dialogverhaltens zeigten die Lehramtsstudierenden in der Trainingsgruppe einen signifikant geringeren prozentualen Gesprächsanteil als die Lehramtsstudierenden der Kontrollgruppe, $F(2,32) = 10.52$, $p < .01$, $\eta^2 = .24$. Außerdem gaben die Lehramtsstudierenden der Trainingsgruppe signifikant weniger Erklärungen als die Lehramtsstudierenden der Kontrollgruppe, $F(2,32) = 8.19$, $p < .01$, $\eta^2 = .20$. Diese Befunde zeigen, dass die Dominanz der angehenden Lehrkräfte durch das Training erfolgreich reduziert werden konnte. Bezogen auf die Diagnosegenauigkeit zeigte sich, dass Lehramtsstudierende mit vielen vertiefenden interaktiven Episoden im Dialog akkurater diagnostizierten, $F(1,33) = 3.19$, $p < .05$, $\eta^2 = .09$, als Lehramtsstudierende mit wenigen vertiefenden interaktiven Episoden. Dieses Ergebnis weist darauf hin, dass vertiefende Interaktionen mit einer akkurateren Diagnose einhergehen und ein solches schülerzentriertes Verhalten daher gefördert werden sollte.

Die Befunde unserer Studie zeigen, dass eine relativ kurze und sparsame Intervention mit dem Ziel der Vermittlung kommunikativer Strategien deutliche Effekte auf das Verhalten angehender Lehrkräfte im Dialog mit Lernenden hat. Solche Trainings könnten daher in der Lehrerbildung und -fortbildung zukünftig berücksichtigt werden, um Fähigkeiten zur formativen Diagnose im Dialog gezielt zu fördern.

Donnerstag
12. März

15:30 Uhr

GB 03/49

Chair(s)
Gast, Anna

3

Wahle, Caroline Verena; Praetorius, Anna-Katharina; Hochdoerffer, Katrin; Schrader, Friedrich-Wilhelm

Einmal gut, immer gut? Untersuchungen zur Stabilität der diagnostischen Kompetenz von Lehramtsstudierenden

Die diagnostische Kompetenz von Lehrkräften wird als notwendige Bedingung für adaptive Lehr-Lernprozesse angesehen und als Kernkompetenz von Lehrkräften definiert (Corno & Snow, 1986). In Untersuchungen zur diagnostischen Kompetenz von Lehrkräften konnten jedoch deutliche Unterschiede in der Urteilsgenauigkeit von Lehrkräften nachgewiesen werden (z. B. Südkamp et al., 2012). Bislang erst unzureichend geklärt sind die Ursachen der Unterschiede in der Fähigkeit von Lehrpersonen, Schüler/-innen sowohl in Bezug auf kognitive, als auch motivational-emotionale Merkmale genau einzuschätzen. Als mögliche Ursachen wurden in bisherigen Studien vielfach Persönlichkeitsmerkmale (z.B. Intelligenz) angeführt. Die Voraussetzung dafür ist jedoch, dass es sich bei der diagnostischen Kompetenz um ein zeitstabiles Merkmal handelt. Hierzu gibt es erste Untersuchungen (z. B. Lorenz & Artelt, 2009). Jedoch wurden die bisherigen Studien im natürlichen Schulkontext durchgeführt, wo sich die Schülermerkmale im Rahmen der Untersuchung verändern. Die vorliegende Studie analysiert die Stabilität der diagnostischen Kompetenz von Physiklehramtsstudierenden (N = 26; davon 77 % Gymnasiallehramt, 23 % Realschullehramt) in Bezug auf das Fähigkeitsselbstkonzept in Physik und metakognitive Urteile (Anstrengungsbereitschaft und wahrgenommene Aufgabenschwierigkeit) von Schüler/-innen der 10. Klasse (N = 218). Mit Lehramtsstudierenden in der Stichprobe wurde somit der frühestmögliche Zeitpunkt gewählt, zu dem man Lehrkräfte untersuchen kann. Im Rahmen der Untersuchung unterrichteten die Lehramtsstudierenden die ihnen bis dahin unbekanntes Schüler/-innen 90 Minuten lang. Im Anschluss an ihren Unterricht schätzten sie die Schüler/-innen in Bezug auf die unterschiedlichen Merkmale ein. Nach drei Monaten wurden die Einschätzungen wiederholt. Zwischen den beiden Messzeitpunkten erfolgte keine weitere Interaktion zwischen den Studierenden und den Schüler/-innen. Die Datenauswertung erfolgte basierend auf den drei Akkuratheitskomponenten nach Schrader (1989). Zu beiden Messzeitpunkten finden sich mittlere Zusammenhänge zwischen Lehrerurteil und den Schülerratings. Ergebnisse der ersten Erhebung zeigen, dass die Lehramtsstudierenden die mittlere Ausprägung der Schülermerkmale (Niveauelemente) bei allen Merkmalen sehr genau einschätzen können. In Bezug auf die Differenzierungskomponente tendieren sie zu einer leichten Überschätzung der Merkmalsstreuung. Die Rangkomponente weist eine mittlere Korrelation von $r = .40$ für das Fähigkeitsselbstkonzept auf. Bei den metakognitiven Urteilen finden sich jedoch nur geringe Werte für die Rangkomponente ($r = .14$ bei der wahrgenommenen Anstrengungsbereitschaft bzw. $r = .32$ bei der wahrgenommenen Aufgabenschwierigkeit). Ein sehr ähnliches Bild zeichnet sich bei den Ergebnissen des zweiten Messzeitpunktes ab. Zur Berechnung der Stabilität wurden die Genauigkeitswerte der beiden Messzeitpunkte in Bezug auf alle drei Komponenten miteinander korreliert. Die Ergebnisse zeigen bei allen eingeschätzten Merkmalen mittelhohe bis hohe Stabilitäten der Urteilsgenauigkeit über einen Zeitraum von drei Monaten hinweg ($r = .32$ bis $.86$). Insbesondere das Fähigkeitsselbstkonzept der Schüler/-innen in Physik haben die Lehramtsstudierenden zu beiden Zeitpunkten vergleichbar genau eingeschätzt. Die höchsten Korrelationen finden sich hier in Bezug auf die Differenzierungskomponente (Übereinstimmung zwischen eingeschätzter und tatsächlicher Merkmalsstreuung innerhalb einer Schülergruppe mit $r = .80$, $p = .01$) und die Rangkomponente (Korrelation zwischen Lehrerurteil und Kriterium mit $r = .76^{**}$, $p = .01$). Auch die Merkmalsstreuung der wahrgenommenen Anstrengungsbereitschaft wird von den Lehramtsstudierenden zu beiden Zeitpunkten ähnlich akkurat eingeschätzt ($r = .86^{**}$, $p = .01$). Die Niveauelemente (mittlere Ausprägung der Schülermerkmale) stellt sich in Bezug auf das Fähigkeitsselbstkonzept in Physik ($r = .57^{**}$, $p = .01$) und die wahrgenommene Schwierigkeit ($r = .32^{**}$, $p = .01$) der Schüler/-innen jedoch als etwas weniger stabil heraus. Die Ergebnisse der Studie deuten darauf hin, dass diagnostische Kompetenz unabhängig vom Untersuchungszeitpunkt eine hohe zeitliche Stabilität aufweist. Dies ist bereits bei Lehramtsstudierenden ohne Berufserfahrung der Fall. Die Studie liefert zudem Hinweise darauf, dass bei der Analyse von möglichen Determinanten diagnostischer Kompetenz stabile Merkmale von Schüler/-innen (z.B. Fähigkeitsselbstkonzept) auch zeitlich überdauernd stabiler eingeschätzt werden, als weniger stabile Merkmale (z.B. metakognitive Urteile), wenngleich die Ergebnisse dieser Untersuchungen aufgrund der recht geringen Stichprobengröße noch einmal repliziert werden sollten.

Donnerstag
12. März

15:30 Uhr

GB 03/49

Chair(s)
Gast, Anna

Kunina-Habenicht, Olga; Hautz, Wolf; Knigge, Michel; Spies, Claudia; Ahlers, Olaf

ASCLIRE: Test zur computerbasierten Erfassung diagnostischer Kompetenz angehender Ärzt(inn)e(n)

Theoretischer Hintergrund: Medizinische Ausbildungsforschung bildet einen wichtigen Teil der empirischen Hochschulforschung, wengleich diese selten in der erziehungswissenschaftlichen oder psychologischen Literatur thematisiert wird. Im Rahmen der medizinischen Ausbildung wird der diagnostischen Kompetenz von Ärzt(inn)en eine wichtige Bedeutung beigemessen (Epstein & Hundert, 2002). Diagnostische Kompetenz von Ärzt(inn)en wird in der Literatur unter verschiedenen Begriffen thematisiert (beispielsweise "clinical reasoning", "medical problem solving" oder "medical decision making") (für Reviews siehe Elstein, 2009; Eva, 2004; Norman, 2005). Im medizinischen Universitätsstudium werden zur Erfassung der diagnostischen Kompetenz neben Multiple-Choice-Fragen unter anderem auch Wissenstests (bspw. Script Concordance Test; Charlin, Roy, Brailovsky, Goulet & van der Vleuten, 2000) und checklistenbasierte Prüfungsverfahren (z.B. Objective Structure Clinical Examinations; Boursicot, Roberts & Burdick, 2010) eingesetzt, die eine handlungsnahere Erfassung der diagnostischen Kompetenz von Ärzt(inn)en im klinischen Setting erlauben sollen. Diese Verfahren weisen jedoch einige methodische Probleme auf, die im Vortrag kurz erläutert werden (bspw. hoher administrativer Aufwand oder mangelnde Vergleichbarkeit über verschiedene Testsituationen).

Fragestellung: Um die methodischen Probleme bereits vorliegender Verfahren zu umgehen, wurde in einem interdisziplinären Projekt zwischen Psycholog(inn)en und Mediziner(inn)en ein normreferenzierter computerbasierter Test zur Erfassung diagnostischer Kompetenz angehender Ärzt(inn)e(n) (ASCLIRE) entwickelt, welcher es erlaubt, diagnostische Kompetenz in authentischem klinischen Setting ökonomisch zu erfassen. Er erweitert bisherige Verfahren in zweierlei Hinsicht: Zum einen werden neben der Akkuratheit der Diagnose auch die Reaktionszeiten und die herangezogenen diagnostischen Maßnahmen erfasst. Zum anderen bildet der Test den gesamten Entscheidungsprozess ab – von der Sammlung diagnostischer Information bis zur Diagnosefindung. Bezug nehmend auf das Validitätskonzept von Messick (1989) wurde die Validität der Testwerte in diesem neu entwickelten Testverfahren untersucht, indem unter anderem die Testleistung der Studierenden mit der der Expert(inn)en verglichen wurde.

Methoden: An der Studie nahmen 287 Studierende im Reformstudiengang der Charité – Universitätsmedizin Berlin und 20 Expert(inn)en mit mindestens dreijähriger Berufserfahrung teil. Zur Testerprobung wurden sechs Fälle aus der Wissensdomäne „Luftnot“ entwickelt, drei davon akut, drei subakut. Die Bearbeitung des Tests erfolgte in drei Schritten: Im ersten Schritt wurde den Teilnehmer(inn)en anhand einer Fallvignette und eines Patientenvideos ein Patient mit Luftnot vorgestellt. Die Teilnehmer(inn)en sollten im zweiten Schritt aus 36 möglichen diagnostischen Maßnahmen in beliebiger Reihenfolge alle von ihnen für notwendig erachteten Maßnahmen auswählen, um eine Diagnose zu stellen und erhielten eine multimediale Darstellung der erhobenen Befunde. Im dritten Schritt wurden die Teilnehmer(inn)en gebeten, eine Diagnose zu stellen. Aufgezeichnet wurden die gestellte Diagnose, alle gewählten diagnostischen und therapeutischen Maßnahmen sowie die entsprechenden Reaktionszeiten. Für statistische Analysen wurden drei Leistungsindikatoren berechnet: Akkuratheit der Diagnose (0 falsch/1 richtig), benötigte Gesamtzeit zur Lösung der einzelnen Fälle und die Anzahl der relevanten ergriffenen Maßnahmen. Eine Maßnahme wurde als relevant angesehen, wenn mindestens 50% der Expert(inn)en diese Maßnahme ergriffen haben.

Ergebnisse: Die latente Reliabilität für die Akkuratheit der Diagnose über sechs Fälle war ausreichend ($\Omega = .63$), während latente Reliabilitäten für die Bearbeitungszeit und Anzahl der relevanten Maßnahmen zufriedenstellend waren (jeweils $\Omega = .85$). Für die Validität der Testwerte lassen sich unter anderem folgende Argumente finden: Im Einklang mit der allgemeinen Literatur zur Expertiseforschung (Bohle Carbonell, Stalmeijer, Könings, Segers & Van Merriënboer, 2014; Ericsson, Charness, Feltovich & Hoffman, 2006) wiesen Expert(inn)en höhere Akkuratheit ($d = 1.6$) und niedrigere Bearbeitungszeiten auf als Studierende ($d = 1.4$) auf. Die Akkuratheit der Diagnose nahm im Studienverlauf mit der Semesteranzahl signifikant zu. Die empirische Struktur des Tests wurde untersucht mit Strukturgleichungsmodellen im MPlus 7.2 (Muthén & Muthén, 1998–2010). Das Modell mit drei latenten Faktoren für die Akkuratheit der Diagnose, Entscheidungszeit und Wahl relevanter Maßnahmen weist einen sehr guten Modellfit auf ($\chi^2 = 152.2$, $df = 132$, $p = 0.11$; CFI = 0.983; RMSEA = 0.023), wobei die Akkuratheit der Diagnose nicht signifikant mit der Entscheidungszeit korrelierte. Diese und weitere Befunde werden diskutiert im Hinblick auf die Validität der Testwerte im ASCLIRE.

Donnerstag
12. März

15:30 Uhr

GB 03/49

Chair(s)
Gast, Anna

Brauch, Nicola; Bihrer, Andreas

Indikatoren-Entwicklung zur Messung von Fragestellungskompetenz künftiger Geschichtslehrkräfte

Forschungsstand und Theorie: Eine Anforderung an Lehrkräfte im Unterricht besteht darin, kognitive Aktivierung der Schüler/-innen zu leisten, indem z.B. fachspezifische Probleme zum Gegenstand des Unterrichts werden. Für die Ausbildung künftiger Geschichtslehrkräfte spielt die Einsicht in das domänenspezifische Konstruktionsprinzip von problematisierenden Leitfragen eine besondere Rolle, weil sie diese im kompetenzorientierten Geschichtsunterricht für die Gestaltung lernprozessförderlicher Unterrichtsarrangements selbständig konzipieren müssen (Schreiber 2008). Gleichzeitig geben Studien zur Untersuchung der Aufgabenkultur im Geschichtsunterricht Anlass zur Annahme, dass Lehrkräfte kaum problemorientiert unterrichten (Mägdefrau & Michler 2012). D.h. künftige Lehrkräfte müssen eine problematisierende Herangehensweise an Geschichte im Studium erwerben (Brauch & Bihrer 2011). Die Entwicklung von Indikatoren für die Identifizierung und Bewertung historischer Fragestellungen bei der Problematisierung von Geschichtsdarstellungen im Studium ist bislang kaum erforscht, wohingegen ein breiter Konsens in Einführungswerken für das Geschichtsstudium über die zentrale Bedeutung dieser Kompetenzdimension besteht (Lingelbach & Rudolph 2005).

In unserer explorativen Studie gehen wir davon aus, dass fachspezifische Fragestellungen die Erkenntnisziele der Geschichtswissenschaft repräsentieren und zur Beantwortung einer fachspezifischen Methodik bedürfen (Maggioni, Van Sledright & Alexander 2009).

Im Kontext der BMBF-Förderlinie zur Verbesserung der Lehre (QPL) wurde diesem Forschungsanliegen in einem standardisierten Seminar zur mittelalterlichen Geschichte für Studienanfänger nachgegangen. Erstes Ziel war die Bestimmung von Ankerbeispielen für Fragestellungen in Bezug auf theoretisch vorab definierte Indikatoren und für die Repräsentation individuell unterschiedlicher Niveaus im Rahmen eines jeden Indikators.

Methode: Im interdisziplinären Diskurs zwischen Geschichtswissenschaft und Geschichtsdidaktik und Einbindung der Erziehungswissenschaft (QPL/PerLe-Team der Universität Kiel) wurden theoriebasierte Indikatoren zur Identifizierung und Konstruktion historischer Fragestellungen entwickelt. Im Rahmen von drei Proseminaren (2013 bis 2014) wurde mit den Studierenden an der Entwicklung ihrer Fragekompetenz im Kontext übergeordneter historischer Problemstellungen gearbeitet. Dabei wurde die Indikatoren basierte Entwicklung von Fragestellungen für die Hausarbeiten ins Zentrum der Seminararbeit während des Semesters gestellt.

Die Hausarbeiten (Titel, Gliederung, Einleitung, Schluss) von 43 Studierenden (18m/25w) wurden inhaltsanalytisch untersucht. Als Basismerkmal zur Klassifizierung einer Fragestellung kam die Prüfung von deren domänenspezifischer Spezifizierung zur Anwendung: a) Zeitverlaufsbezug mit dem Erfordernis der empirischen Analyse, b) Quellenprüfungs-Bezug mit der Erfordernis der Kontextualisierung des empirischen Materials oder c) Theoriebezug mit empirisch-kontextualisierender Prüfung. Darüber hinaus wurden Indikatoren für die Relevanz des Erkenntnisziels und die Angemessenheit der Methode entworfen.

Ergebnis: Derzeit sind 31 Arbeiten bereits ausgewertet, die Ergebnisse der dritten Kohorte befinden sich noch in Bearbeitung. Der übergeordnete Indikator der domänenspezifischen Validität ließ sich durch den Bezug auf die erkenntnisleitenden Prinzipien der Geschichtswissenschaft in drei Fragetypen (vgl. oben a bis c) differenzieren. Für die Indikatoren (Relevanz: R 1 Fach und R 2 Gesellschaft; Methodik: M 1 Art des Quellenbezugs, M 2 Funktion des empirischen Fallbeispiels, M 3 Ausformulierung der Leitfrage) und die Unterscheidung von Niveaus ließen sich Ankerbeispiele identifizieren.

Schlussfolgerung: Die Annahme, dass sich durch die Geschichtstheorie Indikatoren für die Klassifizierung historischer Fragestellungen entwickeln lassen, hat sich auch in der kommunikativen Validierung im interdisziplinären Diskurs bewährt. Für die derzeitige Vorbereitung der Interraterreliabilitäts-Prüfung scheint die zur Beantwortung der Frage erforderliche Komplexität der historiographischen Analyse eine gute Ausgangsbasis zur Präzision der Beschreibungen der Niveau-Unterschiede der Fragestellungen zu sein.

Donnerstag
12. März

15:30 Uhr

GB 03/49

Chair(s)
Gast, Anna

Bürgermeister, Anika; Ringeisen, Tobias; Seifert, Matthias; Raufelder, Diana

Moderationskompetenz bei Studierenden: Zum Zusammenspiel von sozialer Eingebundenheit, Kompetenzerleben und Lehrverhalten

Theoretischer Hintergrund: Entsprechend der Selbstbestimmungstheorie (SDT; Deci & Ryan, 2000) stellen die Autonomie, das Kompetenzerleben und die soziale Eingebundenheit psychologische Grundbedürfnisse dar, deren Befriedigung im Kontext von Schule und Studium zu höherer intrinsischer Motivation, mehr Spaß und Interesse am Lernen sowie einer positiven Kompetenzentwicklung beiträgt. Empirische Studien mit Schülern und Studierenden konnten zeigen, dass bestimmte Formen des Lehrverhaltens und der Unterrichtsgestaltung die Befriedigung dieser Grundbedürfnisse fördern und mit den postulierten Vorteilen für Lernen und Leistung einhergehen (z.B. Haerens et al., 2013). Die Mehrzahl der Arbeiten konzentriert sich dabei auf autonomieförderliches Lehrverhalten (z.B. Reeve, 2009). Kaum untersucht wurde hingegen, welche Effekte eine simultane Förderung von Kompetenzerleben und sozialer Eingebundenheit auf den Kompetenzaufbau von Lernenden hat und ob eine Wechselwirkung zwischen beiden Faktoren besteht (Jang, Reeve, & Deci, 2010; Vansteenkiste et al., 2012).

Fragestellung: Anhand einer quasi-experimentellen Feldstudie wurde untersucht, ob ein nach den Prämissen der SDT systematisch variiertes Unterrichtskonzept das Kompetenzerleben und die soziale Eingebundenheit Studierender beim Moderieren von Arbeitsgruppen differenziell fördert und welche (wechselseitigen) Wirkungspfade zwischen der Erfüllung beider Grundbedürfnisse im Längsschnitt bestehen. Auf Basis der o.g. Studien zum Lehrverhalten wurde angenommen, dass autonomieförderlicher Unterricht mit transparenten Leistungserwartungen einen stetigen Zuwachs an Moderationskompetenz begünstigt (Kontrollgruppe = KG). Autonomieförderlicher Unterricht mit zusätzlichem begleitenden Feedback und emotionaler Unterstützung durch den Dozenten sollte mit einer noch stärkeren Kompetenzzunahme einhergehen (Experimentalgruppe = EG). Weiterhin wurden für beide Gruppen jeweils positive Zusammenhänge für den wahrgenommenen Kompetenzzuwachs bzw. das Gefühl sozialer Eingebundenheit zwischen aufeinanderfolgenden Messzeitpunkten erwartet. Für die EG wurden zusätzlich wechselseitige Wirkungspfade zwischen dem Kompetenzerleben und der Eingebundenheit postuliert.

Methode: Durchgeführt wurde eine quasi-experimentelle Feldstudie mit vier Messzeitpunkten innerhalb eines Semesters. Die Stichprobe umfasste Studierende zweier Hochschulen (N = 168; MAlter = 24.35; SD = 4.00), die einer EG (n = 87; MAlter = 24.44; SD = 3.66) bzw. einer KG (n = 81; MAlter = 23.23; SD = 4.32) zugeordnet wurden. Die Studie wurde im Rahmen von Seminaren zur Förderung der Moderationskompetenz realisiert, wobei jeweils vier Kurse mit ca. 20 Studierenden die KG und die EG bildeten. Beiden Gruppen lag ein autonomieförderliches Unterrichtskonzept zugrunde (vgl. Vansteenkiste et al., 2012; Reeve, 2009). Die EG zeichnete sich zusätzlich durch zahlreiche Übungsphasen mit Simulation einer Moderation unter Prüfungsbedingungen sowie durch prozessbegleitendes, individuelles Feedback und emotionale Unterstützung durch den Dozenten aus, um gezielt das Gefühl sozialer Eingebundenheit und des Kompetenzerlebens zu fördern. In der KG wurden die Inhalte weniger intensiv geübt und das Dozentenverhalten beschränkte sich auf informationelle Unterstützung und eine individuelle Rückmeldung nach den Moderationsprüfungen.

Mittels Fragebogen wurde die wahrgenommene soziale Eingebundenheit im Seminar (Friebe, 2005) sowie der erlebte Zuwachs der Moderationskompetenz (Braun, Gusy, Leidner, & Hannover, 2008) erfasst. Um Unterschiede der Zusammenhangsmuster zwischen beiden Gruppen zu analysieren, wurden Multigruppen-Pfadmodelle mittels Mplus 7.0 (Muthén & Muthén, 1998-2012) berechnet.

Ergebnisse: In beiden Gruppen war ein stetiger Zuwachs der Moderationskompetenz zu verzeichnen, der für die EG stärker ausfiel. Weiterhin waren in beiden Gruppen der erlebte Kompetenzzuwachs bzw. die erlebte soziale Eingebundenheit zwischen den benachbarten Messzeitpunkten hoch miteinander korreliert. Nur in der EG hingegen zeigte sich, dass der Kompetenzzuwachs und die soziale Eingebundenheit der Studierenden über verschiedene Messzeitpunkte hinweg in einer positiven Wechselbeziehung miteinander stehen. Eine Kombination aus Moderationsübungen unter Prüfungsbedingungen, begleitendem Feedback und ausgeprägtem Unterstützungsverhalten durch den Dozenten begünstigt demnach ein nachfolgend hohes Kompetenzerleben bei den Studierenden, welches wiederum mit einer höheren sozialen Eingebundenheit zum folgenden Zeitpunkt einhergeht. Die Befunde werden dahingehend interpretiert, dass ein autonomieförderlicher Unterricht, der zusätzlich durch prozessbegleitendes Feedback, ausreichend Trainingsphasen und eine sozio-emotionale Unterstützung der Studierenden gekennzeichnet ist, dafür verantwortlich ist, ob sich ein Lernender als sozial eingebunden und gleichzeitig kompetent wahrnimmt.

Donnerstag
12. März

15:30 Uhr

GB 03/49

Chair(s)
Gast, Anna

3

Merk, Samuel

Epistemologie (in) der Lehrerbildung?

Lehramtsstudierende lernen im Laufe ihrer akademischen Ausbildung eine Vielzahl an wissenschaftlichen Disziplinen und damit eine Vielfalt an methodischen Programmen der Erkenntnisgewinnung kennen (Hofer, 2000). Dabei decken sie nicht nur interindividuell mit dem Studium der jeweiligen Fachwissenschaften nahezu die gesamte erkenntnistheoretische Bandbreite einer Universität ab, auch intraindividuell variieren die wissenschaftlichen Methoden, mit denen Lehramtsstudierende konfrontiert werden (unabhängig von ihrer Fächerkombination), beispielsweise im bildungswissenschaftlichen Studium in Abhängigkeit von den jeweiligen Themen und Dozierenden. Forschungsarbeiten legen nahe, dass es für Lehramtsstudierende kaum zu leisten ist, aus diesen vielfältigen Erkenntnissen global-kohärente Wissensstrukturen aufzubauen, vielmehr sind in sich kohärente, lokale Wissensmodule zu erwarten (Bromme, 1992).

Diese Zergliederung hat Eingang in aktuelle Modelle der ‚Professionellen Kompetenz‘ gefunden (Baumert & Kunter, 2006). Es existiert allerdings kaum Forschung, die die erkenntnistheoretische Vielfalt des Lehramtsstudiums dezidiert zum Gegenstand hat. Dies ist erstaunlich, weil das Konstrukt der epistemologischen Überzeugungen (Überzeugungen bzgl. der Natur und Genese wissenschaftlichen Wissens) ein seit Jahrzehnten interdisziplinär bearbeiteter Forschungsgegenstand ist (Hofer 2000). Diese zumeist grundlagenorientierte Forschung hat lehr- lerntheoretische Relevanz: Epistemologische Überzeugungen präzisieren academic performance und conceptual change (Qian & Alvermann, 1995; Schommer, 1993), sind bei der Quellenwahl, Internetsuche und dem Onlinelernen relevant (Hofer, 2004; Porsch, 2011) und interagieren mit dem selbstreguliertem Lernen (Bromme, Kienhues & Stahl, 2008). Epistemologische Überzeugungen können metaphorisch, also sowohl als Katalysator (lehr- lerntheoretische Relevanz) als auch als Produkt (Teil professioneller Kompetenz) der Lehrerbildung beschrieben werden.

Das Symposium zielt darauf ab, sich der geschilderten Herausforderung der heterogenen wissenschaftstheoretischen Ansätze der Lehrerbildung mithilfe aktueller Forschungsergebnisse zu epistemologischen Überzeugungen aus vier Disziplinen (pädagogische Psychologie, Fachdidaktik, empirische Professionalisierungsforschung, Bildungstheorie) anzunehmen. Ein Beitrag aus pädagogisch-psychologischer Perspektive zum Verhältnis von Metakognition und epistemologischen Überzeugungen, sowie dessen Relevanz für das selbstregulierte Lernen (im Lehramtsstudium) stellt den Ausgangspunkt dar. Es folgt ein Beitrag über die Domänenspezifität epistemologischer Überzeugungen Lehramtsstudierender am Beispiel der Geschichtsdidaktik, der die Heterogenität der Epistemologie in der Lehrerbildung explizit aufnimmt. Der dritte Beitrag ist (nach der Topologie der GEBF 2015) disziplinär in der Lehrerbildung verankert und untersucht prädiktive Effekte epistemologischer Überzeugungen auf die Bedeutsamkeitseinschätzung bildungswissenschaftlicher Anteile, beschäftigt sich also mit den motivationalen Folgen der komplexen epistemologischen Situation in der Lehrerbildung. Abschließend ist ein bildungstheoretischer Beitrag geplant, da es u.E. günstig ist, die Betrachtung der „Epistemologie der Lehrerbildung“ aus der Perspektive empirischer Wissenschaften (Beitrag eins bis drei) um reflexive Denkformen zu ergänzen.

Donnerstag
12. März

15:30 Uhr

HGB 20

Chair(s)
Merk, Samuel

Pieschl, Stephanie

Wissen ist subjektiv, aber für die Klausur brauche ich nur die richtige Antwort! Epistemologische Überzeugungen und die Kalibrierungshypothese.

In diesem theoretischen Beitrag werden epistemologische Überzeugung (EÜ) und Kalibrierung definiert und ihr Zusammenhang anhand empirischer Beispiele aufgezeigt. EÜs sind persönliche Überzeugungen von Lernern über die Natur des Wissens. Sie beschäftigen sich mit Fragen nach dem Ursprung, der Natur und den Grenzen des Wissens und den Methoden, mit denen Menschen Wissen beweisen und rechtfertigen. Traditionelle Konzeptionen von EÜs gehen davon aus, dass diese sich von absolutistisch (z.B. Wissen ist eine Ansammlung von sicheren Fakten) zu relativistisch (z.B. Wissen ist ein komplexes Netzwerk veränderbarer und sozial ausgehandelter Bausteine) weiterentwickeln. Dabei wird angenommen, dass es sich bei EÜs um relativ stabile Personenmerkmale handelt, die in vielen verschiedenen Kontexten relevant sind. Empirische Befunde zeigen hingegen, dass sich EÜs zwischen Domänen unterscheiden und dass nicht durchgängig relativistische Überzeugungen besseren Lernerfolg vorhersagen. Aus diesem Grund beschreiben neuere Konzeptionen von EÜs diese einerseits als „Linse“ durch die alle Lernaufgaben und -inhalte wahrgenommen werden (Bromme, Pieschl, & Stahl, 2010). Andererseits wird betont, dass EÜs nicht in allen Kontexten gleichermaßen eine Rolle spielen, sondern flexibel „angeschaltet“ werden (Greene, Muis, & Pieschl, 2010). In diesem Sinne sind sophistische EÜs eine Voraussetzung dafür, dass Lernaufgaben und -inhalte genau wahrgenommen werden und kontextspezifisch über die Natur des Wissens reflektiert wird.

Theoretisch müsste sich der Einfluss von EÜs in vielen verschiedenen Lehr-Lern-Kontexten untersuchen lassen. Vor allem dort sollte sich die Überlegenheit von sophistischen EÜs zeigen, wo Lerner ihre Lernprozesse an Kontextbedingungen anpassen.

Solch eine Adaptivität an Kontextbedingungen ist ein Kernbestandteil vieler Modelle des selbstregulierten Lernens. Beispielsweise sollten gute selbstregulierte Lerner ihre Lernprozesse an die Komplexität der vorliegenden Lernaufgaben anpassen (Bromme, Pieschl, & Stahl, 2010; Pieschl, Stahl, & Bromme, 2013; Pieschl, Stahl, Murray, & Bromme, 2012). Aufgabenkomplexität in diesem Zusammenhang ist nicht gleichzusetzen mit Aufgabenschwierigkeit, sondern bezieht sich auf die Komplexität der kognitiven Operationen, die zur Lösung einer Aufgabe notwendig sind. Einfache Aufgaben in diesem Sinne erfordern eine Reproduktion von Lernmaterialien während komplexere Aufgaben beispielsweise erfordern, dass das Gelernte angewendet oder bewertet wird. Im Rahmen des selbstregulierten Lernprozesses sollten Lerner zunächst die unterschiedliche Komplexität von Aufgaben wahrnehmen und daraufhin angemessene Lernstrategien planen und diese durchführen. Wenn Lernern dies erfolgreich gelingt – beispielsweise verwenden sie viel Zeit und kognitive Ressourcen auf komplexe Aufgaben und wenig Zeit und kognitive Ressourcen auf einfache Aufgaben – dann spricht man auch von einer gelungenen Kalibrierung (Pieschl, 2009): Die Lerner passen ihre Lernprozesse genau an die Anforderungen der jeweiligen Aufgabe an.

Bringt man dies mit EÜs in einen Zusammenhang, so kann man Folgendes annehmen: Lernern mit sophistischen EÜs sollte es besser gelingen, ihre Lernprozesse an die Aufgaben anzupassen (Bromme, Pieschl, & Stahl, 2010; Pieschl, Stahl, & Bromme, 2013). Sie nehmen schon die Aufgabenanforderungen angemessener wahr und sie sollten besser dazu in der Lage sein, jeweils an angemessener Stelle und in angemessenem Ausmaß über die Natur des zugrunde liegenden Wissens zu reflektieren. Dies sollte sich vor allem bei komplexen Lernaufgaben zeigen. Wenn es beispielsweise darum geht, verschiedene Lerninhalte vergleichend zu bewerten, dann sollten EÜs eine wichtige Rolle spielen. Lerner, die von sicherem unveränderlichem Wissen ausgehen, werden sich vermutlich auf wenige Informationsquellen verlassen und diese nicht unbedingt kritisch hinterfragen. Lerner, die von veränderbarem Wissen ausgehen, sollten dagegen einerseits mehrere Informationsquellen hinzuziehen, andererseits mehr Sorgfalt auf die Auswahl der bestmöglichen Quellen verwenden. Dies sollte sich nicht nur im Prozess der Aufgabenbearbeitung sondern auch in der Qualität der Aufgabenlösung widerspiegeln.

Diese theoretischen Annahmen werden in diesem Beitrag durch empirische Beispiele illustriert (Bromme, Pieschl, & Stahl, 2014; Pieschl, Bromme, Porsch, & Stahl, 2008; Pieschl, Stahl, Murray, & Bromme, 2012; Pieschl, Stallmann, & Bromme, 2014; Stahl, Pieschl, & Bromme, 2006) und ihre Implikationen für die Lehr-Lern-Praxis werden diskutiert.

Donnerstag
12. März

15:30 Uhr

HGB 20

Chair(s)

Merk, Samuel

Nitsche, Martin; Waldis, Monika

Die Erfassung von epistemologischen und lehr-lerntheoretischen Beliefs von Schweizer Geschichtslehrpersonen - qualitative und quantitative Messinstrumente

Der professionellen Kompetenz von Lehrpersonen kommt Bedeutsamkeit für das Lernen von Schülerinnen und Schüler zu. Dabei kommt nicht nur dem Fachwissen und dem fachdidaktischen Wissen sondern auch den Überzeugungen zur Wissensstruktur und zu Lernprozessen im jeweiligen Fach eine wichtige Bedeutung zu (Voss, Kleickmann, Kunter & Hachfeld 2011). Im Fach Geschichte wird von einem engen Zusammenhang zwischen fachdidaktischem Wissen und epistemologischen und lerntheoretischen Überzeugungen ausgegangen (Waldis, Nitsche, Marti, Hodel & Wyss 2014). Hierbei wird an die psychologische Forschung angeknüpft, die zeigte, dass epistemologischen Beliefs Einfluss auf Erkenntnis- und Lernprozesse zukommt (Hofer & Pintrich 1997). Für das Fach Geschichte liegt die Herausforderung darin, epistemologische und fachdidaktische Beliefs empirisch reliabel und valide zu erfassen. Dies liegt u.a. daran, dass Versuche der Erfassung domänenspezifischer Beliefs im Fach Geschichte mittels geschlossener Fragebogenitems bisher nur teilweise erfolgreich waren. Obwohl in Anlehnung an King und Kitchener (1994) bzw. Kuhn und Weinstock (2002) von wenigstens vier Formen ausgegangen wird (criterialists, relativists, naïve realists, or dichotomous reasoners), konnten bislang nur zwei Formen mittels geschlossener Items abgefragt und faktoranalytisch als eigenständige Dimensionen modelliert werden (Maggioni, VanSledright & Alexander 2009). Im Rahmen des Projekts „VisuHist – Ausprägung und Genese professioneller Kompetenz von Geschichtslehrpersonen“ wird erneut die Erfassung domänenspezifischer Beliefs angegangen. Im Vortrag wird danach gefragt a) wie epistemologische Beliefs von Geschichtslehrpersonen erfasst werden können, b) welche Formen epistemologischer Beliefs im Fach Geschichte bei angehenden Geschichtslehrpersonen in der Schweiz zu finden sind und c) wie diese mit lerntheoretischen Überzeugungen zusammenhängen. Dafür wurden zunächst im Rahmen eines Paper & Pencil-Tests bei rund 100 Geschichtslehrerstudierenden die Frage erhoben, wie historisches Wissen entsteht und wodurch es charakterisiert ist. Mittels inhaltsanalytischer Verfahren konnte in Anlehnung an die Geschichtstheorie (z.B. Lorenz 1997) drei Typen bei den Studierenden unterschieden werden (Positivist, Indifferentes Denken, Konstruktivist). Davon ausgehend wurde auf der Grundlage der vorliegenden Versuche zu den epistemologischen Beliefs im Fach Geschichte (Maggioni 2010) und den lerntheoretischen Beliefs (Messner & Buff 2007; Schlichter 2012) geschlossene Fragebogenitems entwickelt und bei angehenden Geschichtslehrpersonen erprobt. Während für die lerntheoretischen Beliefs die theoretisch angedachte Struktur faktorenanalytisch unterschieden werden konnte (Instruktional, Konstruktivistisch, Sozialkonstruktiv), war dies bei den epistemologischen Beliefs nicht der Fall. Auf der Basis dieser offen und geschlossen erfassten empirischen Daten können erste Aussagen zur Ausprägung lerntheoretischer und epistemologischer Beliefs bei Schweizer Geschichtslehrpersonen gemacht werden.

Donnerstag
12. März

15:30 Uhr

HGB 20

Chair(s)
Merk, Samuel

Merk, Samuel; Cramer, Colin

Prädiktive Effekte epistemologischer Überzeugungen auf die Bedeutsamkeitseinschätzung bildungswissenschaftlicher Studienanteile Lehramtsstudierender

Theoretischer Hintergrund und Fragestellung: Den bildungswissenschaftlichen Studienanteilen wird von Lehramtsstudierenden eine vergleichsweise geringe Bedeutung zugeschrieben (Terhart 2009). In der Literatur lassen sich dafür mindestens drei Erklärungsansätze identifizieren. Der erste Erklärungsansatz rekurriert auf das wissenschaftstheoretisch komplexe Verhältnis von Theorie und Praxis in der Lehrerbildung (Patry 2014). Die bildungswissenschaftlichen Disziplinen, die auf Vermehrung wissenschaftlicher Kenntnis zielen, können nicht unmittelbar für die Profession nutzbares Wissen, das unter Handlungszwängen angewendet werden muss, generieren. Ein zweiter Erklärungsansatz betrachtet die Differenzen wissenschaftlichen und handlungsleitenden Wissens aus kognitionspsychologischer Perspektive. Während in akademischen Lehrveranstaltungen zumeist deklaratives oder prozedurales Wissen gelehrt wird, wäre für effektives Handeln vernetztes konzeptuelles Wissen notwendig (Kunter 2011). Ein dritter Ansatz schließlich besteht darin, empirisch (proximale) Prädiktoren für die geringe Bedeutsamkeitseinschätzung, wie z.B. die Diskrepanz zwischen erwarteten und erfahrenen Lehrinhalten, zu identifizieren und diese über diverse Institutionen der Lehrerbildung oder über verschiedene Professionen hinweg vergleichend zu betrachten (Cramer 2013). Beide theoretischen Erklärungsansätze postulieren also eine für die Lehrerbildung genuine und wissenschaftstheoretisch nicht überwindbare Diskrepanz zwischen akademischer Betrachtung und handelnder Anwendung der Gegenstände bildungswissenschaftlicher Studienanteile.

Die vorliegende Arbeit leitet aus dieser Diskrepanz die Hypothese einer Assoziation der subjektiven wissenschaftstheoretischen Deutung bildungswissenschaftlichen Wissens und der Bedeutsamkeitseinschätzung desgleichen ab. Eine empirische Prüfung dieser Hypothese scheint insbesondere deshalb machbar, da die Konzeptualisierung und Erfassung subjektiver (Laien-)Wissenschaftstheorie unter dem Schlagwort der epistemologischen Überzeugungen in den letzten zwei Jahrzehnten zunehmend Aufmerksamkeit der (pädagogisch-psychologischen) Forschung gewonnen hat (Trautwein & Lüdtke 2004). So liegt z.B. erste Evidenz für die Funktionalität epistemologischer Überzeugungen bei der Zielsetzung und Bewertung von Lernaufgaben vor (Bromme, Pieschl, Stahl 2010). Im Beitrag wird geprüft, inwiefern epistemologische Überzeugungen prädiktiv für die Bedeutsamkeitseinschätzung allgemeinen pädagogischen Wissens (GPK) und fachdidaktischen Wissens (PCK) ist.

Methode: Bei der untersuchten Stichprobe handelt es sich um eine Substichprobe (Studienreferendarinnen und -referendare) einer quasi-längsschnittlichen Kohortenstudie. Sie weist rotierte Skalen zu epistemologischen Überzeugungen bzgl. diverser Domänen sowie motivationaler und emotionaler Konstrukte auf (Ntot = 1028, Nsub = 198). Epistemologische Überzeugungen bzgl. GPK und PCK wurden mithilfe der Connotative Aspects of Epistemological Beliefs (CAEB) (Stahl & Bromme 2007) und des Fragebogens zum Entwicklungsniveau epistemologischer Überzeugungen (FREE) (Krettenauer 2005) erfasst. Für die Erfassung der Bedeutsamkeitseinschätzung wurde eine intraindividuelle mit einer interindividuellen Perspektive kombiniert: Von allen Befragten wurde die Bedeutsamkeitseinschätzung allgemeinen pädagogischen und fachdidaktischen Wissens, sowie die selbsteingeschätzte Kompetenz bzgl. aller acht KMK-Standards zur Lehrerbildung aus den Bereichen Unterrichten, Erziehen und Beurteilen erfasst. Für diese drei Variablen liegen also genestete Daten vor (in jeder Person acht Ratings; eines je Kompetenz).

Deshalb wurden zunächst Intra-Klassenkorrelationen geschätzt, um das Ausmaß der „Kompetenzspezifität“ der Ratings der Bedeutsamkeitseinschätzung zu eruieren. Die Schätzung der prädiktiven Effekte erfolgte mit Zwei-Ebenen-Strukturgleichungsmodellen (Mehta & Neale, 2005) um mehrere abhängige Variablen modellieren zu können und die Nestung der Daten zu berücksichtigen. Als Prädiktoren des wahrgenommenen Nutzens GPK und des wahrgenommenen Nutzens PCK wurden auf der Zwischen-Personen Ebene Dummyvariablen für die einzelnen Kompetenzen sowie die selbsteingeschätzte Kompetenz spezifiziert. Auf der Zwischen-Personen Ebene wurden die erfassten domänenspezifischen epistemologischen Überzeugungen, also die Dimensionen des CAEB und des FREE als Prädiktoren herangezogen.

Ergebnisse: Die Intra-Klassenkorrelationen (hier eher: „Intrapersonenkorrelationen“) liegen zwischen .14 und .24, weisen also auf eine starke Kompetenzspezifität dieser Variablen hin. Diese Kompetenzspezifität lässt sich auch mit Friedman-Tests nachweisen.

Der Identifikation prädiktiver Effekte auf die Bedeutsamkeitseinschätzungen wurde in einer Serie von random intercepts (Zwei-Ebenen-)Strukturgleichungsmodellen nachgegangen. Zunächst wurde ein Modell spezifiziert, das lediglich Indikatorvariablen für die entsprechende KMK-Kompetenz, sowie das Ausmaß der selbsteingeschätzten Kompetenz als Prädiktoren auf der Inner-Personen-Ebene aufwies. Dieses Modell zeigt eine sehr hohe Anpassungsgüte und bedeutsame (Pseudo-)determinationskoeffizienten. Ein Modell das außerdem auf Zwischen-Personen-Ebene noch die Dimensionen des CAEBs und des FREEs als Prädiktoren für die Zwischen-Personen-Komponenten der abhängigen Variablen Bedeutsamkeitseinschätzung-GPK und Bedeutsamkeitseinschätzung-PCK enthielt, wies eine befriedigende Modellpassungsgüte und hohe Determinationskoeffizienten auf.

Donnerstag
12. März

15:30 Uhr

HGB 20

Chair(s)

Merk, Samuel

Harant, Martin

Bildungstheoretische Reflexionen zum Wissenschaftlichen in der Lehrerbildung

Der Begriff der (prinzipiell unbestimmten) Bildsamkeit als Grundbegriff der Pädagogik geht von Beginn an mit einem idealistischen Wissenschaftsverständnis einher, das sich – anders als der Begriff der modernen science a) durch den Gedanken einer auf Anlage- und Umweltbedingungen irreduziblen transzendental begründeten Subjektkonstitution auszeichnet und b) einen emphatischen Praxisbegriff mit einschließt, der auf der praktischen Philosophie aufruhend systematisch zwischen Handeln und Verhalten unterscheidet, sowie handlungsorientierende Gewissheiten in Form von praktischen Ideen beinhaltet (vgl. Herbart 1982 & Schleiermacher 1983). Der Bildungsbegriff im Wortkompositum Lehrerbildung bewegt sich seit der Verwissenschaftlichung der Pädagogik im 19. Jahrhundert in diesem idealistisch geprägten Verstehenshorizont.

Spätestens seit den 1960er Jahren gewinnt in Deutschland demgegenüber eine forschungsbasierte Verwissenschaftlichung der Lehrerbildung an Bedeutung und tritt an die Stelle des im obigen Sinne bildungstheoretisch begründeten und implizit ‚science-kritisch‘ ausgerichteten Einflusses der Pädagogik (vgl. Litt 1921/1952; vgl. Spranger 1920/1970), die sich als Handlungswissenschaft bzw. als hermeneutische Wissenschaft begreift (Benner 1994).

Die Tatsache der Exklusion der dem Bildungsverständnis inhärenten praktisch-philosophischen Reflexionsanteile aus dem genuin Wissenschaftlichen, das nunmehr eine empirisch-analytische Zuspitzung erfährt, führt im Gefolge auf philosophischer Seite zu der Feststellung:

„Die philosophische Überzeugung des Deutschen Idealismus, dass Wissenschaft bilde, trifft auf die empirisch-analytischen Verfahrensweisen nicht mehr zu“ (Habermas 1963, 362).

Kann unter dem Primat von science noch von Lehrerbildung im originären Sinne gesprochen werden? Gibt es eine Lehrerbildung jenseits des idealistisch konnotierten Wissenschaftsbegriffs?

Das veränderte Wissenschaftsverständnis führt in jedem Fall zu einer Neujustierung des Verhältnisses vom Wissenschaftlichen zur Bildung und zu einem veränderten Verständnis von Bildung. Die Annäherung an das Wissenschaftliche im Sinne der science erfolgt dabei interessanterweise sowohl im Horizont einer im obigen Sinne bildungstheoretisch abstinenter empirisch-analytischer Neuausrichtung von Erziehungswissenschaft (Brezinka 1978) und Bildungsforschung als auch im Gefolge bildungstheoretischer Reflexionen im Bereich der Kritischen Erziehungswissenschaft, in der die Forderung nach uneingeschränkter Verwissenschaftlichung von Lehr-Lern-Kontexten in kritisch-emanzipatorischer Absicht erhoben wird (Blankertz 1971). Dass die bildungstheoretisch motivierte Problematisierung des Verhältnisses von Bildung und Wissenschaft, wie sie von der idealistisch konnotierten Pädagogik vorgetragen wurde gleichwohl bleibend fortbesteht, verdeutlicht zum einen die im Rahmen seiner Studie Hermeneutik und Bildung von Buck gegenüber der Wissenschaftsentwicklung hin zur modernen science vorgebrachten Einwände (1981, 185) „die Bildung durch Wissenschaft unmöglich erscheinen lassen“ (Buck 1981, 185). Zum anderen verdeutlicht sie sich in der von Benner & Brüggem (2011, 327) erhobenen Grundsatzkritik gegenüber der Kritischen Erziehungswissenschaft, dass deren am Science-Verständnis ausgerichtete Verwissenschaftlichung von Bildungsprozessen nicht an den durch die geisteswissenschaftliche Pädagogik im Rahmen des Szientismusproblems eröffneten Diskurs „um das prekär gewordene Verhältnis von Bildung und Wissenschaft anknüpft“.

Damit stehen sich historisch und systematisch bedingt bis in die Gegenwart drei grundsätzlich verschiedene und idealtypisch rekonstruierbare Verstehenshorizonte gegenüber, das Verhältnis von Bildung und Wissenschaft im Sinne von science zu fassen, die jeweils zu unterschiedlichen Verständnissen dessen führen, was das Wissenschaftliche der Lehrerbildung im Grunde sein kann und was es nicht sein kann: Dabei lässt sich eine problematisierende, eine synthetisierende und eine szientifisch ausgerichtete Verhältnisbestimmung unterscheiden, wobei letztere sich begriffsgeschichtlich betrachtet der idealistischen Konnotationen des Bildungsbegriffs und seiner theoretischen Reflexion prinzipiell entledigt.

Der Teilbeitrag geht von der These aus, dass diese idealtypisch rekonstruierbaren Verstehenshorizonte die Verhältnisbestimmung des Wissenschaftlichen zur Lehrerbildung bereits im Vorfeld präformieren.

Er beschreibt dazu in einem ersten Schritt das Spannungsfeld der Gedankenlinien zum Verhältnis von Bildung und Wissenschaft, indem er auf Grundlage einer hermeneutischen Analyse die drei dargelegten idealtypischen Verstehenshorizonte dieser Verhältnisbestimmung rekonstruiert und dabei die Grundlinien der Lehrerbildung in die Rekonstruktion der möglichen Verstehenshorizonte der Verhältnisbestimmung von Bildung und Wissenschaft einzeichnet. (vgl. Herrmann 2000)

In einem zweiten Schritt wird resümiert, ob und ggf. inwiefern genuin bildungstheoretische Reflexionen zum Wissenschaftlichen aus dem Horizont der idealistischen Philosophie und aus dem Horizont der Kritischen Erziehungswissenschaft für empirische Bildungsforschungskontexte, etwa im Rahmen der Expertiseforschung (vgl. Bromme 2008b), als anschlussfähig gelten können.

Donnerstag
12. März

15:30 Uhr

HGB 20

Chair(s)

Merk, Samuel

3

von Kotzebue, Lena

Zum Einfluss des Professionswissens auf das Unterrichtshandeln (ProwiN-Video)

Das Professionswissen von Lehrkräften wird als eine wesentliche Voraussetzung für erfolgreiches Unterrichtshandeln betrachtet (z.B. Abell, 2007; Bransford et al., 2005; Kunter et al., 2011; Wahl, 1991). In Anlehnung an die von Shulman (1986, 1987) vorgenommene Konzeptualisierung hat sich im Bereich der Bildungsforschung die Unterscheidung von fachlichem (CK), fachdidaktischem (PCK) und pädagogisch Professionswissen (PK) durchgesetzt (Abell, 2007; König & Blömeke, 2009; König & Seifert, 2012; Kunter et al., 2011; Lipowsky, 2006). Empirische Evidenz für den Einfluss des Professionswissens auf das Unterrichtshandeln gibt es bislang wenig. Erste Befunde im Rahmen der COACTIV-Studie deuten auf einen positiven Einfluss von PCK auf die kognitive Aktivierung sowie von PK auf die konstruktive Lernerunterstützung und die Klassenführung in Bezug auf das Fach Mathematik hin (Baumert et al., 2010). Das Unterrichtshandeln wurde im Rahmen der COACTIV-Studie allerdings eher indirekt über Unterrichtsdokumente oder Schülerfeedback erfasst. In diesem Symposium werden Studien aus dem BMBF-geförderten Projekt ProwiN (Professionswissen in den Naturwissenschaften) vorgestellt, welche das Unterrichtshandeln in den Fächern Biologie, Chemie und Physik anhand von Unterrichtsvideos analysieren und mit dem Professionswissen sowie der Schülerleistung in Bezug setzen.

Zentrale Fragestellung des Symposiums ist die Frage nach dem Zusammenhang zwischen dem Professionswissen der Lehrkräfte, ausgewählten Facetten des Unterrichtshandelns und der Schülerleistung.

Alle im Symposium beinhalteten Teilprojekte verwenden zur Erfassung des Professionswissens Paper-Pencil-Tests, welche im Rahmen der ersten Projektphase auf Basis eines gemeinsamen Modells entwickelt und validiert wurden (Dollny, 2011; Jüttner et al. 2013; Kirschner, 2013; Lenske et al., in Vorb.). In den Einzelbeiträgen wird sowohl auf unterschiedliche Fächer als auch unterschiedliche Dimensionen des Professionswissens fokussiert.

Der erste Beitrag widmet sich dem Zusammenhang des Unterrichtshandelns im Chemieunterricht mit CK und PCK. Im zweiten Beitrag wird ebenfalls der Zusammenhang zwischen dem fachspezifischen Professionswissen und dem Unterrichtshandeln untersucht, wobei auf den Umgang mit Modellen im Biologieunterricht fokussiert wird. Der dritte Beitrag betrachtet inwieweit CK und PCK das kognitive Aktivierungspotential des Physikunterrichts und den Leistungszuwachs der Schüler beeinflusst und hinterfragt kritisch die Testinstrumente. Abschließend wird unter Verwendung der gleichen Stichprobe des vorherigen Beitrags über den Einfluss von PK auf das Unterrichtshandeln mit Blick auf die Qualitätskriterien Klassenführung, kognitive Aktivierung und Umgang mit Heterogenität berichtet. Darüber hinaus wird überprüft, ob sich das Professionswissen vermittelt über das Unterrichtshandeln auf die Schülerleistung auswirkt.

Die Ergebnisse der Einzelbeiträge werden im Hinblick auf theoretische und praktische Implikationen jeweils von den Vortragenden diskutiert. Abschließend werden alle Einzelbeiträge des Symposiums diskutiert, wobei auf einzelne Schwächen und Stärken eingegangen wird und Forschungsdesiderata abgeleitet werden.

Donnerstag
12. März

15:30 Uhr

HGB 30

Chair(s)

von Kotzebue, Lena
Lenske, Gerlinde

Tröger, Holger; Strübe, Martina; Sumfleth, Elke; Tepner, Oliver

Fachspezifisches Professionswissen von Chemielehrkräften – ProwiN Videostudie

Aus Arbeiten zum Professionswissen von Lehrkräften lassen sich Hinweise auf einen Zusammenhang zwischen dem Professionswissen von Lehrkräften und der Unterrichtsqualität sowie dem Lernerfolg der Schülerinnen und Schüler ableiten (Abell, 2007; Krauss & Kunter, 2008). Von den nach Shulman (1987) sieben Dimensionen des Professionswissens, die für das Lehren von Bedeutung sind, werden in der gegenwärtigen Forschung vornehmlich drei diskutiert: pädagogisches Wissen (PK), Fachwissen (CK) und fachdidaktisches Wissen (PCK) (Baumert et al., 2010). Der Zusammenhang zwischen dem fachspezifischen Professionswissen (CK & PCK) von Lehrkräften (Dollny, 2011), ihrem Unterrichtshandeln und dem Lernerfolg ihrer Schülerinnen und Schüler ist für das Fach Chemie jedoch weitgehend ungeklärt.

Im Fach Chemie wird dieser Zusammenhang anhand des Unterrichts von gymnasialen Chemielehrkräften der 8. oder 9. Jahrgangsstufe zum Thema „Atombau und Periodensystem der Elemente“ untersucht. Im Rahmen von zwei Studien werden der Umgang der Chemielehrkräfte mit Fachsprache und Schülerfehlern (Studie 1) sowie mit Modellen und Experimenten (Studie 2) erforscht und überprüft, ob ein Zusammenhang zwischen dem fachspezifischen Professionswissen der Lehrkräfte hinsichtlich der jeweiligen Aspekte, ihrem diesbezüglichen Unterrichtshandeln sowie dem Lernerfolg der Lernenden besteht. Die zugrundeliegenden Hypothesen lauten, dass sich Lehrkräfte mit hohem fachspezifischen Professionswissen in ihrem Unterrichtshandeln von Lehrkräften mit niedrigem fachspezifischen Professionswissen unterscheiden und dass Schülerinnen und Schüler von Lehrkräften mit hohem fachspezifischen Professionswissen einen höheren Lernzuwachs aufweisen als die Lernenden von Lehrkräften mit niedrigem fachspezifischen Professionswissen.

Diese Hypothese wird unter Verwendung eines Multi-Methoden-Designs erforscht. Das theoretische fachspezifische Professionswissen der Chemielehrkräfte sowie ihre fachspezifischen Einstellungen und der Lernerfolg der Lernenden werden mithilfe schriftlicher Testinstrumente erhoben. Dabei wird auf Testinstrumente, die im Rahmen der ersten Projektphase von ProwiN entwickelt wurden (Dollny, 2011), sowie auf zusätzlich entwickelte Testinstrumente zurückgegriffen. Das Unterrichtshandeln der Chemielehrkräfte wird mithilfe von Videoanalysen untersucht. Die Analyse der Videodaten erfolgt unter Verwendung selbst konstruierter Videokodiermanuale mit den Schwerpunkten auf dem Umgang der Lehrenden mit Fachsprache, Schülerfehlern, Modellen und Experimenten.

Bei der Konstruktion der Manuale wurde auf Videodaten vergangener Studien (Mutke, in Vorbereitung; Pollender, in Vorbereitung) und zum Umgang mit Experimenten auf ein evaluiertes Videokodiermanual zu experimentierspezifischen Qualitätsmerkmalen im Chemieunterricht (Schulz, 2011) zurückgegriffen. Der Umgang mit Schülerfehlern wird in Anlehnung an Pollender (in Vorbereitung) analysiert und der Umgang der Lehrenden mit Modellen wird anhand modellspezifischer Qualitätsmerkmale, in Anlehnung an Schulz (2011) theoriebasiert entwickelt, beurteilt. Der Umgang mit Fachsprache wird in Anlehnung an etablierte Modellierungen inhaltlicher Komplexität (Kauertz, Fischer, Mayer, Sumfleth, & Walpuski, 2010; Lau, 2011) hinsichtlich inhaltlicher und sprachlicher Komplexität untersucht.

Erste Analysen der schriftlichen Testdaten, die im Schuljahr 2013/2014 erhoben wurden, zeigen einen signifikanten Zuwachs des Fachwissens der Lernenden ($n = 385$) zwischen Prätest ($\alpha = .40$, $M = 10.85$, $SD = 3.13$) und Posttest ($\alpha = .72$, $M = 16.35$, $SD = 4.52$), $t(384) = 25.32$, $p < .001$, $d = 1.42$. Die Analyse zeigt weiterhin, dass sich die Lerngruppen hinsichtlich des Wissenszuwachses hoch signifikant unterscheiden $F(1,16) = 8.44$, $p < .001$, $\eta_p^2 = .27$. Die schriftlichen Testinstrumente weisen gute Reliabilitäten ($\alpha_{CK}\text{-Test} = .71$, $\alpha_{PCK}\text{-Test} = .87$) auf, zeigen ein hohes Fachwissen der Lehrkräfte ($n = 13$) und eine hohe Varianz hinsichtlich des generellen fachdidaktischen Wissens.

Um die Aussagekraft der Analysen zu erhöhen, wurde beschlossen, weitere Daten im Schuljahr 2014/2015 zu erheben. Die Daten dieser Erhebungen werden voraussichtlich bis Ende des Kalenderjahres vorliegen und werden bis zur Tagung ausgewertet.

Im Vortrag werden die dann vorliegenden Ergebnisse der Videoauswertung präsentiert und erste Ergebnisse zum Zusammenhang zwischen dem theoretischen fachspezifischen Professionswissen der Lehrkräfte, ihrem erfassten Unterrichtshandeln sowie dem Lernzuwachs der Schülerinnen und Schüler vorgestellt.

Donnerstag
12. März

15:30 Uhr

HGB 30

Werner, Sonja; Förtsch, Christian; von Kotzebue, Lena; Neuhaus, Birgit J.

Zusammenhang zwischen Professionswissen von Lehrkräften, Modelleinsatz und Schülervariablen – eine quantitative Videostudie im Biologieunterricht

Unterrichtsqualität und Professionswissen von Lehrkräften werden als zentrale Bedingungen für die Lernleistung von Schülerinnen und Schüler angenommen (vgl. u.a. Neuhaus, 2007). Das Professionswissen kann durch die drei Dimensionen Fachwissen (CK), fachdidaktisches Wissen (PCK), sowie pädagogisches Wissen (PK) operationalisiert werden und besitzt eine zentrale Bedeutung für das professionelle Handeln einer Lehrkraft (Abell, 2007; Baumert et al., 2010). Der Zusammenhang zwischen dem Professionswissen, der Unterrichtsqualität und den Schülervariablen wurde bereits in einigen Studien im Fach Mathematik untersucht (vgl. Baumert et al., 2010; Hill, Rowan & Loewenberg Ball, 2005). Hierbei erfassten Baumert et al. (2010) jedoch den Unterricht nur indirekt über Aufgaben. Hill et al. (2007) hingegen konnten durch Analyse von Unterrichtsvideos Zusammenhänge zwischen allen drei Ebenen nachweisen, allerdings nur bei einer kleinen Stichprobe (N = 10). Unterrichtsqualität wird dabei durch verschiedene Unterrichtsaspekte erfasst. Ein wichtiges biologiespezifisches Unterrichtsqualitätsmerkmal stellt der Einsatz von Modellen dar (Wüsten et al., 2011). Bisherige Forschung in diesem Bereich fokussiert jedoch lediglich die Erfassung des Modellverständnisses von Lernenden und Lehrenden (vgl. Grosslight et al., 1991; Justi & Gilbert, 2002), sowie die Operationalisierung von Modellkompetenz und deren empirische Überprüfung (vgl. Grünkorn et al., 2014; Upmeyer zu Belzen & Krüger, 2010). Basierend auf den dargelegten theoretischen Ausführungen untersucht die vorliegende im ProWiN-Projekt eingebettete Studie folgende Fragestellung: Welche Zusammenhänge bestehen zwischen dem Professionswissen einer Biologielehrkraft, dessen Modelleinsatz im Unterricht und verschiedenen Schülervariablen? Hierfür wurden 42 Biologielehrkräfte (Berufserfahrung ohne Referendariat in Jahren: M = 5,5, SD = 5,6; Alter in Jahren: M = 35,0, SD = 8,1) für jeweils zwei Unterrichtsstunden in der 9. Jahrgangsstufe (Thema: Neurobiologie) videographiert (N = 84 Videos). Von allen teilnehmenden Lehrkräften wurde mit Hilfe eines Fragebogens (CK und PCK in Biologie: Jüttner et al., 2013; PK: Lenske et al., in Vorb.) das Professionswissen erhoben. Nach jeder videographierten Stunde füllten die Lernenden einen Fragebogen zum situationalem Interesse (N = 3 Items; vier-stufige Likert-Skala) aus. Alle Schülerinnen und Schüler erhielten außerdem sowohl einen Fragebogen zu ihrem Interesse am Fach Biologie (N = 14 Items; vier-stufige Likert-Skala) als auch einen Prä-Post-Leistungstest. Um den Modelleinsatz im Unterricht valide zu erfassen, wurde theoriebasiert ein event-basiertes Kategoriensystem mit den Skalen Modellarbeit und Modellkritik entwickelt. Die Skala Modellarbeit umfasst u.a. die Subskalen Zweck des Modells und sinnvoller Einsatz des Modells. Die Skala Modellkritik wird u.a. durch Intensität weiter differenziert. Zusätzlich dazu wurde jede identifizierte Modellphase bezüglich des Modellverständnisses (Grosslight et al., 1991) geratet. Die Tests zum fachdidaktischen Wissen (N = 9 Items) und Fachwissen (N = 12 Items) gelten als validiert. Sie wurden Rasch-skaliert und die Daten wiesen auf eine produktive Messung hin (alle Infit-/Outfit-MNSQ < 1,3). In einer Teilstichprobe von 64 Unterrichtsvideos konnten bisher 141 Phasen der Modellarbeit und 25 Phasen der Modellkritik identifiziert werden. In den meisten Fällen (81 %) diente das Modell zur Veranschaulichung, in lediglich 19 % der Phasen als Mittel der Erkenntnisgewinnung. Insgesamt wurden in 62,4 % der Modelle passend zum Stundenziel eingesetzt. In den 25 vorhandenen Modellkritik-Phasen wurde lediglich in 16 % dieser Phasen die Modellkritik als eigener Unterrichtspunkt bzw. in einer Diskussion tiefer gehend behandelt. In einer Teilstichprobe konnten signifikante Zusammenhänge zwischen dem fachdidaktischen Wissen und dem Modellverständnis im Unterricht ($r = 0.4$; $p = .02$) identifiziert werden. Ergänzend dazu sollen Zusammenhänge zwischen dem Modelleinsatz im Unterricht und Schülervariablen untersucht werden. Ergebnisse der Analyse werden mit der Gesamtstichprobe von 42 Lehrkräften und 84 Videos auf der GEBF-Tagung vorgestellt. Die Ergebnisse der Studie sollen zur Optimierung der Ausbildung von Lehrkräften an den Universitäten im Bereich der Biologie genutzt werden.

Donnerstag
12. März

15:30 Uhr

HGB 30

Chair(s)

von Kotzebue, Lena
Lenske, Gerlinde

Zum Einfluss des Professionswissens auf das Unterrichtshandeln (ProWiN-Video)

Cauet, Eva; Borowski, Andreas; Fischer, Hans

Der ProwiN Professionswissenstest Physik - Erfassen wir handlungsrelevantes Wissen?

Das Professionswissen von Lehrkräften wird seit langem als wichtige Voraussetzung für effektives Unterrichten diskutiert (Peterson, Carpenter & Fennema, 1989; Abell, 2007). Motiviert durch die COACTIV Studie, die zeigen konnte, dass das fachspezifische Professionswissen von Mathematiklehrkräften über das kognitive Aktivierungspotenzial von im Unterricht eingesetzten Aufgaben auf Schülerleistung wirkt (Baumert et al., 2010), wurden in den vergangenen Jahren mehrere Tests zur Erhebung des fachspezifischen Professionswissens von Physiklehrkräften entwickelt (z.B. Riese, 2009; Kulgemeyer et al., 2012; Kröger, Neumann & Peterson, 2013). Die Validierung der Tests erfolgt meist über den Vergleich bekannter Gruppen mit zu erwartenden Fähigkeitsunterschieden oder durch Zusammenhangsanalysen zwischen den Dimensionen des Professionswissen: Fachwissen (CK), fachdidaktisches (PCK) und pädagogisches Wissen (PK). Bisher existieren nur wenige Studien, die Bezüge zu Unterrichtsqualität und Schülervariablen herstellen, um zu entscheiden, ob das erhobene Wissen relevant für effektives Unterrichten ist. Auch liefern diese Studien bisher keine einheitlichen Ergebnisse zu dieser Fragestellung (vergl. Lange, 2010; Ohle, 2010; Fischer, Labudde, Neumann & Viiri, 2014). Diese Lücke weiter zu schließen, ist ein Ziel des ProwiN Projektes. In der ersten Phase des Projektes wurde ein Testinstrument zur Erfassung des fachspezifischen Professionswissens von Physiklehrkräften entwickelt und wie oben beschrieben validiert (Kirschner, 2013). Um die Relevanz des erhobenen Wissens für effektives Unterrichten zu untersuchen, wird in der zweiten Projektphase der Zusammenhang zwischen fachspezifischem Professionswissen, Unterricht und Schülerleistungszuwachs analysiert. Die hier vorgestellte Studie untersucht die Fragestellung, inwieweit Unterschiede im kognitiven Aktivierungspotenzial des Physikunterrichts und in den Leistungszuwächsen von Schulklassen im Rahmen einer Unterrichtseinheit über Unterschiede im PCK und CK der Lehrkräfte erklärt werden können. Hierfür wurden das Professionswissen von 23 Gymnasiallehrkräften (34.8% weiblich, MAlter=44 Jahre, SDAlter=12 Jahre) und das Fachwissen der Lernenden (N=610, MAlter=14 Jahre, SDAlter=1 Jahr) in ihren Klassen (Jahrgangsstufe 8/9) vor und nach der Unterrichtseinheit Mechanik erhoben und je zwei Unterrichtsstunden in Mechanik videographiert. Bisher wurde nur die erste Unterrichtsstunde mit Hilfe eines Ratingmanuals zur Kognitiven Aktivierung (6 Subskalen mit insg. 25 Handlungsindikatoren, Bewertung auf 3-stufiger Likertskala, $\alpha_C=.93$, adapt. nach Vogelsang, 2014) ausgewertet. Die Professionswissenstests beinhalteten offene und Multiple-Choice Aufgaben (CK: 11 Aufgaben zum Schulwissen und vertieften Schulwissen in Mechanik, ICC(2,1)unjust ≥ 0.96 , Rasch Personen Rel.=.73; PCK: 11 Aufgaben zu Schülervorstellungen, Umgang mit Experimenten und Konzepten in der Mechanik, ICC(2,1)unjust ≥ 0.77 , Rasch Personen Rel.=.59). Der Schülerfachwissenstest bestand aus 34 Multiple-Choice Aufgaben zum Fachwissen in Mechanik (Rasch Personen Rel. Prä-/Post-Test=.51/.61). Auf Schülerebene gab es vom Prä- zum Post-Test einen signifikanten Leistungszuwachs ($t(610)= 10.50$, $p<.001$, $d=.43$). Mehr als 10% der Gesamtvarianz im Post-Test lag zwischen den Klassen (ICC(1,1)=.104). Um die Varianz in der Länge der Unterrichtseinheit (Dauer zwischen 9-30 Wochen) zu berücksichtigen, wurde diese Variable neben dem Prä-Test, den kognitiven Fähigkeiten der Lernenden und der zuhause gesprochenen Sprache (Prädiktoren auf Schülerebene) als Prädiktor auf Klassenebene in ein Mehrebenenmodell zur Erklärung der Varianz im Post-Test aufgenommen (Schülerebene: $R^2=.31$, $p<.001$, Klassenebene: $R^2=.63$, $p<.001$). Die kognitive Aktivierung klärt weitere 15% der Varianz auf Klassenebene auf ($\beta_{\text{Stand.}}=.40$, $p<.05$) wogegen weder CK noch PCK signifikant zur Varianzaufklärung beitragen ($\beta_{\text{Stand.},CK}=.14$, $\beta_{\text{Stand.},PCK}=-.19$). Während sich für CK ein Zusammenhang zur kognitiven Aktivierung zumindest andeutet ($r=.37$, $p=.079$, signifikant bei einseitiger Testung), kann für PCK kein Zusammenhang gefunden werden ($r=.15$, $p=.504$). Für eine Stichprobe von N=23 wären signifikante Zusammenhänge allerdings auch erst für Korrelationen von $r>.47$ zu erwarten. Während die Ergebnisse für den CK Test aufgrund der Stichprobengröße noch nicht eindeutig interpretierbar sind, deuten die Ergebnisse bezüglich des PCK Tests darauf hin, dass der Test das handlungsrelevante Wissen für effektives Unterrichten nicht erfasst, obwohl er auf in der Community breit akzeptierten (aber normativ gesetzten) Wissensfacetten basiert. Außerdem kann geschlussfolgert werden, dass die anfangs beschriebene herkömmliche Validierung von Testinstrumenten zur Erfassung des Professionswissens nicht ausreichend zu sein scheint, um zu behaupten, dass die Testinstrumente unterrichtsrelevantes Wissen valide messen.

Donnerstag
12. März

15:30 Uhr

HGB 30

Chair(s)

von Kotzebue, Lena
Lenske, Gerlinde**Zum Einfluss des Professionswissens auf das Unterrichtshandeln (ProwiN-Video)**

Lenske, Gerlinde; Cauet, Eva; Leutner, Detlev; Wirth, Joachim

Zum Einfluss des pädagogisch-psychologischen Professionswissens auf das Unterrichtshandeln von Lehrkräften im Physikunterricht und auf den Lernerfolg der Schülerinnen und Schüler

Ausgangslage und theoretischer Hintergrund: Zahlreiche Befunde aus der empirischen Bildungsforschung verdeutlichen den bedeutsamen Einfluss des Lehrerhandelns auf die Schülerleistung (vgl. zusammenfassend Hattie, 2009). Obwohl seit Jahrzehnten die Annahme besteht, dass das Professionswissen von Lehrkräften eine wesentliche Voraussetzung für erfolgreiches Unterrichtshandeln ist (Abell, 2007; Bransford et al., 2005; Fischler, 2008; Khourey-Bowers & Fenk, 2009; Kunter et al., 2011; Wahl, 1991), gilt die Forschungslage insbesondere in Bezug auf das pädagogisch-psychologische Professionswissen (PPK) als besonders dünn (Voss & Kunter, 2011).

PPK umfasst das fächerübergreifende Wissen um Strategien und Mittel zur Genese und Erhaltung lernförderlicher Bedingungen (Baumert & Kunter, 2006; Shulman, 1986, 1987). Im Rahmen des Projekts COACTIV fanden sich erste Hinweise dafür, dass PPK die von der Schülerschaft wahrgenommene Effizienz der Klassenführung und die Qualität der Schülerorientierung im Fach Mathematik beeinflusst (Voss et. al., 2014). Inwieweit die Befunde auf andere (naturwissenschaftliche) Fächer übertragbar sind, ist bislang unklar. Des Weiteren wurde bisher das Unterrichtshandeln eher indirekt über die Analyse von Unterrichtsdokumenten oder die Schülerwahrnehmung erfasst; Videoanalysen wurden nicht einbezogen.

Fragstellung: Die vorliegende Studie setzt genau hier an und untersucht videobasiert, ob sich PPK auf das Unterrichtshandeln im Fach Physik auswirkt (F1). Angesichts der theoretischen Annahme und der bisherigen Befunde, wird vermutet, dass sich zwischen PPK und der prozessualen Unterrichtsqualität positive Zusammenhänge zeigen (H1).

Methode.

Stichprobe: Insgesamt nahmen rund 40 Sekundarschullehrkräfte aus Nordrhein-Westfalen an der Studie teil. Zu einer Teilstichprobe von 23 Lehrkräften (34.8% weiblich, MAAlter=44 Jahre, SDAAlter=12 Jahre), welche zum Zeitpunkt alle an einem Gymnasium in einer achten oder neunten Klasse Physikunterricht gaben, sind die Daten bereits ausgewertet. Die dazugehörige Schülerstichprobe umfasst 610 Schüler (MAAlter=14 Jahre, SDAAlter=1 Jahr).

Prozeduren und Instrumente: Das Professionswissen wurde per Paper-Pencil-Test erfasst. Das hierzu eingesetzte Testinstrumentarium wurde im Rahmen des Projekts ProwiN auf Basis des in Coactiv verwendeten Tests entwickelt und validiert (Lenske et al., in Vorb.). Um die prozessuale Unterrichtsqualität zu erfassen, wurden jeweils zwei Unterrichtssequenzen aus dem Bereich Mechanik videographiert und bezüglich der Qualität der Klassenführung, der Individualisierung und der kognitiven Aktivierung analysiert. Für die Analyse der Klassenführung und der Individualisierung wurden neue Ratingmanuale entwickelt; für das Rating der kognitiven Aktivierung wurde auf ein im Physikunterricht bewährtes Ratingmanual rekurriert (Vogelsang, 2014). Die Schülerleistung in Mechanik wurde vor und nach der Einheit erhoben. Auf Schülerebene gab es zwischen Prä- und Posttest einen signifikanten Leistungszuwachs ($t(610) = 10.50, p < .001, d = .43$). Da die Dauer der Einheit zwischen den Klassen variierte, wurde neben dem Prätest und der kognitiven Fähigkeiten auch für die Anzahl an Unterrichts-stunden in Mechanik kontrolliert. 10% der Gesamtvarianz im Post-Test lag zwischen den Klassen ($ICC(1,1) = .104$). Der effektive Leistungszuwachs der Schüler diente als Kriterium zur Validierung der Ratingmanuale.

Ergebnisse: Die neu konzipierten Ratingmanuale weisen eine gute Inter-Raterreliabilität auf ($r \geq .8$). Das Rating zur kognitiven Aktivierung basiert auf Konsensentscheid. Die beobachtete Qualität der Klassenführung und der kognitiven Aktivierung stehen in einem positiven mittleren Zusammenhang ($r > .4$) mit dem Leistungszuwachs der Schüler – was als ein Hinweis auf die Validität der Manuale und der damit verbundenen Messungen interpretiert werden kann. In Bezug auf die beobachtete Individualisierung im Physikunterricht liegt ein Bodeneffekt vor (d.h. es waren nahezu keine Anzeichen von individualisiertem Unterricht sichtbar). Bedingt hierdurch ließ sich weder ein Zusammenhang mit der Schülerleistung noch mit PPK nachweisen. Erwartungsgemäß ergaben sich positive mittlere Zusammenhänge zwischen dem PPK der teilnehmenden Lehrkräfte und ihrer Unterrichtsqualität in Bezug auf ausgewählte Facetten der Klassenführung ($r = .53$) und der kognitiven Aktivierung ($r = .43$). Der Effekt von PPK auf den Leistungszuwachs der Schüler ($\beta = .37$) wird vollständig durch das Lehrhandeln bezüglich Klassenführung und kognitiver Aktivierung mediiert. Zusammenfassend lässt sich festhalten, dass die Studie für eine weitverbreitete, aber bislang nicht hinreichend untersuchte Annahme (H1) empirische Evidenz liefert.

Donnerstag
12. März

15:30 Uhr

HGB 30

Rakoczy, Katrin

Interventionsstudien im Unterricht - Differenzierte Analyse der Effekte auf den Lernprozess im Kontext heterogener Lerngruppen

Gegenwärtige Diskussionen zum Thema Heterogenität fordern im Kern Lernarrangements, die der Unterschiedlichkeit der Lernenden stärker gerecht werden und auf ihre jeweiligen Bedürfnisse stärker eingehen. Damit einher geht auch die Forderungen nach einem Paradigmenwechsel weg von Heterogenität als Belastung oder besondere Herausforderung hin zu Heterogenität als Chance und Bereicherung (z.B. Solzbacher, 2008; Wischer, 2007). Die empirische Befundlage zu den teils normativ gefärbten Diskussionen ist jedoch ausbaufähig (z. B. Bellin, 2009; Gröhlich, Scharenberg & Bos, 2009). Inzwischen wird weniger dem Ausmaß von Heterogenität selbst als vielmehr dem produktiven Umgang mit ihr in der konkreten Unterrichtsgestaltung ein großer Stellenwert für das Lernen und die Motivation von Lernenden beigemessen.

Im geplanten Symposium werden Interventionsstudien vorgestellt, die verschiedene Varianten der Unterrichtsgestaltung differenziert auf ihre Wirksamkeit hin prüfen.

Der erste Beitrag greift zunächst die Heterogenitätsdebatte auf und nimmt eine differenzierte Betrachtung verschiedener Leistungsindikatoren zur Beschreibung von Heterogenität vor. Hieran anschließend wird die Bedeutung des Leseverständnisses als Heterogenitätsmaß herausgestellt und mit Hilfe einer Interventionsstudie die Rolle einer auf Heterogenität ausgelegten Unterrichtsmethode für den Zusammenhang von Leistungsheterogenität und späterem Lernerfolg von Grundschulkindern untersucht. Die nächsten drei Beiträge aus Interventionsstudien knüpfen hieran an und nehmen eine differenziertere Analyse der Effekte auf Lernen und Motivation von Lernenden vor. Der zweite Beitrag setzt ebenfalls am Leseverständnis von Grundschulkindern an und prüft in drei aufeinanderfolgenden Interventionsstudien, welche Bedeutung das schrittweise Ausblenden von Implementationskontrolle für das Leseverständnis und die Lesegeschwindigkeit hat. Der dritte Beitrag bezieht sich auf eine Interventionsstudie zu Formativem Assessment im Mathematikunterricht der Sekundarstufe. Hier werden die über wahrgenommene Nützlichkeit von Leistungsrückmeldungen und die Selbstwirksamkeit der Lernenden vermittelten Effekte auf die Lern- und Interessensentwicklung von Lernenden geprüft. Auch der vierte Beitrag bezieht sich auf den Mathematikunterricht in der Sekundarstufe. Hier wird die Bedeutung von multiplen Lösungen einer Aufgabe für Freude und Langeweile von Lernenden im Unterricht geprüft sowie deren vermittelnde Rollen auf das Interesse untersucht.

Das Symposium verbindet fachdidaktische, erziehungswissenschaftliche und pädagogisch-psychologische Perspektiven, um Interventionen differenziert auf ihre Wirksamkeit hin zu analysieren. Während der differenzierte Blick auf Wirkungen im ersten Beitrag durch die theoretische und empirische Berücksichtigung der Heterogenität der Lernenden erfolgt, wird er im zweiten Beitrag durch die Diskussion der Bedeutung der Implementationskontrolle für die Wirksamkeit von Interventionen umgesetzt. Im dritten und vierten Beitrag werden Effekte von Interventionen aufgeschlüsselt, indem das Erleben der Lernenden als vermittelnder Prozess einbezogen wird, was wiederum weitere Erkenntnisse für das Lernen in heterogenen Gruppen bietet.

Abschließend werden die Beiträge von Christine Pauli reflektiert und diskutiert.

Donnerstag
12. März

15:30 Uhr

HGB 40

Chair(s)

Rakoczy, Katrin
Decristan, Jasmin**Interventionsstudien im Unterricht - Differenzierte Analyse der Effekte auf den Lernprozess im Kontext heterogener Lerngruppen**

Decristan, Jasmin; Adl-Amini, Katja; Naumann, Alexander; Fauth, Benjamin; Rieser, Svenja; Büttner, Gerhard; Hardy, Ilonca; Klieme, Eckhard

Leistungsheterogenität und Lernen: Differenzierte Befunde zur Wahl von Leistungsindikatoren und zur Rolle der Unterrichtsgestaltung

Das Thema Heterogenität hat gegenwärtig einen großen Stellenwert in Bildungsforschung, Schulpraxis und bildungspolitischen Diskussionen. Ein produktiver Umgang mit Heterogenität, wie oftmals gefordert, sollte sich in einem positiven Zusammenhang zwischen Leistungsheterogenität in Schulklassen und Lernergebnissen von Schülerinnen und Schüler widerspiegeln. Längsschnittdaten im Grundschulbereich bieten hierzu keine klare Aussage (vgl. Bellin, 2009; Gröhlich, Scharenberg & Bos, 2009; Künsting, Post, Greb, Faust & Lipowsky, 2010). Die Studien verwenden jedoch zum einen synonym ganz unterschiedliche Leistungsbereiche als Heterogenitätsindikatoren (z.B. kognitive Grundfähigkeiten oder Leseverständnis), zum anderen wird die Unterrichtsgestaltung als entscheidend für den Zusammenhang zwischen Leistungsheterogenität und Lernen angesehen (Gröhlich et al., 2009). Eine Unterrichtsmethode, die die Unterschiedlichkeit der Schülerinnen und Schüler aktiv nutzt, ist tutorielles Lernen. Der vorliegende Beitrag soll anhand einer differenzierteren Betrachtung verschiedener Leistungsheterogenitätsindikatoren und der Bedeutung der Unterrichtsgestaltung weitere Erkenntnisse zur Relevanz von Leistungsheterogenität für das Lernen liefern.

Forschungsfragen und Hypothese: (1) Welche Zusammenhänge zeigen sich zwischen verschiedenen Leistungsheterogenitätsmaßen in Grundschulklassen – lässt sich synonym von „der“ Leistungsheterogenität sprechen? (2) Welche Leistungsindikatoren fokussieren Grundschullehrkräfte, wenn sie das Ausmaß von Leistungsheterogenität ihrer Klasse beurteilen? (3) Welche Bedeutung hat die Unterrichtsgestaltung für den Zusammenhang zwischen (der von Lehrkräften fokussierten) Leistungsheterogenität in Klassen und dem Lernen der Kinder? Dazu wird angenommen, dass eine Unterrichtsgestaltung, die auf einen produktiven Umgang mit Heterogenität abzielt, den Zusammenhang zwischen Leistungsheterogenität und Lernergebnissen positiv moderiert, und dass somit Klassen mit vergleichsweise hoher Leistungsheterogenität besonders von dieser Unterrichtsgestaltung profitieren.

Methode.

Design: Die Daten stammen aus einer Interventionsstudie in der Grundschule. Anhand eines quasi-experimentellen Gruppenvergleichs wurden verschiedene Unterrichtsmethoden, jeweils im Vergleich zu einer Kontrollgruppe in ihren Wirkungen für Schülerinnen und Schülern evaluiert. Dieser Beitrag fokussiert auf die Unterrichtsmethode des tutoriellen Lernens; hierbei wird Heterogenität aktiv einbezogen, indem Kinder in leistungsheterogenen Tandems unter hohem Strukturierungsgrad der Zusammenarbeit gemeinsam lernen. Lehrkräfte erhielten Fortbildungen im tutoriellen Lernen und in einem Inhaltsbereich des naturwissenschaftlichen Sachunterrichts. Die Kontrollgruppe erhielt die gleiche inhaltspezifische Fortbildung sowie eine Fortbildung in Elternberatung. Anschließend setzten die Lehrkräfte die Inhalte in ihren Grundschulklassen um.

Stichprobe: 1070 Kinder (mittleres Alter: 9 Jahre, 49% Mädchen) sowie 54 Lehrkräfte (mittleres Alter: 43 Jahre, 89% Lehrerinnen). Der Vergleich tutorielles Lernen vs. Kontrollgruppe beinhaltet eine Substichprobe mit 25 Klassen und 513 Kindern.

Instrumente: vor der Intervention erfasste Leistungsheterogenitätsmaße: kognitive Grundfähigkeit (CFT 20-R; Weiß, 2006), naturwissenschaftliche Kompetenz (aus TIMSS adaptierte Items; Martin, Mullis & Foy, 2008), und Leseverständnis (ELFE 1-6; Lenhard & Schneider, 2006). Dazu beurteilten Lehrkräfte das Ausmaß von Leistungsunterschieden in ihrer Klasse. Das Lernergebnis bezieht sich auf einen Test zum konzeptuellen Verständnis (adaptiert nach Hardy et al., 2006) nach der Intervention.

Ergebnisse: (1) Es fanden sich keine signifikanten Übereinstimmungen (Spearman's $\rho < .16$, $p > .05$) zwischen den klasseninternen Streuungen in kognitiven Grundfähigkeiten, Leseverständnis oder naturwissenschaftlicher Kompetenz. (2) Logistische Regressionsanalysen zeigten, dass klasseninterne Leistungsstreuungen im Leseverständnis (nicht aber in kognitiven Grundfähigkeiten oder naturwissenschaftlicher Kompetenz) die Einschätzungen von Lehrkräften zum Ausmaß von Leistungsunterschieden in Klassen vorhersagte; Lehrkräfte fokussierten somit insbesondere diesen Leistungsbereich bei ihren Heterogenitätseinschätzungen. (3) Daten einer Mehrebenen-Regressionsanalyse bestätigten eine positive Interaktion aus Leistungsheterogenität im Leseverständnis und tutoriellem Lernen ($\beta = .34$, $SE = .19$, $p < .05$): Schülerinnen und Schüler in Klassen mit tutoriellem Lernen und hoher Leistungsheterogenität zeigten die im Mittel besten Lernergebnisse.

Implikationen: (1) Von „der“ Leistungsheterogenität einer Klasse lässt sich kaum sprechen. Die Verwendung unterschiedlicher Leistungsindikatoren könnte somit zur uneinheitlichen Befundlage bei der Vorhersage von Schülerleistungen durch Leistungsheterogenitätsmaße beigetragen haben. (2) Lehrkräfte fokussieren bei Heterogenitätseinschätzungen von Klassen insbesondere das Leseverständnis als einen eng mit der Unterrichtssprache gekoppelten und vergleichsweise gut beobachtbaren Leistungsbereich. (3) Ob Leistungsheterogenität in Klassen zu positiven Lernergebnissen führt, hängt wesentlich von der Unterrichtsgestaltung und dem von Lehrkräften fokussierten Leistungsindikator ab. Dies könnte ein weiterer Erklärungsansatz für die uneinheitliche Befundlage sein.

Donnerstag
12. März

15:30 Uhr

HGB 40

Chair(s)

Rakoczy, Katrin
Decristan, Jasmin

Interventionsstudien im Unterricht - Differenzierte Analyse der Effekte auf den Lernprozess im Kontext heterogener Lerngruppen

Seifert, Susanne; Gasteiger-Klicpera, Barbara; Schwab, Susanne; Paleczek, Lisa

Leseförderung in der Grundschule – Interpretation der Ergebnisse unter Berücksichtigung der Implementationsqualität

Theoretischer Hintergrund: Im Jahr 2011 konnten in Österreich 20% (Vogtenhuber et al., 2012) und in Deutschland etwa 15% (Bos et al., 2012) der Kinder am Ende der vierten Grundschulklasse, nur grundlegendste Leseanforderungen erfüllen, d.h. sie waren höchstens in der Lage, einfachen Texten explizite Einzelinformationen zu entnehmen. Um die Entwicklung der schulischen und beruflichen Laufbahn von leseschwachen Kindern positiv zu beeinflussen, sollten Schwächen frühestmöglich diagnostiziert und intensive, individuell abgestimmte Fördermaßnahmen angeboten werden (Gasteiger-Klicpera & Fischer, 2008; Torgesen, 2005). Um die Stabilität der Effekte langfristig zu gewährleisten, ist es empfehlenswert, die Förderung in den Schulalltag zu implementieren und somit zu einem veränderten Unterrichtskonzept im Sinne einer differenzierten, den Leistungen der SchülerInnen entsprechenden Förderung zu gelangen (Gasteiger-Klicpera & Fischer, 2008; Strickland, 2002).

Ziel/ Fragestellung: Innerhalb des Projekts LARS (Language And Reading Skills) wurde ein differenziertes Leseförderprogramm in den Schulalltag implementiert und begleitend evaluiert. Es wurde untersucht, ob Kinder, die mit dem Förderprogramm unterrichtet wurden, eine bessere Leistungsentwicklung im Leseverständnis und in der Lesegeschwindigkeit zeigten als Kinder, die regulär beschult wurden.

Untersuchungsmethode: Das Gesamtprojekt umfasste drei Studien, die jeweils in einem quasi-experimentellen Prä-Posttest-Design mit Vergleichsgruppe angelegt waren. Vor und nach den Interventionsstudien wurden die literalen Fähigkeiten mittels standardisierter Testverfahren erhoben (Leseverständnis: Wort-, Satz-, Textebene – ELFE 1-6, Lenhard & Schneider, 2006; Lesegeschwindigkeit: Wort- und Pseudowortlesen – SLRT II, Moll & Landerl, 2010).

Die Kinder der Interventionsgruppen wurden in jeder Studie zwei Mal wöchentlich im Rahmen der Deutsch- und Sachunterrichtsstunden, abhängig von ihrem Leistungsniveau mit einem in vier Schwierigkeitsstufen differenziertem Fördermaterial unterstützt. Innerhalb der drei Interventionsstudien wurde die Intensität der Implementationskontrolle sukzessive reduziert. Zunächst wurden in einer ersten Studie ZweitklässlerInnen (n=55; Vergleichsgruppe n=50) drei Monate lang durch Projektmitglieder gemeinsam mit dem/der KlassenlehrerIn gefördert. Anschließend wurden in einer zweiten Studie ZweitklässlerInnen (n=159; Vergleichsgruppe n=218) über ein gesamtes Schuljahr hinweg von den KlassenlehrerInnen selbst gefördert, die zuvor umfangreich eingeschult wurden und die ersten 2 Monate in der Förderung durch Projektmitglieder eng begleitet wurden. In der dritten Studie wurden DrittklässlerInnen (n=303; Vergleichsgruppe n=339) über ein gesamtes Schuljahr hinweg von den KlassenlehrerInnen gefördert, die ebenfalls zunächst eine Einschulung erhielten, dann jedoch lediglich einmal im Monat von einem Projektmitglied supervidiert wurden.

Ergebnisse: Für die erste Interventionsstudie zeigte sich, dass die geförderten Kinder einen deutlich größeren Lernzuwachs in der Lesegeschwindigkeit (Wortlesen: $F(3,99)=40.14$, $p<.01$, $\eta^2=.28$; Pseudowortlesen: $F(3,99)=72.78$, $p<.01$, $\eta^2=.42$) und im Leseverständnis (Gesamt-Prozentrang: $F(3,99)=16.82$, $p<.01$; $\eta^2=.14$) zeigten als Kinder der Vergleichsgruppe.

In der zweiten Interventionsstudie zeigten die geförderten Kinder im Vergleich zur Vergleichsgruppe im Leseverständnis auf Satzebene eine signifikant bessere Entwicklung ($F(1,356)=8.63$, $p<.01$; $\eta^2=.02$), allerdings nicht auf Wortebene. Die Lernzuwächse in der Lesegeschwindigkeit unterschieden sich nicht signifikant zwischen Interventions- und Vergleichsgruppe.

Für die dritte Interventionsstudie konnten lediglich für die durchschnittlichen LeserInnen ($F(1,225)=5.37$; $p<.05$) und Kinder mit Deutsch als Zweitsprache ($F(1,570)=4.46$; $p<.05$) der Interventionsgruppe größere Lernzuwächse in der Lesegeschwindigkeit nachgewiesen werden.

Ausblick: Den Lehrenden wurde von Interventionsstudie zu Interventionsstudie mehr Freiheit in der Implementation gelassen, damit sie die Intervention nach ihren Vorstellungen einfacher in den Unterricht integrieren konnten. Die geringere Implementationskontrolle könnte die Implementationsqualität beeinflusst haben. Auf welche Weise dies der Fall war, soll durch weitere Auswertungen geklärt werden. Nach Souvignier und Antoniou (2007) können bei Interventionen, die über einen kürzeren Zeitraum sehr kontrolliert durchgeführt werden, größere Effekte erzielt werden als bei ökologisch validen Interventionen, die über einen längeren Zeitraum und unter natürlicheren Unterrichtsbedingungen implementiert werden. Der Einfluss der Implementationskontrolle auf die Effekte von Interventionen soll im vorliegenden Beitrag diskutiert werden.

Donnerstag
12. März

15:30 Uhr

HGB 40

Chair(s)

Rakoczy, Katrin
Decristan, Jasmin

Interventionsstudien im Unterricht - Differenzierte Analyse der Effekte auf den Lernprozess im Kontext heterogener Lerngruppen

Rakoczy, Katrin; Klieme, Eckhard; Pinger, Petra; Harks, Birgit; Blum, Werner; Hochweber, Jan

Formatives Assessment im Mathematikunterricht – indirekte Effekte auf den Lernerfolg über wahrgenommene Nützlichkeit und Selbstwirksamkeit

Theoretischer Hintergrund & Fragestellung: In der Anglo-Amerikanischen Literatur wird formatives Assessment häufig als vielversprechender Ansatz zur Verbesserung von Unterricht und Lernen beschrieben (z.B. Black & Wiliam, 2009; Wiliam & Thompson, 2008). Als zentraler empirischer Beleg hierfür gilt die Übersichtsarbeit von Black und Wiliam (1998). Die darin berichtete Effektstärke von .4 bis .7 ist jedoch aufgrund der Heterogenität der einbezogenen Studien zu hinterfragen (Bennett, 2011), und zahlreiche Autoren (z.B. Dunn & Mulvenon, 2009) fordern mehr empirische Forschung zur Wirkung formativen Assessments. Diese sollte nach Bennett (2011) auf einer Definition von formativem Assessment beruhen, die beschreibt, welche Elemente formatives Assessment ausmachen und wie, d.h. durch welche „underlying mechanisms“ (Bennett, 2011), es zu den berichteten Effekten auf den Lernerfolg kommt (Wiliam & Thompson, 2008).

Als zentrale Elemente formativen Assessments haben sich die Erfassung von diagnostischer Information mittels psychometrisch fundierter Messinstrumente und ihre lernförderliche Rückmeldung erwiesen (Wiliam & Thompson, 2008). Eine Rückmeldeart, die versucht, die aus der Literatur bekannten Kriterien lernförderlicher Rückmeldung in sich zu vereinen, und die sich empirisch unter Laborbedingungen als lernförderlich erwiesen hat, stellt die lösungsprozessbezogene Rückmeldung dar (Besser, Leiss, Harks, Rakoczy, Klieme & Blum, 2010; Rakoczy, Harks, Klieme, Blum & Hochweber, 2013). Bezüglich der Wirkungsweise formativen Assessments ist davon auszugehen, dass Handlungen und Prozesse auf Schülerseite initiiert werden, die wiederum mit dem Lernerfolg verbunden sind (Bennett, 2011). Dazu zählt, dass Lernende die Rückmeldung, die sie erhalten, als nützlich ansehen (Rakoczy et al., 2013; Stiggins, 2006) und dass sie sich als selbstwirksam erleben (Stiggins, 2006).

Anknüpfend an diese Befunde wurde eine Intervention zum formativen Assessment im Mathematikunterricht entwickelt, indem lösungsprozessbezogene Rückmeldung in den Mathematikunterricht implementiert wurde. Ziel des Beitrags ist es, der Wirkungsweise formativen Assessments auf die Interessen- und Leistungsentwicklung der Lernenden nachzugehen, indem indirekte Effekte über die wahrgenommene Nützlichkeit der Rückmeldung und die Selbstwirksamkeit auf die Interessen- und Leistungsentwicklung geprüft werden.

Methode: Es wurde eine Interventionsstudie mit Prä- und Posttest in einer 13-stündigen Unterrichtseinheit zum „Satz des Pythagoras“ durchgeführt. An der Studie nahmen 26 Lehrkräfte mit ihren Klassen (620 Lernende) teil und wurden zufällig zur Interventions- oder Kontrollgruppe zugewiesen. Die beiden Bedingungen wurden durch Fortbildungen der Lehrkräfte realisiert. Allen Lehrkräften wurden die fachlichen Unterrichtsinhalte und die zu behandelnden Aufgaben vermittelt. In der Interventionsgruppe wurden die Lehrkräfte darüber hinaus darin geschult, zu drei festgelegten Zeitpunkten (5., 8., und 11. Stunde) die Leistung der Lernenden anhand von Diagnoseaufgaben zu erfassen und schriftliche lösungsprozessbezogene Rückmeldung zu geben.

In den Unterrichtsstunden vor und nach der Intervention wurde jeweils das Interesse und die Selbstwirksamkeit der Lernenden in Bezug auf den bevorstehenden Test per Fragebogen (vgl. Rakoczy, Buff & Lipowsky, 2005), ihre Mathematikleistung (vgl. Klieme et al., 2010), sowie die wahrgenommene Nützlichkeit der Rückmeldungen im Unterricht in einem zweiten Fragebogen (in Anlehnung an Dresel & Ziegler, 2007) erfasst.

Ergebnisse: Mehrebenenmediationsanalysen nach Pituch und Stapleton (2012) zeigen indirekte Effekte der Intervention zum formativen Assessment im Vergleich zur Kontrollgruppe über die wahrgenommene Nützlichkeit auf die Selbstwirksamkeit sowohl auf individueller als auch auf Klassenebene (individuell: $\beta = .03$, $p < .05$, Klassenebene: $\beta = .20$, $p < .01$). Dieser Pfad zeigt sich auf individueller Ebene als relevant für die Interessenentwicklung ($\beta = .01$, $p < .05$) und auf Klassenebene für die Leistungsentwicklung ($\beta = .24$, $p < .05$).

Diskussion: Die Ergebnisse zeigen, dass es mit der Intervention gelungen ist, die Selbstwirksamkeit der Lernenden zu stärken. Sie stellt einen wichtigen Faktor dar, durch den Lehrkräfte die Entwicklung von Lernenden beeinflussen können (Brookhart, 1997). Darüber hinaus haben Lernende, die ihre Rückmeldung als nützlich ansehen und sich in der Folge selbstwirksamer erleben, mehr Interesse an einem weiteren bevorstehenden Test. Für die Leistungsentwicklung erwies sich die Komposition der Nützlichkeitswahrnehmung und Selbstwirksamkeit als bedeutsam.

Donnerstag
12. März

15:30 Uhr

HGB 40

Chair(s)

Rakoczy, Katrin
Decristan, Jasmin

Interventionsstudien im Unterricht - Differenzierte Analyse der Effekte auf den Lernprozess im Kontext heterogener Lerngruppen

Schukajlow, Stanislaw; Rakoczy, Katrin

Freude und Langeweile als vermittelnde Variablen für die Interessenentwicklung im Mathematikunterricht mit multiplen Lösungen

Theoretischer Hintergrund und Fragestellungen: Die Entwicklung von multiplen Lösungen wird in der Mathematikdidaktik schon lange gefordert und ist in den Bildungsstandards fest verankert (National Council of Teachers of Mathematics, 2000). Ihr Effekt auf motivational-affektive Merkmale und mögliche vermittelnde Variablen im Unterricht ist bisher jedoch wenig untersucht worden. In einem von der DFG geförderten Projekt „Multiple Lösungen in einem selbstständigkeitsorientierten Mathematikunterricht“ wurden diese Fragen im Rahmen einer Interventionsstudie untersucht. Es konnte eine positive Wirkung der Behandlung von multiplen Lösungen auf das Interesse von Lernenden festgestellt werden (Schukajlow & Krug, 2014). Im vorliegenden Beitrag wird daran anschließend untersucht, inwiefern die Entwicklung multipler Lösungen das emotionale Erleben (Freude und Langeweile) im Unterricht beeinflusst und inwieweit das emotionale Erleben den bereits bekannten Effekt auf das Interesse vermittelt. Theoretischer Hintergrund sind Theorien der Interessenentwicklung (Hidi & Renninger, 2006; Krapp, 2002) und die Control-Value-Theory of Achievement Emotions (Pekrun, 2006). Im Rahmen beider Theorien wurde eine Reihe von empirischen Befunden zu einzelnen Wirkungsbeziehungen (wie z.B. positiver Einfluss von Freude auf Interesse), die im Model angenommen werden, hervorgebracht (Goetz, Frenzel, Hall, & Pekrun, 2008; Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010).

Die Hauptfragen des vorliegenden Beitrags sind: (1) Wie beeinflusst die Behandlung von multiplen Lösungen Freude und Langeweile im Unterricht? (2) In wie weit wird der Einfluss der Unterrichtsmethode auf Interesse im Posttest über Freude und Langeweile im Unterricht vermittelt?

Methode: An der Untersuchung nahmen 144 Lernende aus 6 Realschulklassen der 9. Jahrgangsstufe (43% weiblich, im Mittel 15.1 Jahre alt) teil. Innerhalb der Klasse wurden die Lernenden zu der Gruppe „Multiple Lösungen“ (ML) vs. „Eine Lösung“ (EL) zugewiesen. Die beiden Gruppen wurden auf Klassenebene über Leistungen und Geschlecht parallelisiert und von geschulten Lehrkräften in separaten Räumen fünf Schulstunden lang unterrichtet. Jede Lehrkraft hatte die gleiche Anzahl von ML- und EL-Gruppen instruiert, um den Einfluss der Lehrerpersönlichkeit in beiden Bedingungen konstant zu halten. Die beiden Experimentalgruppen haben parallele Versionen der gleichen Aufgaben bearbeitet. In der ML-Gruppe sollte jeder Lernende zwei Lösungen zu einer offenen Modellierungsaufgabe entwickeln. In der EL-Gruppe wurde eine geschlossene Version dieser Aufgabe gestellt und eine Lösung verlangt (Schukajlow & Krug, 2014). Die Analyse der Materialien bestätigte eine adäquate Umsetzung beider Treatments.

Vor und nach der Untersuchung wurde das Interesse an Mathematik (3 Items, Cronbachs-Alpha= .80; .74) mit Hilfe der 5-stufigen Likert-Skala von (Frenzel, Pekrun, Dicke, & Goetz, 2012) abgefragt. An zwei Messzeitpunkten während der Unterrichtseinheit haben die Probanden angegeben, wie viele Lösungen sie zu der jeweiligen Unterrichtsaufgabe entwickelt haben (0=keine Lösung bis 2=mehr als zwei Lösungen), wie viel Freude (3 Items, Cronbachs-Alpha=.84; .82) und Langeweile (3 Items, Cronbachs-Alpha=.85; .86) sie im Unterricht empfunden haben (Emotionen wurden adaptiert aus Pekrun et al., 2005). Das Antwortformat war ebenfalls 5-stufig („von stimmt gar nicht“ bis „stimmt genau“).

Ergebnisse: Freude und Langeweile von Lernenden in beiden Experimentalgruppen unterscheiden sich nicht voneinander. Somit hat die Behandlung von multiplen Lösungen zu realitätsbezogenen Aufgaben keinen Einfluss auf die Emotionen der Schüler während der Unterrichtseinheit. In einem Pfadmodell (CFI=.97; Chi-Quadrat(5)=9.5, $p > .05$; SRMR=.04) zeigt sich jedoch ein indirekter Effekt des Treatments über die Anzahl der tatsächlich im Unterricht entwickelten Lösungen und die erlebte Freude auf das Interesse ($\beta = .07$; $p < .05$). Langeweile wird zwar ebenfalls von der Anzahl der multiplen Lösungen beeinflusst, zeigt aber keine Wirkungen auf Interesse im Nachtest. Abschließend wird die Rolle von Freude und Langeweile für die Interessenentwicklung in Mathematik diskutiert.

Donnerstag
12. März

15:30 Uhr

HGB 40

Chair(s)

Rakoczy, Katrin
Decristan, JasminInterventionsstudien im Unterricht - Differenzierte Analyse der Effekte auf
den Lernprozess im Kontext heterogener Lerngruppen

Lütje-Klose, Birgit

Kompetenzen von SchülerInnen mit sonderpädagogischem Förderbedarf in inklusiven und exklusiven Settings - neue Befunde aus inter- und transdisziplinärer Perspektive

Im Zentrum der aktuellen Inklusions-Debatte in Deutschland steht die Frage, wie die sehr heterogene Gruppe der SchülerInnen mit sonderpädagogischem Förderbedarf oder weiteren Risikofaktoren in der Entwicklung bestmöglich in der Ausbildung basaler Kulturtechniken unterstützt werden können. Die bislang vorliegenden nationalen (z.B. Wocken 2005) und internationalen Befunde (z.B. Gebhardt et al. 2012) sprechen mehrheitlich für Vorteile inklusiver Settings für die Leistungsentwicklung (zsf. Schnell, Sander & Federolf, 2011), jedoch mangelt es an repräsentativen und längsschnittlichen Studien. Die Beiträge dieses Symposiums liefern hierzu aktuelle Ergebnisse und sind so ausgewählt, dass sie aus erziehungswissenschaftlicher und pädagogisch-psychologischer Sicht ein inhaltlich facettenreiches Bild zeichnen.

In dem Beitrag von Neumann et al. werden im Rahmen des NRW-weiten und interdisziplinär angelegten BiLieF-Projekts gewonnene, längsschnittliche Befunde zur differentiellen Entwicklung schriftsprachlicher Kompetenzen von GrundschülerInnen in drei verschiedenen Fördersettings vorgestellt, die sich speziell auf Kinder mit Förderbedarf Lernen beziehen (Schwinger et al. 2014).

Die Analysen von Janke et al. zielen ebenfalls auf eine vergleichende Betrachtung von Grundschülerinnen, fokussieren hierbei auf Kinder mit spezifischen Sprachentwicklungsstörungen und sprachunauffälligen Gleichaltrigen. Sie satteln auf Daten einer in ganz Baden-Württemberg durchgeführten Erhebung des interdisziplinären Ki.SSES-Proluba-Projekts und beleuchten, inwiefern bei der Förderung von sprachauffälligen Kindern auch andere (Folge-)Probleme in Form von Lese-Rechtschreib- und Rechenschwächen sowie Verhaltensauffälligkeiten zu berücksichtigen sind (Forschungsgruppe Ki.SSES-Proluba 2014).

Die von Kocaj et al. berichteten Ergebnisse zum Einfluss der Klassenkomposition auf die schulischen Leistungen von Kindern mit sonderpädagogischem Förderbedarf basieren auf Daten des IQB-Ländervergleichs 2011. Hier werden die bundesweit ermittelte Leistungen Sekundarstufenschüler/-Innen mit unterschiedlichen Förderschwerpunkten adressiert, die in Relation zu den durchschnittlichen Anforderungsbedingungen analysiert werden (Kocaj et al. 2014).

Palaczek et al. berichten Ergebnisse aus der Längsschnittstudie LARS (Language And Reading Skills). Dabei werden literale und sprachliche Fähigkeiten von Kindern mit Deutsch als Erst- und Zweitsprache im Grundschulalter verglichen. Die Stichprobe umfasst Schulen aus Einzugsgebieten, welche als soziale Brennpunkte gelten. In diesem Beitrag werden die Leistungsentwicklungen der SchülerInnen mit Daten des Elternfragebogens ergänzt und familiäre Einflussfaktoren auf die Schulleistung fokussiert.

Die gewählte Bündelung der Beiträge eröffnet die Chance zu eruieren, (1) ob ein über Klassenstufen und Förderbedarfe generalisierendes "Determinationsmodell" der Kompetenzentwicklung von Heranwachsenden mit sehr heterogenen Lernvoraussetzungen wie sonderpädagogischem Förderbedarf und anderen Risikofaktoren denkbar ist, (2) inwiefern hierbei an in der empirischen Bildungswissenschaft etablierte Angebot-Nutzen-Modelle (Helmke 2009) angeschlossen werden kann und (3) welche Adaptionen im Licht der inter- und transdisziplinär gewonnenen Erkenntnisse vorzunehmen wären. Diese und andere Fragen sollen im abschließenden Diskussionsbeitrag von Claudia Mähler behandelt werden.

Donnerstag
12. März

15:30 Uhr

HGB 50

Chair(s)

Lütje-Klose, Birgit
Schwinger, Malte

Kompetenzen von SchülerInnen mit sonderpädagogischem Förderbedarf in inklusiven und exklusiven Settings - neue Befunde aus inter- und transdisziplinärer Perspektive

Neumann, Phillip; Stranghöner, Daniela; Gorges, Julia; Lütje-Klose, Birgit; Wild, Elke; Schwinger, Malte

Längsschnittliche Lese- und Rechtschreibentwicklung von Schülerinnen und Schülern mit einem sonderpädagogischen Förderbedarf im Lernen in inklusiven und exklusiven Förderarrangements – Ergebnisse aus der BiLieF-Studie

In Deutschland steigt der Anteil der integrativ beschulten SchülerInnen mit einem SFB seit der Ratifizierung der UN-Behindertenrechtskonvention im Jahr 2009. Die größte Gruppe sind weiterhin SchülerInnen mit einem SFB im Förderschwerpunkt Lernen (SFB-L) (vgl. Bildungsbericht 2014). Die historisch gewachsene Separierung von leistungsschwachen SchülerInnen im deutschen Schulsystem beruht auf der Annahme, dass SchülerInnen in Förderschulen durch kleinere und (leistungs-)homogenere Gruppen besser zu fördern seien. Dagegen sprechen Forschungsergebnisse, welche mehrheitlich einen höheren Bildungserfolg von integrativ beschulten SchülerInnen mit SFB konstatieren (vgl. zsf. Schnell, Sander & Federolf, 2011). Ein Desiderat stellen dabei jedoch Längsschnittstudien dar, welche den Entwicklungsverlauf von SchülerInnen über mehrere Schuljahre schulformübergreifend vergleichen können. Die „Bielefelder Längsschnittstudie zum Lernen in inklusiven und exklusiven Förderarrangements“ (BiLieF) begleitet SchülerInnen mit einem SFB-L von Mitte der dritten Klasse bis zum Anfang der fünften Klasse in Nordrhein-Westfalen (NRW). Dabei werden drei verschiedene Beschulungsmodelle für SchülerInnen mit einem SFB-L vergleichend untersucht, die in NRW zu Studienbeginn vorzufinden waren und zugleich Grundformen sonderpädagogischer Fördermodelle in Deutschland repräsentieren: Förderschulen, Grundschulen mit Gemeinsamen Unterricht (GU) und Grundschulen in Kooperation mit einem Kompetenzzentrum für sonderpädagogische Förderung (KsF). Die beiden inklusiv ausgerichteten Modelle unterscheiden sich hinsichtlich der Zuweisung der sonderpädagogischen Ressource und Umsetzung der Förderung: In Schulen mit GU ist die sonderpädagogische Lehrkraft Teil des Kollegiums und der Unterricht ist durch einen hohen Anteil an Doppelbesetzung mit klassenintegrierter sonderpädagogischer Förderung geprägt. In Schulen in Kooperation mit einem KsF wird sonderpädagogische Unterstützung in der Regel als externe „Serviceleistung“ von Form von Beratung und individueller Förderung umgesetzt. Theoretisches Rahmenmodell der Studie ist ein adaptiertes Angebot-Nutzungs-Modell, basierend auf den Arbeiten von Helmke (2009) und Ditton (2000). Im Zentrum stehen leistungsbezogene und psycho-soziale Merkmale der SchülerInnen, welche zu vier Messzeitpunkten erhoben und in Bezug zu schulischen und familialen Bedingungen analysiert werden.

In diesem Beitrag werden erste längsschnittliche Befunde zu der Frage berichtet, ob sich GrundschulInnen in den verschiedenen Fördersettings (FS/GU/KsF) hinsichtlich ihrer schriftsprachlichen Kompetenzentwicklung unterscheiden.

Für die Ausgangsstichprobe wurden 425 SchülerInnen aus 202 Klassen und 159 Schulen gewonnen, davon 49 Förderschulen, 84 Grundschulen mit GU und 26 Grundschulen, die mit einem KsF kooperieren. Ein hoher Prozentsatz der zu t1 teilnehmenden Kinder und Eltern konnte zu t2 wieder gewonnen werden, zudem wurden weitere, erstmalig die Einschlusskriterien erfüllende Kinder aufgenommen. Insgesamt nahmen 426 Viertklässler aus 167 Schulen und 224 Klassen teil (davon 51 FS-, 94 GU- und 22 KsF-Kinder).

Die Lese- und Rechtschreibleistung der SchülerInnen wurde zweimal (T1: Mitte der dritten Klasse; T2: Anfang der vierten Klasse) mittels standardisierter Testverfahren erhoben. Für die Leseleistung wurde der Leseverständnistest für Erst- bis Sechstklässler (ELFE 1-6; Lenhart & Schneider, 2006), für die Rechtschreibleistung die Hamburger Schreibprobe 2 (HSP 1-9; May 2010) eingesetzt. In den gruppenvergleichenden Analysen wurde die Grundintelligenz, die zum ersten Messzeitpunkt mittels zweier Untertests aus dem sprachfreien Grundintelligenztest Skala 1 (CFT 1; Cattell, Weiß & Osterland, 1997) erhoben wurde, als Hintergrundvariable kontrolliert, da die SchülerInnen an Grundschulen in Kooperation mit einem KsF signifikant höhere Werte als SchülerInnen an Grundschulen mit GU erreichten und diese wiederum höhere Werte als SchülerInnen an Förderschulen.

Im Längsschnitt zeigt sich für SchülerInnen aller Schulformen ein signifikanter Zuwachs in der Lese- und Rechtschreibkompetenz in dem betrachteten Zeitfenster. Die Ergebnisse der MANOVAs machen dabei deutlich, dass die SchülerInnen an Förderschulen einen geringeren Leistungszuwachs vor allem als die GU-SchülerInnen erzielen. Allerdings fallen die kompetenzbezogenen Unterschiede zwischen den Settings deutlich geringer aus als im Querschnitt, die ermittelten Effektstärken weisen auf eher kleine Beschulungseffekte hin. Diese und weitere Befunde werden vor dem Hintergrund unseres adaptierten AN-Modells diskutiert.

Donnerstag
12. März

15:30 Uhr

HGB 50

Janke, Bettina; Glück, Christian; Becker, Elisa; Teichert, Katherine; Hatz, Hubertus; Berg, Margit

Lernschwierigkeiten und Verhaltensauffälligkeiten bei Kindern mit Spezifischer Sprachentwicklungsstörung. Herausforderungen für die „inklusive“ Beschulung?

Kinder mit Spezifischen Sprachentwicklungsstörungen zeigen Defizite in der Sprachproduktion und / oder im Sprachverständnis, in der Wortschatzentwicklung sowie der phonologischen Informationsverarbeitung bei unauffälliger Intelligenz. Vorangegangene Studien zeigen, dass sie häufig auch Probleme beim Schriftspracherwerb und dem Erwerb mathematischer (Basis-)kompetenzen aufweisen. Häufig finden sich auch Berichte über Verhaltensauffälligkeiten.

Die Längsschnittstudie Ki.SSES-Proluba untersucht die sprachliche, soziale und Schulleistungsentwicklung von Kindern mit einer Spezifischen Sprachentwicklungsstörung (Ki.SSES) und vergleicht diese mit altersgleichen Kindern mit typischer Sprachentwicklung (Ki.TSE) ab der Einschulung bis zum Ende von Klasse 2. Die Ki.SSES besuchen bei Einschulung entweder eine Sprachheil- oder eine Regelgrundschule. Der Beitrag untersucht zunächst den Zusammenhang zwischen Rechtschreibleistung, Leseleistung und Mathematikleistung einerseits und Internalisierendem bzw. Externalisierendem Verhalten andererseits am Ende von Klasse 1. Es werden Ergebnisse von 170 Kindern (55 Mädchen) aus Baden-Württemberg dargestellt, die nach Erfüllung der Einschlusskriterien den beiden Gruppen (Ki.SSES oder Ki.TSE) zugewiesen werden konnten. Mit allen Kindern wurde vor oder kurz nach der Einschulung ein Sprachtest (SETK/HSET) und ein Intelligenztest durchgeführt, in dem sie einen mindestens durchschnittlichen IQ aufweisen ($IQ \geq 85$) mussten. Die Rechtschreibleistung wurde mit dem WRT 1+, die Leseleistung mit der WLLP-R, die Mathematikleistung mit dem DEMAT 1+ getestet. Der rezeptive Wortschatz wurde mit einer modifizierten Version des PPVT-4 (Dunn & Dunn 2007) erfasst. Die Lehrkräfte beurteilten Verhaltensprobleme am Ende von Klasse 1 mit dem SDQ (Goodman, 1997).

Die Ergebnisse einfaktorieller Varianzanalysen ergaben signifikante Unterschiede in den Bereichen Rechtschreibleistung, Lesen und Mathematik für alle Ki.SSES (unabhängig von der Beschulungsart) zugunsten der Ki.TSE. Die Ki.SSES wurden zudem von ihren Lehrern in allen SDQ-Skalen als auffälliger und weniger prosozial als die Ki.TSE eingeschätzt. Für weitere Analysen wurden die Kinder einer der folgenden Lerngruppen zugewiesen: Lese- und Rechtschreibschwäche (L+S), Rechenschwäche (MS), kombinierte Lernschwäche (LS+MS) und ohne Schwierigkeiten (OS). Die RS-Gruppe musste aufgrund von zu geringen Zellohäufigkeiten ausgeschlossen werden ($N=14$). Chi-Quadratanalysen ergaben einen signifikanten Zusammenhang zwischen den SSES-Gruppen und Lernschwierigkeitsgruppen, $\chi^2(2, N=156)=61.90, p<.001$. Von der Ki.SSES-Gruppe wurden 46% der Gruppe mit einer kombinierten Schwäche zugewiesen, von der Ki.TSE-Gruppe dagegen nur 18%. Um dem Zusammenhang zwischen Schulleistungen und Verhaltensauffälligkeiten nachzugehen, wurden nur für Ki.SSES univariate Varianzanalysen durchgeführt. Die drei Gruppen unterschieden sich im Hinblick auf die eingeschätzte Hyperaktivität ($F(2,80)=7.74, p<.01, \eta^2=.16$) und den Gesamtproblemwert im SDQ ($F(2, 80)=4.36, p<.05, \eta^2=.10$). Post-hoc-Tests zeigten, dass die LS+MS-Gruppe als am auffälligsten beurteilt wurde. Kinder der LS+RS-Gruppe wurden wiederum als auffälliger beurteilt als die Kinder ohne Schwierigkeiten.

Die sonderpädagogische Förderung für Ki.SSES verfolgt unabhängig vom Beschulungsort die Ziele der sprachlichen Rehabilitation und Sicherstellung eines für den Schulerfolg ausreichenden Leistungsniveaus. Mit dem Erreichen beider Ziele erlischt der sonderpädagogische Förderbedarf. Für knapp 10% der Ki.SSES an der Sonderschule trifft dies am Ende der 1. Klasse zu. Sie sind in die Regelgrundschule zurückgeschult worden.

Während in den sprachlichen Bereichen Aussprache und Grammatik der sonderpädagogische Förderbedarf dann abgebaut ist, wenn die Lautsprache (weitestgehend) korrekt verwendet wird, gilt als Erfolgskriterium in Bezug auf die Wortschatz- und Leseentwicklung die durchschnittliche Leistung des ersten Quartils der Kinder ohne Sprachentwicklungsstörungen (Ki.TSE). Dieses wurde im rezeptiven Wortschatz nur von etwa einem Viertel und in der Leseleistung nur etwa von der Hälfte der Ki.SSES zum Ende der Klasse 1 erreicht.

Die Ergebnisse der Analysen bestätigen, dass Kinder mit SSES neben sprachlichen Risiken und einer Kovariation mit Verhaltensauffälligkeiten auch eine Risikogruppe für die Leistungsentwicklung im Fach Deutsch und Mathematik darstellen und auch nach Rücküberweisung auf die Regelgrundschule die kriteriale Bezugsnorm in den sprachlichen Bereichen rezeptiver Wortschatz und Leseleistung überwiegend nicht erreichen, was die Leistungsheterogenität in inklusiven Settings deutlich erhöht. Die Bedeutung für die Entwicklung der Schülerinnen und Schüler in inklusiven und separierenden Unterrichtssettings wird im Hinblick auf Anforderungen für die schulische Förderung und die Unterstützung der psychosozialen Entwicklung diskutiert.

Donnerstag
12. März

15:30 Uhr

HGB 50

Chair(s)

Lütje-Klose, Birgit
Schwinger, Malte

Kompetenzen von SchülerInnen mit sonderpädagogischem Förderbedarf in inklusiven und exklusiven Settings - neue Befunde aus inter- und transdisziplinärer Perspektive

Kocaj, Aleksander; Kuhn, Poldi; Rjosk, Camilla; Jansen, Malte; Pant, Hans Anand; Stanat, Petra

Der Zusammenhang zwischen sozio-kultureller Herkunft, Klassenkomposition und Schulleistung von Kindern mit sonderpädagogischem Förderbedarf

Kinder mit sonderpädagogischem Förderbedarf (SPF) stehen im Zuge aktueller Inklusionsbemühungen zunehmend im Fokus der empirischen Bildungsforschung. Ziel inklusiver Bemühungen ist es, Schülerinnen und Schülern mit SPF bessere Bildungschancen zu ermöglichen. Darüber hinaus sollen Bildungsbenachteiligungen abgebaut werden, die mit der sozio-kulturellen Herkunft verbunden sind. Vorliegende Arbeiten geben Hinweise darauf, dass Schülerinnen und Schüler in Förderschulen besonders ungünstige Eingangsvoraussetzungen aufweisen (Bos, Müller, & Stubbe, 2010). Kinder in Förderschulen kommen häufiger aus bildungsfernen oder zugewanderten Familien (Autorengruppe Bildungsberichterstattung, 2010; Kemper & Weishaupt, 2011; Wocken, 2005). Gleichzeitig weisen sowohl Schülerinnen und Schüler aus bildungsfernen Schichten als auch Kinder mit Zuwanderungshintergrund Nachteile in ihren schulischen Leistungen auf (Stanat et al., 2012). Somit könnten Unterschiede in sozio-kulturellen Merkmalen zwischen Kindern mit SPF in Förderschulen und allgemeinen Grundschulen zu Unterschieden in ihren schulischen Leistungen beitragen. Darüber hinaus übt auch die sozio-kulturelle und fähigkeitsbezogene Komposition der Schülerschaft einen Einfluss auf die Leistungsentwicklung aus (Baumert, Stanat, & Watermann, 2006; Rauer & Schuck, 2007), wobei insbesondere leistungsschwächere Schülerinnen und Schüler von der Klassenkomposition beeinflusst zu werden scheinen (De Fraine et al., 2003). Durch die Verteilung von Schülerinnen und Schülern mit bestimmten sozio-kulturellen Merkmalen auf unterschiedliche Schularten und die damit verbundenen institutionellen Unterschiede (z. B. in Lehrplänen und Unterrichtsmethoden) können in Förderschulen und allgemeinen Grundschulen differenzielle Lern- und Entwicklungsmilieus entstehen, die über individuelle Lernvoraussetzungen hinaus die Lernentwicklung der Schülerschaft in Abhängigkeit der Art der Beschulung beeinflussen (Baumert et al., 2006).

In einer vorangegangenen Untersuchung (Kocaj et al. 2014) konnten wir feststellen, dass Schülerinnen und Schüler mit SPF in allgemeinen Grundschulen bessere schulische Leistungen im Lesen, Zuhören und in Mathematik erzielen als vergleichbare Kinder mit SPF in Förderschulen. Ungeklärt ist bislang, inwieweit Merkmale der Klassenkomposition zur Erklärung der Leistungsunterschiede beitragen.

Im vorliegenden Beitrag wurde auf Grundlage der Daten IQB-Ländervergleichs Primarstufe 2011 der Frage nachgegangen, ob Förderschulen im Vergleich zu allgemeinen Grundschulen für Kinder mit SPF differenzielle Lern- und Entwicklungsmilieus bereitstellen. Die Beantwortung der Fragestellung wurde in zwei Schritten vollzogen. Im ersten Schritt wurde geprüft, ob sich Kinder mit SPF in Abhängigkeit von der Art ihrer Beschulung hinsichtlich sozio-kultureller und demografischer Merkmale unterscheiden. Dabei wurden sowohl strukturelle Aspekte (berufliche Situation der Eltern, sozialer Status, höchster Bildungsabschluss der Eltern) als auch prozessuale Merkmale (Bildungsaspirationen der Eltern, häuslicher Buchbestand, Lesezeit) der sozio-kulturellen Herkunft in den Blick genommen. Weiterhin wurden Alter, Schuleintrittsalter, Geschlecht, Zuwanderungshintergrund der Eltern und zu Hause gesprochene Sprache der Schülerinnen und Schüler untersucht. Im zweiten Schritt wurde mit Hilfe von Mehrebenenmodellen geprüft, ob die soziale und leistungsbezogene Klassenzusammensetzung (Kompositionseffekt) sowie die Art der Beschulung (Institutionseffekt) über individuelle Schülermerkmale hinaus mit Unterschieden in den schulischen Leistungen (Leseverständnis, Mathematik) von Kindern mit SPF zusammenhängen.

Für die Analysen wurden Kinder mit SPF in Förderschulen (N = 413) und allgemeinen Grundschulen (N = 658) berücksichtigt, die eindeutig den Förderschwerpunkten Lernen, Sprache, oder emotionale und soziale Entwicklung zugeordnet werden konnten. Die Ergebnisse zeigen, dass Kinder mit SPF in allgemeinen Grundschulen insgesamt aus günstigeren sozio-kulturellen Verhältnissen stammen als Kinder in Förderschulen. Signifikante Zusammenhänge zwischen kognitiven Grundfähigkeiten auf Klassenebene und schulbezogenen Kompetenzen auf Individualebene weisen darauf hin, dass Kinder mit SPF in Klassen mit höherem Fähigkeitsniveau bessere Leistungen erzielen. Im Leseverständnis zeigte sich zusätzlich ein positiver Effekt der Beschulung in allgemeinen Grundschulen. Die soziale Klassenzusammensetzung wies keinen eigenständigen Effekt auf die Schulleistungen auf. Somit scheinen allgemeine Grundschulen im Vergleich zu Förderschulen für Kinder mit SPF bessere Lern- und Entwicklungsmöglichkeiten bereitzustellen.

Donnerstag
12. März

15:30 Uhr

HGB 50

Paleczek, Lisa; Gasteiger-Klicpera, Barbara; Seifert, Susanne; Schwab, Susanne

Die Entwicklung von Sprach-, Lese- und Rechtschreibfähigkeiten von L1 und L2 Grundschulkindern

Theoretischer Hintergrund: Lesefähigkeiten sind eng mit sprachlichen Fähigkeiten verknüpft. Dies gilt insbesondere für das Lesesinnverstehen, das in hohem Maße von den sprachlichen Fähigkeiten (v.a. Wortschatz und Grammatik) abhängt (z.B. Skebo et al., 2013). Die mangelnde Beherrschung der deutschen Sprache stellt einen der Gründe dafür dar, dass Kinder mit Deutsch als Zweitsprache (L2) schwächere Leistungen im Lesen zeigen als Gleichaltrige mit Deutsch als Erstsprache (L1) (Nusche, 2009). Defizite in der Bildungssprache können zudem nicht nur im Leseunterricht zu Schwierigkeiten führen, sondern auch in den anderen Schulfächern (Roberts, Sorajani, & Vaughn, 2010).

Fragestellung: Die Interdependenz zwischen Sprach- und Lesefähigkeiten von L1 und L2 Kindern sollte genauer beleuchtet werden, um Lernumgebungen und präventive Strategien, die die individuellen Voraussetzungen aller Kinder berücksichtigen, zu entwickeln. Innerhalb der Längsschnittstudie LARS (Language And Reading Skills) wurden die literalen und sprachlichen Fähigkeiten von L1 und L2 Grundschulkindern (2. und 3. Schulstufe) erhoben. Im Rahmen dieser Studie wurden, unter anderem, folgende Fragestellungen untersucht: (1) In welchem Ausmaß unterscheiden sich L1 und L2 Kinder in ihren Lese-, Sprach- und Rechtschreibfähigkeiten (i) zu Beginn der 2. Klasse (T1) sowie (ii) zu Beginn der 3. Klasse (T3)? (2) Wie entwickeln sich Lese-, Sprach- und Rechtschreibfähigkeiten von L1 und L2 Kindern (i) in der 2. Klasse (T1 bis T2) und (ii) in der 3. Klasse (T3 bis T4)? (3) Unterscheidet sich die Entwicklung dieser Fähigkeiten von L1 und L2 Kindern, bei Kontrolle der Ausgangsfähigkeiten?

Untersuchungsmethode: Die Dekodierfähigkeiten (SLRT II: Moll & Landerl, 2010), das Lesesinnverständnis (Elfe 1-6: Lenhard & Schneider, 2006), das Wortschatzwissen (modifizierte Form des WWT 6-10: Glück, 2007) und die Grammatikfähigkeiten (P-ITPA: Esser et al., 2010) der Kinder wurden zu Beginn (NT1=377; NT3=606) und am Ende (NT2=218; NT4=273) der zweiten und dritten Schulstufe erhoben. Die Erstsprache der Kinder wurde mittels LehrerInnenfragebögen erhoben.

Ergebnisse: Am Beginn der 2. Klasse unterscheiden sich L1 und L2 Kinder im Lesesinnverstehen ($F[2,374]=12.58, p<.001, \eta^2=.06$), dem Wortschatzwissen ($F[1,375]=17.18, p<.01, \eta^2=.04$) und den Grammatikfähigkeiten ($t[303.72]=12.65, p<.001$), wobei L2 Kinder die schlechteren Ausgangsleistungen aufweisen. Keine Unterschiede konnten jedoch in den Rechtschreib- und Dekodierfähigkeiten nachgewiesen werden. Zu Beginn der 3. Klasse spiegeln die Ergebnisse ein leicht verändertes Bild wider: L1 und L2 Kinder unterscheiden sich im Dekodieren von Wörtern ($F[1,602]=9.22, p<.001, \eta^2=.02$), aber nicht im Dekodieren von Pseudowörtern. Die Unterschiede im Lesesinnverständnis ($F[3,602]=25.63, p<.001, \eta^2=.11$), im Wortschatzwissen ($t[422.73]=14.70, p<.001$), den grammatikalischen ($t[424.54]=11.58, p<.001$) und Rechtschreibfähigkeiten ($t[402]=2.67, p<.01$) sind auch in dieser Klassenstufe signifikant, wobei erneut in allen Fähigkeiten die Ausgangsleistungen der L2 Kinder schlechter sind als die der L1 Kinder.

Bezüglich der Entwicklung von L1 und L2 Kindern während der 2. Klasse konnten Unterschiede in der Grammatik (Zeit x L1/L2: $F(1,215)=5.50, p<.05, \eta^2=.03$) und im Wortschatz (Zeit x L1/L2: $F(1,216)=11.39, p<.001, \eta^2=.05$) ermittelt werden. In der 3. Klasse hingegen konnten keine Unterschiede in der Entwicklung von L1 und L2 Kindern nachgewiesen werden. In einer Regressionsanalyse konnte gezeigt werden, dass die Erstsprache der Kinder, nach Kontrolle der Ausgangsfähigkeiten, nur auf die Entwicklung des Wortschatzes einen signifikanten Einfluss hatte ($\beta=-.10, t=-2.02; p\leq.05$).

Diskussion und Ausblick: Bei der Interpretation der Ergebnisse ist zu berücksichtigen, dass die untersuchten Schulen in Einzugsgebieten, die als soziale Brennpunkte gelten, situiert sind. Dies zeigte sich darin, dass auch die Ergebnisse von L1 Kindern in einigen der untersuchten Bereiche unterdurchschnittlich waren. Es ist anzunehmen, dass die Familien der Kinder über geringes kulturelles Kapital (Bourdieu, 1983) verfügen und nur wenige (schrift-)sprachliche Anregungen im häuslichen Umfeld vorhanden sind. Die Familie ist jedoch neben der Schule für den Bildungserfolg bedeutsam (Brake & Büchner, 2006). Um familiäre Einflussfaktoren zu untersuchen, werden in der Präsentation die Ergebnisse eines Elternfragebogens dargestellt. Zusätzlich wird die hierarchische Struktur der Daten berücksichtigt und mehrebenenanalytische Verfahren angewendet.

Donnerstag
12. März

15:30 Uhr

HGB 50

Chair(s)

Lütje-Klose, Birgit
Schwinger, Malte

Kompetenzen von SchülerInnen mit sonderpädagogischem Förderbedarf in inklusiven und exklusiven Settings - neue Befunde aus inter- und transdisziplinärer Perspektive

Hartmann, Stefan

Validierungspraxis in der empirischen Bildungsforschung

In der empirischen Bildungsforschung werden zur Erfassung von Kompetenzen häufig schriftliche Tests eingesetzt. Geschlossenen Antwortformaten kommt dabei eine besondere Bedeutung zu, da diese im Vergleich zu anderen Testverfahren eine reliable Testung bei geringen Kosten ermöglichen (Stecher & Klein, 1997). Beim Umgang mit den Ergebnissen stellt sich immer wieder aufs Neue die Frage, inwieweit es berechtigt ist, die mittels Ankreuz-Antworten erhobenen Testleistungen als gültiges Maß für die beschriebenen Kompetenzen zu interpretieren. Die Maßnahmen zur Beantwortung dieser Frage werden unter dem Begriff "Validierung" zusammengefasst. Ihr Ziel ist das Sammeln von Evidenz dafür, dass die Interpretation der Testwerte hinsichtlich eines vorab definierten Zwecks gerechtfertigt ist. Gemäß der Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association & National Council on Measurement in Education [AERA, APA & NCME], 2014) kommen als Quellen dieser Evidenz fünf Bereiche infrage: Testinhalt, Antwortprozesse, interne Struktur, Beziehung zu anderen Variablen sowie Konsequenzen der Testung. Die vier Beiträge des Symposiums greifen mit unterschiedlichen Methoden auf diese Evidenzquellen zu:

Buchholz, Hartig, Mehren und Rempfler nutzen konfirmatorische Faktorenanalysen, Strukturgleichungsmodelle und manifeste Regressionen, um interne Struktur und prognostizierte Itemschwierigkeiten eines Tests zur geografischen Systemkompetenz zu beurteilen und Hypothesen zur Beziehung der Testergebnisse mit anderen Variablen zu testen. Friedrich und Heise untersuchen mit einem faktorenanalytischen Ansatz die interne Struktur eines neu entwickelten Instruments zur Messung von Textverständlichkeit und prüfen Hypothesen zur prädiktiven Validität der Verständlichkeitsmerkmale als Vorhersagekriterien für Emotionen beim Lesen und für Verstehensleistungen. Mathesius, Upmeier zu Belzen und Krüger nutzen die Methode des lauten Denkens, um zu beurteilen, inwiefern Multiple-Choice-Aufgaben zu Kompetenzen der naturwissenschaftlichen Erkenntnisgewinnung in Biologie vorhergesagte Antwortprozesse initiieren. Hartmann, Krüger, Upmeier zu Belzen und Pant stellen am Beispiel eines Tests zu Kompetenzen aus dem Bereich des wissenschaftlichen Denkens ein Verfahren zur Optimierung von Known-Groups-Vergleichen vor, mit dem sich konfundierende Effekte minimieren lassen, wenn theoriebasiert vorhergesagte Leistungsunterschiede zwischen Extremgruppen geprüft werden sollen.

Die vier Beiträge verdeutlichen, dass Validierung in den empirischen Bildungswissenschaften kein Lippenbekenntnis ist. Sie liefern Beispiele dafür, wie das Sammeln von Evidenzen für die Validität der Interpretation von Testleistungen in der Praxis auf vielfältige Art und Weise, mit hohem Aufwand und großer Sorgfalt sowie unter Anwendung eines breiten Spektrums innovativer aber auch bekannter Verfahren betrieben werden kann. Die Heterogenität der vorgestellten Methoden ist dabei nicht zuletzt ein positiver Synergieeffekt der interdisziplinären Vernetzung von Fachdidaktik, Psychologie und Methodenforschung in den einzelnen Projekten.

Donnerstag
12. März

15:30 Uhr

HGC 20

Chair(s)

Upmeier zu Belzen, Annette
Hartmann, Stefan

Buchholz, Janine; Hartig, Johannes; Mehren, Rainer; Rempfler, Armin

Validierung eines Tests zur geografischen Systemkompetenz – Abgrenzung zu Lesefähigkeit und fluider Intelligenz

Theoretischer Hintergrund: Erstmals wurde ein theoretisch fundiertes und empirisch überprüfbares Modell der Systemkompetenz im Geografieunterricht für den Mittleren Schulabschluss entwickelt, wobei die Analysen zur Validität den Fokus dieses Vortrags bilden. Dem vorausgegangen ist die Ableitung eines dreidimensionalen Struktur- und Stufenmodells geografischer Systemkompetenz aus der Literatur, sowie die Entwicklung von 154 Items unterschiedlicher Stufung (d.h. monokausale, lineare und komplexe Vernetzung) zu den Dimensionen Systemorganisation, Systemverhalten und Systemadäquater Handlungsintention.

Fragestellung: Bei GeoSysKo handelt es sich um ein papierbasiert administriertes, sprachlich dargebotenes Testinstrument zur Erfassung geografischer Systemkompetenz. Zur Gewährleistung valider Schlussfolgerungen der mit dem Testinstrument ermittelten Testscores stellt sich, neben einer Prüfung der angenommenen Struktur und Stufung, die Frage, ob sich die gemessene geografische Systemkompetenz von allgemeiner Lesefähigkeit sowie fluider Intelligenz abgrenzen lässt.

Methode: Die Daten der Hauptuntersuchung stammen von 1926 Schülerinnen und Schülern der 9. Jahrgangsstufe in Bayern. Zur Untersuchung der Validität wurden Analysen (a) zum Strukturmodell, (b) zum Stufenmodell, sowie (c) zur Abgrenzbarkeit der geografischen Systemkompetenz von Lesefähigkeit und Intelligenz durchgeführt.

Für (a) wurden ein- und mehrdimensionale IRT Modelle geschätzt, hinsichtlich ihrer latenten Korrelationen inspiziert und ihre Modellgüte verglichen (Evidenzquelle der Validität basierend auf interner Struktur; AERA/APA/NCME, 2014). Für (b) wurden die empirischen Itemschwierigkeiten anhand der bei der Itementwicklung angenommenen Stufung vorhergesagt (Hinweis auf Validität basierend auf Antwortprozessen; AERA/APA/NCME, 2014). Für (c) liegen Daten eines simultan erhobenen Lesegeschwindigkeitstests (LGVT; Schneider, Schlagmüller & Ennemoser, in Vorb.) sowie eines Tests zur Erfassung fluider Intelligenz (BEFKI; Wilhelm, Schroeders & Schipolowski, in Vorb.) vor. In einem Strukturgleichungsmodell wurde untersucht, in welchem Ausmaß diese die GeoSysKo-Testscores vorherzusagen vermögen (Hinweis auf Validität basierend auf Beziehungen zu anderen Variablen; AERA/APA/NCME, 2014). Sollten sich dabei keine allzu hohen Zusammenhänge zeigen, kann dies als Hinweis für die Abgrenzbarkeit der mit dem Testinstrument erfassten Kompetenz dienen.

Ergebnisse: Bezogen auf die Modellgüte erwies sich das postulierte dreidimensionale Modell gegenüber einem eindimensionalen Modell als überlegen. Allerdings führte die hohe latente Korrelation, die noch bessere Modellgüte des alternativen Modells sowie inhaltliche Überlegungen (bspw. Köck, 2004) dazu, die Dimensionen zum Wissenserwerb (Systemorganisation und Systemverhalten) zusammenzufassen und einer Teilkompetenz gegenüberzustellen, die sich stärker auf Wissensanwendung bezieht (Systemadäquate Handlungsintention). Im Folgenden wurde daher von einem zweidimensionalen Modell geografischer Systemkompetenz ausgegangen.

Die aus diesem Modell resultierenden empirischen Itemschwierigkeiten wurden anschließend in einer OLS-Regression auf manifester Ebene durch die theoretisch angenommene Stufung der Items vorherzusagen versucht (vgl. Hartig, Frey, Nold & Klieme, 2012). Demnach können 29% der Varianz in den Schwierigkeiten der Items, die die erste Dimension (Systemorganisation, Systemverhalten) messen, durch die angenommene Stufung erklärt werden. Für Dimension 2 (Systemadäquate Handlungsintention) beträgt der Wert sogar 57,2%. Dieser differentielle Befund stärkt die den Analysen zugrundeliegende Faktorstruktur.

Schließlich wurden die Testscores der beiden GeoSysKo-Dimensionen durch LGVT- und BEFKI-Score vorhergesagt. Etwa jeweils die Hälfte der Varianz lassen sich erklären, wobei der Zusammenhang der GeoSysKo-Dimensionen mit Lesegeschwindigkeit etwas stärker war. Dieser Befund stärkt die Annahme, dass das Testinstrument mehr als die mit dem Lesegeschwindigkeits- und Intelligenztest gemessenen Fähigkeiten erfasst. Allerdings erhöhte sich im Gesamtmodell die Residualkorrelation gegenüber der im Strukturmodell ermittelten latenten Korrelation zwischen den GeoSysKo-Dimensionen. Dies impliziert, dass die Dimensionalität des Konstrukts aus unterschiedlichen Anforderungen der Items an Lesefähigkeit und Intelligenz resultiert, bedarf bis zur Konferenz jedoch noch vertiefender Analysen.

Zusammenfassend lässt sich zunächst festhalten, dass die verschiedenen hier vorgenommenen Analysen eine reichhaltige Informationsgrundlage zur Diskussion der Validität der GeoSysKo-Testwertinterpretationen bereitstellen. Die Analysen zur Dimensionalität liefern Hinweis auf das aus der Literatur abgeleitete Strukturmodell, während das Stufenmodell durch die gute Vorhersagekraft der theoretisch angenommenen Aufgabenanforderungen empirisch unterstützt wird. Es wurde auch Evidenz gesammelt, dass das schriftliche Instrument zur Erfassung der geografischen Systemkompetenz von Lesefähigkeit und Intelligenz, zumindest in Teilen, trennbar ist.

Donnerstag
12. März

15:30 Uhr

HGC 20

Chair(s)

Upmeier zu Belzen, Annette
Hartmann, Stefan

Friedrich, Marcus C. G.; Heise, Elke

Validierung eines Fragebogens zur Erfassung der Textverständlichkeit

Textverständlichkeit ist ein wichtiger Prädiktor für Textverstehen (Benjamin, 2012). Langer, Schulz von Thun und Tausch (2006) fanden zudem einen Einfluss der Textverständlichkeit auf das emotionale Erleben beim Lernen mit Texten. Gegenwärtig fehlt es allerdings an Instrumenten zur theoretisch fundierten und ökonomischen Erfassung der Textverständlichkeit. Ein Vergleich des Konzepts von Kintsch und Vipond (1979) mit anderen Konzepten der Textverständlichkeit ergab 11 Merkmale der Textverständlichkeit: Wortschwierigkeit, Satzschwierigkeit, Textlänge, Propositionsdichte, Argumentdichte, Aufwand zur Inferenzbildung, Aufwand für Reinstatements, Aufwand für Reorganisationen, Variation der Sprache, Hervorhebungen, und Anschaulichkeit.

Zu jedem dieser Merkmale wurden 5 Items generiert und zu einem Fragebogen zusammengestellt. In einer früheren Untersuchung zeigte sich, dass mit Hilfe dieses Fragebogens sieben Faktoren reliabel erfasst werden können, und zwar Wortschwierigkeit, Satzschwierigkeit, Argumentdichte, Propositionsdichte, Aufwand für Reorganisationen, Variation der Sprache und ein Faktor, der als „Klarheit der Vorstellung vom Inhalt“ interpretiert wird. Die vorliegende Studie geht der Frage nach, wie konstruktvalide die Skalen des Fragebogens Textverständlichkeit sind und wie gut mit Hilfe der Skalen Textverstehen und Emotionen beim Lesen vorhergesagt werden können (prognostische Validität).

Dazu wurden von einem Ausgangstext aus einem Geographie-Lehrbuch für die 11. Jahrgangsstufe sieben weitere Versionen erzeugt, die jeweils auf die Variation eines Merkmals der Textverständlichkeit abzielten. In einer Online-Studie mit $N = 174$ Teilnehmenden wurde jeder Person per Zufall jeweils eine Version des Textes vorgelegt. Im Anschluss bearbeiteten die Befragten einen Fragebogen zur Intensität positiver und negativer Emotionen beim Lesen, den Fragebogen zur Textverständlichkeit und einen Verstehenstest zum Inhalt des Textes.

Zwischen dem tatsächlichen Textverstehen und den Skalen Wortschwierigkeit, Satzschwierigkeit, Argumentdichte und Aufwand für Reorganisationen zeigten sich Korrelationen zwischen $r = .08$ und $.11$; Zwischen dem tatsächlichen Textverstehen und den Skalen Propositionsdichte, Variation der Sprache und Klarheit der Vorstellungen zeigten sich Korrelationen zwischen $r = .24$ und $.31$.

Zwischen der Skala zur Erfassung positiver Emotionen und den Skalen Wortschwierigkeit, Satzschwierigkeit, Argumentdichte und Aufwand für Reorganisationen zeigten sich Korrelationen zwischen $r = .03$ und $.14$. Zwischen dem tatsächlichen Textverstehen und den Skalen Propositionsdichte, Variation der Sprache und Klarheit der Vorstellungen zeigten sich Korrelationen zwischen $r = .36$ und $.78$.

Zwischen der Skala zur Erfassung negativer Emotionen und den Skalen Variation der Sprache, Klarheit der Vorstellung und Propositionsdichte zeigten sich Korrelationen zwischen $r = .01$ und $.19$. Zwischen den Skalen zur Erfassung negativer Emotionen und den Skalen Wortschwierigkeit, Satzschwierigkeit, Argumentdichte und Aufwand für Reorganisationen zeigten sich Korrelationen zwischen $r = .23$ und $.35$.

Zur Konstruktvalidierung der Skalen wurde in t-Tests geprüft, inwiefern die Manipulation der Texte sich auf die Bewertungen der Texte durch die Lesenden auswirkte. Dazu wurden die verschiedenen Versionen des Textes jeweils hinsichtlich des manipulierten Merkmals mit dem Original-Text verglichen. Hinsichtlich der Skalen Wortschwierigkeit, Satzschwierigkeit, Argumentdichte und Aufwand für Reorganisationen ergaben sich signifikante Effekte zwischen $d = .51$ und $.84$. Hinsichtlich der Skalen Propositionsdichte, Variation der Sprache und Klarheit der Vorstellung ergaben sich nicht signifikante Effekte zwischen $d = .02$ und $.12$.

Die Ergebnisse sprechen für die Konstruktvalidität von vier der sieben Skalen. Zur Validierung der Skala Propositionsdichte soll nun die tatsächliche Propositionsdichte der verschiedenen Versionen des Textes ermittelt werden, um die Skala mit dem objektiven Außenkriterium zu korrelieren. Der weitere Umgang mit den Skalen „Variation der Sprache“ und „Klarheit der Vorstellung“ wird diskutiert.

Die Zusammenhänge zwischen den Skalen und dem tatsächlichen Textverstehen sind niedriger ausgefallen, als bei einem Fragebogen erwartet wurde, der die Reibungslosigkeit der Textverarbeitung erfassen soll. Eine mögliche Ursache für diese niedrigen Zusammenhänge kann darin bestehen, dass die Lesezeit nicht erfasst wurde. Lesende, die den Text weniger verständlich fanden, haben sich möglicherweise mehr Zeit genommen, um zum selben Verständnis zu gelangen. Dies soll in einer Folgestudie geprüft werden.

Donnerstag
12. März

15:30 Uhr

HGC 20

Chair(s)

Upmeier zu Belzen, Annette
Hartmann, Stefan

Mathesius, Sabrina; Upmeier zu Belzen, Annette; Krüger, Dirk

Lautes Denken bei der Bearbeitung von Multiple-Choice-Aufgaben zur Erkenntnisgewinnung im Projekt Ko-WADiS: Ein Ansatz zur Validierung von Testergebnissen in der Biologie.

Naturwissenschaftliche Erkenntnisgewinnung kann als komplexer Problemlöseprozess aufgefasst werden, der die Anwendung von methodischem und inhaltlichem Konzeptwissen vereint (Mayer, 2007). Dies beinhaltet den Umgang mit charakteristischen Methoden und stellt somit einen Teil der naturwissenschaftlichen Grundbildung dar (Bybee, 2002). Im interdisziplinären Forschungsprojekt Ko-WADiS gilt es, für die Methoden „Untersuchen“ (Mayer, 2007; Teilkompetenzen: Fragestellungen generieren, Hypothesen formulieren, Untersuchungen planen, Untersuchungen auswerten) und „Modelle nutzen“ (Upmeier zu Belzen & Krüger, 2010; Teilkompetenzen: Zweck von Modellen, Testen von Modellen, Ändern von Modellen) ein Messinstrument zur domänenübergreifenden Erfassung und Modellierung von Kompetenzen Lehramtsstudierender der Biologie, Chemie und Physik zu entwickeln, zu validieren und zu evaluieren. Das langfristige Ziel ist es, eine Mehrkohorten-Längsschnittstudie über den gesamten Studienverlauf hinweg durchzuführen und Implikationen zur Optimierung der lehramtsbezogenen Studiengänge abzuleiten.

In der ersten Projektphase wurde basierend auf bestehenden Kompetenzmodellen und inhaltlich am universitären Curriculum orientiert ein Messinstrument mit 123 Multiple-Choice-Aufgaben (42 mit biologischen Kontexten) konstruiert. Um die Validität der aus den Testleistungen ableitbaren Interpretationen zu erhöhen, wurden die Antwortalternativen induktiv auf der Basis von im Vorfeld offen erhobenen Studierendenantworten konstruiert, sodass diese plausibel und zielgruppenorientiert formuliert sind (vgl. Sadler, 1998). Innerhalb des Entwicklungsprozesses wurden sowohl die theoretische Fundierung als auch die psychometrische Qualität als Kriterien bei der Auswahl geeigneter Aufgaben herangezogen.

Weiterführend stellen sich jedoch die Fragen, inwiefern die Aufgaben den erwarteten kognitiven Problemlöseprozess initiieren und inwieweit die Testergebnisse für die Multiple-Choice-Aufgaben als manifeste Indikatoren zur Erfassung des latenten Konstrukts naturwissenschaftliche Erkenntnisgewinnung interpretiert werden können (vgl. AERA, APA & NCME, 2014).

Zur Untersuchung dieser Fragen wurde bei einer Teilstichprobe von Lehramtsstudierenden der Naturwissenschaften ($n = 48$) aus unterschiedlichen Semestern und mit unterschiedlichen Fähigkeitsniveaus während der Bearbeitung von jeweils fünf bis 14 ausgewählten Aufgaben mit biologischen Kontexten die Methode des lauten Denkens angewendet (Ericsson & Simon, 1980). Das Fähigkeitsniveau wurde durch eine vorhergehende Bearbeitung eines Testhefts mit 14 Aufgaben (jeweils zwei Aufgaben aus jeder der sieben Teilkompetenzen) erfasst. Die Audiografien des lauten Denkens bei der Bearbeitung der Multiple-Choice-Aufgaben wurden entsprechend dem Vorgehen bei der qualitativen Inhaltsanalyse (Mayring, 2010) transkribiert, redigiert und deduktiv unter Berücksichtigung der theoretisch fundierten Konstruktion sowie erweiternd induktiv codiert; die Zweitkodierung ist zur Tagung abgeschlossen. Die Analyse lässt erkennen, dass bei der Bearbeitung der Aufgaben kognitive Problemlöseprozesse aktiviert werden, die auf ein Denken innerhalb der anvisierten Teilkompetenzen schließen lassen. Die Studierenden argumentierten dabei innerhalb des Wissensbereichs der naturwissenschaftlichen Methoden „Untersuchen“ bzw. „Modelle nutzen“. Es besteht bezogen auf 385 Aufgabenbearbeitungen eine Passung von 81,6% zwischen der Auswahl eines Distraktors bzw. Attraktors im Testheft und innerhalb der Situation beim lauten Denken. Die Abweichungen hin zu bzw. weg von einer elaborierten Antwortalternative lassen keine positiven oder negativen Effekte der Methode ausmachen. Weiter ist festzustellen, dass die mit dem schriftlichen Test erfassten Fähigkeiten der Probanden mittlere bis hohe Korrelationen mit der durchschnittlichen Lösefähigkeit beim lauten Denken aufweisen. Die Entscheidung für eine Antwortalternative sowie die Begründung dieser lassen sich in Kreuztabellen darstellen; diese verdeutlichen, inwiefern der Attraktor (nicht) gewählt und dies (nicht) kompetent begründet wurde. Ferner lassen sich aus den dokumentierten Antwortprozessen beim lauten Denken die genutzten Strategien zur Lösung der Problemstellungen in jeder Teilkompetenz qualitativ einschätzen und es lassen sich daraus Rückschlüsse auf die zur Lösung der Aufgaben eingesetzten Kompetenzen ableiten. Es zeigt sich, dass die Studierenden methodisches und inhaltliches Konzept- sowie Strategiewissen zur Lösung der Aufgaben nutzen. Bezüge zu bekannten Struktur- und Niveaubeschreibungen entsprechend der bei der Konstruktion als theoretische Basis genutzten Kompetenzmodelle sind erkennbar. Auf der Tagung werden ausgewählte Kategorien zur Strukturierung der angewendeten Konzepte exemplarisch vorgestellt. Die Eignung der Maßnahme als mögliche Evidenzquelle zur Validierung der Testergebnisse im Messinstrument wird diskutiert und es wird die Relevanz der Verknüpfung quantitativer und qualitativer Ansätze entsprechend eines Mixed-Method-Designs in den Blick genommen.

Donnerstag
12. März

15:30 Uhr

HGC 20

Chair(s)

Upmeier zu Belzen, Annette
Hartmann, Stefan

Hartmann, Stefan; Krüger, Dirk; Upmeier zu Belzen, Annette; Pant, Hans Anand

Known-Groups-Vergleiche mit gematchten Stichproben zur Validierung der Ergebnisse eines schriftlichen Tests zum wissenschaftlichen Denken

Ziel des vorgestellten Projektes ist es, Kompetenzen des wissenschaftlichen Denkens (Mayer, 2007) bei Lehramtsstudierenden der Fächer Biologie, Chemie und Physik zu erfassen. Als Instrument kommt ein schriftlicher Test mit Multiple-Choice-Aufgaben zum Einsatz, der mit Methoden der Item-Response-Theorie ausgewertet wird. Aus validitätstheoretischer Sicht stellt sich die Frage, inwieweit es gerechtfertigt ist, die mit diesem Instrument erhobenen Testleistungen als Maß für Kompetenzen im Bereich des wissenschaftlichen Denkens zu interpretieren. Zur Beantwortung dieser Frage gilt es, empirische Evidenz für die Validität der Interpretation zu sammeln. Als Evidenzquellen kommen in Betracht: Testinhalt, Antwortprozesse, interne Struktur, Beziehung zu anderen Variablen sowie Konsequenzen der Testung (AERA, APA & NCME, 2014).

Dem Bereich "Relation zu anderen Variablen" lassen sich Known-Groups-Vergleiche (Cronbach & Meehl, 1955) zuordnen. Hierbei wird geprüft, ob sich theoriebasiert vorhergesagte Kompetenzunterschiede zwischen bestimmten Personengruppen empirisch in Form unterschiedlicher Testleistungen zeigen. Bezüglich des wissenschaftlichen Denkens wurden mit der Anzahl der Lerngelegenheiten, der inhaltlichen Breite und dem didaktischen Ansatz bei der Kompetenzvermittlung drei Einflussfaktoren auf den Kompetenzstand identifiziert. Die Anzahl der Lerngelegenheiten nimmt im Studienverlauf zu. Dementsprechend wurde die Hypothese aufgestellt, dass Studierende in der Masterphase über höhere Kompetenzen verfügen als Studierende in der Bachelorphase. Von einer höheren inhaltlichen Breite profitieren Studierende, die in Kombinationsstudiengängen zwei naturwissenschaftliche Fächer studieren. Dementsprechend wurde die Hypothese aufgestellt, dass Studierende mit zwei naturwissenschaftlichen Fächern über höhere Kompetenzen verfügen als Studierende mit nur einem naturwissenschaftlichen Fach. Hinsichtlich des didaktischen Ansatzes bei der Kompetenzvermittlung kann zwischen "teaching science" und "teaching about science" unterschieden werden, wobei das Zweite die effizientere Methode zur Vermittlung von Kompetenzen des wissenschaftlichen Denkens darzustellen scheint (Hodson, 2014). Dementsprechend wurde die Hypothese aufgestellt, dass Lehramtsstudierende, die in Lehrveranstaltungen über die Natur der Naturwissenschaften unterrichtet werden, über höhere Kompetenzen verfügen als Studierende der Naturwissenschaften, denen wissenschaftstheoretische Inhalte nicht explizit vermittelt werden.

Um eine Known-Groups-Validierung in Form einer Prüfung der genannten Hypothesen vorzunehmen, wurde der vorab normierte schriftliche Test zum wissenschaftlichen Denken (Hartmann, Mathesius, Stiller, Straube, Krüger & Upmeier zu Belzen, in Druck) in einer Stichprobe von 2065 Studierenden eingesetzt. Zur Schätzung von Personenfähigkeiten wurde ein eindimensionales IRT-Modell skaliert, als Personenschätzer dienten Plausible Values. Die Gruppenunterschiede wurden zunächst in Form eines multiplen Regressionsmodells direkt im IRT-Modell geschätzt. Hierbei erwiesen sich die Einflüsse konfundierender Personenmerkmale, Multikollinearität der Prädiktoren sowie ungleiche Zellenbesetzungen als problematisch. Aufgrund der für quasi-experimentelle Studien typischen Bildung von Gruppen anhand vorab gegebener Merkmale war es nicht möglich, diese Einflüsse durch Randomisierung zu kontrollieren. Diesem Problem wurde mit Propensity Score Matching (Rosenbaum & Rubin, 1983) begegnet. Mit dem Verfahren lassen sich aus Versuchs- und Kontrollgruppen Teilstichproben ziehen, die hinsichtlich mutmaßlich konfundierender Hintergrundmerkmale gleich zusammengesetzt sind. Mittelwertsunterschiede zwischen gematchten Stichproben lassen sich infolgedessen zuverlässiger auf die Gruppenvariable zurückführen.

Für zwei der drei Hypothesen wurden signifikante Effekte in der vorhergesagten Richtung gefunden: Masterstudierende schnitten besser ab als Bachelorstudierende und Studierende mit zwei naturwissenschaftlichen Fächern schnitten besser ab als Studierende mit einem naturwissenschaftlichen Fach. Dabei zeigte sich jeweils, dass die Effektstärken beim Vergleich gematchter Teilstichproben höher ausfallen als in der Gesamtstichprobe. Im latenten Regressionsmodell wurden die tatsächlichen Effekte demnach unterschätzt.

Für die dritte Hypothese wurde ein Effekt entgegen der vorhergesagten Richtung gefunden: Naturwissenschaftsstudierende schnitten besser ab als Lehramtsstudierende. In weiterführenden Analysen deutete sich an, dass die Ursache hierfür in einer Interaktion aus Studienphase und Studientyp zu suchen ist: Lehramtsstudierende profitieren beim Erwerb von Kompetenzen des wissenschaftlichen Denkens offenbar insbesondere von Lehrveranstaltungen in der Masterphase ihres Studiums.

Insgesamt stützen die Ergebnisse die Annahme, dass die Testleistungen als Maß für Kompetenzen des wissenschaftlichen Denkens interpretiert werden können. Sie stehen damit im Einklang mit weiteren Validierungsmaßnahmen des Projekts (s.a. den Vortrag von Sabrina Mathesius in diesem Symposium).

Donnerstag
12. März

15:30 Uhr

HGC 20

Chair(s)

Upmeier zu Belzen, Annette
Hartmann, Stefan

Schult, Johannes

Intelligenz.Wert.Schätzen. – Intelligenz und Bildungserfolg

Die allgemeine Intelligenz weist nicht nur enge Zusammenhänge mit unterschiedlichen Erfolgsindikatoren in Schule, Hochschule und Leben in westlichen Gesellschaften auf, sondern erlaubt auch gute Vorhersagen späteren Erfolgs (vgl. z.B. Jensen, 1998; Rost, 2013). Trotz dieser vielfach dokumentierten Bedeutsamkeit wird Intelligenz jedoch oftmals einseitig kritisiert oder einfach ignoriert – so wird Intelligenz in der Bildungsforschung nur sporadisch wertgeschätzt. Aufbauend auf der breiten Befundlage der (pädagogisch-)psychologischen Intelligenzforschung informiert das Symposium über aktuelle Ergebnisse zu differenziellen und spezifischen Zusammenhängen zwischen einerseits Intelligenz und andererseits schulischen Leistungen und späteren Bildungsrenditen. Dabei werden die Struktur von Intelligenz und kognitiven Leistungen sowie die Konsequenzen für die Unterrichts- und Beratungspraxis thematisiert.

In den einzelnen Vorträgen werden neben (klassischen) Intelligenzmaßen insbesondere Konstrukte betrachtet, die konzeptuell zwischen der psychometrischen Intelligenz und (hoch-)schulischen Performanz verortet sind. Dazu zählen komplexes Problemlösen (als eine Kernkompetenz des 21. Jahrhunderts) und Studierfähigkeit (der in Hochschul-Auswahlverfahren große Bedeutung zukommt). Disziplinär sind die Beiträge in der (Pädagogischen) Psychologie und Empirischen Erziehungswissenschaft verortet. Neben der allgemeinen Bedeutung von Intelligenz werden somit spezifische Auswirkungen auf bildungsrelevante Merkmale betrachtet und ein Brückenschlag zwischen psychologischer Theorie und bildungswissenschaftlicher Forschungspraxis präsentiert.

Im Beitrag von Valerius und Sparfeldt wird der Zusammenhang zwischen Intelligenz und Schulerfolg näher beleuchtet. Sowohl die kognitiven Fähigkeiten, als auch die Schulleistung werden mit genesteten, inhaltspezifischen Faktoren sowie mit Generalfaktoren modelliert. Auf der Ebene von Schulfächern zeigen sich dabei konvergent-diskriminante Validitätsmuster, die für die Relevanz von Sub-g-Faktoren in der Pädagogischen Psychologie sprechen.

Ähnlich differenziert betrachten Schult und Sparfeldt latente Modelle zur Erklärung von Studienerfolg mit Hilfe von allgemeinen und inhaltspezifischen Intelligenzfaktoren in einer größeren US-amerikanischen Stichprobe. Hier zeigen sich ähnliche Beziehungsmuster.

Brunner, Fischbach, Baudson, Preckel und Martin widmen sich anschließend der Betrachtung von Intelligenz und Bildungserfolg aus einer weite Teile der Lebensspanne abdeckenden Perspektive. Einerseits wird dabei die Bedeutsamkeit von Intelligenz für erfolgreiche Bildungskarrieren unterstrichen; andererseits sprechen die Ergebnisse für die Notwendigkeit von angemessener diagnostischer Kompetenz auf Seiten der Lehrkräfte.

Dass die prognostische Validität von Intelligenz hinsichtlich Erfolg im Studium durch die Erfassung von komplexen Problemlösefähigkeiten (KPL) gesteigert werden kann, zeigen Stadler, Becker, Niepel und Greiff. Studienerfolg beschränkt sich in diesem Zusammenhang nicht nur auf die Prüfungsleistungen, sondern auch um die subjektive Selbsteinschätzung. In letzterem Bereich zeigt KPL höhere Validitäten als Aufgaben zum schlussfolgernden Denken.

Ricarda Steinmayr diskutiert abschließend die Befunde, die u.a. verschiedene zeitliche, geographische und biographische Bereiche abdecken. Dabei werden die Beiträge der Intelligenz im bildungswissenschaftlichen und pädagogisch-psychologischen Kontext kritisch-wertschätzend verbunden.

Donnerstag
12. März

15:30 Uhr

HGC 30

Chair(s)

Schult, Johannes
Sparfeldt, Jörn

Valerius, Sonja; Sparfeldt, Jörn

Zusammenhänge zwischen Intelligenz und Schulleistungen: allgemein (g) und spezifisch (Gruppenfaktoren)?

Die besondere Bedeutung der allgemeinen Intelligenz (im Sinne des Spearmanschen g-Faktors) als Korrelat und Prädiktor schulischer Leistungen ist vielfach belegt (z.B. Jensen, 1998; Rost, 2013); im Mittel korrelieren allgemeine Intelligenz und Schulleistungen um $.5 \leq r \leq .6$. Weniger Einigkeit herrscht hingegen darüber, ob und wenn ja welche Intelligenz(gruppen)faktoren systematisch mit spezifischeren Schulleistungen in einzelnen Schulfächern oder Schulfachgruppen kovariieren. Auch wenn theoretisch differenzierte Zusammenhänge der z.B. hinsichtlich der Inhalte oder Materialarten korrespondierenden Intelligenz- und Schulleistungsfaktoren erwartet werden können, haben sich bisher nur wenige Studien der entsprechenden Forschungsfrage systematisch genähert. Hierfür bieten sich insbesondere nested-factor-Modelle an (z.B. Brunner et al., 2012), die – über die Korrelation eines Intelligenzgeneralfaktors mit einem Schulleistungsgeneralfaktor hinaus – die Möglichkeit der Analyse der Zusammenhänge der jeweils vom (Intelligenz- bzw. Schulleistungs-)Generalfaktor unabhängigen, domänenspezifischen (Intelligenz- bzw. Schulleistungs-)Gruppenfaktoren bieten. Die Forschungsfrage des vorliegenden Beitrags bezieht sich entsprechend auf die Klärung des differenzierten Zusammenhangsgefüges von allgemeinen und spezifischen Intelligenzfaktoren sowie allgemeinen und spezifischen Schulleistungen (Zensuren).

Daher bearbeiteten 562 Gymnasiasten der 9. Klassenstufe insgesamt 21 Subtests aus zwei Intelligenzstrukturtestbatterien (Kurzformen des KFT 4-12+R und des BIS-4). Die Schulleistungen in Deutsch, Mathematik, Englisch, Französisch, Physik und Geschichte wurden über die Zensuren des letzten Halbjahreszeugnisses erfasst. In kombinierten Nested-Factor-Modellen wurden die Zusammenhänge des allgemeinen (g) sowie der spezifischen Intelligenzfaktoren (verbal, numerisch, figural) mit einem allgemeinen Schulleistungsfaktor sowie spezifischen (verbal, mathematisch/naturwissenschaftlich) Schulleistungsfaktoren untersucht.

Der allgemeine Schulleistungsfaktor hing mit dem allgemeinen Intelligenz-g-Faktor mittelhoch und substantiell zusammen. Zusätzlich zeigten sich konvergente Korrelationen der spezifischen verbalen und spezifischen numerischen Intelligenz-Inhaltsfaktoren mit den spezifischen verbalen und mathematisch/naturwissenschaftlichen Schulleistungsfaktoren in mittlerer bis großer Größe sowie geringere divergente Korrelationskoeffizienten. Der spezifische figurale Intelligenz-Inhaltsfaktor korrelierte positiv sowohl mit dem spezifisch verbalen als auch spezifisch mathematisch/naturwissenschaftlichen Schulleistungsfaktor, in absoluter Höhe jedoch niedriger als die entsprechenden konvergenten Intelligenz-Inhaltsfaktoren.

Die Ergebnisse werden hinsichtlich der Relevanz einer systematischen Betrachtung der Zusammenhänge zwischen einerseits der allgemeinen Intelligenz und spezifischen Intelligenzfaktoren mit andererseits den allgemeinen und spezifischen Schulleistungen diskutiert.

Donnerstag
12. März

15:30 Uhr

HGC 30

Chair(s)

Schult, Johannes
Sparfeldt, Jörn

Schult, Johannes; Sparfeldt, Jörn

Korrespondieren Begabungs- und Studierfähigkeitsschwerpunkte mit entsprechenden Schulleistungen?

Einleitung: Während der Zusammenhang zwischen Intelligenz und Schulleistung weithin belegt ist (Jensen, 1998; Rost, 2013), scheint die Betrachtung der Ebene unterhalb des Generalfaktors g eher selten zusätzliche Erkenntnisse zu liefern (Leeson et al., 2008). Ein Grund dafür ist die hohe g -Sättigung von interkorrelierten Sub- g -Faktoren in hierarchischen Intelligenzstrukturmodellen - sog. Higher Order (HO) Modelle. Im Gegensatz dazu bieten sog. Nested Factor (NF) Modelle eine Alternative mit einem Generalfaktor sowie davon unabhängigen inhaltspezifischen Gruppenfaktoren (vgl. Brunner et al., 2012). Die genesteten Gruppenfaktoren für verschiedene Inhaltsfaktoren (z.B. numerisch, verbal) korrespondieren über verschiedene Testbatterien hinweg (Valerius & Sparfeldt, 2014). Darauf aufbauend soll nun geprüft werden, ob und wie weit sich die Struktur der Intelligenz sowie der Studierfähigkeit (Generalfaktor und Gruppenfaktoren) in einer entsprechenden Schulleistungsstruktur widerspiegelt (basierend auf den Zensuren: ein allgemeiner Schulleistungsfaktor sowie davon unabhängige, spezifische sprachliche und numerische Gruppenfaktoren). Wir erwarteten, dass der allgemeine Schulleistungsfaktor deutlich mit dem Generalfaktor der Intelligenz sowie mit dem Generalfaktor der Studierfähigkeit korreliert, dass der verbale Fähigkeitsgruppenfaktor mit dem sprachlichen Schulleistungsfaktor positiv zusammenhängt, sowie dass der numerische Fähigkeitsgruppenfaktor mit dem mathematisch-naturwissenschaftlichen Schulleistungsfaktor korreliert. Insgesamt sollte sich ein konvergent-diskriminantes Beziehungsmuster der spezifischen Gruppenfaktoren ergeben.

Methode: Ausgangspunkt war ein Datensatzes einer größeren US-amerikanischen Stichprobe ($n > 4000$), der eine Reihe von Intelligenztest- und Studierfähigkeitstest-Scores sowie die Schulnoten in zentralen Fächern enthält. Es wurden kombinierte NF-Modelle der Intelligenz, der Studierfähigkeit und der Schulleistung gerechnet und mit zusätzlichen Alternativ-Modellen (z.B. nur Generalfaktoren) verglichen. Dies waren einerseits ein NF-Modell mit allgemeinem und domänenspezifischen Fähigkeitsfaktoren (numerisch, verbal) und andererseits (basierend auf den Schulnoten) entsprechend ein allgemeiner Schulleistungsfaktor sowie genestete Gruppenfaktoren (sprachlich, mathematisch-naturwissenschaftlich).

Ergebnisse: Erwartungsgemäß korrelierten Intelligenz- und Studierfähigkeits-Generalfaktoren eng mit dem Schulleistungsgeneralfaktor. Nach Kontrolle für diese Generalfaktoren zeigte sich für den verbalen Bereich trotz einer breiten Indikatorisierung keine konvergente Korrelation. Für die numerischen Fähigkeiten hingegen resultierte ein starker Zusammenhang mit der mathematisch-naturwissenschaftlichen Schulleistung. Zwischen den verbalen und numerischen Fähigkeits- und Schulleistungsgruppenfaktoren zeigten sich niedrigere negative Zusammenhänge.

Diskussion: Die guten Fit-Kennwerten sprechen für die Gültigkeit der vorgestellten Ergebnisse. Die Modellierbarkeit einer analogen Fähigkeits- und Schulleistungsstruktur und das konvergent-diskriminante Zusammenhangsmuster stützen im Sinne einer kriteriumsorientierten Validierung die Bedeutung einerseits der allgemeinen Intelligenz (im Sinne des absoluten Niveaus) und andererseits der Gruppenfaktoren (z.B. im Sinne eines einander entsprechenden Begabungs- und Schulleistungsprofils).

Donnerstag
12. März

15:30 Uhr

HGC 30

Chair(s)

Schult, Johannes
Sparfeldt, Jörn

Stadler, Matthias; Becker, Nicolas; Niepel, Christoph; Greiff, Samuel

Die Rolle komplexer Problemlösefähigkeit in der Vorhersage von Studienerfolg

Die Zeit an der Universität ist eine kritische Phase im Leben vieler Studierenden, die viele Chancen aber auch Herausforderungen mit sich bringt. Basierend auf der Annahme, dass der erfolgreiche Abschluss eines Hochschulstudiums eine komplexe Anforderung mit vielfältigen, sich gegenseitig dynamisch beeinflussenden Aspekten darstellt, ist es das Ziel dieser Studie, den Nutzen komplexer Problemlösefähigkeiten (KPL) für die Vorhersage objektiven und subjektiven Studienerfolgs zu untersuchen.

Frühere Arbeiten konnten wiederholt zeigen, dass KPL ein valider Prädiktor für den akademischen Erfolg von Schülerinnen und Schülern verschiedener Altersstufen ist, der auch über schlussfolgerndes Denken hinaus inkrementelle Validität besitzt (z.B. Wüstenberg, Greiff & Funke, 2012; Greiff et al., 2014). Dennoch hat nach unserem Wissen bisher keine andere, veröffentlichte Studie die Beziehung zwischen KPL und Studienerfolg untersucht. Um diese Forschungslücke zu schließen wurden in dieser Studie KPL und schlussfolgerndes Denken hinsichtlich ihrer Validität in der Vorhersage von Studienerfolg verglichen.

Im Rahmen dieser Studie bearbeiteten N = 150 deutsche Studierende aus drei verschiedenen Studiengängen (Biologie, Sportwissenschaften und Psychologie) eine computergestützte Testbatterie bestehend aus dem IST-Screening (Liepmann, Beauducel, Brocke & Nettelstroth, 2012), als Maß für schlussfolgerndes Denken (erfasste Subfacetten: Verbal, Numerisch und Figural), sowie einen Test aus mehreren kurzen aber komplexen Mikrowelten zur Erfassung von KPL (MicroDYN; Greiff, Wüstenberg & Funke, 2012). Zusätzlich wurden die Durchschnittsnoten der Studierenden als Indikator für objektiven Studienerfolg in den entsprechenden Prüfungsämtern sowie einer subjektive Bewertung des Studienerfolgs der Studierenden mittels fünf Fragebogen-items erhoben. Die Beziehungen zwischen den beiden Prädiktoren und den Indikatoren des Studienerfolgs wurden latent modelliert wobei mehrere hierarchische Strukturgleichungsmodelle verglichen wurden.

Eines der wichtigsten Ergebnisse war, dass KPL einen signifikanten Beitrag zur Vorhersage der Durchschnittsnoten der Studierenden leistete ($R^2 = .18$) und auch in Übereinstimmung mit unseren Hypothesen inkrementelle Validität über schlussfolgerndes Denken hinaus aufwies ($\Delta R^2 = .09$). Außerdem erklärte KPL substantiell Varianz in der subjektiven Bewertung der Studierenden bezüglich ihres Studienerfolgs ($R^2 = .10$) mit substantieller inkrementeller Validität ($\Delta R^2 = .09$) gegenüber schlussfolgerndem Denken ($R^2 = .02$).

Diese Ergebnisse haben verschiedene theoretische sowie praktische Implikationen. Zunächst legen sie nahe, dass neben schlussfolgerndem Denken auch andere kognitive Fähigkeiten, wie die Fähigkeit komplexe Probleme zu lösen, relevant für den erfolgreichen Abschluss eines Hochschulstudiums sein können. Darüber hinaus zeigen sie, dass es wichtig ist, objektiven und subjektiven Studienerfolg konzeptuell zu trennen. Wenngleich diese beiden Konstrukte viele Überschneidungen aufweisen, beinhaltet subjektiver Studienerfolg neben der tatsächlichen Leistung auch eine Gewichtung der Bedeutung dieser Leistung. Besonders angesichts des konsistenten Befunds, dass nur wenige Studierende wegen objektiv schlechten Leistungen ihr Studium beenden, scheint es wichtig, auch diese subjektive Komponente von Studienerfolg besser zu verstehen und vorhersagen zu können. Abschließend wird diskutiert, ob KPL zusätzlich zu schlussfolgerndem Denken in Studieneignungstests integriert werden sollte.

Donnerstag
12. März

15:30 Uhr

HGC 30

Chair(s)

Schult, Johannes
Sparfeldt, Jörn

Brunner, Martin; Fischbach, Antoine; Baudson, Tanja Gabriele; Preckel, Franzis; Martin, Romain

40 Jahre später: Kann die Einschätzung von Grundschullehrkräften zur Intelligenz ihrer Schülerinnen und Schülern deren Bildungsrenditen als Erwachsene vorhersagen?

Eine wichtige Aufgabe von Lehrkräften ist es, die kognitiven Leistungen ihrer Schülerinnen und Schüler einzuschätzen, um die kognitiven Anforderungen ihres Unterrichts an die Heterogenität der Schülerschaft in ihren Klassen bestmöglich anzupassen. Zu Fragen wie gut Lehrkräfte die Intelligenz ihrer Schülerinnen und Schüler einschätzen können und wie sich solche Einschätzungen auswirken können, sind zwei Forschungsfelder relevant. Die Forschung zur diagnostischen Kompetenz zeigt, dass es Lehrkräften relativ gut gelingt, die Rangreihe der kognitiven Leistungen für die Schülerinnen und Schüler ihrer Klasse zu bestimmen (Südkamp, Kaiser, & Möller, 2012). Die Forschung zu selbsterfüllenden Prophezeiungen (Jussim & Harber, 2005) beschäftigt sich seit der grundlegenden Pygmalion-Studie (Rosenthal & Jacobson, 1968) mit den Folgen der Über- oder Unterschätzung der Intelligenz von Schülerinnen und Schülern für deren weitere schulische und außerschulische Entwicklung. In beiden Forschungsfeldern ist dabei bislang ungeklärt, inwiefern korrekte aber auch Fehleinschätzungen der Intelligenz von Schülerinnen und Schülern deren schulische Entwicklung oder Bildungsrenditen langfristig beeinflussen. Dies ist die übergreifende Forschungsfrage für den vorliegenden Beitrag.

Als Datengrundlage greifen wir auf die repräsentative Stichprobe der Luxemburger MAGRIP-Studie zurück, an der 731 Kinder im Alter von etwa 12 Jahren am Ende der Grundschulzeit und dann 40 Jahre später im Alter von etwa 52 Jahren teilnahmen. Für diese Personen lagen Einschätzungen zur Intelligenz durch Lehrkräfte, die Ergebnisse aus einem standardisierten Intelligenztest (Leistungsprüfsystem L-P-S, Horn, 1983) und Informationen zu wichtigen Kovariaten (Schulnoten, sozioökonomischer Familienhintergrund in Form des HISEI, Migrationshintergrund, Geschlecht) aus der Grundschulzeit vor. Für das Erwachsenenalter verwendeten wir Daten zum Bildungserfolg (höchster erreichter Bildungsabschluss, Anzahl an Jahren in schulischer oder universitärer Bildung) sowie einen breiten Kranz an Bildungsrenditen wie kognitive Leistungsfähigkeit im Erwachsenenalter (gemessen mit dem L-P-S), subjektive und funktionale Gesundheit, sozioökonomischer Status (HISEI), Einkommen, Selbstwert und Lebenszufriedenheit.

Unsere statistischen Analysen zeigten erstens, dass die Einschätzungen der Lehrkräfte zur Intelligenz ihrer Schülerinnen und Schüler deren Bildungserfolg und Bildungsrenditen (ausgenommen Lebenszufriedenheit und Selbstwert) signifikant über einen Zeitraum von 40 Jahren vorhersagen konnten. Zweitens, wiesen die Ergebnisse pfadanalytischer Auswertungen daraufhin, dass diese Vorhersagekraft auch bei Kontrolle der Intelligenz im Grundschulalter zu beobachten war. Drittens, wenn in den Pfadmodellen noch zusätzlich für die Durchschnittsnoten der Kinder kontrolliert wurde, sank die Vorhersagekraft der Lehrereinschätzungen deutlich ab. Dieser Befund unterstreicht, dass die Vorhersagekraft der Intelligenzeinschätzungen der Lehrkräfte vor allem darauf zurückzuführen ist, dass diese Einschätzungen in hohem Maße die erzielten Noten ihrer Schülerinnen und Schüler widerspiegeln. Viertens, zeigte sich, dass die Intelligenzeinschätzungen der Lehrkräfte auch nach Kontrolle der Intelligenz und der Kovariaten mit dem Bildungserfolg (gemessen über die Anzahl an Jahren in schulischer und universitärer Bildung) zusammenhing. Dies deutet darauf hin, dass sich über den Bildungserfolg von Schülerinnen und Schülern kleine Effekte selbsterfüllender Prophezeiungen langfristig auf die psychosoziale Entwicklung und Bildungsrenditen entfalten könnten. Fünftens, der Zusammenhang zwischen Intelligenz im Grundschulalter und Bildungsrenditen wird für die kognitive Leistungsfähigkeit im Erwachsenenalter teilweise und für mehrere Bildungsrenditen (funktionale und subjektive Gesundheit, Einkommen, sozioökonomischer Status) nahezu vollständig über den Bildungserfolg mediiert.

In der Zusammenschau zeigen unsere Ergebnisse, dass – unabhängig davon, ob Intelligenz durch Lehrkräfteeinschätzungen oder Intelligenztests erfasst wird – die Intelligenz im Grundschulalter bedeutsam wichtige Bildungsrenditen vorhersagen kann, und dass diese Vorhersage zu großen Teilen über den Bildungserfolg mediiert wird. Damit unterstreichen die vorliegenden Ergebnisse empirisch die besondere Bedeutung der diagnostischen Kompetenz von Lehrkräften, da Fehleinschätzungen der kognitiven Leistungsfähigkeit langfristig, nachteilige Konsequenzen für die Entwicklung von Schülerinnen und Schülern haben können.

Donnerstag
12. März

15:30 Uhr

HGC 30

Chair(s)

Schult, Johannes
Sparfeldt, Jörn

Düggeli, Albert

Perspektiven zu Bildungsübergängen im Fokus von Erwartungs- Nutzungs- und Gerechtigkeitsaspekten

Bildungsübergänge strukturieren schulische Ausbildungsverläufe. Sie stehen mit ausbildungs- bzw. arbeitsmarktlicher Teilhabe genauso in Zusammenhang wie mit individuellen Identitätsprozessen und prägen damit berufsbiografische Entwicklungen junger Menschen substantiell (z. B. Bergmann et al., 2012; Drewek, 2010). In gegliederten Ausbildungssystemen bestehen Übergänge innerhalb des obligatorischen Schulsystems oder zwischen der obligatorischen und der postobligatorischen Ausbildung. Diese Gliederung wird aus bildungspolitischer Sicht damit legitimiert, Lernende jenen Ausbildungswegen zuzuordnen, die mit ihren individuellen Voraussetzungen oder ihren Zielsetzungen möglichst gut korrespondieren (Bildungsbericht Schweiz, 2014). Dies hat zur Folge, dass die Heterogenität in den entstandenen Gruppen abnimmt, zwischen den Gruppen jedoch zunimmt. Die Realisierung einer solchen Ausbildungssteuerung ist jedoch mit teilweise problematischen Selektionsmechanismen verbunden und wirft Fragen zur Partizipationsgerechtigkeit und zu Entwicklungschancen von Heranwachsenden auf (Bildungsbericht Schweiz, 2014; Maaz, 2009; Baumert, 2006). Im vorgeschlagenen Symposium soll der Frage nach der Bildungsgerechtigkeit bei Bildungsübergängen auf der Basis von aktuellen Studien zu Erwartungs-, Nutzungs-, und Gerechtigkeitsaspekten aus dem deutschen oder schweizerischen Bildungskontext vertiefend nachgegangen werden. Dabei steht zum einen der Übergang von der Primar- in die Sekundarstufe I, zum anderen der Übertritt von der obligatorischen Schule in die postobligatorische Ausbildung im Fokus. Der erste Beitrag untersucht, inwiefern Eltern- oder Lehrpersonenerwartungen die schulische Leistung bereits vor dem Übertritt von der Primar- in die Sekundarstufe I beeinflussen. Im zweiten Beitrag wird ebenfalls dieser Übergang analysiert, und zwar unter der Frage, ob sich wert-erwartungstheoretische Dimensionen auch in individuellen Argumentationsmustern von Eltern finden lassen. Die beiden anderen Beiträge haben den Übergang ins postobligatorische Ausbildungssystem im Blick. Der erste nimmt die Frage auf, ob sich in unterschiedlich offenen Bildungssystemen auch unterschiedliche Nutzungsmuster von berufsbildenden Ausbildungsangeboten zeigen. Der zweite fokussiert die Fragestellung, wie das Gerechtigkeitsempfinden bei Lernenden beim Übergang vom obligatorischen ins postobligatorische Ausbildungssystem determiniert ist. Die vorliegenden Ergebnisse machen deutlich, dass die Übergangsthematik von mittlerweile breit erforschten und stabilen Befunden zur Wirkung von Selektionsmechanismen innerhalb von Bildungssystemen weiter ausdifferenziert und mit Bezug auf mögliche Massnahmen der Bildungssteuerung diskutiert werden kann. Weiter zeigen die Resultate, dass der Übergang von der obligatorischen in die postobligatorische Ausbildung unter dem Aspekt einer höheren Durchlässigkeit beim Übergang von der Schule in anschliessende Ausbildungsangebote der Sekundarstufe II weiter zu untersuchen ist; vielleicht mit Bezug auf spezifische lokalstrukturelle Gegebenheiten, zusätzlich vielleicht auch im Versuch, unterschiedliche Ausbildungssysteme vergleichend zu betrachten.

Donnerstag
12. März

15:30 Uhr

HGC 40

Chair(s)

Düggeli, Albert
Trautwein, Ulrich
Maaz, Kai

Neuenschwander, Markus

Perspektiven zu Bildungsübergängen im Fokus von Erwartungs- Nutzungs- und Gerechtigkeitsaspekten

Niederbacher, Edith; Neuenschwander, Markus P.

Schereneffekt vor dem Übertritt in die Sekundarstufe I? – Der Einfluss von Eltern- und Lehrpersonenerwartungen

Bildungssysteme beeinflussen mit der Organisation von Leistungsniveaus (resp. Schulformen) die Leistungsentwicklung von Schülerinnen und Schülern (Baumert, Stanat & Watermann, 2006). Die Leistungszunahme von Jugendlichen in der Sekundarstufe I variiert je nach zugewiesener Schulform (Schereneffekt) (z.B. Becker et al., 2006; Pfof, Karing, Lorenz & Artelt, 2010), was darauf zurückzuführen ist, dass die Schulformen unterschiedlich leistungsförderliche Lern- und Entwicklungsmilieus bilden (Baumert et al., 2006). Außerdem sind Erwartungen von Eltern und Lehrpersonen höchst wirksam für die Leistungsentwicklung (Eccles et al., 1983; Rosenthal & Jacobson, 1974).

Im vorliegenden Beitrag wird untersucht, ob bereits vor dem Übertritt in die Sekundarstufe I ein Schereneffekt auftritt. Wir gehen davon aus, dass Eltern und Lehrpersonen bereits zwei Jahre vor Übertritt ein Leistungsniveau antizipieren (Ditton & Krüsken, 2009). Vorstellungen und Normen, die mit den unterschiedlichen Leistungsniveaus verbunden sind, werden somit bereits am Ende der Primarschule über die Erwartungen an die Kinder vermittelt. Das zukünftige Leistungsniveau müsste sich dadurch – vermittelt über die Leistungserwartungen – bereits vor dem Übertritt differenziell auf die Leistungsentwicklung der Kinder in Deutsch und Mathematik auswirken. Wenn der Schereneffekt zwischen dem 5. und 7. Schuljahr aufgrund der Leistungserwartungen von Eltern und Lehrpersonen entsteht, müsste er unabhängig davon auftreten, ob der Übertritt nach dem 5. Schuljahr oder nach dem 6. Schuljahr stattfindet.

Folgende Forschungsfragen stehen im Zentrum: a) Beeinflusst das zugewiesene Leistungsniveau der Sekundarstufe I die Leistungsentwicklung in Deutsch und Mathematik zwischen dem 5. und 7. Schuljahr im Sinne eines Schereneffekts? b) Beeinflusst der Übertrittszeitpunkt in die Sekundarstufe I das Auftreten dieses Schereneffekts? c) Wird die Wirkung des Leistungsniveaus auf die Leistungsentwicklung in Deutsch und Mathematik durch Eltern- und Lehrpersonenerwartungen vor dem Übertritt mediiert?

Die Schweizer Längsschnittstudie Wirkungen der Selektion WiSel untersucht die Leistungsentwicklung von Schülerinnen und Schülern in vier deutschsprachigen Kantonen vom 5. zum 7. Schuljahr. Zu allen drei Messzeitpunkten wurden standardisierte Leistungstests in Deutsch und Mathematik eingesetzt. Zudem wurden die Schülerinnen und Schüler, ihre Eltern und Lehrpersonen mit einem standardisierten Fragebogen befragt. Zur Prüfung der Hypothesen wurden zwei Gruppen verglichen, die sich im Übertrittszeitpunkt in die Sekundarstufe I unterscheiden. Schülerinnen und Schüler aus den Kantonen Aargau und Basel-Landschaft (N=212) traten nach dem 5. Schuljahr in die Sekundarstufe I ein und befanden sich während des Erhebungszeitraums ein Jahr in der Primarschule und zwei Jahre in der segregierten Sekundarstufe I. In den Kantonen Bern und Luzern (N=386) fand der Übertritt nach dem 6. Schuljahr statt, wodurch sich die Kinder im Erhebungszeitraum zwei Jahre in der Primarschule befanden und ein Jahr in einem Leistungsniveau der Sekundarstufe I.

Multiple stufenweise Regressionsanalysen zeigten, dass die Leistungen der Kinder in Deutsch und Mathematik beim Übertritt in weniger anspruchsvolle Leistungsniveaus weniger zunahm als beim Übertritt in anspruchsvolle Leistungsniveaus (Schereneffekt). Die Schereneffekte zeigten sich bereits vor dem Eintritt in die Sekundarstufe I und traten unabhängig vom Übertrittszeitpunkt auf. Der Einfluss des antizipierten Leistungsniveaus (hohe vs. mittlere und tiefe Anforderungen sowie tiefe vs. mittlere und hohe Anforderungen) auf die Leistungsveränderung in Mathematik wurde vollständig durch die Erwartungen von Eltern mediiert, nicht aber durch die Erwartungen von Lehrpersonen. Der Übertrittszeitpunkt spielte dabei keine Rolle. Im Fach Deutsch hatten die Erwartungen von Eltern und Lehrpersonen in beiden Gruppen keinen vermittelnden Effekt.

In der Diskussion wird den Gründen für den Schereneffekt vor dem Eintritt in die Sekundarstufe I nachgegangen.

Donnerstag
12. März

15:30 Uhr

HGC 40

Chair(s)

Düggeli, Albert
Trautwein, Ulrich
Maaz, Kai

Neuenschwander, Markus

Perspektiven zu Bildungsübergängen im Fokus von Erwartungs- Nutzungs- und Gerechtigkeitsaspekten

Klinge, Denise; Dumont, Hanna; Maaz, Kai

Die elterliche Entscheidung beim Übergang von der Grundschule in die Sekundarstufe I – Eine qualitative Betrachtung von Wert-Erwartungstheorien

Soziale Ungleichheiten beim Übergang von der Grundschule in die Sekundarstufe I lassen sich nach Boudon (1974) durch primäre und sekundäre Herkunftseffekte erklären. Während sich primäre Effekte in unterschiedlichen schulischen Leistungen der Schülerinnen und Schüler manifestieren, beschreiben sekundäre Herkunftseffekte soziale Unterschiede in der Bildungsbeteiligung, die aus unterschiedlichen Bildungsaspirationen und einem unterschiedlichen Entscheidungsverhalten von Eltern resultieren (zum empirischen Forschungsstand siehe Dumont, Maaz, Neumann & Becker, 2014). Um diese Unterschiede im Entscheidungsverhalten zu erklären, werden in der Übergangsforschung häufig Wert-Erwartungstheorien herangezogen, die davon ausgehen, dass sich Eltern vor dem Hintergrund von Kosten-Nutzen-Bilanzen für eine Schule bzw. Schulform entscheiden (Breen & Goldthorpe, 1997, Erikson & Jonsson, 1996, Esser 1999). Dazu werden die Komponenten der Theorien – elterliche Werte in Bezug auf Bildung und Bildungsabschlüsse, Erwartungen bzgl. des Erfolgs der Kinder an einer Schulform und Kosten des Besuchs einer Schulform – in der Regel mittels standardisierter Fragebögen erfasst und anschließend entsprechend der Vorhersagen der jeweiligen Theorie statistisch modelliert.

Im vorliegenden Beitrag wird untersucht, ob sich die Komponenten der Wert-Erwartungstheorien auch in den subjektiven Entscheidungslogiken von Eltern beim Übergang von der Grundschule auf die Sekundarschule ihrer Kinder wiederfinden lassen. Dazu wurden 25 narrative Interviews mit Berliner Eltern direkt nach der Übergangentscheidung im Schuljahr 2010/2011 durchgeführt. Die gewählte Erhebungsmethode sollte die direkte Anknüpfung an das Erlebte der Interviewten gewährleisten. Die Stichprobe umfasste dabei sowohl Eltern, die sich für das Gymnasium oder die im Zuge der Berliner Schulstrukturereform neu eingeführte Integrierte Sekundarschule (ISS) als auch für oder gegen die Förderprognose der Grundschule entschieden.

Die Auswertung erfolgte mittels quantitativer Inhaltsanalyse. Dazu wurde zunächst ein umfassendes Kategoriensystem aller Komponenten der verschiedenen Wert-Erwartungstheorien, die bisher in Studien zum Übergang empirisch untersucht wurden, erstellt. Anschließend wurden die Interviews mit Hilfe von MAXQDA hinsichtlich des Auftauchens dieser Komponenten untersucht. Es konnte gezeigt werden, dass sich fast alle bisher im Rahmen von Fragebogenstudien erfassten Wert- und Erwartungskomponenten innerhalb der Elterninterviews abbilden ließen: familiäre Ressourcen, das Motiv des Stuserhalts, die Abschluss- und Berufsaspiration der Eltern für ihr Kind, die elterliche Erfolgserwartung für das Bestehen des Kindes an einer Schulform/Schule bzw. für das Erreichen eines bestimmten Schulabschlusses, die Belastbarkeit des Kindes, die Schulformempfehlung der Grundschule, die Leistungsmotivation des Kindes, die Schulnoten und die subjektiv wahrgenommenen Leistungsfähigkeit. Die finanziellen Kosten, welche in Fragebogenstudien üblicherweise erfasst werden, fanden in den Elterninterviews keine Erwähnung. Es ließen sich jedoch subjektive Kosten, vor allem Kosten bezüglich des zeitlichen Aufwandes zur Erreichung der Schule, finden. Außerdem zeigte sich, dass einzelne Komponenten wesentlich häufiger genannt wurden als andere: Während beispielsweise das Motiv des Stuserhalts in nur wenigen Interviews Erwähnung fand, spielten die familiären Ressourcen, wie das Wissen über das Schulsystem, und die Fähigkeit, sich Informationen zu beschaffen, um sich im Schulsystem selbstbestimmt bewegen zu können, eine große Rolle bei der Übergangentscheidung.

Neben den Komponenten der Wert-Erwartungstheorien, konnten in den Interviews auch neue Aspekte, die für die Schulentscheidung der Eltern bedeutsam waren, gefunden werden. So spielte beispielsweise das elterliche Erfahrungswissen – basierend auf eigenen Schulerfahrungen, der älterer eigener Kinder, oder von anderen Eltern – für die Schulentscheidung eine entscheidende Rolle. Auch erwähnten alle Befragten den Wunsch des Kindes als handlungsleitend für die Schulentscheidung. Darüber hinaus standen häufig Faktoren der Einzelschule, wie beispielsweise das Profil und Angebot der Schule, die Wirkung der Schule und der Lehrerschaft, die Größe, die Nähe und die Erreichbarkeit der Schule, bei der Entscheidung im Vordergrund.

Damit lässt sich zusammenfassend festhalten, dass es über die zentralen Komponenten der Wert-Erwartungstheorien, die bisher mittels standardisierter Fragebögen erfasst wurden, hinaus weitere wichtige Aspekte – insbesondere der Einzelschule – gibt, die einen großen Einfluss auf die Schul- bzw. Schulformentscheidung von Eltern haben und in zukünftige Studien zu Untersuchung des Übergangs von der Grundschule in die Sekundarstufe I einfließen sollten.

Donnerstag
12. März

15:30 Uhr

HGC 40

Oesch, Dominique; Düggeli, Albert

Öffnungsoptionen und Nutzungsmuster - Zur Durchlässigkeit in verschiedenen Bildungssystemen

Jugendliche begehen individuell Bildungswege. Sie kommen dabei in Bildungssystemen an gewisse Verzweigungen, welche ihnen unterschiedlich begehbare Wege eröffnen. Schulische Übergänge, aber auch die Möglichkeit der Umstufung von einem Schultyp in einen anderen innerhalb einer Schulstufe, können solche Verzweigungen darstellen. Dass auch der Übertritt von der obligatorischen Schule in die nachobligatorischen Bildungsangebote erfolgreich verläuft und jeder Abschluss mit einem Anschluss verbunden ist, ist dabei ein bildungspolitischer Grundsatz der Schweiz (EDK, 2006). Dass kein Bildungsweg "in einer Sackgasse" enden darf (Bildungsrat, 1972), gilt ebenso in Deutschland.

Regelungen bezüglich Offenheit und Umfang von Übergängen im Bildungssystem sowie die Umstufungsmöglichkeiten zwischen verschiedenen Schultypen einer Schulstufe bestimmen die sogenannte "theoretische Durchlässigkeit" eines Bildungssystems. Diese gilt als ein Bedingungsfaktor für erfolgreiche Bildungswege auf systemischer Ebene (Trautwein, Nagy & Maaz, 2011). Die Förderung der Durchlässigkeit im Schweizer Bildungswesen ist in der aktuellen Harmonisierungs-Reform (HarmoS), welche teilweise kantonsübergreifend versucht Bildungssysteme zu vereinheitlichen, ein zentrales Anliegen (EDK, 2007). Durch kantonal (resp. bundeslandweit) festgelegte Regelungen können sich Umstufungsoptionen vor und bei Übergängen unterscheiden. So können beim Übergang ins postobligatorische Ausbildungssystem die Anschlussoptionen auf Sekundarstufe II unterschiedlich geregelt sein. Beispielsweise können sie vom zuvor besuchten Schultyp auf der Sekundarstufe I oder aber von der Anzahl bestimmter Angebote (Angebotsvielfalt) auf Sekundarstufe II abhängen.

Aus diesem Grund steht im folgenden Beitrag die Frage im Zentrum, welche Auswirkungen solche systemischen Unterschiede in Bezug auf die Nutzung des Bildungsangebots auf Sekundarstufe II haben? Dabei wird zwischen niedrig-, mittel- und hochzertifizierenden Anschlussangeboten unterschieden und untersucht, inwiefern die auf Sekundarstufe I besuchten Schultypen (mit den Anforderungsniveaus niedrig (Realschule, WBS A), mittel (Sekundar B, WBS E) und hoch (Progymnasium, Gymnasium) in die strukturell passenden Anschlussangebote überführen. Unter struktureller Passung der Anforderungsniveaus wird im Zusammenhang mit schultypspezifischen Leistungsanforderungen angenommen, dass ein Schultyp mit niedrigem Anforderungsniveau eher in niedrig-zertifizierende Anschlusslösungen überführt, während die Schülerschaft des mittleren Schultyps eher Angebote mittleren Zertifizierungsniveaus wählen und der Besuch eines hohen Anforderungsniveaus auf der Sekundarstufe I auch hauptsächlich den Besuch eines hohen Zertifizierungsniveaus auf der Sekundarstufe II mit sich bringt.

Die Beantwortung dieser Fragestellung basiert auf Daten der TIDES-Studie (Transitions In Different Educational Systems) und erfolgt in drei Schritten: Erstens werden zwei Bildungssysteme im Hinblick auf ihre theoretische Durchlässigkeit untersucht und beschrieben. Es handelt sich dabei um die beiden Schweizer Kantone oder Kantonsteile Deutsch-Freiburg und Basel-Stadt, gegebenenfalls ergänzt um eine repräsentative Stichprobe Baden-Württembergs. Zweitens werden die Nutzungsmuster beim Übergang nach der obligatorischen Schule analysiert und dabei auch differenziert untersucht, ob sich unterschiedliche Nutzungsmuster in Bezug auf das kognitive Potenzial der Jugendlichen ergeben. Drittens werden die gewonnenen Erkenntnisse über die Nutzungsmuster im Zusammenhang mit der Durchlässigkeit der beiden Bildungssysteme diskutiert.

Mit Hilfe von eigens für diese Studie entworfenen Analyseinstrumenten zur Identifizierung der theoretischen Durchlässigkeit und der Kategorisierung der Bildungsangebote auf Sekundarstufe II werden sowohl deskriptive Darstellungen als auch logistische Regressionsanalysen für die empirische Beantwortung der Fragen verwendet. Erste deskriptive Befunde weisen darauf hin, dass in einem eher durchlässigen Bildungssystem eine grössere Mobilität beim Übertritt in die Sekundarstufe II stattfinden kann. Dabei wurde die Bedeutung der Angebotsvielfalt resp. die Anzahl an Angeboten per Zertifizierungsniveau als relevante Einflussgrößen für systemspezifische Nutzungsmuster sichtbar. Ausserdem zeigt sich im Sinne der strukturellen Passung der Anforderungsniveaus bei differenzierteren Analysen eine Systemabhängigkeit der Nutzungsmuster. Weitere logistische Regressionsanalysen stehen noch aus.

Donnerstag
12. März

15:30 Uhr

HGC 40

Biewer, Caroline

Gerechtigkeitserleben in der Schule

Im Schulkontext hat (Un-)Gerechtigkeitserleben einen großen Stellenwert. Schülerinnen und Schüler wünschen sich fair behandelt zu werden (Tylor, 1962) und auch Lehrerinnen und Lehrern ist es wichtig Kinder und Jugendliche, die sie unterrichten, nach dem Maßstab der Gerechtigkeit zu bewerten (Kanders, 2000). In einer Studie von Biewer, Wandeler & Baeriswyl (2013) konnte gezeigt werden, dass die elternerperzipierte Gerechtigkeit des Übertrittsverfahrens von der Primarschule in die Sekundarstufe I im deutschsprachigen Teil des Kantons Freiburg (CH) vor allem von der elterlich wahrgenommenen Beurteilungsgerechtigkeit der Lehrperson abhängt und ein höherer sozioökonomischer Status der Eltern zu niedrigeren Gerechtigkeitseinschätzungen führt. Es konnten bereits eine Vielzahl positiver Effekte der Lehrpersonsgerechtigkeit für Schulkontextvariablen ermittelt werden. So fasste Dalbert (2011, S.14) zusammen, dass die „individuell und subjektiv erlebte Gerechtigkeit des Lehrerhandelns“ mit besseren Schulleistungen, weniger Arbeitsvermeidung, weniger Bullying, weniger Mogeln, besserem Unterrichts- und Sozialklima, weniger schulischem Belastungserleben und weniger schulischem Exklusionsempfinden zusammenhängt. Da für das (Un-)Gerechtigkeitserleben die individuelle Ausprägung eines Bedürfnisses nach Gerechtigkeit eine zentrale Rolle spielt, werden Zusammenhänge zwischen den schulproximalen Gerechtigkeitskonstrukten und dem Gerechtigkeitsmotiv zu erwarten sein. Die Gerechtigkeitsmotivtheorie geht auf Lerner (1975) zurück. Er postulierte dieses Motiv als ein Bedürfnis oder einen Wunsch eines jeden Individuums nach Gerechtigkeit im Verhalten anderer, aber auch im eigenen Verhalten. Dabei wird angenommen, dass es ein implizites und ein explizites Gerechtigkeitsmotiv gibt. Während das implizite Motiv das individuelle, intuitive und unbewusste Streben nach Gerechtigkeit umfasst, versteht man unter dem expliziten oder auch selbstattribuierten Motiv eine bewusste Selbstbeschreibung eines Individuums über den Grad an Wichtigkeit von Gerechtigkeit für die Person selbst. Mit dem impliziten Motiv werden spontane Verhaltensweisen und Reaktionen auf erlebte Ungerechtigkeit erklärt, mit dem selbstattribuierten Motiv wird vor allem kontrolliertes Verhalten assoziiert (Dalbert, 2001).

Im Kontext der TIDES-Studie wurden im Februar und März 2013 die Schülerinnen und Schüler der 9. Jahrgangsstufe im Kanton Basel-Stadt und im deutschsprachigen Teil Freiburgs (CH) zu schulproximalen Gerechtigkeitskonstrukten Schulgerechtigkeit und Lehrpersonengerechtigkeit befragt. Zur Erhebung des Gerechtigkeitsmotivs wurden die Konstrukte Gerechte-Welt-Glaube und Gerechtigkeits sensitivität eingesetzt. Der Gerechte-Welt-Glaube dient als Indikator für das implizite Gerechtigkeitsmotiv, die Gerechtigkeits sensitivität aus der Nutznießerperspektive dient der Operationalisierung des expliziten Gerechtigkeitsmotivs. Zudem wurde das Wahrnehmen sozialer Disparitäten im Schulkontext im Sinne von Ungleichheitserleben erhoben.

Ziel ist es wahrgenommene Gerechtigkeit in der Schule aus der Perspektive der Schülerinnen und Schüler aufzuschlüsseln. Hierbei sollen Einflüsse des Gerechtigkeitsmotivs und des Ungleichheitserlebens ermittelt werden. Die Datenstruktur verlangt die Berücksichtigung der Clusterung in Schulklassen und lässt einen Vergleich der beiden Kantone zu. Es wird angenommen, dass das implizite Gerechtigkeitsmotiv negativ mit dem Ungleichheitserleben der Schülerinnen und Schüler zusammenhängt. Je höher der Gerechte-Welt-Glaube, umso weniger soziale Disparitäten werden wahrgenommen. Weiter wird erwartet, dass ein hohes implizites Gerechtigkeitsmotiv der Schülerinnen und Schüler mit höheren Werten in der Einschätzung der Lehrpersonsgerechtigkeit und der Schulgerechtigkeit einhergeht. Die Wahrnehmung von sozialen Disparitäten hingegen wirkt sich vermutlich negativ auf die schulproximalen Gerechtigkeitskonstrukte aus. Insgesamt besteht die Annahme, dass das implizite Motiv stärkere Zusammenhänge mit der erlebten Lehrperson- und Schulgerechtigkeit und dem Ungleichheitserleben aufweist. Bezüglich möglicher Unterschiede im Gerechtigkeitserleben zwischen den beiden Kantonen bestehen keine vorhergehenden Annahmen. Hier wird rein explorativ vorgegangen.

Donnerstag
12. März

15:30 Uhr

HGC 40

Bertram, Christiane

Wie man Leistung und Motivation steigern kann: Befunde aus vier randomisierten, fachdidaktisch fundierten Interventionsstudien

Randomisierte kontrollierte Studien (RCT: randomized controlled trial) – aus der medizinischen Forschung kommend (Sackett, Rosenberg, Gray, Haynes & Richardson, 1996) – gelten in der empirischen Bildungsforschung als „Goldstandard“ eines evidenzbasierten Forschungsdesigns. In den USA werden RCT-Standards im „What works Clearing House“ (<http://ies.ed.gov/ncee/wwc/>) definiert und eingefordert, in Großbritannien im „Consort 2010“ (Schulz, Altman & Moher, 2009). In Deutschland sind randomisierte Interventionsstudien eher selten, da die Durchführung aufwendig und teuer ist (Spiewak, 2013). Notwendig ist eine interdisziplinäre Zusammenarbeit zwischen Fachdidaktikern, Schulpraktikern und empirischen Bildungsforschern, um fachdidaktisch und pädagogisch sinnvolle Intervention entwickeln und ihre Wirksamkeit mithilfe von Testinstrumenten, die psychometrischen Gütekriterien entsprechen, zu überprüfen.

Das beantragte Symposium zeichnet sich aus durch die Heterogenität der beteiligten Fachdisziplinen (Mathematik, Naturwissenschaft, Geschichte und Englisch) und durch die Interdisziplinarität innerhalb der Projekte, die fachdidaktische und methodische Expertise zusammenführen, um Interventionen zu gestalten und ihre Wirkung in einem experimentellen Setting zu überprüfen. In den vier Interventionsstudien wurden eindeutige Fragestellungen formuliert, mindestens zwei Versuchsgruppen gebildet und randomisiert den Bedingungen zugewiesen, Störfaktoren minimiert und für die Ausgangsvoraussetzungen der Schüler/innen kontrolliert. Die Daten wurden in einem Prä-Post- und Follow-Up-Design erhoben. Durch die enge Zusammenarbeit von Fachdidaktikern und empirischen Bildungsforschern war es möglich, für die Fachdidaktik relevante Fragestellungen mit Hilfe inferenzstatistischer Analysen und unter Berücksichtigung der Multi-Level-Struktur der Daten zu beantworten.

Christiane Bertram et al. stellen die Ergebnisse einer randomisierten Interventionsstudie mit 30 Interventionsklassen und 5 Kontrollklassen (N = 900) im Geschichtsunterricht vor, bei der der Frage nachgegangen wurde, welche differenziellen Effekte die Arbeit mit Zeitzeugeninterviews (live – Video – Text) im Hinblick auf ausgewählte historische Kompetenzen wie auch Faktenkenntnisse und die Einschätzung der Unterrichtseinheit hatte. Nicole Piesche et al. berichten aus ihrer randomisierten Interventionsstudie zum bilingualen naturwissenschaftlichen Unterricht, in der die Wirkung der Unterrichtssprache (in 15 Klassen englisch – in 15 Klassen deutsch) (N = 722) auf den sachfachlichen Lernerfolg und die Motivation der Schülerinnen und Schüler untersucht wurde. Ute Sproesser et al. überprüften innerhalb eines experimentellen 2x2 Designs, welche Wirkung reflexionsorientierte Prompts in einer schülerzentrierten Lernumgebung auf das statistikbezogene Selbstkonzept und Interesse der Lernenden (N = 503) hatten. Brigitte Schreier et al. berichten von einer cluster-randomisierten Studie mit 82 Klassen (N = 1916), in der die Effekte von zwei Nützlichkeitsinterventionen in Mathematik auf Kompetenzüberzeugungen, Anstrengungsbereitschaft sowie Leistung der Schüler/innen untersucht wurde.

Das Symposium stellt vor, wie fachdidaktisch motivierte Interventionen die Leistung und Motivation von Lernenden steigern – und wie dies empirisch überprüft werden kann.

Donnerstag
12. März

15:30 Uhr

HGC 50

Chair(s)

Bertram, Christiane

Wie man Leistung und Motivation steigern kann: Befunde aus vier randomisierten, fachdidaktisch fundierten Interventionsstudien

Bertram, Christiane; Wagner, Wolfgang; Trautwein, Ulrich

Mehr Spaß – weniger gelernt. Ergebnisse einer randomisierten Interventionsstudie zur Wirksamkeit von Zeitzeugenbefragungen im Geschichtsunterricht

Theoretischer Hintergrund: Zeitzeugenbefragungen werden häufig als geeignetes Mittel gesehen, um Schüler/innen für Geschichte zu interessieren und sie ihre Eingebundenheit in die Geschichte erleben zu lassen. Der Bildungsplan 2004 für das Gymnasium des Landes Baden-Württemberg geht davon aus, dass durch „Zeitzeugen [...] historische Wirklichkeit konkret erfahrbar wird“ (Ministerium Baden-Württemberg, 2004, 219). Im Geschichtsunterricht stellen Zeitzeugenbefragungen jedoch ein „riskantes“ Terrain dar, um Schülerinnen und Schülern Geschichte nahe zu bringen. Die unmittelbare Begegnung mit gelebter, authentischer Geschichte (aus „erster Hand“) kontrastiert mit der Gefahr, dass Zeitzeugen als „bessere Historiker“ (Sauer, 2006, 238) wahrgenommen werden, denn Schülerinnen und Schüler neigen dazu, den Aussagen von Zeitzeugen mehr Vertrauen zu schenken als Schulbuchdarstellungen oder Historikeraussagen (Angvik & von Borries, 1997).

Fragestellung, Design und Instrumente: Die Vor- und Nachteile von Zeitzeugenbefragungen wurden auf der Basis eines experimentellen Ansatzes anhand des Themas „Friedliche Revolution in der DDR“ untersucht. Basierend auf einer Stichprobe von insgesamt 35 Klassen (N = 900), von denen jeweils 10 zufällig einer der drei Interventionsgruppen zugewiesen wurde und weitere fünf Klassen als Kontrollgruppe lediglich getestet wurden, wurde der Fragestellung nachgegangen, welche differenziellen Effekte die Arbeit mit Zeitzeugeninterviews (live – Video – Text) im Hinblick auf (a) die Schüler-Einschätzung der Effektivität der Unterrichtseinheit, (b) die historischen Sachkompetenzen und (c) die „Faktenkenntnisse“ der Schüler hat. Die Instrumente zur Messung der Kompetenz-Outcomes wurden in vorher gehenden Studien auf Basis des „Historischen Kompetenz-Strukturmodells“ (Körber, Schreiber & Schöner, 2007) entwickelt und im Hinblick auf die psychometrischen Gütekriterien validiert (Bertram, Wagner & Trautwein, 2013; Bertram, Wagner & Trautwein, 2014). Die vier Kompetenz-Skalen erfassten die Einsicht, dass Geschichte notwendigerweise eine Rekonstruktion ist („Sachkompetenz ‚Rekonstruktion‘“), die Einsicht in die Notwendigkeit, Darstellungen kritisch zu dekonstruieren („Sachkompetenz ‚Dekonstruktion‘“) und die Einsicht in die (Retro-) Perspektivität von Zeitzeugen („Sachkompetenz ‚Zeitzeuge‘“), wie auch das Vermögen, zwischen Quellen und Darstellungen unterscheiden zu können („Sachkompetenz ‚Quelle oder Darstellung?‘“). Die Faktenkenntnisse wurden mit einem Lückentext ($t_3: \alpha = .83$) gemessen. Die Effektivität der Unterrichtseinheit bezogen auf die Inhalte, die Methoden und das Interesse schätzten die Lernenden selbst ein ($t_3: .65 < \alpha < .80$).

Methode: Die Skalierung der Kompetenztests erfolgte mithilfe von ConQuest (Wu, Adams, Wilson & Haldane, 2007) und auf der Basis von weighted likelihood-Schätzern (WLEs; Warm, 1989), wobei sich für die insgesamt vier Sub-Kompetenzskalen akzeptable Itemfits für sämtlich Indikatoren ergaben (weighted mean square-Statistiken: Skala „Sachkompetenz ‚Rekonstruktion‘“: $0.82 < WMSNQ < 1.34$; Skala „Sachkompetenz ‚Dekonstruktion‘“: $0.89 < WMSNQ < 1.24$; Skala „Sachkompetenz ‚Zeitzeuge‘“: $0.88 < WMSNQ < 1.29$; Skala „Sachkompetenz ‚Quelle oder Darstellung?‘“: $0.86 < WMSNQ < 1.12$). Fehlende Werte wurden durch multiple Imputation (10 Imputationen) auf der Grundlage eines Datensatzes mit sämtlichen abhängigen Variablen, Prädiktoren und Kovariaten für alle Erhebungszeitpunkte ersetzt. Bei der Schätzung der Standardfehler wurden sowohl die Besonderheiten der Stichprobenziehung (Cluster-Sampling) sowie die Verwendung multipler Imputation berücksichtigt. Gruppenunterschiede für die jeweilige abhängige Variable wurden auf der Basis von Dummy- bzw. Kontrastkodierungen berechnet.

Ergebnisse: Unter Kontrolle der Eingangsvoraussetzungen (z.B. KFT, Lesetempo, SES, Noten, Leistungen im Vortest) und der Leistungsbereitschaft (Testbarometer) erreichten die Schüler/innen der Treatmentgruppen im Vergleich zur Kontrollgruppe in den Kompetenzskalen „Sachkompetenz ‚Zeitzeuge‘“, „Sachkompetenz ‚ReKonstruktion‘“ und „Sachkompetenz ‚Quelle versus Darstellung‘“ wie auch im Lückentext statistisch signifikant höhere Werte. Bezogen auf die Unterschiede zwischen den Treatmentgruppen, schnitten die Lernenden in der Live-Gruppe in den Subskalen „Sachkompetenz ‚Zeitzeuge‘“ wie auch „Sachkompetenz ‚Dekonstruktion‘“ statistisch signifikant schlechter ab als die der Video- und Textgruppe, doch schätzten sie ihre Lerneffekte durch die Arbeit mit Zeitzeugeninterviews in inhaltlicher, methodischer und motivationaler Hinsicht statistisch signifikant höher ein. Während die Kompetenzunterschiede zwischen den Interventionsgruppen zum Follow-Up-Test nicht mehr gemessen wurden, blieb die deutlich positivere Einschätzung der Lernmöglichkeiten durch die Unterrichtseinheit von den Schüler/innen der Live-Gruppe auf etwas geringeren Level bestehen.

Donnerstag
12. März

15:30 Uhr

HGC 50

Chair(s)

Bertram, Christiane

Wie man Leistung und Motivation steigern kann: Befunde aus vier randomisierten, fachdidaktisch fundierten Interventionsstudien

Piesche, Nicole; Jonkmann, Kathrin; Fiege, Christiane; Keßler, Jörg

Ergebnisse einer randomisierten Interventionsstudie zur Wirksamkeit des bilingualen naturwissenschaftlichen Unterrichts auf das sachfachliche Lernen und die Motivation der Schülerinnen und Schüler an Realschulen

Theoretischer Hintergrund: Bilingualer Unterricht etabliert sich immer stärker als alternative Form des fremdsprachlichen und inhaltlichen Lernens und ist europaweit seit Jahren ein beforschtes Gebiet (Pérez-Cañado, 2012). Lange Zeit fanden der bilinguale Unterricht und die damit verbundene Forschung fast ausschließlich an Gymnasien (Hollm et al., 2013b) in positiv selektierten Klassen statt (Bruton, 2011). In letzter Zeit ist eine Ausweitung auf andere Schulformen, z.B. Realschulen (Hollm et al., 2013a), zu beobachten. Wie sich der bilinguale Unterricht dort in nicht positiv selektierten Klassen auswirkt, ist bislang jedoch weitgehend unerforscht (Küppers & Trautmann, 2013). Ein noch größerer Forschungsbedarf besteht hinsichtlich der Effekte auf die sachfachlichen Lernzuwächse (Hollm et al. 2013b), deren Prädiktoren und die Motivation (Breibach & Viebrock, 2012), insbesondere im naturwissenschaftlichen Unterricht (Bohn & Doff, 2010).

Bisherige Studien kommen bezüglich des Lernzuwachs zu uneindeutigen Ergebnissen (positive Effekte: Koch & Bündler, 2008; keine signifikanten Unterschiede: Haagen-Schützenhöfer et al., 2011; negative Effekte: Hartmannsgruber, 2014). Mit Hinblick auf die Leistungsprädiktoren ist lediglich gezeigt worden, dass sich gute Englischleistungen im bilingualen Unterricht positiv auswirken (Vockrodt-Scholz & Zydatis, 2007). Der Rolle der Sachfachkompetenzen wurde bisher keine Aufmerksamkeit geschenkt. Bezüglich der Effekte des bilingualen naturwissenschaftlichen Unterrichts auf die Motivation im Sachfach ist die Befundlage ebenfalls sehr dünn (keine Unterschiede: Bonnet, 2004; tendenziell negative Effekte: Hartmannsgruber, 2014). Der Frage, ob das Geschlecht oder die Kompetenzen in der Sprache oder dem Sachfach die Ergebnisse moderieren, wurde ebenfalls selten Aufmerksamkeit geschenkt (Abendroth-Timmer, 2007).

Fragestellung: Vor diesen Hintergründen beantwortet die Studie folgende Fragen: (1) Wie wirkt sich bilingualer naturwissenschaftlichen Unterricht in einer unselektierten Stichprobe an Realschulen auf den Wissenszuwachs im Sachfach aus? (2) Unterscheiden sich sprachliche und sachfachliche Kompetenzen in ihrer Rolle als Prädiktor für den Lernzuwachs in den beiden Bedingungen? (3) Wie wirkt sich der Unterricht auf die Motivation im Sachfach aus? Moderieren das Geschlecht und die Kompetenzen in der Fremdsprache und dem Sachfach die Ergebnisse?

Methode: Die Stichprobe umfasste 30 6. Klassen (722 Teilnehmende). Diese stammten aus 10 Realschulen, die in den Klassen 5 und 6 über kein bilinguales Angebot verfügten, um eine positive Selektion der Schülerschaft zu vermeiden, und wurden randomisiert den beiden Bedingungen zugeteilt. Diese bestanden aus einer fünf Doppelstunden umfassenden Unterrichtseinheit zum Thema „Sinken und Schwimmen“ (Möller, 2005), die entweder auf Deutsch oder bilingual deutsch-englisch von derselben Lehrkraft unterrichtet wurde. Zu Beginn der Studie fand ein Prä-Test zur Kontrolle der Ausgangslage statt, dem sich die 3-wöchige Unterrichtsintervention anschloss. Im Anschluss an jede der fünf Doppelstunden füllten die Teilnehmenden einen Kurzfragebogen zum Unterrichtserleben aus. Die Studie wurde durch einen Post-Test direkt im Anschluss an die Intervention und einen Follow-up-Test 6 Wochen später komplettiert.

Ergebnisse: Multiple Regressionsanalysen zeigten, dass die monolingual deutsch unterrichtete Schülerschaft im Posttest und im Follow-up-Test signifikant höhere Werte erzielten als die bilinguale Vergleichsgruppe.

Mit Hinblick auf die Leistungsprädiktoren kann gesagt werden, dass sich im bilingualen Unterricht ein positives Englischselbstkonzept signifikant positiv auf die Testleistung auswirkte. Das sachfachliche Vorwissen sagte den Lernzuwachs in beiden Bedingungen signifikant voraus, wobei es im monolingualen Unterricht noch bedeutender war. Weitergehende Untersuchungen brachten Moderationseffekte des Geschlechts und des naturwissenschaftlichen Interesses zu Tage.

Bezüglich der Motivation wurde deutlich, dass sich der bilinguale Unterricht leicht nachteilig auf das Kompetenz- und das Angsterleben der Lernenden auswirkte. Als Moderatoren erwiesen sich die Englischleistung und das Englischselbstkonzept in Bezug auf das Kompetenzerleben und das Selbstkonzept sowie das Interesse in den Naturwissenschaften in Bezug auf die kognitive Mitarbeit.

Donnerstag
12. März

15:30 Uhr

HGC 50

Chair(s)

Bertram, Christiane

Wie man Leistung und Motivation steigern kann: Befunde aus vier randomisierten, fachdidaktisch fundierten Interventionsstudien

Sproesser, Ute; Engel, Joachim; Kuntze, Sebastian

Förderung von domänenspezifischem Selbstkonzept und Interesse im Bereich von Statistical Literacy

Theoretischer Hintergrund: Grundfähigkeiten im Verstehen quantitativer Information sind notwendig, um an der Gesellschaft teilhaben und um Entscheidungen unter Unsicherheit treffen zu können (z.B. Kuntze, Lindmeier & Reiss, 2008; Wallman, 1993). Auch bezogen auf den Bereich Statistik werden motivationale Variablen wie das Selbstkonzept und Interesse inzwischen als bedeutsame Zielvariablen von Unterricht angesehen. Dies liegt nicht zuletzt daran, dass sie Lernprozesse, Bildungsentscheidungen sowie die Bereitschaft, Statistik im täglichen Leben zu nutzen, beeinflussen können (Gal, Ginsburg & Schau, 1997; Nagy, Trautwein, Baumert, Köller & Garrett, 2006; Pekrun & Zirngibl, 2004). Aus Inhaltsbereichen wie der Mathematik ist bekannt, dass Selbstkonzept und Interesse im Verlauf der Schulzeit signifikant abnehmen (Frenzel, Goetz, Pekrun & Watt, 2010; Jacobs, Lanza, Osgood, Eccles & Wigfield, 2002; Watt, 2006). Andererseits konnte die Förderbarkeit dieser Variablen durch gezielte Maßnahmen gezeigt werden: Die Förderung scheint besonders dann erfolgsversprechend zu sein, wenn sich Förderkonzepte auf einen konkreten Inhaltsbereich beziehen und soziales Lernen sowie das Erleben von Autonomie ermöglichen (Hidi & Renninger, 2006; O'Mara, Marsh, Craven & Debus, 2006; Schukajlow et al., 2012). Da es im Bereich der Statistikdidaktik im Vergleich zur Mathematikdidaktik bislang nur wenige Studien gibt, die sich mit diesen Variablen beschäftigen (z.B. Bond, Perkins & Ramirez, 2012; Ramirez, Schau & Emmioglu, 2012), fehlt insbesondere empirische Evidenz zu Förderansätzen bezüglich des statistikbezogenen Selbstkonzepts und Interesses. Diese Studie nimmt in den Blick, inwieweit das bereichsspezifische Selbstkonzept und Interesse durch eine statistikbezogene und lernerzentrierte Intervention in der achten Realschulklasse gefördert werden konnte. Ergebnisse dieser Studie können über die Verbreiterung der fachdidaktischen Basis bei der Erstellung wirksamer Lernmaterialien helfen.

Fragestellung: Die Studie zielte auf die folgenden Forschungsfragen ab: (1) Sind mathematik- und statistik-bezogenes Selbstkonzept und Interesse empirisch trennbar? (2) Kann das domänenspezifische Selbstkonzept und Interesse durch eine statistikspezifische Intervention gefördert werden? Gibt es Unterschiede zwischen verschiedenen Treatmentgruppen? (3) Welche Rolle spielt Kompetenz im Bereich von Statistik für die Entwicklung von statistik-bezogenem Selbstkonzept und Interesse?

Methode und Untersuchungsdesign: Dieser Studie liegen die Daten von 503 Achtklässlern aus 29 Klassen zu Grunde. Die Lernenden der 25 Interventionsklassen wurden randomisiert vier Treatmentgruppen zugeteilt, die sich in der fachlichen Schwerpunktsetzung, jedoch nicht im methodischen Vorgehen unterschieden. Während vier Unterrichtsstunden bearbeiteten sie statistikbezogene Lernmaterialien in einer schülerzentrierten, handlungsorientierten Arbeitsform. Vor und nach der Intervention absolvierten sie einen fachlichen Kompetenztest zum „Nutzen von Darstellungen und Modellen in statistischen Kontexten“ (vgl. Kuntze, 2013). Außerdem wurden das Selbstkonzept und Interesse bezogen auf Mathematik und Statistik über fünfstufige Likert-Skalen erhoben (vgl. Pekrun et al., 2002; Gundlach, Kuntze, Engel & Martignon, 2010). Die Lernenden der vier Kontrollklassen bearbeiteten die gleichen Tests und dienten so als Referenzpunkt über die Zeit. Um der Nestung der Daten gerecht zu werden, wurden durch den „type-is-complex“-Befehl von Mplus (Muthén & Muthén, 1998 – 2012) die Standardfehler adjustiert.

Ergebnisse und Diskussion: Eine konfirmatorische Faktorenanalyse zeigte, dass die Skalen mathematik- bzw. statistikbezogenes Selbstkonzept und Interesse empirisch vier separate Faktoren bildeten. Dies bestätigte ihre Eignung für die vorliegende Studie.

Während die Werte der mathematikbezogenen Skalen relativ stabil blieben, erhöhten sich die Mittelwerte der statistikbezogenen Skalen im Verlauf der kurzen Intervention im Vergleich zur Kontrollgruppe signifikant mit kleinen bis mittleren Effektstärken. Dass keine signifikanten Unterschiede zwischen den vier Treatments festgestellt werden konnten, deutet darauf hin, dass die grundsätzliche Auseinandersetzung mit statistischen Inhalten in einer schülerzentrierten Arbeitsform zur beobachteten Erhöhung geführt hat.

Insbesondere Lernende mit relativ hoher Fachkompetenz im Vortest konnten durch die Intervention in ihrem Selbstkonzept gefördert werden. Eine ähnliche Tendenz zeigte sich auch bezogen auf das statistikbezogene Interesse, wobei diese knapp nicht signifikant ausfiel ($p=0.056$). Aptitude-Treatment-Interaktionen dieser Art unterstreichen die wechselseitige Beziehung zwischen Leistungs- und motivationalen Variablen und somit die Bedeutung der Förderung beider Bereiche durch geeignete Materialien.

Donnerstag
12. März

15:30 Uhr

HGC 50

Chair(s)

Bertram, Christiane

Wie man Leistung und Motivation steigern kann: Befunde aus vier randomisierten, fachdidaktisch fundierten Interventionsstudien

Schreier, Brigitte; Dicke, Anna-Lena; Gaspard, Hanna; Häfner, Isabelle; Flunger, Barbara; Nagengast, Benjamin; Trautwein, Ulrich
Kurze Intervention, lange Effekte: Förderung der Kompetenzüberzeugungen, Anstrengungsbereitschaft und Leistung in Mathematik durch eine Nützlichkeitsintervention

Theoretischer Hintergrund: Laut Erwartungs-Wert-Theorie (Eccles et al., 1983) und empirischen Befunden (im Überblick: Fredricks, Blumenfeld, & Paris, 2004) ist die Betonung der Relevanz der Lerninhalte für das Leben der Schülerinnen und Schüler ein zentraler Aspekt motivations- und leistungsfördernden Unterrichts. Vor diesem Hintergrund wurden Interventionsprogramme erarbeitet, in denen Lehrkräfte im Unterricht Verbindungen zwischen dem Lernstoff und Karrieremöglichkeiten herstellten (Woolley, Rose, Orthner, Akos, & Jones-Sanpei, 2013) oder die Lernenden Texte über die persönliche Relevanz der Lerninhalte verfassten (Hulleman, Godes, Hendricks, & Harackiewicz, 2010; Hulleman & Harackiewicz, 2009). Durch diese Interventionen wurden Nützlichkeitsüberzeugungen, Interesse und Noten der Lernenden in Mathematik oder Naturwissenschaften gefördert.

Da es allerdings noch keine Studien gibt, in denen die Wirksamkeit mehrerer Interventionsansätze im Schulkontext vergleichend überprüft wurde, bleibt unklar, ob bisher wirklich die effektivsten Nützlichkeitsinterventionen eingesetzt wurden. Auch über die Wirksamkeit von feldbasierten Nützlichkeitsinterventionen über das manipulierte Konstrukt hinaus sowie über die Beständigkeit der Effekte ist noch wenig bekannt. Laut der Eccles et al. (1983) Erwartungs-Wert-Theorie und Korrelationsstudien (z.B. Chouinard, Karsenti, & Roy, 2007; Husman & Hilpert, 2007) sind Nützlichkeitsüberzeugungen positiv mit Kompetenzüberzeugungen und leistungsbezogenem Verhalten von Schülerinnen und Schülern assoziiert. Dadurch liegt die Annahme nahe, dass durch Nützlichkeitsinterventionen auch Kompetenzüberzeugungen wie das Selbstkonzept und die Selbstwirksamkeit sowie behaviorale Maße wie die Anstrengungsbereitschaft von Schülerinnen und Schülern gefördert werden könnten. Im Gegensatz zu Laborstudien (z.B. Shechter, Durik, Miyamoto, & Harackiewicz, 2011, Studie 1) wurden solche Effekte jedoch noch nicht im Schulkontext untersucht. Außerdem sind die Befunde zur Förderung von Schulnoten durch Nützlichkeitsinterventionen inkonsistent, weshalb Effekte auf standardisierte Leistungstests untersucht werden sollten.

Fragestellung: Die vorliegende Studie beschäftigt sich mit der Frage, ob Kompetenzüberzeugungen (Selbstkonzept, Hausaufgabenselbstwirksamkeit), Anstrengungsbereitschaft und Testleistung in Mathematik durch verschiedene Nützlichkeitsinterventionen im Klassenzimmer längerfristig gefördert werden können. Konkret wurden die Wirksamkeit eines neuen Interventionsansatzes (Beurteilen von Zitaten) basierend auf Theorien der sozialen Kognition (z.B. Markus & Nurius, 1986) und eines etablierten Interventionsansatzes (Verfassen eines Textes, vgl. Hulleman et al., 2009; 2010) untersucht.

Methode: 1916 Schülerinnen und Schüler aus 82 baden-württembergischen Gymnasialklassen der 9. Jahrgangsstufe nahmen an der Studie teil. Zu Studienbeginn wurden die 73 Mathematiklehrkräfte und deren Klassen randomisiert einer von drei Versuchsbedingungen (Zitate, Text, Wartekontrollgruppe) zugewiesen. Vor der Intervention (T1) sowie sechs Wochen (T2) und fünf Monate (T3) danach berichteten die Schülerinnen und Schüler in Fragebögen ihr Selbstkonzept und ihre Hausaufgabenselbstwirksamkeit in Mathematik. Zu T3 bearbeiteten die Schülerinnen und Schüler zusätzlich einen Mathematiktest. Die Mathematiklehrkräfte schätzten zu allen drei Testzeitpunkten die Anstrengungsbereitschaft ihrer Schülerinnen und Schüler ein.

Die Interventionen bestanden jeweils aus einer Präsentation zum Nutzen der Mathematik und einer individuellen Schreibaufgabe, die je nach Interventionsbedingung variierte. Schülerinnen und Schüler in Klassen der Zitatebedingung beurteilten Interviewzitate zum Nutzen der Mathematik. Schülerinnen und Schülern in Klassen der Textbedingung verfassten einen Text zum persönlichen Nutzen von Mathematikkenntnissen in ihrem Leben. Schülerinnen und Schüler in Klassen der Wartekontrollbedingung erhielten erst nach Abschluss der Datenerhebungen eine Nützlichkeitsintervention.

Zur Beantwortung der Fragestellung wurden hierarchisch lineare Regressionsanalysen für jedes Outcome als abhängige Variable berechnet. Die Interventionsbedingungen wurden in Form einer Dummyvariablen auf Klassenebene als Prädiktoren ins Modell gefügt, so dass die Wartekontrollklassen die Referenzgruppe darstellten. Für den Ausgangswert des jeweiligen Outcomes wurde auf beiden Analyseebenen kontrolliert.

Ergebnisse: Durch die Zitatebedingung wurden das Selbstkonzept sowie die Anstrengungsbereitschaft der Schülerinnen und Schüler in Mathematik zu T2 und die Hausaufgabenselbstwirksamkeit zu T2 und T3 gefördert. Weiter zeigte sich ein positiver Effekt der Zitatebedingung auf die Leistung im Mathematiktest. Durch die Textbedingung wurde die Hausaufgabenselbstwirksamkeit in Mathematik zu T3 gefördert, wohingegen keine Effekte auf das Selbstkonzept, die Anstrengungsbereitschaft und die Testleistung der Schülerinnen und Schüler gefunden wurden.

Zusammenfassend zeigt diese Studie das Potential einer bestimmten kurzen Nützlichkeitsintervention zur Förderung von Kompetenzüberzeugungen, Anstrengungsbereitschaft und Testleistung im realen Schulkontext.

Donnerstag
12. März

15:30 Uhr

HGC 50

Chair(s)

Bertram, Christiane

Wie man Leistung und Motivation steigern kann: Befunde aus vier randomisierten, fachdidaktisch fundierten Interventionsstudien

Korbach, Andreas; Brünken, Roland; Park, Babette

Der Einfluss von Seductive Details auf die kognitive Belastung und das Blickverhalten beim Lernen mit einem multimedialen Lernprogramm.

Das Design computerbasierter multimedialer Lernprogramme bietet zahlreiche Möglichkeiten Zusatzinformationen zu integrieren, die zwar hochinteressant, für das Lernziel jedoch nicht von Bedeutung sind. Solche Zusatzinformationen werden mit der Absicht eingesetzt, Lerninhalte attraktiver zu gestalten und die Lernvoraussetzungen zu verbessern. Tatsächlich hat die Verwendung solcher Zusatzinformationen jedoch einen negativen Effekt auf den Lernerfolg. Solche Zusatzinformationen werden als Seductive Details bezeichnet und der negative Effekt auf den Lernerfolg als Seductive Details Effekt. Harp und Mayer (1998) formulieren drei Annahmen zur Entstehung des Seductive Details Effektes, wonach Seductive Details entweder durch Aktivierung ungeeigneten Vorwissens eine ungünstige Schemakonstruktion fördern, den Informationsverarbeitungsprozess der relevanten Informationen unterbrechen oder die Aufmerksamkeit des Lerners von der Informationsverarbeitung der relevanten Informationen ablenken. Das Ziel der vorliegenden Studie ist es, den Informationsverarbeitungsprozess mithilfe von Eye Tracking Verfahren genauer zu betrachten und mit Maßen kognitiver Belastung zu kombinieren, um detaillierte Aussagen über die Wirkweise von Seductive Details treffen zu können.

Die Studie wurde als einfaktorielles Design mit Studenten der Universität des Saarlandes durchgeführt (N=50), der Zwischensubjektfaktor bestand in der zusätzlichen Präsentation von Seductive Details. Beide Gruppen bearbeiteten ein elfseitiges Lernprogramm zum Thema ATP Synthese, wobei die Experimentalgruppe die Seductive Details Version bearbeitete, die zusätzliche Informationen auf 4 der 11 Seiten beinhaltete. Die Augenbewegungen der Versuchspersonen wurden während der Bearbeitung des Lernprogramms aufgezeichnet (Tobii TX300). In Anlehnung an die Eye-Mind Hypothese (Rayner, 1998) konzentrierte sich die Auswertung der Blickdaten auf die Fixationsdauer, die Anzahl der Fixationen und die direkten Wechsel der Fixationen zwischen den relevanten Aereas of Interest (AOIs). Die kognitive Belastung wurde mithilfe der Rhythmus-Methode (Park, 2010; Park & Brünken, 2011; Park & Brünken, resubm.), dem Index of Cognitive Activity (ICA) (Marshall, 2007), sowie den Items zur Selbsteinschätzung der mentalen Belastung nach Paas (1992) erhoben. Der Lernerfolg wurde mit einem Nachtest im Anschluss an die Bearbeitung des Lernprogramms erfasst.

Die Ergebnisse bestätigen den Seductive Details Effekt, mit einem signifikant negativen Einfluss der Seductive Details auf den Lernerfolg $F(1,48)=6.578$, $p<.05$, $d=.74$. Die Analyse der Augenbewegung zeigt signifikante Unterschiede für die gesamte Fixationsdauer der relevanten Bild AOIs, $F(1,45) = 11.543$, $p<.01$, $d=1.02$, die gesamte Anzahl der Fixationen innerhalb der relevanten Bild AOIs, $F(1,45)=12.750$, $p<.01$, $d=1.07$ und die Wechsel von den relevanten Text AOIs zu den relevanten Bild AOIs, $F(1,37)=6.792$, $p<.05$, $d=.86$. Dabei werden die relevanten Abbildungen in der Gruppe ohne Seductive Details insgesamt sowohl länger, als auch häufiger fixiert und es wird häufiger zwischen den relevanten AOIs gewechselt. Für die gesamte Fixationsdauer der relevanten Text AOIs, $F<1$, sowie die Anzahl der Fixationen innerhalb der relevanten Text AOIs zeigten sich keine signifikanten Unterschiede, $F(1,44)=1.818$, n.s.. Darüber hinaus zeigte die Rhythmus-Methode eine signifikant höhere kognitive Belastung für die Seductive Details Gruppe $F(1,44)=4.103$, $p<.05$, $d=.61$. In der subjektiven Einschätzung der mentalen Belastung und der Aufgabenschwierigkeit zeigten sich keine signifikanten Gruppenunterschiede, $F(1,48)=1.433$, n.s. bzw. $F<1$. Der ICA zeigte ebenfalls keinen signifikanten Unterschied zwischen den Gruppen für die Verarbeitung der relevanten Bild-, $F<1$ und Textinformation AOIs, $F(1,47)=1.063$, n.s.. Die geringere Anzahl der Fixationen sowie die kürzere Fixationszeit innerhalb der Bild AOIs und besonders die geringere Anzahl der Wechsel zwischen den Text und Bild AOIs geben Grund zu der Annahme, dass Seductive Details zu einer insgesamt oberflächlichen Informationsverarbeitung der relevanten Informationen führen. In Kombination mit den Ergebnissen der Rhythmus Methode ist die Annahme naheliegend, dass die höhere kognitive Belastung in der Seductive Details Gruppe auf Prozesse wie das Aufrechterhalten der relevanten Information oder die Informationsselektion zurückzuführen sind. Somit stützen die Ergebnisse zunächst die Annahme einer Unterbrechung der relevanten Informationsverarbeitungsprozesse. Um hier präzisere Aussagen treffen zu können, soll im weiteren Verlauf eine Blickpfadanalyse durchgeführt werden, um zu überprüfen inwiefern die Schemakonstruktion durch Wechsel zwischen relevanten Informationen und Seductive Details beeinflusst wird.

Donnerstag
12. März

15:30 Uhr

GC 03/142

Chair(s)

Korbach, Andreas

Dicke, Theresa; Stebner, Ferdinand; Kühl, Tim; Leutner, Detlev; Wirth, Joachim

Ein Modell zum Zusammenspiel von Motivation, mentaler Belastung und mentaler Anstrengung beim Lernen mit Multimedia

Ein erfolgreicher Lerner zeichnet sich dadurch aus, dass er im Sinne der Cognitive Load Theory (CLT; Sweller, 1988) so viel mentale Anstrengung in einen Lernvorgang investiert wie nötig. Dabei stellt sich die Frage, welche Faktoren das Investieren mentaler Anstrengung beim Lernen beeinflussen. Ist der Lernende motiviert, investiert er durchschnittlich womöglich mehr mentale Anstrengung. Lineare Zusammenhänge zwischen der Motivation und der mentalen Anstrengung konnten bislang aber nur in wenigen Studien konsistent nachgewiesen werden (vgl. Zander, 2010). Ein zweiter relevanter Faktor ist die mentale Belastung (Paas, Tuovinen, Tabbers & van Gerven, 2003), wobei zunehmende mentale Belastung durch eine höhere Schwierigkeit einer Lernaufgabe das Investieren mentaler Anstrengung begünstigen könnte. Empirische Befunde können allerdings auch diese Annahmen nicht eindeutig bestätigen (Stebner, 2012). Vielmehr wird deutlich, dass das Zusammenspiel der drei Faktoren Motivation, mentale Belastung und mentale Anstrengung beim multimedialen Lernen bisweilen nur wenig erforscht wurde (Paas, Tuovinen, von Merrienboer & Darabi, 2005) und es an einem theoretischen Modell mangelt.

In der vorliegenden Studie werden zunächst Mediationsmodelle überprüft, in denen Motivation und mentale Belastung den Lernerfolg jeweils indirekt über die mentale Anstrengung vorhersagen. Anschließend wird ein theoretisches Modell überprüft, das von Stebner (2012) auf Basis der CLT entwickelt wurde und die mentale Anstrengung beim Lernen als Resultat einer Interaktion von Motivation und mentaler Belastung ansieht.

Die Stichprobe besteht aus 274 nordrhein-westfälischen Gymnasiasten (47.4% männlich) mit einem Durchschnittsalter von $M = 13.05$ Jahren ($SD = 0.77$, $Min = 11$, $Max = 16$). Die Probanden lernten in einer computerbasierten multimedialen Lernumgebung die chemischen Prozesse des Waschvorgangs (Höffler, Prechtl & Nerdel, 2010). Es wurde das Vor- und Nachwissen sowie die mentale Anstrengung (Paas, 1992), die mentale Belastung (Kalyuga, Chandler & Sweller, 1999) und die Motivation (u.a. Rheinberg, Vollmeyer & Burns, 2001) zu mehreren Zeitpunkten erhoben.

Die Pfadmodelle wurden mit Mplus (Muthén & Muthén, 2010) berechnet. Die moderierte Mediationsanalyse zur Berechnung eines konditionalen indirekten Effekts wurde mit dem MODMED-Verfahren von Preacher et al. (2007) durchgeführt. Alle Ergebnisse basieren auf bias-korrigierten Konfidenzintervallen (95% CI).

Zunächst wurden die angenommenen Mediationsmodelle überprüft, indem a) Motivation den Lernerfolg via mentaler Anstrengung (Modell A) und b) mentale Belastung den Lernerfolg via mentaler Anstrengung vorhersagt (Modell B). Dabei zeigt sich für Modell A lediglich signifikante direkte Effekte, bei dem höhere Motivation eine höhere Anstrengung sowie einen größeren Lernerfolg vorhersagt. In Modell B zeigten sich ebenfalls signifikante direkte Effekte, in denen eine höhere mentale Belastung einen geringeren Lernerfolg und eine höhere Anstrengung hervorsagt. Höhere Anstrengung sagt zudem einen höheren Lernerfolg vorher. Der angenommene indirekte Effekt von mentaler Belastung über mentaler Anstrengung zu Lernerfolg war ebenfalls signifikant und positiv.

Im Anschluss wurde eine mögliche Interaktion von Motivation und mentaler Belastung überprüft. Da sich lediglich für Modell B ein signifikanter indirekter Effekt zeigte, wurde eine moderierte Mediation getestet, in welcher die Stärke der Mediation von Modell B vom Niveau der Motivation abhängt. Das Modell zeigte eine gute Passung auf die Daten und in Übereinstimmung mit den vorherigen Ergebnissen zeigten sich signifikante direkte Effekte von mentaler Belastung auf Anstrengung und Lernerfolg sowie von Motivation auf Anstrengung und von dieser wiederum auf den Lernerfolg. Zudem zeigte sich ebenso ein signifikanter indirekter Effekt von mentaler Belastung auf den Lernerfolg via Anstrengung. Dieser blieb jedoch für alle Werte des Moderators (Motivation) konstant und damit unabhängig vom Niveau des Moderators. Dementsprechend wurde ein finales Modell aufgestellt, in dem Modell B um einen direkten Pfad von Motivation auf die Anstrengung ergänzt wurde. Das Modell zeigte eine sehr gute Passung und alle Pfade wurden signifikant.

Zusammenfassend können die Ergebnisse als erster erfolgreicher Schritt zur Validierung eines theoretischen Modells, welches das Zusammenspiel von Motivation, mentaler Belastung und mentaler Anstrengung beim multimedialen Lernen beschreibt, angesehen werden.

Donnerstag
12. März

15:30 Uhr

GC 03/142

Chair(s)
Korbach, Andreas

Hahnel, Carolin; Goldhammer, Frank; Naumann, Johannes; Kröhne, Ulf

Zur Bedeutung exekutiver Arbeitsgedächtnisfunktionen beim Lesen digitaler Texte

Theoretischer Hintergrund: Exekutive Arbeitsgedächtnisfunktionen spielen insgesamt beim Lesen und Verstehen von Texten eine zentrale Rolle (z.B. Kendeou et al., 2014). Sie befähigen Lesende dazu, ihre Informationsverarbeitung während des Leseprozesses zu steuern und zu kontrollieren. Digitale Texte werden allerdings häufig in Form von Hypertexten gelesen, die zusätzliche Anforderungen an das Arbeitsgedächtnis stellen (DeStefano & LeFevre, 2007). Hypertexte, wie sie vor allem im Internet zu finden sind, zeichnen sich durch eine verzweigte Verbindungsstruktur zwischen einzelnen Textteilen aus. Textteile oder Seiten sind dabei untereinander durch Hyperlinks verbunden, über die sie angesteuert werden können. Für Lesende bedeutet das, dass sie neben dem eigentlichen Leseprozess zusätzlich kognitive Ressourcen aktivieren müssen, um mit der vorliegenden Verknüpfungsstruktur angemessen umgehen zu können (z.B. Naumann et al., 2008).

Unsere Studie befasst sich mit der Bedeutung der exekutiven Arbeitsgedächtnisfunktion Updating für das Lesen digitaler Texte (Digital Reading). Updating beschreibt, wie effektiv mentale Textrepräsentationen im Arbeitsgedächtnis überwacht und aktualisiert werden, wenn neue aufgabenrelevante Informationen hinzukommen (Miyake et al., 2000). Wenn beim Hypertextlesen Informationen von mehreren Seiten benötigt werden, sollten effektive Arbeitsgedächtnismechanismen Lesende dabei unterstützen, bereits gelesene Informationen mental zu speichern und diese erneut abzurufen, um sie mit neuerer Information abgleichen und gegebenenfalls aktualisieren zu können. Wir erwarten daher, dass eine bessere Updating-Fähigkeit die Leistung von PISA Leserinnen und Lesern beim Lösen digitaler Textaufgaben vorhersagt.

Weiterhin sollten entsprechend der cognitive load theory insbesondere kognitiv anspruchsvolle Aufgaben Arbeitsgedächtnisanforderungen erhöhen (DeStefano & LeFevre, 2007). PISA Leseaufgaben werden anhand ihres primär angesprochenen kognitiven Leseprozesses klassifiziert (Aspekte des Lesens; OECD, 2013). Dabei wird angenommen, dass die Leseaspekte „Informationen suchen und extrahieren“, „Kombinieren und Interpretieren“, „Reflektieren und Bewerten“ und – der Digital Reading-eigene Aspekt – „Komplex“ hierarchisch organisiert sind. Ein hypertextspezifischeres Merkmal stellt außerdem die Gesamtanzahl aufrufbarer Seiten in einer digitalen Leseaufgabe dar. Wir erwarten, dass digitale Leseaufgaben (1) schwieriger werden, wenn sie komplexere kognitive Prozesse fokussieren und verhältnismäßig viele Seiten umfassen, und dass (2) die Bedeutung exekutiver Arbeitsgedächtnisfunktionen vor allem in schwierigen Aufgaben zum Tragen kommt.

Methode: Aus der PISA 2012-Stichprobe wurde eine Substichprobe aus 293 Schülerinnen und Schülern gezogen, die sowohl an der PISA Haupterhebung als auch an einer zusätzlichen Begleituntersuchung zum Einsatz computerbasierter Assessments in PISA teilnahmen. Aus der PISA Haupterhebung wurden die Schülerantworten aus den 19 Aufgaben der Digital Reading-Skala verwendet. Als schwierigkeitsbestimmende Merkmale der Digital Reading-Aufgaben wurden der primäre Leseaspekt sowie die Gesamtanzahl der Hypertextseiten in einer Digital Reading-Aufgabe berücksichtigt. Im Rahmen der Begleituntersuchung wurde das Arbeitsgedächtnis mithilfe einer numerischen Updating-Aufgabe und als Kovariate das allgemeine Leseverständnis erfasst.

Ergebnisse: Erste Ergebnisse aus der Analyse generalisierter gemischter linearer Modelle (GLMM; De Boeck et al., 2011) zeigten, dass Updating ($B = 0.24$, $p < .001$) über den Effekt des allgemeinen Leseverständnisses hinaus ($B = 0.75$, $p < .001$) die Lösungswahrscheinlichkeit der digitalen Leseaufgaben positiv vorhersagt. SchülerInnen profitieren folglich von effizienten exekutiven Arbeitsgedächtnisfunktionen – unabhängig davon, ob sie gute Leser sind oder nicht. Weiterhin sank die Lösungswahrscheinlichkeit von Aufgaben (d.h. die Aufgaben wurden schwerer), wenn sie hierarchiehöheren Leseaspekten zugeordnet wurden (Information suchen und extrahieren: $B = 1.29$, $p < .001$; Kombinieren und Interpretieren: $B = 0.43$, $p < .001$; Reflektieren und Bewerten: $B = 0.57$, $p < .001$; Komplex: $B = -0.41$, $p < .001$) oder ihre Hypertextumgebungen viele Seiten umfassten ($B = -0.67$, $p < .001$). In einem zweiten Modell wurden zusätzlich Interaktionen zwischen Updating, den Leseaspekten und der Gesamtseitenanzahl spezifiziert. Es fand sich eine positive Dreifachinteraktion zwischen Updating, dem Leseaspekt „Komplex“ und der Seitenzahl ($B = 0.92$, $p = .006$). Folglich profitieren SchülerInnen insbesondere dann von effizienten exekutiven Arbeitsgedächtnisfunktionen, wenn sie umfangreiche und kognitiv anspruchsvolle Hypertextleseaufgaben bearbeiten. Die Ergebnisse werden hinsichtlich ihrer Relevanz für den Umgang und das Lernen mit Hypertexten diskutiert.

Donnerstag
12. März

15:30 Uhr

GC 03/142

Chair(s)

Korbach, Andreas

Knörzer, Lisa; Brünken, Roland; Park, Babette

Beeinflussen Lernermerkmale die Wirkung von Emotionen auf multimediales Lernen?

Theoretischer Hintergrund und Fragestellung: Die Cognitive-Affective Theory of Learning with Media (CATLM; Moreno, 2006) Modell umfasst die Grundannahmen des Lernens mit Multimedia der Cognitive Theory of Multimedia Learning (CTML; Mayer, 2005) und fügt drei weitere Zusatzannahmen hinzu: Die affective mediation assumption postuliert, dass sowohl affektive und motivationale Faktoren wie beispielsweise Emotionen oder die aktuelle Lernmotivation das Lernen und somit den Lernerfolg beeinflussen. Ebenso wirken metakognitive und selbstregulative Fähigkeiten (metacognitive mediation assumption) sowie Lernermerkmale (individual differences assumption) auf den Lernerfolg.

Diese Studie fokussiert basierend auf dem Circumplex Model of Affect (Russel, 2003) als zugrundeliegendes Emotionskonstrukt einerseits die affective mediation assumption und untersucht die Wirkung von Emotionen unterschiedlicher Valenz (positiv, negativ, plus eine neutrale Kontrollgruppe) auf das Lernen mit Multimedia. Andererseits werden in Anlehnung an die individual differences assumption Lernermerkmale bezüglich kognitiver Ressourcen (Vorwissen, Arbeitsgedächtniskapazität), Emotionsregulationsstrategien als Trait (kognitive Neubewertung, expressive Unterdrückung; Gross, 1998) und Persönlichkeitsmerkmalen (Neurotizismus und Offenheit; Borkenau & Ostendorf, 2007) als Moderatorvariablen geprüft.

Methode: Die Studie wurde als einfaktorielles experimentelles Design mit dem dreifach gestuften Faktor Emotionsvalenz (positiv, neutral, negativ) realisiert. Es nahmen N=75 Studierende teil (69% weiblich, Alter: M=24.01, SD=3.44), welche zufällig einer der drei Gruppen zugewiesen wurden.

Zur Emotionsinduktion wurde eine im Vorfeld getestete (Knörzer, Brünken, & Park, eingereicht), kombinierte Methode mit Musik und autobiographischem Erinnern eingesetzt (Eich et al., 2007; Kuhbandner et al., 2011). Als multimediale Instruktion wurde ein computerbasiertes Biologie-Lernprogramm in visueller Modalität (Text + Bild) und mit programmgesteuerter Lernzeit verwendet. Die Lernleistung wurde mit einem Test bestehend aus 21 Items gemessen (max. 30 Punkte). Zur Erhebung einer Emotionsbaseline und als Manipulationscheck diente die Valenz-Skala aus den PANAVA-Kurzskalen (Schallberger, 2005). Als Kontrollvariablen bzw. mögliche Moderatoren wurden ein Vorwissenstest, ein Arbeitsgedächtniskapazitätstest (Corsi-Block-Aufgabe; Kessels et al., 2008), ein Fragebogen zur Erfassung der Emotionsregulationsstrategien Neubewertung und Unterdrückung (ERQ; Abler & Kessler, 2009) sowie die Skalen Offenheit und Neurotizismus als Persönlichkeitsmerkmale (NEO-ffi; Borkenau & Ostendorf, 2007) verwendet. Alle Skalen besaßen zufriedenstellende Reliabilität.

Die Studienteilnehmer wurden individuell getestet. Zunächst wurden die Emotionsbaseline, Kontroll- bzw. Moderatorvariablen erhoben, es folgten die experimentelle Manipulation und der Manipulationscheck. Anschließend wurde das Lernprogramm präsentiert und abschließend der Lernerfolgstest durchgeführt.

Ergebnisse: In der Emotionsbaseline sowie den Kontroll- bzw. Moderatorvariablen wurden keine signifikanten Gruppenunterschiede festgestellt. Der Manipulationscheck nach der Emotionsinduktion bestätigte eine erfolgreiche Emotionsinduktion mit signifikanten Emotionsunterschieden zwischen allen drei Gruppen. Bezüglich der Lernleistung schnitt die positive Gruppe am schlechtesten, die negative Gruppe am besten ab.

Zur Untersuchung der angenommenen Moderationen wurde das PROCESS-Makro für SPSS (Hayes, 2012) verwendet, wobei aufgrund von Softwareanforderungen lediglich die positive und negative Gruppe in die Analyse einbezogen werden konnten. Es wurden drei Moderatormodelle mit Emotionsvalenz als unabhängiger und Lernleistung als abhängiger Variable sowie jeweils zwei Moderatorvariablen berechnet. Als Moderatoren wurden paarweise die erhobenen Lernermerkmale kognitive Ressourcen (Vorwissen und Arbeitsgedächtniskapazität), Emotionsregulationsstrategien (Unterdrückung und Neubewertung) sowie Persönlichkeitsmerkmale (Offenheit und Neurotizismus) aufgenommen. Die Wahl des Moderatormodells (doppelte einfache Moderation vs. moderierte Moderation) wurde aufgrund des Grades der zusätzlichen Varianzaufklärung durch die Dreifachinteraktion bei einer moderierten Moderation (Emotionsvalenz*Moderator1*Moderator2) getroffen. Statistische Daten werden im Rahmen des Konferenzbeitrages detailliert präsentiert und diskutiert. Zusammengefasst zeigte sich, dass Lerner mit geringem Vorwissen unabhängig von ihrer Arbeitsgedächtniskapazität eine geringere Lernleistung erzielten, welche bei offenen Lernern (unabhängig ihrer Neurotizismuswerte) weniger von Emotionen beeinflusst wurde. Bei Lernern mit einer deutlich bevorzugten, stark ausgeprägten Emotionsregulationsstrategie (z.B. starke Neubewertung und geringe Unterdrückung) verschwand der Einfluss der Emotionsvalenz auf den Lernerfolg bzw. kehrte sich um. Die Ergebnisse unterstreichen deutlich die Wichtigkeit von Lernermerkmalen bei der Untersuchung des Einflusses von Emotionen auf multimediales Lernen.

Donnerstag
12. März

15:30 Uhr

GC 03/142

Chair(s)

Korbach, Andreas

Harks, Birgit; Klieme, Eckhard; Hartig, Johannes; Leiss, Dominik

Gendereffekte in Mathematik unter Berücksichtigung von kognitiven und inhaltlichen Teilbereichen

Theoretischer Hintergrund: Obwohl sich Jungen und Mädchen kaum hinsichtlich ihrer durchschnittlichen Mathematikleistung unterscheiden (Lindberg, Hyde, Petersen & Linn, 2010), zeigen sich für bestimmte mathematische Inhaltsbereiche geschlechtsspezifische Leistungsunterschiede. Befunde hierzu sind jedoch widersprüchlich: Die Metaanalyse von Hyde, Fennema und Lamon (1990) zeigt z.B. keinen Leistungsunterschied in Algebra, aber einen Leistungsvorteil von Jungen in Geometrie. Mullis, Martin und Foy (2008) hingegen beschreiben einen Leistungsvorteil von Mädchen in Algebra und Geometrie. Widersprüchliche Ergebnisse wie diese sind möglicherweise auch durch unterschiedliche kognitive Anforderungen der verwendeten Items (z.B. Rechnen vs. Problemlösen) bedingt. So gibt es beispielsweise Hinweise dafür, dass Jungen Mädchen in Problemlöse- oder Textaufgaben überlegen sind, während Mädchen in Rechenaufgaben relativ gut abschneiden (Hyde et al., 1990).

Der vorliegende Beitrag soll zu einem besseren Verständnis von Geschlechtsunterschieden in Mathematik beitragen, indem bei der Analyse von Gendereffekten kognitive und inhaltliche Teilbereiche zugleich berücksichtigt werden. Anknüpfend an Harks, Klieme, Hartig und Leiss (im Druck) konzentrieren wir uns dabei auf die inhaltlichen Bereiche Pythagoras und lineare Gleichungssysteme sowie auf die kognitiven Bereiche Modellierungskompetenz und technische Kompetenz. Die empirische Trennbarkeit inhaltspezifischer Modellierungs- und technischer Kompetenzen wurde von Harks et al. bereits demonstriert.

Forschungsfragen: Unterscheiden sich Jungen und Mädchen hinsichtlich ihrer... (1) Leistung in den Inhaltsbereichen (a) Pythagoras und (b) lineare Gleichungssysteme? (2) (a) Modellierungskompetenz und (b) technischen Kompetenz? (3) (a) Modellierungskompetenz im Inhaltsbereich Pythagoras, (b) technischen Kompetenz im Bereich Pythagoras, (c) Modellierungskompetenz im Inhaltsbereich lineare Gleichungssysteme und (d) technischen Kompetenz im Bereich lineare Gleichungssysteme?

Methode: Es wurden 122 Items konstruiert, die dem Inhaltsbereich lineare Gleichungssysteme oder dem Inhaltsbereich Pythagoras zugeordnet werden können und primär technische Kompetenz oder Modellierungskompetenz erfassen. Die Items wurden von insgesamt 1 570 Schüler(inne)n (51,4% weiblich) der neunten Jahrgangsstufe (Realschule) bearbeitet. Zur statistischen Auswertung wurden drei zweiparametrische mehrdimensionale Item-Response-Theorie Modelle mit Einfachladungsstruktur analysiert, in denen jeweils Geschlecht (0 = weiblich, 1 = männlich) als Prädiktor der latenten Leistungsdimensionen modelliert wurde. Modell 1 differenziert zwischen den Dimensionen Pythagoras und lineare Gleichungssysteme, Modell 2 zwischen technischer und Modellierungskompetenz und Modell 3 zwischen technischer Kompetenz und Modellierungskompetenz jeweils spezifisch für die Inhaltsbereiche Pythagoras und lineare Gleichungssysteme. Zur Beantwortung der Fragestellungen werden die standardisierten Betakoeffizienten betrachtet. Ergänzend wurde pro Modell ein Waldtest durchgeführt, der die Äquivalenz der standardisierten Betakoeffizienten prüft. Die Analysen erfolgten in MPlus 7.0 (Muthén & Muthén, 1998-2012).

Ergebnisse: Die Ergebnisse zu Modell 1 zeigen einen Leistungsvorteil von Jungen gegenüber Mädchen in den Inhaltsbereichen (a) Pythagoras ($\beta = .33$, $p < .001$) und (b) lineare Gleichungssysteme ($\beta = .35$, $p = .039$), wobei sich die beiden Gendereffekte nicht signifikant voneinander unterscheiden (Wald = 0.01, $df = 1$, $p = .910$). Die Befunde zu Modell 2 deuten darauf hin, dass Schüler (a) eine höhere Modellierungskompetenz ($\beta = .56$, $p < .001$) aufweisen als Schülerinnen, (b) nicht aber eine höhere technische Kompetenz ($\beta = .21$, $p = .283$). Die beiden Gendereffekte unterscheiden sich signifikant voneinander (Wald = 3.96, $df = 1$, $p = .047$). Die Ergebnisse zu Modell 3 zeigen, dass Jungen im Vergleich zu Mädchen über eine höhere (a) Modellierungskompetenz ($\beta = .48$, $p < .001$) und (b) technische Kompetenz ($\beta = .21$, $p = .010$) im Inhaltsbereich Pythagoras verfügen sowie über eine höhere (c) Modellierungskompetenz ($\beta = .53$, $p < .001$) im Bereich lineare Gleichungssysteme; (d) die technische Kompetenz im Bereich lineare Gleichungssysteme ist aber nicht höher ausgeprägt als die der Mädchen ($\beta = .05$, $p = .679$). Die Gendereffekte für die vier Dimensionen unterscheiden sich signifikant voneinander (Wald = 13.60, $df = 3$, $p = .004$).

Die Ergebnisse verdeutlichen den Mehrwert einer differenzierten Betrachtung von Gendereffekten in Mathematik unter Berücksichtigung von inhaltlichen und kognitiven Teilbereichen. Mögliche Implikationen, auch unter dem Aspekt der individuellen Förderung, werden diskutiert.

Donnerstag
12. März

15:30 Uhr

GC 03/149

Schroeders, Ulrich; Penk, Christiane; Jansen, Malte; Pant, Hans Anand

Geschlechtsunterschiede in mathematisch-naturwissenschaftlichen Kompetenzen am Ende der Sekundarstufe I

Theoretischer Hintergrund: In der öffentlichen und bildungspolitischen Diskussion um eine faire Beteiligung aller Schülergruppen an schulischen Bildungsprozessen kommt geschlechtsbezogenen Leistungsunterschieden in den MINT-Fächern besonders großes Interesse entgegen (Hannover, 2004; Hannover & Kessels, 2011), auch deshalb, weil eine gute Ausbildung in diesen Fächern als bedeutsam für die wirtschaftliche Entwicklung und für den Wohlstand eines Landes angesehen wird (KMK, 2009; Rindermann & Thompson, 2011). An den Hochschulen in der Bundesrepublik Deutschland ist der Anteil der Frauen unter den Absolvierenden der MINT-Fächer mit 31 Prozent relativ gering. Inwieweit diese Unterrepräsentation auf schulische Kompetenzunterschiede zwischen den Geschlechtern zurückführen lässt, ist nicht hinreichend geklärt. Wie Geschlechterunterschiede in mathematisch-naturwissenschaftlichen Kompetenzen ausfallen, hängt von einer Vielzahl von Faktoren ab: vom untersuchten Kompetenzbereich (Hyde, 2005), der Altersgruppe (Bloom, Hill, Black & Lipsey, 2008) sowie der methodischen Herangehensweise (Brunner, Krauss & Martignon, 2011). In großen Schulleistungsstudien wurden für die mittleren Kompetenzstände von Schülerinnen und Schülern an deutschen Schulen im mathematischen Bereich signifikant höhere Kompetenzwerte für die Jungen berichtet (Prenzel et al., 2007a; Bos et al., 2003; Winkelmann, van den Heuvel-Panhuizen & Robitzsch, 2008), wohingegen sich für die Naturwissenschaften keine Kompetenzunterschiede zwischen den Geschlechtern finden ließen (Prenzel et al., 2007b; Rönnebeck, Schöps, Prenzel, Mildner & Hochweber, 2010).

Fragestellung: Im vorliegenden Beitrag werden Befunde aus dem IQB-Ländervergleich 2012 berichtet, in dem die mittleren Kompetenzstände von Jungen und Mädchen in Mathematik und drei naturwissenschaftlichen Fächern – Biologie, Chemie, Physik – verglichen wurden. Durch die separate Fächerbetrachtung erlaubt die Studie damit ein differenziertes Abbild der Geschlechtsunterschiede in den Naturwissenschaften. Des Weiteren werden die Mittelwertsvergleiche nach Inhaltsbereichen und nach Schulart (Gymnasium vs. Nicht-Gymnasium) differenziert. Neben Mittelwertsunterschieden wurden auch Unterschiede in der Kompetenzverteilung untersucht, da eine stärkere Leistungsvariabilität bei den Jungen bereits seit Jahrzehnten in der psychologischen Forschung diskutiert wird (vgl. Greater-male-variability-Hypothese; Feingold, 1992; Brunner, Gogol, Sonnleitner, Keller, Krauss, Preckel, 2014).

Methode: Am IQB-Ländervergleich 2012 nahm eine für Deutschland repräsentative Stichprobe von rund 45.000 Jugendlichen der neunten Jahrgangsstufen teil. Im mathematischen Bereich wurde außer einem Globalmaß, die Leistung in Bezug auf die fünf Leitideen a) Zahl, b) Messen, c) Raum und Form, d) Funktionaler Zusammenhang und e) Daten und Zufall erhoben; im naturwissenschaftlichen Bereich entlang zweier Kompetenzbereiche (Erkenntnisgewinnung, Fachwissen) in Biologie, Chemie und Physik gemäß der Bildungsstandards (KMK, 2004, 2005a, 2005b, 2005c, 2005d).

Ergebnisse: Sowohl für die globale mathematische Kompetenz als auch für die fünf mathematischen Leitideen ergaben sich leicht höhere und statistisch signifikante Kompetenzwerte für die Jungen im Vergleich zu den Mädchen. Der Unterschied in der Globalskala entspricht am Ende der Sekundarstufe I einer Effektstärke von $d = 0.16$ und einem Lernzuwachs von ungefähr zwei Drittel Schuljahren. Die größten Unterschiede für die inhaltlichen Leitideen zeigten sich für a) Zahl und b) Daten und Zufall. Für die naturwissenschaftlichen Fächer wiesen mit Ausnahme des Kompetenzbereichs Physik Fachwissen die Mädchen Leistungsvorteile gegenüber den Jungen auf. Im Fach Biologie waren die Unterschiede mit $d = 0.22$ für Fachwissen und Erkenntnisgewinnung am deutlichsten. Dieser Vorteil der Mädchen im Fach Biologie entspricht näherungsweise einem Leistungsvorsprung von einem Schuljahr. In den Fächern Chemie und Physik ließen sich wesentlich geringere geschlechtsbezogenen Kompetenzunterschiede zugunsten der Mädchen feststellen. Bzgl. der Leistungsstreuung ergab sich folgendes Bild: Während in allen mathematischen Skalen die Leistungen beider Geschlechter gleich stark streuten, fiel die Streuung in allen naturwissenschaftlichen Kompetenzbereichen für die Jungen höher aus. Die Aufschlüsselung der Ergebnisse hinsichtlich der Schulart ergab für Mathematik keine schulartspezifischen Geschlechtseffekte. In den Naturwissenschaften hingegen ließen sich systematische Verschiebungen zugunsten der Jungen feststellen, so dass sich die geringen Unterschiede in den Fächern Chemie und Physik sogar umkehrten (vgl. Simpson-Paradoxon). Die Ergebnisse werden hinsichtlich der praktischen Bedeutsamkeit der gefundenen Effekte, den potenziellen Ursachen und der Debatte von Jungen als den neuen „Bildungsverlierern“ diskutiert.

Donnerstag
12. März

15:30 Uhr

GC 03/149

Chair(s)

Harks, Birgit

Hörmann, Markus

Leistungsbeurteilung im (virtuellen) Klassenzimmer: Über den Einfluss von Gender-Stereotypisierung auf die Validität von Lehrerurteilen

Theoretischer Hintergrund: Die diagnostische Kompetenz ist aus wissenschaftlicher, politischer und didaktischer Perspektive eine wesentliche Facette der pädagogisch-psychologischen Expertise von Lehrenden. Trotz der weitreichenden Auswirkung von Beurteilungen auf die Bildungsbiographien von Schülerinnen und Schülern (SuS) wird seit Dekaden die mangelnde Validität dieser Urteile und ein unbefriedigender Forschungsstand beklagt (Dünnebieber, Gräsel & Krolak-Schwerdt, 2009; Artelt & Gräsel, 2009).

Bisherige Ergebnisse zur Beurteilungsvalidität von Lehrenden berichten Abweichungen von etwa 20% zu standardisierten Testverfahren (Südkamp & Möller, 2009), eine generelle Tendenz zur Mitte (Fiedler, Walther, Freytag & Plessner, 2002) sowie eine validere Beurteilung der Rangreihenfolge im Vergleich zu absoluten Einschätzungen (Schrader & Helmke, 2002).

Das sozialkognitive duale Prozessmodell liefert, übertragen auf den Unterricht, theoretische Hinweise auf eine mögliche Einschränkung der Validität durch die geschlechterspezifische Stereotypisierung von SuS durch Lehrende. Abhängig von der Falltypikalität und der Verantwortlichkeit werden Informationen nach diesem Modell entweder heuristisch oder systematisch verarbeitet (z.B. Chen & Chaiken, 1999). Gender-Stereotype können für Lehrende als solche Heuristiken dienen und die Validität bei Beurteilungen einschränken.

In diesem Zusammenhang steht, dass seit langem eine sinkende offene und explizite Gender-Stereotypisierung der MINT-Bereiche zu beobachten ist (z.B. Halpern et al., 2007) und sogar die weitgehende explizite Ablehnung des Stereotyps, dass Männer eher für Mathematik geeignet sind als Frauen, dokumentiert werden kann (Hyde, Fennema, Ryan, Frost & Hopp, 1990). Anders verhält es sich allerdings mit impliziten Gender-Stereotypen, deren Existenz und Bedeutung immer deutlicher durch empirische Befunde belegt werden. Bezüglich Mathematik und Sprache zeigen sich deutliche implizite Gender-Stereotype (z.B. Kiefer & Sekaquaptewa, 2007). Diese werden von Männern und Frauen gleichermaßen geteilt (Nosek & Smyth, 2011) und treten bereits im Vorschulalter auf (Steffens, Jelenec & Noack, 2010).

Fragestellung: Fundierte empirische Belege für kausale Zusammenhänge zwischen Geschlecht und Laufbahneempfehlungen als ein wichtiger Bereich der diagnostischen Kompetenz von Lehrenden fehlen bislang. Einer der Gründe für diese Lücke ist die schwere Kontrollierbarkeit von Schülermerkmalen wie Intelligenz, Motivation oder Fähigkeit. Daher wird in dieser Studie untersucht, ob SuS aufgrund ihres Geschlechts und unabhängig von ihrer tatsächlichen Fähigkeit Motivation unterschiedliche Empfehlungen für die Studiengänge Maschinenbau und Germanistik erhalten. Durch die Verwendung des virtuellen Klassenzimmers als Untersuchungsparadigma können Fähigkeit und Motivation der SuS kontrolliert werden.

Zudem wird nach den Hintergründen dieser möglicherweise stereotypen Empfehlungen gefragt. Die bisherigen Befunde lassen eine implizite Stereotypisierung vermuten, die sich auf Mathematik und Sprache bezieht. Konkret wird also eine Assoziation zwischen Geschlecht und Domäne gemäß traditionellen Gender-Stereotypen vermutet.

Wir nehmen an, dass es ebenso übereinstimmend mit aktuellen Erkenntnissen zu keinem Zusammenhang zwischen expliziten und impliziten Gender-Stereotypen kommt.

Methodik: Neunundfünfzig weibliche und 22 männliche Probanden, die über Interessenvertretungen Lehrender akquiriert wurden, nahmen an dieser experimentellen Online-Studie teil. Die abhängigen Variablen Validität der Einschätzungen (bzgl. Fähigkeit und Motivation) sowie die Empfehlung für einen Studiengang (Maschinenbau und Germanistik) wurden mit dem virtuellen Klassenzimmer (z.B. Fielder, Wöllert, Tauber, Hess, 2013) erhoben. In einer virtuell-simulierten Unterrichtssituation interagieren die Probanden in der Rolle des Lehrenden mit sechs fiktiven SuS. Die Probanden stellen Fragen aus verschiedenen Bereichen der Fächer Mathematik und Deutsch. Auf Grundlage der jeweiligen videobasierten Antworten der fiktiven SuS nehmen die Probanden nach der Interaktion jeweils Beurteilungen bezüglich Fähigkeit, Motivation und Studieneignung von allen im Klassenzimmer sitzenden SuS vor.

Zudem wurde die Variable implizite Gender-Stereotype als Assoziationsstärke mit dem Sample Paired Features-Task (SPF; Bar-Anan, Nosek & Vianello, 2009) erhoben. Explizite Gender-Stereotype wurden mit sechs Items erfasst (Steffens et al. 2010). Alle Instrumente wurden für das Experiment auf browserbasierte Umgebungen portiert.

Ergebnisse: Es ergeben sich signifikante, geschlechterbedingte Unterschiede gemäß traditioneller Gender-Stereotype in den Empfehlungen für die Studiengänge Germanistik (d -Cohen = 0.29) und Mathematik (d -Cohen = 0.44). Es zeigen sich implizite Gender-Stereotype Assoziationen zwischen den Domänen Mathematik und Sprachen. Die Unabhängigkeit von expliziten und impliziten Gender-Stereotypen wird nachgewiesen.

Donnerstag
12. März

15:30 Uhr

GC 03/149

Chair(s)
Harks, Birgit

3

Moser, Franziska

Wo sind die Schülerinnen? Eine Bild- und Textanalyse zu Geschlechter(un)gerechtigkeit in deutschen Schulbüchern

Kinder- und Schulbücher erreichen eine große Anzahl von Leserinnen und Lesern. Sie dienen nicht nur der Vermittlung von Wissen, sondern transportieren auch gesellschaftliche Werte, Normen und Stereotype über verschiedene soziale Gruppen (z.B. Markom & Weinhäupl, 2007). Vor diesem Hintergrund haben sich seit mehr als 40 Jahren viele Studien der Frage gewidmet, inwiefern Kinder- und Schulbücher zum gesellschaftlichen Konsens von Geschlechter(un)gerechtigkeit beitragen (z.B. Fichera, 1996; Finsterwald & Ziegler, 2007; Lindner & Lukesch, 1994). Die Darstellung von traditionellen Rollenbildern sowie die Unterrepräsentation von weiblichen im Vergleich zu männlichen Charakteren gaben dabei Anlass zur Kritik, weil Kindern dadurch nur eingeschränkte Identifikationsmöglichkeiten zur Verfügung stehen (z.B. Diekman & Murnen, 2004). Seit den ersten Studien zu geschlechtergerechten Darstellungen in Büchern hat sich die gesellschaftliche und soziale Situation von Frauen und Männern und der gesellschaftliche Diskurs über Geschlecht sehr verändert (Wilde & Diekman, 2005). Deshalb stellt sich im vorliegenden Beitrag die Frage, ob diese Veränderungen in neuen, in Deutschland zugelassenen Schulbüchern auch gefunden werden.

Mittels Inhaltsanalyse (Mayring, 2003) wurden jeweils neun Schulbücher für die Fächer Deutsch und Mathematik der Klassenstufen eins, drei und fünf, die nach dem Inkrafttreten des Bundesgleichstellungsgesetzes (2001) herausgegeben wurden, analysiert und verglichen. Mittels Zufallsauswahl wurden jeweils 10% der Gesamtanzahl an Seiten eines jeden Buches untersucht. Der Fokus der Analyse lag dabei auf den beiden am meisten kritisierten Aspekten vergangener Studien: (1) den Häufigkeiten der Darstellung von weiblichen und männlichen Charakteren sowie (2) deren Aktivitäten und Rollen. Darüber hinaus schließt die Analyse auch ein in den vergangenen Jahren kaum untersuchtes subtileres Merkmal geschlechter(un)gerechter Darstellungen mit ein: (3) die Verwendung von geschlechter(un)gerechter Sprache (Moser, Hannover, & Becker, 2013). Wie eine Reihe von Studien zeigte, führt die Verwendung von generisch männlichen Personenbezeichnungen dazu, dass sich Rezipierende weniger weibliche Personen vorstellen, als wenn geschlechtergerechte Sprachformen verwendet werden (z.B. Stahlberg, Braun, Irmen, & Sczesny, 2007). Gerade im Schulkontext sollte jedoch die Verwendung von geschlechtergerechter Sprache von besonderer Bedeutung sein, um stereotypische Vorstellungen abzubauen.

Die Ergebnisse unserer Studie (Moser & Hannover, 2014) zeigen, dass auch in neuen Mathematikbüchern männliche Charaktere häufiger vertreten sind als weibliche. Für die Deutschbücher fand sich dieses Ergebnis nur für die Darstellung erwachsener Personen. Für beide Schulfächer zeigte sich, dass insgesamt weibliche und männliche Charaktere gleich oft vielen verschiedenen und nicht nur stereotypischen Aktivitäten nachgingen. Anlass zur Kritik geben aber, wie bereits in früheren Studien, die familiären und beruflichen Rollen erwachsener Charaktere. Männer wurden im Vergleich zu Frauen häufiger in beruflichen Tätigkeiten dargestellt, während Frauen im Vergleich zu Männern häufiger in ihrer familiären Rolle als Mutter oder Großmutter dargestellt waren. Geschlechtergerechte Sprache wurde in den Büchern zu beiden Schulfächern angewendet, jedoch wurde die Anwendung nicht konsequent umgesetzt. Zusammengefasst zeigen die Ergebnisse, dass neue Schulbücher im Vergleich zu früheren geschlechtergerechter geworden sind, sie aber nach wie vor auf subtile Weise zur sozialen Konstruktion der Ungleichheit der Geschlechter beitragen.

Donnerstag
12. März

15:30 Uhr

GC 03/149

Chair(s)
Harks, Birgit

Wille, Eike; Hübner, Nicolas; Oschatz, Kerstin; Nagengast, Benjamin; Trautwein, Ulrich

Geschlechtsspezifische kognitive und nicht-kognitive Effekte der Oberstufenreform in Baden-Württemberg

Theoretischer Hintergrund: Ein wesentliches Ziel der Oberstufenreform in Baden-Württemberg im Schuljahr 2003/04 betraf die Stärkung der Kompetenzen von Abiturientinnen und Abiturienten im Kernbereich Mathematik bei gleichzeitiger Reduzierung der Fachleistungsunterschiede innerhalb der Schülerschaft (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2002). Dieses Ziel sollte durch die Aufhebung der Niveaudifferenzierung in Grund- und Leistungskurse und die Einführung eines für alle Schülerinnen und Schüler einheitlichen, vierstündigen Mathematikunterrichts erreicht werden. Vom Anspruchsniveau und seiner Intensität, nicht aber vom Umfang des Stoffes orientiert sich dieser an den bisherigen Leistungskursen (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2002).

Verschiedene Studien zum Vergleich von Leistungsentwicklung in gruppiertem und nichtgruppiertem Unterricht zeigen ein uneinheitliches Befundmuster (z.B. Feldhusen, 1989; Hoffer, 1992; Lou et al., 1996; Slavin, 1990). Untersuchungen von Nagy, Neumann, Trautwein und Lüdtke (2010) legen für die Reform der Oberstufe in Baden-Württemberg nahe, dass sich die Mathematikleistung der Abiturientinnen und Abiturienten im Mittel leicht erhöhte und sich die Leistungsstreuung verringerte.

Bisher konnten unterschiedliche Studien aufzeigen, dass zwischen Lernenden am Ende der Schulzeit große Differenzen im Hinblick auf ihre mathematischen Leistungen bestehen (z.B. Baumert, Bos, & Lehmann, 2000; Köller, Baumert, & Schnabel, 1999), insbesondere zwischen Jungen und Mädchen (Köller & Klieme, 2000; Schnabel & Gruehn, 2000). Als Ursache hierfür wird unter anderem das Kurswahlverhalten in der Oberstufe angeführt, da Jungen häufiger Mathematikleistungskurse belegen als Mädchen (Köller & Klieme, 2000; Schnabel & Gruehn, 2000; Zwick & Renn, 2000). In Folge einer Nivellierung des leistungsdifferenzierten Unterrichts, wie durch die Neuordnung der gymnasialen Oberstufe in Baden-Württemberg, können somit insbesondere geschlechtsspezifische Effekte vermutet werden, welche zu einer Verringerung der Leistungsunterschiede zwischen männlichen und weiblichen Abiturienten führen sollten. Eine empirische Unterstützung solcher geschlechtsspezifischer Effekte bei der Umstellung von gruppiertem zu nichtgruppiertem Unterricht steht bislang jedoch aus.

Schulische Leistungen in Mathematik stehen in einem hohen Zusammenhang mit fachspezifischen Selbstkonzepten und Interessen (z.B. Ackerman & Heggstad, 1997; Baumert, Schnabel, & Lehrke, 1998; Eccles, 1983; Köller, Trautwein, Lüdtke, & Baumert, 2006; Marsh, 2005). Es werden jeweils reziproke Wirkrichtungen zwischen Selbstkonzept bzw. Interesse und Leistungen vermutet (Eccles, 2009; Marsh, Trautwein, & Lu, 2005; Möller & Trautwein, 2009; Nagy, 2007). Forschungsergebnisse legen auch hier Geschlechtsunterschiede nahe und zeigen, dass Mädchen im Mittel ein geringeres mathematisches Selbstkonzept (Marsh, 1989; Nagy, Watt, et al., 2010) und ein geringeres Interesse an praktisch-technischen und intellektuell-forschenden Tätigkeiten (Eccles, 1994; Su, Rounds, & Armstrong, 2009) äußern.

Fragestellung: Forschungsleitende Fragestellung der vorliegenden Studie ist die Untersuchung geschlechtsspezifischer Effekte der Oberstufenreform in Baden-Württemberg auf mathematische Leistungen, mathematisches Selbstkonzept und praktisch-technische sowie intellektuell-forschende Interessen.

Methode: In den Analysen wurden Schülerinnen und Schüler, die ihr Abitur vor der Oberstufenreform absolvierten (TOSCA 2002; N =4.730), mit Schülerinnen und Schülern verglichen, die ihr Abitur nach der Oberstufenreform erwarben (TOSCA 2006; N =5.016). Zur Beantwortung der forschungsleitenden Fragestellung wurden multiple Regressionen spezifiziert, in denen jeweils die moderierende Funktion des Geschlechtes in Bezug auf den Zusammenhang zwischen der Reform und dem entsprechenden Outcome überprüft wurde. Zur Berücksichtigung der hierarchischen Struktur der Daten wurden robuste Schätzverfahren verwendet. Fehlende Werte wurden in den Analysen mittels Full-Information-Maximum-Likelihood-(FIML)Verfahren berücksichtigt. Zur Vergleichbarkeit der Ergebnisse der vorliegenden Studie mit bereits bestehenden Untersuchungen (Nagy, Neumann, et al., 2010) wurden die Analysen zur Mathematikleistung auf Basis von multiple imputierten Werten durchgeführt. In allen Analysen wurden GewichtungsvARIABLEN berücksichtigt.

Ergebnisse: Es zeigte sich nach der Reform eine bedeutsame Verringerung des Geschlechtsunterschieds in der Mathematikleistung. Darüber hinaus fanden sich bedeutsame negative Interaktionseffekte zwischen der Reform und dem Geschlecht im Hinblick auf das Selbstkonzept sowie die beiden untersuchten Facetten des Interesses, die eine Vergrößerung der Geschlechtsunterschiede in diesen Bereichen nach der Reform nahelegen.

Donnerstag
12. März

15:30 Uhr

GC 03/149

Chair(s)
Harks, Birgit

3

Lüdmann, Mike; Dimitrova, Vasilena; Boeger, Annette

Zur Relevanz von Zielorientierungen und Selbstwirksamkeitserwartungen für die Anwendung von Lernstrategien im Lehramtsstudium

Im Vortrag soll die Frage behandelt werden, inwieweit Zielorientierungen und Selbstwirksamkeitserwartungen für die Anwendung von adäquaten vs. dysfunktionalen Lernstrategien im Lehramtsstudium relevant sind. Es wird sich hierbei zeigen, dass die Selbstwirksamkeitserwartung im Falle von Annäherungsleistungszielen ein entscheidender Mediator in Bezug auf die Wahl adäquater Lernstrategien ist.

Wichtig ist zunächst, dass man in Bezug auf Lernstrategien zwischen Tiefen- und Oberflächenstrategien differenziert (z.B. Marton & Saljö, 1976, 1984). Bei Oberflächenstrategien steht das Behalten von Informationen mittels Wiederholung im Mittelpunkt. Es erfolgt keine Umstrukturierung und Verknüpfung des Vorwissens mit den neuen Inhalten. Tiefenstrategien wie Elaboration und Organisation führen zu besserem und intensiverem Verständnis durch eine sinnvolle Ordnung des Lernstoffes, eine Reduktion auf das Wesentliche und vielfältige Verknüpfungen mit dem Vorwissen. Adäquate Lernstrategien sind von entscheidender Bedeutung im Studium (Schiefele, Streblow, Ermgassen & Moschner, 2003; Friedrich & Mandl, 2006). Bei prozessnaher und handlungsorientierter Erfassung des Strategieeinsatzes zeigt sich eine deutliche Beziehung zum Lernerfolg (Artelt, 2000; Leutner & Leopold, 2006; Friedrich & Mandl, 2006). Nach gegenwärtiger Forschungslage hängt die dominierende Anwendung von Tiefen- vs. Oberflächenstrategien von motivationalen Faktoren ab, wobei vor allem die Zielorientierungen der Lernenden eine entscheidende Rolle spielen. Bei einer Lernzielorientierung findet verstärkt eine Anwendung von Tiefenstrategien statt (Ames & Archer, 1988; Streblow & Schiefele, 2006), während bei Leistungszielen eher Oberflächenstrategien herangezogen werden (Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002).

Problematisiert werden muss allerdings, dass in diesem Zusammenhang meist keine Differenzierung der Leistungszielorientierung in Annäherungs- und Vermeidungstendenz vorgenommen wird. So bleibt die genaue Beziehung zwischen der motivationalen Orientierung und der Wahl von Lernstrategien unklar. Insbesondere in Bezug auf die Annäherungsleistungszielenorientierung, die in der Forschung ein ambivalentes Bild im Hinblick auf die Leistungsprognose zeigt, bleibt es uneinsichtig, wann sie zur einer Bevorzugung von Tiefenstrategien und wann zu einer Bevorzugung von Oberflächenstrategien führt. Ebenso ist aufgrund der bestehenden Datenlage nicht klar, ob zwischen den verschiedenen Schularten Unterschiede bei der Wahl von Lernstrategien (und ihrer Abhängigkeit von Zielorientierungen und Selbstwirksamkeitserwartungen) bestehen.

Angenommen wird, dass eine Lernzielorientierung einen systematischen positiven Einfluss die Anwendung von Tiefenstrategien hat, während eine Vermeidungsleistungszielorientierung einen systematischen negativen Einfluss die Anwendung von Tiefenstrategien hat (sodass hierdurch die Oberflächenstrategien dominieren). In beiden Fällen sollte diese Beziehung unabhängig von der Ausprägung der Selbstwirksamkeitserwartung bestehen. Eine Annäherungsleistungszielorientierung sollte hingegen in Abhängigkeit von der Ausprägung der Selbstwirksamkeitserwartung die Wahl von Oberflächen- vs. Tiefenstrategien bedingen. In Bezug auf die Differenzen bezüglich der Anwendung von Lernstrategien zwischen den Studierenden der verschiedenen Schularten wird angenommen, dass zwischen den Oberflächenstrategien keine Unterschiede bestehen, da das Wiederholen der gelernten Inhalte im Lernprozesse meist unverzichtbar ist. Dagegen wird postuliert, dass bei der Anwendung von Tiefenstrategien folgendes Gefälle zwischen den Schularten besteht: HRGe < GS < BK < GyGe. In Bezug auf die Ausprägung der Selbstwirksamkeitserwartung wird angenommen: GS < HRGe < BK < GyGe.

Als Stichprobe diente eine repräsentative Querschnittsuntersuchung der Erstsemester-Lehramtsstudierenden der Universitäten Essen und Aachen, die im Rahmen eines Monitoring-Programms (n=1004) erhoben wurde. Als Instrumente wurden verwendet: (1) Zielorientierungen: SELLMO (Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002) (2) Lernstrategien: LIST (Wild, Schiefele & Winter, 1992; Kurzform von Mayr, 1998) (3) Skalen zur Allgemeinen Selbstwirksamkeitserwartung (Jerusalem & Schwarzer, 1999).

Mittels korrelativer und varianzanalytischer Überprüfung konnten die aufgestellten Hypothesen weitestgehend bestätigt werden. Wie erwartet gibt es einen großen Einfluss der Lernzielorientierung vs. Vermeidungsleistungszielorientierung auf die Anwendung von Tiefenstrategien. Die Ausprägung der Annäherungsleistungszielorientierung zeigt zunächst das erwartete ambivalente Bild, indem es sowohl einen Einfluss auf die Anwendung von Oberflächenstrategien als auch auf die Anwendung von Tiefenstrategien gibt. Es zeigt sich in Bezug auf Oberflächenstrategien jedoch eine Wechselwirkung mit der Selbstwirksamkeitserwartung, sodass diese offenbar eine Mediatorfunktion innehat. Oberflächenstrategien werden erwartungsgemäß von allen Lehramtsstudierenden gleichermaßen angewandt; nur in Bezug auf Tiefenstrategien heben sich die GyGe-Studierenden (und BK-Studierenden) von den HRGe-Studierenden ab. Des Weiteren haben GyGe-Studierende eine höhere Selbstwirksamkeitserwartung als GS-Studierende.

Freitag
13. März

9:00 Uhr

GA 03/49

Chair(s)
Lüdmann, Mike

Merk, Samuel

Epistemologische Überzeugungen als Standards für die Kalibrierung von Lernstrategien im bildungswissenschaftlichen Begleitstudium Lehramtsstudierender

Das bildungswissenschaftliche Studium stellt für Lehramtsstudierende eine wissenschaftstheoretische Herausforderung dar: Im Studium ihrer Fachwissenschaften erhalten sie ausführliche Einführungen in die, dem Fach zugrundeliegende, Epistemologie und erleben diese dann in vielen Veranstaltungen. Im epistemologisch sehr breit gefächerten bildungswissenschaftlichen Begleitstudium dagegen, sollen sie in weit weniger Semesterwochenstunden den State of the Art der (Grundlagen-)Forschung in Anwendungskontexten reflektieren (KMK 2012). Forschungsarbeiten zur Lehrerexpertise legen nahe, dass es Lehramtsstudierenden kaum gelingen dürfte, aus diesen äußerst heterogenen wissenschaftstheoretischen Paradigmen und methodischen Programmen eine kohärente Wissensstruktur aufzubauen, vielmehr sind in sich kohärente, lokale Wissensmodule zu erwarten (Bromme, 1992).

Ein Zugang zu den erkenntnistheoretischen Anteilen dieser Wissensmodule bietet die Forschung zu epistemologischen Überzeugungen, die die Überzeugungen zur Natur wissenschaftlichen Wissens, dessen Genese, Anwendbarkeit und Rechtfertigung beinhalten. Während klassische Konzeptualisierung von einer Entwicklung absolutistischer Überzeugungen (z.B. Wissen sei entweder wahr oder falsch) zu relativistischen Überzeugungen (z.B. Wissen sei immer kontextualisiert) ausgehen, sehen neuere Ansätze epistemologische Überzeugungen erstens als domänen- und zweitens als kontextspezifisch. So wird die Frage nach der Veränderlichkeit mathematischen Wissens von Laien sicher anders beantwortet als die Frage nach der Veränderlichkeit volkswirtschaftlichen Wissens und bildungswissenschaftliches Wissen bzgl. des Gültigkeitsanspruchs in Anbetracht einer Klausur sicher anders, als in Anwendungskontexten bewertet (Muis, Bendixen & Haerle, 2006).

Im Rahmen selbstregulierten Lernens bekommen diese epistemologischen Überzeugungen dann lehr- lerntheoretische Relevanz: Fungieren epistemologische Überzeugungen wie eine Linse durch die die Lernenden eine Aufgabe betrachten (Bromme, Pieschl & Stahl, 2010), setzen sie Standards für die folgende Anpassung der Strategien und Quellen (Kalibrierung), die für eine erfolgreiche Bearbeitung der Aufgabe für notwendig erachtet werden. Dieser „Linsenfunktion“ epistemologischer Überzeugungen und der daraus theoretisch resultierenden Kalibrierung von Strategien und Quellen geht der Beitrag empirisch nach.

Dazu wurden N = 834 Lehramtsstudierenden curricular valide Aufgaben verschiedener Komplexitätsstufen (im Sinne notwendiger kognitiver Operationen) vorgelegt. Zu jeder Aufgabe beantworteten die Studierenden den COPES-Questionnaire mit den Dimensionen superficial processing, multiple perspectives und deep processing (Pieschl, 2009). Die epistemologischen Überzeugungen wurden mit dem Fragebogen zur Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen FREE (Krettenauer, 2005) und den Connotative Aspects of Epistemological Beliefs CAEB (Stahl & Bromme, 2007) erfasst.

Um der Nestung der Daten (Aufgaben in Personen geschachtelt) gerecht zu werden, wurde die Faktorenstruktur des COPES-Questionnaires mithilfe konfirmatorischen Mehrebenenanalysen überprüft (Lüdtke, Trautwein, Schnyder & Niggli, 2007). Die Analyse der prädiktiven Effekte (Kalibrierungsfunktion) epistemologischer Überzeugungen auf die Dimensionen des COPES-Questionnaires erfolgte ebenfalls aufgrund der genesteten Daten anhand hierarchisch linearer Modelle. Hier wurden auf Aufgabenebene (within-person) Dummyvariablen für das Aufgabenniveau und auf der Personenebene (between-person) die epistemologischen Überzeugungen (Dimensionen des FREE und CAEB) als Prädiktoren spezifiziert. Cross-Level-Interaktionsterme, bestehend aus den Dummyvariablen und ausgewählten Dimensionen epistemologischer Überzeugungen stellten die Modellierung der Standardsetzung (Linsenfunktion) dar.

Es zeigten sich sowohl substantielle globale Kalibrierungseffekte über alle drei COPES-Dimensionen (signifikante Level 1 Dummyvariablen der Aufgabenebene) als auch signifikante Cross-Level-Interaktionen des Aufgabenniveaus mit den Dimensionen des CAEBs. Die (Pseudo-) Determinationskoeffizienten weisen auf eine substantielle Varianzaufklärung durch die ausgewählten Prädiktoren hin.

Freitag
13. März

9:00 Uhr

GA 03/49

Chair(s)
Lüdmann, Mike

Wäschle, Kristin; Brauch, Nicola; Lehmann, Thomas; Nückles, Matthias

Unterstützung der selbstgesteuerten Wissensintegration durch Lerntagebücher in der Lehramtsausbildung

Theorie. Reflektierte Entscheidungen im Unterricht erfordern, dass Lehrende auf fachwissenschaftliches, pädagogisches und fachdidaktisches Wissen zurückgreifen können (Shulman, 1986). Typischerweise werden diese Inhalte im Studium unabhängig voneinander erworben und müssen von den Studierenden weitgehend selbständig integriert werden (Darling-Hammond, 2006). Studien zur Integration von Informationen aus verschiedenen Dokumenten zeigen, dass Lernende Strategien präferieren, die zur Entwicklung einer textnahen Repräsentation geeignet sind. Lernende wiederholen die Informationen aus den Texten relativ unverbunden. Textverbindende Elaborations- und Organisationsstrategien, die zur Entwicklung eines integrativen Situationsmodells erforderlich wären, werden hingegen häufig vernachlässigt. Um in Unterrichtssituationen bei didaktischen Entscheidungen auf alle drei Wissensbereiche simultan zugreifen zu können, ist ein integratives Situationsmodell aber erforderlich. Zur Förderung von Integrationsstrategien könnten Lerntagebücher eine hilfreiche Methode sein.

Das Potential von Lerntagebüchern liegt in der Förderung kognitiver und metakognitiver Lernstrategien, welche einen selbstregulierten und nachhaltigen Wissenserwerb begünstigen (Nückles et al., 2009). In Lerntagebüchern werden effektive Lernstrategien gezielt durch Leitfragen und Hinweise (Prompts) unterstützt. Lerntagebücher regen so auch anspruchsvolle Strategien an, die Lernende sonst nicht oder nur sporadisch einsetzen würden.

Im Rahmen unserer Studie haben wir untersucht, welche Lernstrategien angehende Lehrkräfte bei der Integration von fachwissenschaftlichem, pädagogischem und fachdidaktischem Wissen verwenden und inwiefern durch Prompts Integrationsprozesse im Lerntagebuch angeregt werden können. Eine erfolgreiche Integration der rezipierten Inhalte müsste in unterrichtsbezogenen Situationen die Berücksichtigung unterschiedlicher Inhalte erleichtern. Wir gehen daher davon aus, dass angehende Lehrkräfte, die zur Wissensintegration angeregt wurden, eine didaktische Aufgabenstellung ausgewogener unter Berücksichtigung fachwissenschaftlicher, fachdidaktischer und pädagogischer Informationen aus den Texten gestalten können als Studierende ohne Anregung zur Wissensintegration.

Methode. An der experimentellen Studie nahmen 52 Studierende im Lehramt Geschichte teil. Zuerst lasen die Studierenden je einen Text über den Holocaust, über Lernprozesse und über die schulische Vermittlung des Holocaust. Danach verfassten sie über diese drei Texte einen Lerntagebucheintrag. Während die Studierenden in der Kontrollbedingung (n = 27) ihr Lerntagebuch ohne Prompts verfassten, wurden die Studierenden in der Versuchsbedingung (n = 25) durch Prompts dazu angeregt, textübergreifende Organisations- und Elaborationsstrategien zu verwenden. Die Inhalte der Lerntagebücher wurden anhand eines Kategoriensystems analysiert, welches auf Basis eines Kodierschemas für Lerntagebücher (Nückles et al., 2009) und eines Kodierschemas für Prozesse bei der Integration multipler Dokumente (Anmarkrud et al., 2014) erstellt wurde. Die segmentierten Aussagen wurden unter anderem den Kategorien Paraphrasieren (Wiedergeben von Textinformationen) und Integration (Verknüpfung von Informationen aus verschiedenen Texten) zugewiesen. Abschließend wurden die Studierenden gebeten, eine Lernaufgabe für Schüler zum Thema Holocaust zu entwickeln und ihre dabei angestellten didaktischen Überlegungen zu erläutern. Bei der Analyse dieser Aufgabe wurden die Aussagen der Studierenden nach den Wissensdomänen (Fachwissenschaft, Pädagogik, Fachdidaktik) kategorisiert. Die Häufigkeit der Textbezüge gibt an, inwiefern die Studierenden einen bestimmten Text fokussierten oder alle Texte gleichermaßen berücksichtigten.

Ergebnisse. T-Tests zeigten, dass Studierende der Prompts-Gruppe weniger paraphrasierten als Studierende der Kontrollgruppe. Die Studierenden der Prompts-Gruppe wählten signifikant häufiger Integrationsstrategien, bei denen sie Informationen aus mehreren Texten miteinander verknüpften, als die Studierenden der Kontrollgruppe. Weiterhin zeigte sich, dass sich die Studierenden der Kontrollgruppe bei der Gestaltung ihrer Lernaufgabe vor allem auf den fachwissenschaftlichen Text bezogen. Pädagogische und fachdidaktische Inhalte wurden vernachlässigt. Die Studierenden der Prompts-Gruppe hingegen berücksichtigten alle drei Texte in ähnlichem Maße, und thematisierten neben den geschichtswissenschaftlichen Gesichtspunkten beispielsweise auch Lernprozesse ihrer potenziellen Schüler.

Schlussfolgerung. In vielen Studiengängen liegt es meist in der Verantwortung der Lehramtsstudierenden, fachliche, pädagogische sowie fachdidaktische Studieninhalte zu verknüpfen (Darling-Hammond, 2006). Die aktuelle Studie zeigt, dass sie dabei auf Strategien zurückgreifen, welche zu unverbundenen Repräsentationen führen können, die in Lehrsituationen nicht simultan verwendet werden können. Lerntagebücher, welche die Integration interdisziplinärer Lerninhalte durch Prompts gezielt anregen, sind offenbar hilfreich, um Lehramtsstudierende bei der Entwicklung integrativer und flexibel anwendbarer Wissensschemata wirksam zu unterstützen, indem sie die Integration fachwissenschaftlicher, fachdidaktischer und pädagogischer Inhalte fördern.

Freitag
13. März

9:00 Uhr

GA 03/49

Chair(s)
Lüdmann, Mike

Löffler, Simone N.; Bohner, Antonia; Stumpp, Jürgen; Gidion, Gerd

Förderung selbstregulierten Lernens im ingenieurwissenschaftlichen Studium durch ein professionsspezifisches Tutorial und ein interaktives Feedbacksystem

Das akademische Lernen zeichnet sich durch einen hohen Grad an Eigenverantwortlichkeit aus. Somit stellt das selbstregulierte Lernen eine wichtige Schlüsselkompetenz für Studierende aller Fachbereiche dar. Da die Entwicklung von Schlüsselkompetenzen in beruflichen und hochschulischen Bildungsgängen seit dem Bologna-Prozess vermehrt als Ziel ausgewiesen wird, kommt mittlerweile auch vermehrt die Frage nach angemessenen Fördermöglichkeiten für den Hochschulkontext auf. Zur Förderung der Selbstregulation existiert eine Vielzahl von Programmen. Die bestehende Datenlage spricht zwar für positive Effekte von Selbstmanagement-Trainings auf die Arbeitsleistung. Kritisch wird jedoch angemerkt, dass Kontextvariablen sowie die Umsetzung von Zieldefinitionen in konkrete Handlungen auf täglicher Basis oft keine Berücksichtigung finden (Claessens et al., 2009). Ein häufig genannter Kritikpunkt an der nicht-integrativen Förderung von Schlüsselqualifikationen im akademischen Kontext bezieht sich auf den fehlenden Fachbezug bei der Vermittlung überfachlicher Qualifikationen (Dörig, 1996; Arnold & Müller, 2002; Brassard et al., 1994). Beide Aspekte sollten bei der Entwicklung des vorliegenden Interventionsansatzes Berücksichtigung finden. Einerseits sollte der Einsatz elektronischer Tagebücher eine zeitnahe Erfassung des Lernverhaltens im Alltag gewährleisten. Andererseits sollte durch die Verwendung fachspezifischer Anwendungsbeispiele eine Verknüpfung mit spezifischen Fachinhalten erreicht werden (professionsspezifische Förderung).

Zielsetzung der vorliegenden Studie war die Evaluierung eines technikunterstützten Ansatzes zur Förderung selbstregulierter Lernprozesse bestehend aus einem professionsspezifischen Tutorial für den Fachbereich Maschinenbau, das in Kombination mit einem interaktiven Feedbacksystem (für die Begleitung der Lernprozesse im Alltag) eingesetzt wurde. Geprüft werden sollten Effekte auf Strategieeinsatz, kognitive Ressourcen und Lernerfolg.

Das Interventionsprogramm wurde im Rahmen des Projekts „Professing - Nachweisliche Förderung von professionsspezifischer Sozial- und Selbstkompetenz im ingenieurwissenschaftlichen Studium“ (gefördert durch das Ministerium für Wissenschaft, Forschung und Kunst, Baden-Württemberg) von einem interdisziplinär aufgestellten Projektteam aus den Fachbereichen Psychologie, Pädagogik (Technikdidaktik), Maschinenbau und Elektrotechnik entwickelt. Theoretische Grundlage der Interventionsmaßnahmen bildete das Prozessmodell von Schmitz und Wiese (2006) als Weiterentwicklung des zyklischen Modells von Zimmermann (2002).

80 Studierende (2. Semester, Maschinenbau) nahmen im Rahmen einer Lehrveranstaltung zum Erwerb von Schlüsselqualifikationen im Sommersemester 2014 an der Evaluationsstudie teil und wurden randomisiert zwei Gruppen (Kontroll- / Interventionsgruppe) zugeordnet. Alle Studierende absolvierten eine erste Assessmentphase (14 Tage) während der Vorbereitungszeit auf eine Klausur („Höhere Mathematik II“). In dieser Phase beantworteten die Studierenden mittels elektronischer Tagebücher vor und nach jeder Lernphase Fragen. Die Items bezogen sich auf situative Bedingungen, Planung der Lernphasen, kognitive Lernstrategien, metakognitive Strategien, kognitive Ressourcen und den subjektiven Lernerfolg. Die Kontrollgruppe absolvierte in der Vorbereitungsphase (14 Tage) auf eine zweite (hinsichtlich Schweregrad und Frageformat vergleichbare) Klausur („Technische Mechanik II“) das gleiche Assessment wie in der ersten Phase. Für die Interventionsgruppe erfüllte das Monitoringsystem in der Vorbereitungsphase auf die zweite Klausur zusätzlich eine intervenierende Funktion, indem das selbstregulierte Lernen durch automatisiertes Feedback zum individuellen Strategieeinsatz und durch die Festlegung von persönlichen Lernzielen sowie durch die Überprüfung der subjektiven Zielerreichung gefördert werden sollte. Zusätzlich erhielt die Interventionsgruppe ein audio-visuelles Tutorial, in dem verschiedene Elaborations-, Organisations- und Memorierungstechniken vorgestellt und anhand professionsspezifischer Beispiele erläutert wurden. Das Tutorial wurde nach Abschluss der ersten Assessmentphase bearbeitet, sodass die Studierenden die neu erworbenen Strategien erst in der zweiten Lernphase nutzen konnten. Die Kontrollgruppe erhielt erst nach Abschluss der Datenerhebung Zugriff auf das Tutorial (Wartekontrollgruppendesign).

Eine erste deskriptive Datenanalyse deutet auf positive Auswirkungen der Intervention auf den Einsatz der meta-kognitiven Strategien sowie den Lernerfolg und die effektive Lernzeit hin. Hinsichtlich der Anwendung kognitiver Lernstrategien zeigten sich zwischen Kontroll- und Interventionsgruppe in den beiden Assessmentphasen keine Unterschiede (im Vergleich der Mittelwerte über alle Lerneinheiten). In einem nächsten Schritt sind jedoch mittels Mehrebenenanalysen intra- und interindividuelle zeitliche Verläufe über die einzelnen Registriertage zu berücksichtigen, die dann im Rahmen der Tagung präsentiert werden könnten.

Chair(s)

Lüdmann, Mike

Jöstl, Gregor; Klug, Julia; Schober, Barbara; Spiel, Christiane

Wie viel Wissen haben Studierende über SRL und warum wenden sie dieses nicht an? - Ein Ansatz zur Aufklärung widersprüchlicher Befunde

Selbstreguliertes Lernen (SRL) bedeutet „[...] selbstständig einen Lernprozess zu planen, auszuführen und zu evaluieren. Das beinhaltet Entscheidungen in Bezug auf kognitive, motivationale und Verhaltensaspekte des zyklischen Lernprozesses“ (Wirth & Leutner, 2008; S. 103). SRL wird unumstritten als relevant für langfristigen Lernerfolg gesehen (u.a. Streblo & Schiefele, 2006; Dignath & Büttner, 2008). Die bisherige Forschung zu SRL ist umfangreich, vielfältig, heterogen und zum Teil sogar widersprüchlich (Puustinen & Pulkkinen, 2001; Sitzmann & Ely, 2011), insbesondere im Hinblick auf die Korrelation mit Leistung. In manchen Studien findet man keinen oder nur einen geringen Zusammenhang zwischen SRL und Leistung (u.a. Schiefele, Streblo, Ermgassen & Moschner 2003), in anderen einen geringen bis mittleren Zusammenhang, der aber zwischen den einzelnen Strategien variiert (Pintrich & DeGroot, 1990; $r=.18$ für SRL; $r=-.18$ für kognitive Strategien). Studien, in denen SRL trainiert wird, zeigen wiederum Effektstärken bis $.69$ (Dignath & Büttner, 2008). Worin können solche heterogenen Befunde begründet sein? Einen möglichen Erklärungsansatz liefert die ebenso heterogene Art der Untersuchung in den zahlreichen Studien. Vielfältige Variablen des SRL (u.a. Veenman, Van Hout-Wolters & Afflerbach, 2006) werden in verschiedenen Kontexten (u.a. Boekaerts, 2007), auf der Basis unterschiedlicher Modelle (u.a. Schmitz, 2003) erfasst. Vor diesem Hintergrund erscheinen heterogene Ergebnisse plausibel. Ein weiterer Erklärungsansatz kann jedoch in der Art der Messung begründet liegen. Häufig wird mittels Selbsteinschätzungen die SRL Fähigkeit als trait erhoben. Im Sinne eines Produktions- oder Nutzungsdefizites (Miller, 1990; Hasselhorn, 1995) genügt es jedoch nicht, die eigenen SRL Fähigkeiten hoch einzuschätzen. Darüber hinaus sollten Wissen über die Nützlichkeit der Strategien und die Fähigkeit, die Strategien effizient einzusetzen, vorhanden sein. Uns fehlt allerdings bisher ein differenziertes Bild über das Wissen bzgl. SRL und insbesondere darüber, ob die Strategien auch tatsächlich angewendet werden. Durch die Frage nach der tatsächlichen Anwendung im Alltag (vgl. Dupeyrat & Mariné, 2005) könnten wir aufklären, ob die heterogenen Befunde zur Korrelation zwischen SRL und Leistung durch die Nicht-Beachtung des Produktions- und Nutzungsdefizits entstanden sein können. An diesem Desiderat setzen unsere Fragestellung sowie unser Messansatz an. In diesem Beitrag möchten wir klären, (1) wie viel Wissen Studierende über SRL haben, (2) ob sie dieses Wissen auch anwenden, (3) wie groß die Diskrepanzen zwischen Wissen und Anwendung sind, und (4) warum es solche Diskrepanzen gibt. Weiters wird untersucht (5), ob es bezüglich dieser Aspekte Unterschiede zwischen Studierenden verschiedener Studiengängen gibt und (6) inwieweit Zusammenhänge zwischen Wissen, Anwendung, Diskrepanzen und subjektiven wie objektiven Leistungsmaßen bestehen.

Zur Beantwortung dieser Forschungsfragen bearbeitet eine Stichprobe von Wirtschafts- und Psychologiestudierenden ($n=200$) in einer momentan laufenden Erhebung ein innovatives Instrument zur Erfassung von deklarativem, prozeduralem und konditionalem Wissen über günstige SRL-Strategien in relevanten Lernsituationen (ermittelt in zwei Vorstudien, $n=108$ bzw. $n=470$) sowie deren tatsächlicher Anwendung. Bei Diskrepanzen zwischen Wissen und Anwendung wird zusätzlich nach den Gründen für die Nicht-Anwendung gefragt. Außerdem erfolgt eine subjektive und objektive Leistungseinschätzung. Ausgehend von den Ergebnissen einer weiteren Vorstudie ($n=96$ Psychologiestudierende) erwarten wir hohes Wissen auf der deklarativen, etwas geringeres auf der prozeduralen und das geringste Wissen auf der konditionalen Ebene sowie erhebliche Diskrepanzen zwischen Wissen und Anwendung auf allen drei Ebenen, wobei bei Psychologiestudierenden ob der Studieninhalte (insbesondere auf deklarativer Ebene) höheres Wissen erwartet wird. Der Zusammenhang mit Leistungsmaßen wird von deklarativem hin zu konditionalem Wissen als stärker werdend erwartet. Den stärksten Zusammenhang mit Leistungsmaßen erwarten wir bei der tatsächlichen Anwendung von SRL-Strategien.

Durch die systematische Betrachtung der Diskrepanzen zwischen Wissen über und Anwendung von SRL-Strategien sollen widersprüchliche Befunde zu deren Wirksamkeit aufgeklärt werden. Weiters kann das Wissen über die Gründe für deren Nicht-Anwendung als Grundlage für mögliche Förderprogramme dienen.

Schwenk, Christin; Kuhn, Jörg-Tobias; Dehmelt, Vera; Raddatz, Julia; Holling, Heinz; Dobel, Christian

Basale Mengenverarbeitung bei Kindern mit Dyskalkulie - neuronale Korrelate und deren Prädiktivität für die Effektivität eines basisnumerischen Trainings

Dyskalkulie beschreibt eine neurokognitive Lernstörung mit einer Prävalenz von ca. 6%, die nicht nur die schulische Entwicklung und den Alltag betroffener Grundschul Kinder beeinflusst, sondern auch ihre langfristigen Chancen auf dem Arbeitsmarkt sowie ihre psychische und physische Gesundheit. Eine wichtige Aufgabe interdisziplinärer Bildungsforschung besteht deshalb darin, die Ursachen dieser mathematischen Entwicklungsstörung festzustellen. Hierfür werden zunehmend auch neurowissenschaftliche Methoden genutzt. Dadurch gewonnene Daten gewähren Einsicht in basale neuronale Verarbeitungsprozesse, welche der verhaltenmäßigen Ausprägung von Dyskalkulie-Symptomen vorausgehen. Solche Prozesse bilden erstens das Fundament in Entwicklungsmodellen numerischer Kompetenzen (z.B. Butterworth, Varma & Laurillard, 2011; von Aster & Shalev, 2007). Aus einer Förderperspektive können neuronale Befunde zweitens dabei helfen, effektive Trainings für Kinder mit besonderen Schwierigkeiten zu entwerfen, zu evaluieren und potentielle Non-Responder frühzeitig zu identifizieren.

Ein wichtiger Meilenstein in der mathematischen Entwicklung, und gleichzeitig etablierter Kernindikator für Dyskalkulie, ist die biologisch angelegte Fähigkeit, Mengen wahrzunehmen und zu vergleichen. Das Triple-Code-Modell (z.B. Dehaene & Cohen, 1995) integriert diese basisnumerische Kompetenz in einen neurofunktionalen Atlas der Zahlenverarbeitung. Es werden drei miteinander verknüpfte Module postuliert, die durch verschiedene Gehirnregionen gesteuert werden. Die symbolischen Kodierungen von Zahlen als Wort (z.B. „vier“) und Ziffer („4“) werden durch zwei verschiedene Module verarbeitet. Durch den Austausch mit einem dritten Modul können sie in einen analogen Code übersetzt werden, d.h. in eine nicht-symbolische Mengensemantik. Als robuster psychophysikalischer Indikator der Mengenrepräsentation dient der numerische Distanzeffekt (NDE), welcher besagt, dass der Größenvergleich zweier Mengen umso länger dauert, je kleiner die Distanz zwischen ihnen ist. Weil sich beispielweise die Quantitäten „vier“ und „fünf“ auf dem mentalen Zahlenstrahl stärker überlappen als „fünf“ und „neun“, erfordert der erste Vergleich mehr kognitive Ressourcen. Traditionell geht das Triple-Code-Modell davon aus, dass die abstrakte Mengenrepräsentation durch die bilateralen intraparietalen Sulci (IPS) erfolgt. Eine Beeinträchtigung der IPS wird deshalb als ursächlich für Dyskalkulie angenommen und einige bildgebende Studien finden eine atypische Modulation dieser Areale bei dyskalkulischen Kindern (Mussolin, De Volder, Grandin, Schölgel, Nassogne, & Noël, 2010a; Price, Holloway, Räsänen, Vesterinen & Ansari, 2007).

Insgesamt ist die Forschungslage jedoch heterogen und gerade neuere Befunde sprechen gegen ein distinktes abstraktes Zahlenmodul in den IPS und für ein breiteres neurofunktionales Netzwerk, das auch frontale, temporale und okzipitale Bereiche umfasst (Bulthé, De Smedt & Op de Beeck, 2014). Auch in der Vorhersage mathematischer Trainingserfolge dyskalkulischer Grundschul Kinder gibt es erste Hinweise auf die Prädiktivität neuronaler Strukturen jenseits der IPS (Supekar et al., 2013).

Vor diesem Hintergrund verfolgte die vorliegende Studie zwei Ziele: Auf Grundlage symbolischer Ziffernvergleichsaufgaben und mittels magnetoenzephalographischer Messungen wurde erstens der behaviorale und neuronale NDE bei Zweit- bis Viertklässlern mit Dyskalkulie (N=34), im Vergleich zu unbeeinträchtigten Gleichaltrigen (N=31), erhoben: Bezüglich Reaktionszeiten waren die Dyskalkuliker langsamer und zeigten einen signifikant stärkeren NDE als ihre nicht-dyskalkulischen Peers. Neuronal ergaben sich für die dyskalkulische Gruppe acht teilweise überlappende distanzsensitive Regionen im frontalen, temporalen und parietalen Kortex. Die Aktivierungsmuster unterschieden sich von denen der unbeeinträchtigten Gruppe hauptsächlich hinsichtlich der Lateralisierung. Zusammenfassend deuten unsere Befunde darauf hin, dass Mengenverarbeitung bei Kindern mit und ohne Dyskalkulie von multiplen funktionalen Hirnregionen abhängt.

Zweitens erhielten 17 dyskalkulische Kinder ein dreißigtägiges intensives computerbasiertes Training mit basisnumerischem Fokus (Kuhn & Holling, 2014). Für diese Subgruppe wurde untersucht, inwieweit der individuelle neuronale NDE vor Beginn der Intervention die Leistung nach dem Training vorhersagen kann. Die mathematischen Kompetenzen wurden mit einer Screening-Batterie erhoben, die sowohl basisnumerische Verarbeitung als auch Kopfrechnen umfasst. Infolge des Trainings verbesserten sich die dyskalkulischen Kinder in beiden Bereichen. Schrittweise Regressionsanalysen ergaben, dass vor allem der prä-Training NDE in (prä)frontalen und (links-)temporalen Arealen die mathematische Leistung nach dem Training vorhersagen kann. Diese domänenübergreifenden Strukturen sind verantwortlich für Arbeitsgedächtnis, Updating und Monitoring sowie Lernen und semantisches Gedächtnis im Allgemeinen.

Freitag
13. März

9:00 Uhr

GA 03/142

Chair(s)

Schwenk, Christin

Kollar, Ingo; Vogel, Freydis; Fischer, Frank; Ufer, Stefan; Reichersdorfer, Elisabeth; Reiss, Kristina

Adaptierbare Argumentationsskripts: Lieber nur für Lernende mit günstigen Selbststeuerungsfähigkeiten?

Theoretischer Hintergrund: Beim kooperativen Lösen von mathematischen Beweisaufgaben (z.B. in Dyaden) kommt dem gemeinsamen Argumentieren eine zentrale Rolle zu. Hochwertige Argumentationsprozesse finden allerdings nur selten spontan statt, sondern müssen i.d.R. instruktional angeleitet werden, z.B. durch argumentationsbezogene Kooperationsskripts (Noroozi, Weinberger, Biemans, Mulder & Chizari, 2013). Jedoch ist die Effektivität von Argumentationsskripts häufig von individuellen Lernvoraussetzungen abhängig (Kollar et al., 2012). Während fortgeschrittene Lernende möglicherweise besonders von Argumentationsskripts profitieren, die relativ grobe Vorgaben zu den während der Argumentation zu durchlaufenden Phasen machen (z.B. „Verfasse dein Argument!“, „Kritisiere das Argument deines Partners!“), benötigen Novizen vermutlich eher stärker strukturierte Kooperationsskripts, die zusätzlich Vorgaben hinsichtlich während dieser Phasen zu zeigender Aktivitäten machen (z. B. „Begründe dein Argument!“, „Gib an, warum dein Argument für deine Behauptung spricht!“). Erhalten Lernende „unpassende“ Skripts, können Probleme wie „Overscripting“ (Dillenbourg, 2002) oder ein zu geringes Argumentationsniveau (z.B. Stegmann, Weinberger & Fischer, 2007) entstehen. Um unterschiedlichen Lernvoraussetzungen gerecht zu werden, können Argumentationsskripts allerdings adaptierbar gestaltet werden, indem den Lernenden die Möglichkeit gegeben wird, den Grad der Strukturierung der Argumentationsskripts anzupassen. Jedoch erfordern adaptierbare Argumentationsskripts, dass die Lernenden in hinreichendem Maße über die Fähigkeit zur Selbststeuerung verfügen, d.h. die Kompetenz besitzen, das eigene Lernen evaluieren und planen zu können (z.B. Zimmermann & Martinez-Pons, 1990).

Fragestellung: (1) Inwiefern haben der Grad der Strukturierung und die Adaptierbarkeit eines Argumentationsskripts (grob strukturiertes vs. fein strukturiertes vs. adaptierbares Argumentationsskript) einen Einfluss auf den Erwerb von Argumentationsfertigkeiten beim gemeinsamen Lösen mathematischer Beweisaufgaben? (2) Inwiefern beeinflussen die Selbststeuerungsfähigkeiten der Lernenden ihren Erwerb von Argumentationsfertigkeiten beim Lernen mit adaptierbaren und nicht adaptierbaren Argumentationsskripts?

Methode: Die Studie wurde im Rahmen eines zweiwöchigen Vorbereitungskurses für angehende Mathematikstudierende durchgeführt. Die N=111 Teilnehmer wurden randomisiert einer der drei Bedingungen „grob“ bzw. „fein strukturiertes Argumentationsskript“ und „adaptierbares Argumentationsskript“ zugewiesen. In drei Treatmentsitzungen arbeiteten die Teilnehmer in Dyaden (mit wechselnden Lernpartnern) an mathematischen Beweisaufgaben (z.B. „Summiere ungerade Anzahlen von aufeinanderfolgenden Zahlen und versuche Regelmäßigkeiten zu finden. Formuliere eine Vermutung und beweise diese.“). Das jeweilige Argumentationsskript wurde auf zwei verbundenen Laptops angeboten. Das grob strukturierte Argumentationsskript unterteilte die Diskussion der Lernpartner in die Phasen Argument, Gegenargument und Synthese (Leitão, 2000). Das fein strukturierte Argumentationsskript enthielt zusätzlich Hinweise, wie die Argumentation innerhalb der einzelnen Phasen aufgebaut werden soll (Toulmin, 1958). In dem adaptierbaren Argumentationsskript konnten die Lernenden an verschiedenen Stellen wählen, ob sie bei der folgenden Aufgabe das grob oder das fein strukturierte Argumentationsskript erhalten wollten. Zur Messung der Argumentationsfertigkeiten wurde ein Wissenstest vor und nach der Intervention eingesetzt, der anschließend von zwei trainierten Beurteilern ausgewertet wurde (Cohens $\kappa=.82$). Zur Messung der Selbststeuerungsfähigkeit wurde eine Self-Report-Skala aus elf Fragebogen-Items (Cronbachs $\alpha=.72$) verwendet (vgl. Fisher, King & Tague, 2001).

Ergebnisse: Eine ANOVA mit Messwiederholung zeigte einen signifikanten Zuwachs der Argumentationsfertigkeiten aller Lernenden zwischen Prä- und Posttest ($F(1,106)=12.83$, $p<.01$, $\eta^2=.11$). Zwischen den drei Bedingungen unterschieden sich die Lernenden nicht signifikant in ihren erworbenen Argumentationsfertigkeiten ($F(2,106)=1.49$, $p=.23$, $\eta^2=.03$). Folglich lieferten offenbar alle drei Argumentationsskripts hilfreiche Vorgaben zum gemeinsamen Argumentieren. Separate Regressionsanalysen ergaben, dass die individuelle Selbststeuerungsfähigkeit weder beim Lernen mit dem grob strukturierten Argumentationsskript ($b=0.04$, $t(31)=0.11$, $p=.91$, $R^2<.01$) noch mit dem fein strukturierten Skript ($b=0.09$, $t(39)=0.33$, $p=.74$, $R^2<.01$) ein signifikanter Prädiktor für den Erwerb von Argumentationsfertigkeiten war. In der adaptierbaren Skriptbedingung zeigte sich allerdings ein signifikanter und positiver Einfluss der Selbststeuerungsfähigkeiten der Lernenden auf den Wissenserwerb ($b=0.91$, $t(34)=3.126$, $p<.01$, $R^2=.24$). Die Ergebnisse deuten darauf hin, dass die Vorgabe adaptierbarer Argumentationsskripts zumindest für Lernende mit günstigen Selbststeuerungsfähigkeiten eine gute Alternative zur Vorgabe starrer Skripts darstellen kann (Fischer et al., 2013). Um adaptierbare Argumentationsskripts bei Lernenden mit weniger günstigen Selbststeuerungsfähigkeiten effektiv einzusetzen, wäre jedoch ein Angebot zur Unterstützung oder Übernahme der Adaptierung sinnvoll (z.B. Dziol, Walker, Rummel & Koedinger, 2010).

Freitag
13. März

9:00 Uhr

GA 03/142

Chair(s)

Schwenk, Christin

Stemmann, Jennifer; Lang, Martin

Der Einfluss des Vorwissens auf das Explorationsverhalten und die Leistung beim Lösen technischer Probleme

Die Annahme eines Konstrukts allgemeiner (dynamischer) Problemlösekompetenz und das Ziel eine solche computerbasiert zu erfassen, führt zur Konstruktion vermeintlich vorwissensneutraler Problemstellungen, um zu verhindern, dass Probanden mit domänenspezifischen Vorwissen einen Vorteil haben (Greiff, Fischer, 2013). Aber schon Wirth (2004) bezweifelt, dass derart vorwissensfreie Systeme überhaupt existieren können, da bereits die verwendeten Symbole, die räumliche Anordnung von Informationen oder besonders die Untersuchungsumgebung selbst Vorwissen aktivieren können. Ist kein Vorwissen in Bezug auf das zu lösende Problem vorhanden, muss eine Person dieses Defizit ausgleichen und das benötigte Wissen im Prozess des Problemlösens zunächst explorativ erwerben. Je (fach-)spezifischer und umfangreicher das erforderliche Wissen ist, desto schwieriger ist es vermutlich, sich dieses Wissen erst während der Problembearbeitung anzueignen. In experimentellen Untersuchungen, in denen außerdem die zur Verfügung stehende Bearbeitungszeit limitiert ist, führt ein fehlendes fachspezifisches Vorwissen zu einer geringen Problemlöseleistung. Befunde aus Untersuchungen mit Kfz-Mechatronikern und Elektronikern bestätigen Fachwissen als stärksten Prädiktor für fachspezifische Problemlöseleistung (Abele et al., 2012). Problemstellungen, die zur Lösung Wissen voraussetzen, das wenig bereichsspezifisch und demzufolge mehr alltagsbezogen ist, erlauben größere Vorwissensdefizite. Der Erfolg bei der Akquise fehlenden Wissens im Problemlöseprozess hängt hier hingegen von der Verwendung optimaler Strategien des Wissenserwerbs ab (van der Linden et al., 2001). In Bezug auf Problemstellungen, die die Nutzung unbekannter technischer Geräte zum Inhalt haben (z. B. Stemmann, Lang, 2014), liegt der Fokus auf einer Anwendung systematischer Explorationsstrategien. Der Einfluss, den das Explorationsverhalten auf den Wissenserwerb und damit auf die Wissensanwendung beim Problemlösen im Umgang mit computerbasierten Systemen hat, belegen bereits mehrere Studien (z. B. Wüstenberg et al., 2014, Kröner, Leutner, 2002, van der Linden et al., 2001).

Der positive Einfluss des Vorwissens auf die Problemlöseleistung sowie der des Explorationsverhaltens auf den weiteren Wissenserwerb sind unstrittig. Wie sieht es jedoch mit der Wirkung des Vorwissens auf die Vorgehensweise beim Explorieren aus? Hierzu gibt es in der Literatur unterschiedliche (sich nicht ausschließende) Befunde. Einerseits kann Vorwissen einen positiven Effekt auf die Herangehensweise haben. Wer über kein relevantes Vorwissen verfügt, ist bei der Exploration eines Systems zunächst auf Versuch und Irrtum angewiesen, während ein vorhandenes Vorwissen bereits zu Beginn eine systematischere Herangehensweise ermöglicht (Baumert et al., 2003). Süß, Kersting und Oberauer (1993) zeigen jedoch, dass dieser Effekt nur zu Beginn des Wissenserwerbs besteht. Je mehr Wissen während der Exploration erworben wird, desto geringer wird der Einfluss des Vorwissens auf den Lösungsprozess. Dies dürfte aber wohl nur für die Fälle gelten, in denen Defizite im Vorwissen durch in der Exploration erworbenen Wissens ausgeglichen werden können, also nicht für stark fachspezifische Probleme.

Andererseits kann Vorwissen aber auch den weiteren Wissenserwerbsprozess behindern, wenn etwa die neuen Informationen mit dem bestehenden Wissen unvereinbar sind (Wirth, 2004, Buchner, Funke, 1991). In diesem Beitrag überprüfen wir die Annahmen, (1) nach der Probanden mit einem größeren Vorwissen ein unbekanntes technisches System sorgfältiger und systematischer explorieren und (2) dass ein solches Explorationsverhalten zu einer höheren Problemlöseleistung führt.

Zur Beantwortung der Fragestellung werden die Ergebnisse und Logfiledaten aus einem Test, bestehend aus 18 simulierten technischen Geräten auf Basis finiter Automaten sowie einem Fragebogen zur Vorwissenserfassung analysiert (N=28 Teilnehmer einer Kraftwerksmeisterfortbildung). In dieser zur Pilotierung des Testinstrumentes ausgelegten Studie, hatten die Probanden im Anschluss an eine Explorationsphase die simulierten Geräte nach entsprechender Instruktion zu steuern. Die Analyse der Logfile-Daten aus der Explorationsphase erfolgt nach den Kriterien: Vollständigkeit der Systemexploration, Bearbeitungszeiten und Identifikation unterschiedlicher Explorationsstrategien. Die Bewertung der Problemlöseleistungen berücksichtigt sowohl Steuerungsleistungen als auch die benötigte Anzahl an Bearbeitungsschritten.

Der gegenwärtige Analysestand zeigt entgegen anfänglicher Vermutungen keinen direkten Einfluss des Vorwissen auf die Problemlöseleistung. Ob es jedoch einen indirekten Einfluss über das Explorationsverhalten ausübt, wird gegenwärtig untersucht; die Ergebnisse und Herausforderungen des Analyseprozesses sind Inhalt des Vortrags.

Freitag
13. März

9:00 Uhr

GA 03/142

Chair(s)
Schwenk, Christin

3

Rösselet, Stephan; Neuenschwander, Markus P.

Veränderungen des Fähigkeitsselbstkonzepts in Deutsch und Mathematik beim Übertritt in die Sekundarstufe I

Das akademische Fähigkeitsselbstkonzept hängt nicht nur von der Bewertung der eigenen Leistung, sondern auch von den Ergebnissen sozialer Vergleichsprozesse mit Bezugsgruppen ab. Während die eigene Leistung in einem positiven Zusammenhang mit dem Fähigkeitsselbstkonzept steht, ist die Leistung der Klasse, in der sich ein Kind befindet, negativ mit dem Fähigkeitsselbstkonzept verbunden. Ausgehend von der Forschung zum sogenannten Fischteicheffekt (Big-Fish-Little-Pond-Effekt, Marsh 1987, 2005) wird die Veränderung des Fähigkeitsselbstkonzepts Deutsch und des Fähigkeitsselbstkonzepts Mathematik vom 5. ins 6. Schuljahr untersucht. Es werden die Fragen untersucht, wie sich soziale und temporale Vergleiche (vgl. Köller, Trautwein, Lüdtke & Baumert, 2006) zwischen Leistung und Klassenleistung auf die Veränderung des Fähigkeitsselbstkonzepts in Mathematik und Deutsch vom 5. in das 6. Schuljahr auswirken. Ausserdem interessiert, ob sich diese Vergleiche unterscheiden, wenn die Kinder in der Primarstufe verbleiben oder im gleichen Alter in die Sekundarstufe I wechseln (Moderation). Wir vermuten, dass sich beim Verbleib der Kinder in der Primarstufe soziale und temporale Vergleiche überlappen, während mit dem Wechsel der Bezugsgruppe beim Übertritt in die Sekundarstufe I der Vergleich mit den eigenen früheren Leistung und der früheren Klassenleistung an Bedeutung verliert.

Unsere Hypothesen wurden anhand einer Längsschnittstichprobe mit Schülerinnen und Schülern aus vier Schweizer Kantonen getestet. Dabei werden zwei Transitionsgruppen von Schülerinnen und Schülern miteinander verglichen. Die erste Gruppe bilden 502 Schülerinnen und Schüler, die in diesen Schuljahren in der Primarstufe verbleiben (Kantone Bern und Luzern), und die Vergleichsgruppe 345 Schülerinnen und Schüler, die nach dem 5. Schuljahr in die gegliederte Sekundarstufe I übertreten (Kantone Aargau und Basel-Landschaft).

Den Schülerinnen und Schülern wurden im 5. und 6. Schuljahr Leistungstests in Deutsch und Mathematik vorgelegt, die nach der Item-Response-Theorie ausgewertet wurden. Zusätzlich beantworteten sie in einem Schülerfragebogen reliable Items zu den Fähigkeitsselbstkonzepten in Deutsch und Mathematik, die mittels konfirmatorischer Faktorenanalysen zusammengefasst wurden.

Je Fach und Transitionsgruppe wurden Strukturgleichungsmodelle definiert, wonach das Fähigkeitsselbstkonzept im 5. Schuljahr von der Leistung und der Klassenleistung im 5. Schuljahr und das Fähigkeitsselbstkonzept im 6. Schuljahr von der Leistung und der Klassenleistung im 6. Schuljahr abhängt. Zusätzlich wurde überprüft, ob die Veränderung des Fähigkeitsselbstkonzepts vom 5. zum 6. Schuljahr in der Gruppe ohne Übertritt von der Leistung und der Klassenleistung im 5. und im 6. Schuljahr abhängt (Modell 1), während sie in der Gruppe mit Übertritt nur von der der aktuellen Leistung und der aktuellen Klassenleistung abhängt (Modell 2).

Die Ergebnisse bestätigten diese beiden Modelle weitgehend. Modellvergleiche zeigten, dass die frühere Leistung und die frühere Klassenleistung in Deutsch und Mathematik die Veränderung des Fähigkeitsselbstkonzepts beim Verbleib in der Primarstufe erklärten, aber nicht beim Wechsel in die Sekundarstufe I. Es zeigt sich somit, dass beim Verbleib in der Primarstufe sowohl aktuelle als auch frühere Leistungen und Klassenleistungen die Veränderung des Fähigkeitsselbstkonzepts Deutsch und des Fähigkeitsselbstkonzepts Mathematik beeinflussen.

Findet aber ein Übertritt in die Sekundarstufe I mit einem Bezugsgruppenwechsel statt, befinden sich die Schülerinnen und Schüler in einer Situation, in der sie sich neu orientieren müssen. Nach dem Übertritt wird die Veränderung des Fähigkeitsselbstkonzepts Deutsch und des Fähigkeitsselbstkonzepts Mathematik ausschließlich von der aktuellen Leistung bzw. Klassenleistung erklärt. Temporale Vergleiche mit früheren Leistungen und Klassenleistungen haben an Bedeutung verloren, der soziale Vergleich mit der neuen Bezugsgruppe ist aber entscheidend.

Es führt also nicht das Alter bzw. die eintretende Pubertät zu Veränderungen des Selbstkonzepts, sondern der Bezugsgruppenwechsel (vgl. Watermann, Klingebiel & Kurtz 2010).

In der Diskussion werden die wenigen fachspezifischen Abweichungen vom berichteten Ergebnis diskutiert und die überlappenden sozialen und temporalen Vergleichsprozesse zusammenfassend erörtert.

Mazziotti, Claudia; Loibl, Katharina; Nikol, Rummel

Eine quasi-experimentelle Studie zur Bedeutung des kooperativen Lernens im Productive Failure Lernansatz

Lernansätze, bei denen Schülerinnen und Schüler (SuS) durch selbstständiges Aufgabebearbeiten auf zeitlich verzögerte Instruktion vorbereitet werden, sind besonders förderlich für den Erwerb von konzeptuellem Verständniswissen (z.B. Schmidt & Bjork, 1992; Schwartz & Bransford, 1998; Schwartz & Martin, 2004). Zu diesen Lernansätzen zählt der sogenannte Productive Failure Lernansatz (PF; z.B. Kapur, 2012), bei dem SuS zunächst selbstständig in Kleingruppen eine für sie unbekannte Aufgabe bearbeiten und anschließend in einer zweiten Lernphase Instruktion über die kanonische Lösung dieser Aufgabe erhalten. Im Vergleich zu SuS, die zuerst Instruktionen erhalten und danach eine ähnliche Aufgabe bearbeiten (so genannte Direkte Instruktionsbedingung, DI) erwerben SuS aus der PF-Bedingung signifikant mehr Verständniswissen (z.B. Kapur, 2012). Dieser PF-Effekt ist über verschiedene Domänen und Stichproben stabil (z.B. Kapur, 2008). Während der Effekt des Zeitpunkts der Instruktion intensiv untersucht wurde (z.B. Loibl & Rummel, 2013), ist die Rolle der Kooperation, also dem Arbeiten in Kleingruppen, in der Aufgabebearbeitungsphase von PF noch nicht hinreichend erforscht (Collins, 2012). Da beim kooperativen Lernen elaborative Prozesse (Teasley, 1995) und sinnstiftende Aktivitäten (Koedinger, Corbett & Perfetti, 2012) angeregt werden, erwarten wir, dass kooperatives Lernen im Vergleich zum individuellen Lernen in der Aufgabebearbeitungsphase von PF zu mehr Verständniswissen führt. Dies entspricht der Überlegenheit kooperativ-interaktiver Lernaktivitäten gegenüber individuell-konstruktiver Lernaktivitäten gemäß dem ICAP-Rahmenmodell von Chi (2009). Um sicherzustellen, dass elaborative Prozesse und sinnstiftenden Aktivitäten in der kooperativen Lernsituation zum Tragen kommen, legt die Forschung zum kooperativen Lernen (z.B. Fischer et al., 2010) allerdings nahe den Kooperationsprozess zu unterstützen (beispielsweise durch eine Gruppenbelohnung oder ein Rollenscript).

Eine quasi-experimentelle Studie (N = 261), die in der fünften Klasse verschiedener weiterführenden Schulen in Nordrhein-Westfalen durchgeführt wurde, überprüft die Hypothese, dass unterstütztes kooperatives Lernen in der Aufgabebearbeitungsphase von PF im Vergleich zu individuellem Lernen einen förderlichen Effekt hat. Da in vorherigen PF-Studien in der Regel mit älteren SuS bzw. mit Studierenden gearbeitet wurde (z.B. Loibl & Rummel, 2013), sollte diese Studie zusätzlich untersuchen, ob der positive Effekt von PF auch für die vergleichsweise junge Stichprobe greift (Hypothese 2). Es wurden daher vier Bedingungen implementiert, bei denen die Faktoren Sozialform des Lernens (kooperativ vs. individuell) und der Zeitpunkt der Instruktion (PF vs. DI) variierten. Zur Überprüfung der ersten Hypothese werden die zwei PF-Bedingungen verglichen, in denen SuS entweder paarweise oder individuell eine Aufgabe bearbeiteten und anschließend Instruktion erhielten (PF-Koop, PF-Ind). Zur Überprüfung der zweiten Hypothese werden die PF-Bedingungen mit den zwei DI-Bedingungen verglichen, in denen SuS zuerst instruiert wurden und anschließend entweder paarweise oder individuell eine ähnliche Aufgabe bearbeiteten (DI-Koop, DI-Ind). Um die Zusammenarbeit der Paare in den beiden kooperativen Bedingungen (PF-Koop, DI-Koop) zu unterstützen, wurde den Paaren eine Gruppenbelohnung für die stärkste Zusammenarbeit in Aussicht gestellt sowie ein Rollenscript eingeführt. Dieses Rollenscript bestand zum einen aus der Rolle des „Denkers“: Der Schüler bzw. die Schülerin in dieser Rolle wurde angeregt seine Lösungsideen laut zu verbalisieren und zu erklären. Zum anderen beinhaltete das Rollenscript die Rolle des „Fragers“: Der Schüler bzw. die Schülerin wurde angeregt Verständnis- oder Hinweisfragen zu stellen. Die kooperierenden Paare mussten mindestens einmal die Rolle wechseln. Als abhängige Variable wurde zusätzlich zu dem Nachwissenstest, der insbesondere konzeptuelles Verständniswissen erfasste, Kooperationsprozessdaten erhoben. Das heißt, die sprachlichen Äußerungen der Paare in der PF-Koop- und DI-Koop-Bedingungen wurden per Diktiergerät aufgezeichnet. Die so gewonnen Kooperationsprozessdaten geben neben den Ergebnissen des Nachwissenstest auch Auskunft darüber, warum oder unter welchen Bedingungen kooperatives Lernen lernförderlich ist. In Anlehnung an das ICAP-Rahmenmodell (Chi, 2009) erwarten wir, dass insbesondere kooperativ-interaktive Lernaktivitäten, bei denen SuS auf der Lösungsidee des Lernpartners aufbauen, für ein tiefes Verständnis sorgen. Da die Datenanalyse zum jetzigen Zeitpunkt noch nicht abgeschlossen ist, werden die Ergebnisse erst auf der Konferenz präsentiert.

Freitag
13. März

9:00 Uhr

GA 03/142

Chair(s)
Schwenk, Christin

Brückner, Sebastian; Förster, Manuel; Zlatkin-Troitschanskaia, Olga

Der Einfluss der Vorbildung auf das finanzwirtschaftliche Fachwissen von Studierenden in Hochschulen – Eine mehrbenenanalytische Betrachtung

Die valide Erfassung des Fachwissens als wesentlicher Bestandteil fachbezogener Kompetenzen von Studierenden stellt ein stetig bedeutender werdendes Forschungsfeld im Hochschulbereich dar (Blömeke, Zlatkin-Troitschanskaia, Kuhn & Fege, 2013; Tremblay, Lalancette & Roseveare, 2013). Dies gilt insbesondere für die Inhaltsbereiche der Wirtschaftswissenschaften. Während z.B. zum Rechnungswesen (Berger et al., 2013) oder zur Ökonomie (Zlatkin-Troitschanskaia, Förster, Schmidt, Brückner & Beck, 2015) schon erste Befunde vorliegen, ist das finanzwirtschaftliche Fachwissen der Studierenden noch wenig erforscht. Aufgrund seiner zentralen, vernetzenden Funktion im Unternehmen (Preiss, 2005) ist der finanzwirtschaftliche Inhaltsbereich bereits in einer kaufmännischen Berufsausbildung von Bedeutung und wird aufgrund seines algorithmisch-quantifizierenden Schwerpunkts, der u.a. auch mathematisches Wissen erfordert, von Unternehmen verstärkt nachgefragt (z.B. Sailer, 2009). Zur validen Erfassung des finanzwirtschaftlichen Fachwissens sollte neben den curricularen Einflüssen (in Form hochschulischer Lehrveranstaltungen) und den Variationen der Lehr-Lernbedingungen an verschiedenen Hochschulen vor allem die Heterogenität in der Vorbildung der Studierenden angemessen berücksichtigt werden. Der Fokus des vorliegenden Beitrags ist daher darauf gerichtet, zu untersuchen, welchen Einfluss die der Hochschule vorgelagerten formal-institutionellen Bildungsprozesse auf das finanzwirtschaftliche Wissen haben.

Zur Erfassung des finanzwirtschaftliche Fachwissens wurden Aufgaben des Examen General para el Egreso de la Licenciatura en Administración" (EGEL-ADMON; Ceneval, 2010) eingesetzt, die im Rahmen des vom BMBF geförderten Projekts WiwiKom übersetzt und nach den Test Adaption Guidelines (ITC, 2005) und den Standards for Educational and Psychological Testing (AERA, APA & NCME, 2014) u.a. mittels Curriculaanalysen, Experteninterviews, Onlineratings, kognitiven Interviews mit Studierenden und zwei Feldstudien (N=7.295) adaptiert, weiterentwickelt und hinsichtlich der Kriterien „Inhalt“, „Antwortprozesse“, „interne Struktur“ und „Beziehung zu anderen Variablen“ umfassend validiert wurden (Zlatkin-Troitschanskaia, Förster, Brückner & Happ, 2014).

Die hier präsentierten Analysen basieren auf einen Teildatensatz, der im Jahr 2013 bei 3.783 Studierenden an 23 Hochschulen (15 Universitäten und 8 Fachhochschulen) bundesweit erhoben wurde. Hierbei wurde aufgrund des Einsatzes von insgesamt 220 Aufgaben aus verschiedenen wirtschaftswissenschaftlichen Inhaltsbereichen ein Testheftdesign (nested Youden-Square-Design) (Frey, Hartig & Rupp, 2009) verwendet, wobei 773 Studierende aus allen Hochschulen die finanzwirtschaftlichen Aufgaben bearbeiteten.

Die unkonfundierte Schätzung des Einflusses der Vorbildung auf das finanzwirtschaftliche Wissen erfordert aufgrund der komplexen Datenstruktur eine latente Skalierung der Personenfähigkeit z.B. mittels IRT-Modellen (de Ayala, 2009). Weiterhin muss die hierarchische Struktur der Daten mit Studierenden auf Ebene 1 und Hochschulen auf Ebene 2 mittels Mehrebenenanalysen berücksichtigt werden (Raudenbush & Bryk, 2002). Auf Ebene 1 wurden neben den vorbildungsspezifischen Kovariaten (Absolvierung einer kaufmännischen Berufsausbildung und/oder eines Gymnasiums mit Schwerpunkt Wirtschaft), das Geschlecht, die Muttersprache, die Note der Hochschulzugangsberechtigung, die Anzahl besuchter hochschulischer Lehrveranstaltungen mit finanzwirtschaftlichen Inhalten und die Anzahl der Fachsemester der Studierenden aufgenommen. Auf Ebene 2 wurde die Unterscheidung des Hochschultyps (Fachhochschule und Universität) als kontextuelle Kovariate aufgenommen. Methodisch wurden die Daten nach je 5 Imputationen der Kovariaten (Rubin, 1987) und des Testscores mittels Plausible Values (Mislevy, 1991) mehrbenenanalytisch ausgewertet.

Erste Befunde zeigen, dass die Intraklassenkorrelation relativ gering ausfällt, so dass in erster Linie die Kovariaten auf Ebene 1 erklärungsmächtig sind. Ferner zeigt sich, dass die Vorbildung, insbesondere die Berufsausbildung, einen substantiellen Teil des finanzwirtschaftlichen Wissens erklären kann, auch wenn man die hochschulischen Einflüsse, wie z.B. die besuchten Finanz- und Mathematikveranstaltungen, kontrolliert. Da der Effekt auch unter Kontrolle der Fachsemester bestehen bleibt, kann dies als ein Indiz gewertet werden, dass das in der Berufsausbildung und den Gymnasien mit wirtschaftlicher Ausrichtung vermittelte Wissen nicht ohne weiteres durch eine hochschulische Ausbildung kompensiert wird. Um dieser Hypothese genauer nachzugehen braucht es allerdings eine Längsschnittbetrachtung, in der gezielt Studierende mit und ohne finanzwirtschaftliche Vorbildung über ihren Studienverlauf verfolgt werden können. Zugleich finden sich aber auch Unterschiede hinsichtlich des Geschlechts und der kontextuellen Kovariate, die weitere Analysen für das finanzwirtschaftliche Fachwissen notwendig machen.

Freitag
13. März

9:00 Uhr

GA 03/149

Chair(s)
Brückner, Sebastian

Chwallek, Katharina; Jerusalem, Matthias

Evaluation der Wirksamkeit einer Intervention zur Motivationsförderung in der Schule: Vergleich von linearen und diskontinuierlichen Mehrebenenmodellen

In einer quasiexperimentellen über zwei Schuljahre andauernden Interventionsstudie wurden Lehrkräfte trainiert, Förderstrategien in ihren Fachunterricht einzuführen, die auf der theoretischen Basis des Konzepts der Selbstwirksamkeitserwartung nach Bandura entwickelt wurden (z.B. Unterrichtsmaßnahmen zur Verbesserung von Transparenz und Fehlertoleranz). Dies sollte bei Siebt- und Achtklässlern dreier Berliner Oberschulen, verglichen mit einer Kontrollgruppe, zu einer günstigeren Entwicklung von Ressourcen und Risikofaktoren für die schulische Motivation beitragen.

Ziel dieses Beitrags ist die Evaluation der Wirksamkeit der Intervention im Rahmen von Mehrebenenanalysen: Mittels linearer Modelle wird zunächst die lineare Veränderung über beide Projektjahre hinweg geprüft. Des Weiteren sollen diskontinuierliche Modelle Auskunft über sequentielle, jahresspezifische Effekte geben.

Vor und nach dem ersten sowie am Ende des zweiten Projektjahrs wurden Befragungen mit Schülerinnen und Schülern von sechs Interventions- und Kontrollschulen durchgeführt. Der Fragebogen erfasst als motivationale Evaluationskriterien die schulische Selbstwirksamkeit, die Prüfungsängstlichkeit, die wahrgenommene Bedrohung und die Binnendifferenzierung im Unterricht. Insgesamt nahmen 570 Schülerinnen und Schüler an mindestens einer Befragung und 314 an allen drei Befragungen teil. Für den Umgang mit den fehlenden Werten wurde das Maximum-Likelihood-Verfahren eingesetzt, sodass im Rahmen der Mehrebenenanalysen auch die unvollständigen Datensätze berücksichtigt werden können.

Die Ergebnisse der linearen Mehrebenenmodelle weisen darauf hin, dass die wahrgenommene Bedrohung bei den Interventionsschülerinnen und -schülern, verglichen mit der Kontrollgruppe, im Verlauf der beiden Projektjahre signifikant sinkt ($\text{Gamma-101} = -0.125^*$; $\text{SE} = 0.05$). Mittels diskontinuierlicher Modelle lässt sich dieser Effekt für das erste Projektjahr replizieren ($\text{Gamma-101} = -0.211^*$; $\text{SE} = 0.09$). Darüber hinaus belegen die Ergebnisse der diskontinuierlichen Modelle bei den Interventionsschülerinnen und -schülern am Ende des ersten Jahres eine signifikante Reduktion der Prüfungsängstlichkeit ($\text{Gamma-101} = -0.285^*$; $\text{SE} = 0.09$) und einen bedeutsamen Anstieg in der erlebten Binnendifferenzierung ($\text{Gamma-101} = 0.334^*$; $\text{SE} = 0.08$).

Die Ergebnisse weisen auf die motivationale Relevanz der Fördermaßnahmen speziell des ersten Projektjahres hin und verdeutlichen, dass die diskontinuierlichen Modelle bei dem vorliegenden Design der Studie differenziertere Aussagen erlauben als die linearen.

Malin, Lydia

The contextual influences of peers and the competition in local vocational education and training on boys' job aspirations in Germany. A multi-level approach

Horizontal gender segregation in education and labour is a rather persistent phenomenon in all Western societies (e.g. Barone 2011, Smyth & Steinmetz 2008, European Commission 2009, Anker 1998). Whether or not changes in the horizontal (and concomitant vertical) gender segregation occur depends – among various other factors – on occupational preferences at school age.

Previous research shows that boys and girls choose different kinds of subjects (e.g. Charles & Bradley 2009; Correll 2001; Dryler 1998). This so called educational segregation is interrelated with later labour market segregation (e.g. Smyth 2005, Smyth & Steinmetz 2008). However, former research often focuses on women (e.g. Gundert & Mayer 2012; Eccles 1987; Rosenfeldt & Spenner 1992). Therefore, this paper has a closer look on boys' gender-atypical occupational aspirations before they actually enter vocational education and training (VET).

Furthermore, previous research has shown that individual-level predictors such as individual school achievement and family background, have statistical significant but small explanatory power (e.g. Busch 2013; Helbig & Leuze 2012; Buchmann & Kriesi 2012). Descriptive, qualitative and experimental studies give hints for contextual influences for example of peers and labour market conditions which are not considered in quantitative studies until now (e.g. Beierle 2013; Beinke 2006; Sinclair & Carlsson 2013). Therefore, this study extends the previous research by analysing these two contextual factors quantitatively: The first assumption is that parental influences on developing occupational preferences are extended and challenged by the influence of peers (Watt 2010).

Peers have to cope with the same problems and talk for example about their experiences during internships. Through this exchange of information peers can also have a normative influence. They may for example threaten the gender identity if they ridicule a gender-atypical desired occupation (Sinclair & Carlsson 2013). Secondly, assuming that secondary school leavers are often limited to their local VET market as they do not leave home before having finished general and vocational education, a direct influence on their vocational aspirations seems plausible if they adapt preferences to perceived chances of realisation. This assumption is supported by Großkurth and Reißig (2009) who showed that the chance of getting an apprenticeship and later job security are important for the development of occupational preferences of students from lower secondary schools.

Hence, this paper asks: Do peers' orientations and local VET conditions have an impact on the gender-type of boys' occupational preferences over and beyond individual-level predictors?

To answer this question I use data from the German National Educational Panel Study (NEPS), wave 1, starting cohort 4 (youth in grade 9) and the typing of VET markets of the Institute for Employment Research (IAB). Two dependent variables are compared: the first desired job and the realistic job aspiration as a more adjusted measure of occupational preference. The results of the three-level logistic regressions indicate that the career-orientation of peers does only matter for the realistic job aspiration, while the occupational structure and the competition of the regional VET market have an impact on both, boys' first desired job and their realistic job aspiration.

Brückner, Sebastian; Zlatkin-Troitschanskaia, Olga

Eine prozessbezogene Analyse der Validität standardisierter Testaufgaben zur Erfassung wirtschaftswissenschaftlichen Fachwissens von Studierenden in einem Mixed Methods Design

Die valide Erfassung wirtschaftswissenschaftlichen Wissens bei Studierenden stellt in jüngster Zeit ein stetig bedeutender werdendes Forschungsfeld dar und wird durch eine Reihe internationaler und nationaler Forschungsbemühungen befördert (z. B. Berger et al., 2013; Tremblay, Lalancette & Roseveare, 2012; Zlatkin-Troitschanskaia, Förster, Brückner & Happ, 2014). Eine große Zahl an Validierungsstudien fokussiert dabei meist auf multivariate Verfahren, die einen Zusammenhangsnachweis zwischen bspw. einem bestimmten universitären Veranstaltungsbesuch und einem höheren Testwert als hinreichenden Beleg für die Validität eines Messinstruments anführen (Zumbo, 2009). Mit dem Anspruch das Wissen von Studierenden valide zu erfassen, geht jedoch vielmehr die Herausforderung einher, kausale Beziehungen zwischen den Testwerten und denen sie determinierenden mentalen Prozessen herstellen zu können (Borsboom, Mellenbergh & van Heerden, 2004). Die Entwicklung kognitiver Diagnosemodelle, die Hinweise auf mentale Prozesse während der Aufgabenbearbeitung geben können, stellen ein probates Mittel der Prozessvalidierung dar (Leighton & Gierl, 2007). Bislang wurden jedoch nur für vereinzelte Aufgaben in sehr spezifischen, schulischen Kontexten (für schulische Rechnungswesenaufgaben s. Helm, George und Trost, (2014)) solche Modelle entwickelt. Hingegen liegen noch keine inhaltsbereichsübergreifenden Modelle vor, die eine Variation verschiedener Aufgabenformate berücksichtigen und auf das wirtschaftswissenschaftliche Wissen von Studierenden ausgerichtet sind.

Die vorliegende Studie, die Teil des Forschungsprojekts WiwiKom ist, geht diesem Forschungsdefizit nach und verfolgt im ersten Schritt die explorative Analyse der Prozesse, die während der Bearbeitung wirtschaftswissenschaftlicher Testaufgaben von Studierenden auftreten. Im zweiten Schritt werden diese Prozesse hinsichtlich ihrer Konstrukturelevanz sowie hinsichtlich ihrer Eignung zur Entwicklung eines kognitiven Diagnosemodells evaluiert, um eine Validierung der Testaufgaben aus Prozesssicht vornehmen zu können.

Die aus international etablierten Testinstrumenten übersetzten und adaptierten Testaufgaben des Projekts WiwiKom wurden in zwei Feldstudien bei 7.295 Studierenden eingesetzt (s. Zlatkin-Troitschanskaia et al., 2014). Aus einem Itempool von 220 Aufgaben wurden 19 Aufgaben aus verschiedenen Inhaltsbereichen anhand empirischer Itemkennwerte ausgewählt, die auf quantitativer Basis von einer Validität der Testaufgaben zeugen. Die ausgewählten Aufgaben wurden danach von 20 Studierenden mit der Methode des (unmittelbaren und retrospektiven) lauten Denkens bearbeitet. Die Auswahl der Studierenden erfolgte hierbei mittels eines sogenannten „purposeful random sampling“ (Patton, 2002), indem die Verteilung der Validitätskriterien in der Feldstichprobe (z. B. Verteilung der Fachsemester) der Ziehung der Zufallsstichprobe zugrunde gelegt wurde. Die verbalen Protokolldaten wurden im Anschluss transkribiert, mittels eines induktiv entwickelten Kodierleitfadens kodiert und schließlich quantifiziert (Chi, 1997). Durch die Quantifizierung der Daten war es in dem vorliegenden Mixed Methods Design möglich über eine deskriptive Beschreibung hinaus auch eine statistisch zuverlässige Fundierung der Daten vorzunehmen und sie für weitere Analysen zu verwenden, die Hinweise für die Erklärung der Befunde (Testwerte) aus den Feldstudien liefern können.

Die Befunde zeigen, dass 22 verschiedene mentale Prozesse identifiziert werden konnten, die sich in kognitive, metakognitive, affektive und teststrategische Kategorien einordnen lassen. Zu den kognitiven Prozessen zählen u.a. Elaborationen wirtschaftswissenschaftlicher Konzepte auf Basis von Fachwissen, während metakognitive Prozesse z.B. das Planen sowie das Überprüfen der Aufgabenbearbeitung fokussieren. Affektive Prozesse umfassen vor allem emotionale Kommentierungen der eigenen Verstehensprozesse und als Anwendung von Teststrategien wurde der Rückgriff auf Rate- und Ausschlussstrategien kodiert. Insgesamt treten überwiegend kognitive und metakognitive Prozesse bei der Aufgabenbearbeitung auf, wobei diese sich zugleich als besonders bedeutsam für erfolgreiche Aufgabenlösungen herausgestellt haben. Im Hinblick auf zentrale Validierungskriterien wie den Studienfortschritt zeigt sich, dass fortgeschrittenere Studierende nicht nur über eine elaboriertere Wissensbasis verfügen, sondern ebenfalls verstärkt auf hierarchische Inferenzen zurückgreifen, die ihnen sowohl eine höhere Abstraktion als auch eine Konkretisierung der in den Aufgabenstellungen gegebenen Problemsituationen ermöglichen. Auf dieser Befundlage lassen sich vertiefte Erkenntnisse über die Validität der Aufgabenstellungen aus Prozesssicht ableiten, die im Weiteren eine Diskussionsgrundlage für Entwicklung eines kognitiven Diagnosemodells bieten, das zukünftig für geeignete (hochschul-)fachdidaktische Interventionen genutzt werden könnte (Brückner, in Vorbereitung).

Freitag
13. März

9:00 Uhr

GA 03/149

Chair(s)

Brückner, Sebastian

Penk, Christiane; Richter, Dirk

Testteilnahmemotivation im Verlauf eines Large-Scale Assessments – Veränderung und Zusammenhang mit Testleistung unter Verwendung von Second-Order Latent Growth Curve Modeling

Theoretischer Hintergrund: Large-Scale Assessments dienen der Evaluation der Qualität von Bildungssystemen. Um aus den Ergebnissen dieser Tests valide Interpretationen abzuleiten, wird unterstellt, dass die Testteilnehmenden während der Testung ihr Bestes geben und diese Anstrengung auch bis zur letzten Aufgabe aufrechterhalten (Eklöf, 2008, 2010a; Thelk, Sundre, Horst, & Finney, 2009). Daher sind Kenntnisse über die motivationalen Prozesse während der Testsitzung und über den Einfluss der Testteilnahmemotivation auf die Testleistung unabdingbar. Die Forschung im Kontext der Testteilnahmemotivation basiert meist auf der Erwartung-Wert-Theorie (Wigfield & Eccles, 2000), die annimmt, dass die Erfolgserwartungen und die subjektiver Wert eines Tests (z. B. die wahrgenommene Wichtigkeit), die Anstrengungsbereitschaft und Testleistung beeinflussen. Barry und Finney (2014) untersuchten die Veränderung der Testteilnahmemotivation und zeigten, dass Teilnehmende für den zuerst bearbeiteten, kognitiven Tests eine niedrigere Anstrengungsbereitschaft und höhere Wichtigkeit berichten als für die danach bearbeiteten nichtkognitive Tests. Andere Untersuchungen zur motivationalen Veränderung innerhalb eines kognitiven Tests konstatierten eine Abnahme in der Anstrengungsbereitschaft im Verlauf des Tests (Horst, 2010; Wise, Pastor, & Kong, 2009). Keine Studie untersuchte jedoch bislang den Zusammenhang zwischen der Veränderung der Testteilnahmemotivation innerhalb eines kognitiven Tests und der gezeigten Testleistung.

Fragestellung: (1) Verändert sich die Testteilnahmemotivation (Erfolgserwartung, wahrgenommene Wichtigkeit des Tests und Anstrengungsbereitschaft) innerhalb einer zweistündigen Testsitzung? (2) Inwiefern hängt die Veränderung in der Testteilnahmemotivation mit der gezeigten Testleistung zusammen?

Methode: Datengrundlage bildet der IQB-Ländervergleich 2012 (Pant et al., 2013) mit einer national repräsentativen Stichprobe von 42.298 Neuntklässlerinnen und Neuntklässlern. Die Testteilnahmemotivation wurde zu drei Messzeitpunkten erhoben (vor dem Test, nach der Hälfte des Tests und nach dem Test) unter Anwendung des Questionnaire on Current Motivation (Freund, Kuhn, & Holling, 2011), der die Erfolgswahrscheinlichkeit (Erwartung) und Herausforderung (wahrgenommene Wichtigkeit des Tests) auf einer 4-stufigen Likert-Skala erfasst. Zusätzlich wurden Anstrengungsbereitschaft (Eklöf, 2010), Selbstkonzept in Mathematik sowie Kontrollvariablen (Geschlecht, Schulform, sozioökonomischer Status, Zuwanderungshintergrund, Mathematiknote) erhoben.

Second-order latent growth curve modeling wurde angewendet, um die latenten Wachstumsverläufe der Testteilnahmemotivation zu analysieren. Mit dieser Technik kann die anfängliche Testteilnahmemotivation (intercept) und die Veränderung in der Testteilnahmemotivation (slope) modelliert werden.

Für die erste Frage werden drei lineare Wachstumsmodelle (je eins für Erfolgserwartung, wahrgenommene Wichtigkeit des Tests und Anstrengungsbereitschaft) analysiert. Für die zweite Frage werden die intercept und slope Faktoren der drei Konstrukte der Testteilnahmemotivation als Prädiktoren für Testleistung eingesetzt. Außerdem wird Selbstkonzept in Mathematik eingefügt als Prädiktor für den intercept und slope Faktor der Erfolgswahrscheinlichkeit, um für domänenspezifische Motivation zu kontrollieren (Eklöf, 2008). Zusätzliche Prädiktoren der Leistung sind Selbstkonzept und die oben genannten Kontrollvariablen. Als manifeste Indikatoren der Fähigkeit in Mathematik werden Weighted Likelihood Estimates verwendet.

Ergebnisse: Die Analysen zur ersten Fragestellung zeigen, dass sich die Teilnehmenden im Durchschnitt vor dem Test selbstbewusst fühlten, den Test gut zu bewältigen, diesen als wichtig wahrnahmen und bereit waren, Anstrengung zu investieren. Die anfänglich berichtete Erfolgswahrscheinlichkeit wurde im Durchschnitt über den Test hinweg aufrechterhalten. Im Gegensatz dazu verringerten sich die Anstrengungsbereitschaft und die wahrgenommene Wichtigkeit im Verlauf des Tests leicht.

Für die zweite Fragestellung konnte mit den drei Wachstumsmodellen, dem Selbstkonzept und den Hintergrundvariablen der Schülerinnen und Schüler 64% der Varianz in der Mathematikleistung erklärt werden. Neben der Schulform ($\beta = .50$) und der Mathematiknote ($\beta = .20$) waren die stärksten motivationalen Prädiktoren die anfängliche Anstrengungsbereitschaft ($\beta = .17$), die anfängliche Erfolgswahrscheinlichkeit, das mathematische Selbstkonzept sowie die Veränderung der Erfolgswahrscheinlichkeit (je $\beta = .14$). Die anfängliche wahrgenommene Wichtigkeit zeigte ebenfalls einen substantiellen indirekten Effekt auf Testleistung via Anstrengung ($\beta = .25$). Auch nach Kontrolle der Hintergrundvariablen scheint es von Bedeutung, dass die Teilnehmenden den Test mit einer hohen Testteilnahmemotivation beginnen und während der Testsitzung selbstsicher bleiben, die Aufgaben erfolgreich bearbeiten zu können.

Freitag
13. März

9:00 Uhr

GA 03/149

Chair(s)

Brückner, Sebastian

Pohlmann-Rother, Sanna; Kürzinger, Anja

Wie unterstützen Lehrkräfte Grundschul Kinder beim Schreiben eigener Texte im schriftsprachlichen Anfangsunterricht?

Die Gestaltung von Lehrer-Schüler-Interaktionen zur Anregung und Unterstützung des Lernens ist zentrales Thema der aktuellen Lehr-Lernforschung, welche die Unterstützung individueller Lernprozesse als Komponente der Prozessqualität von Unterricht untersucht. Das sozial-konstruktivistische Lehr-Lernverständnis dient als lernpsychologische Begründung für die individuelle Lernunterstützung, wobei der sozialen Interaktion eine hohe Bedeutung für den Aufbau von Wissen zukommt (Krammer, 2009). Mitte der 70er Jahre des 20. Jahrhunderts wurde als Folge der Auseinandersetzung mit Formen und Merkmalen der individuellen Lernunterstützung die Metapher des Scaffoldings als Form der Unterstützung im Lehr-Lerndialog eingeführt (Wood, Bruner & Ross, 1976). Besonders die Phasen der selbstständigen Schülerarbeit eröffnen der Lehrperson die Gelegenheit zur individuellen Unterstützung. Befunde zum Unterricht in den Fächern Physik und Mathematik belegen z.T. allerdings einen niedrigen Anteil an prozessorientierter Lernbegleitung (Krammer, Reusser & Pauli 2010). Für den Anfangsunterricht der Grundschule und speziell für das Fach Deutsch liegen nur sehr wenige Untersuchungen vor, welche die Interaktionen zwischen Lehrkraft und Lernenden während Schülerarbeitsphasen differenziert analysieren (zu Leseübungen vgl. Lotz & Lipowsky 2014). Befunde zur individuellen Lernunterstützung in Schülerarbeitsphasen im Schreibunterricht der Grundschule existieren bislang nicht.

Das DFG-Projekt NaSch1 („Narrative Schreibkompetenz in Klasse 1“) untersucht daher den Schreibunterricht der ersten Jahrgangsstufe anhand von 51 Unterrichtsvideos, die im Rahmen der Längsschnittstudie PERLE („Persönlichkeits- und Lernentwicklung von Grundschulkindern“) erhoben wurden (Lipowsky, Faust & Kastens, 2013). Im Vortrag steht die Schülerarbeitsphase im Fokus sowie die Frage, wie Lehrkräfte die Schüler beim Verfassen eigener Texte unterstützen und anregen. Damit liegt das Augenmerk auf den Aktivitäten der Lehrperson, welche sie zur Anregung der aktiven kognitiven Partizipation der Lernenden in der Schreibphase einsetzt. Die videografierte Schülerarbeitsphase wird niedrig und mittel inferent ausgewertet, wobei im Event-Sampling-Verfahren zunächst jede vorkommende Lehrer-Schüler-Interaktion identifiziert wird. Als individuelle Lernunterstützung wird jede verbale oder nonverbale Lehrer-Schüler-Interaktion zur Unterstützung von Schülern bestimmt, unabhängig von der Form und Qualität dieser. Dazu zählen sowohl direkte Anleitungen z.B. zur Rechtschreibung und Ideenfindung als auch indirekte Hinweise zur Aufgabenbearbeitung. Im Vortrag wird das Kategoriensystem zur Analyse der individuellen Lernunterstützung mit Ankerbeispielen vorgestellt. Im Anschluss werden deskriptive Befunde zur Häufigkeit und Form der Lernunterstützung berichtet und die Verteilung der Unterstützung durch die Lehrkräfte auf einzelne Schüler thematisiert.

Kürzinger, Anja; Pohlmann-Rother, Sanna; Theurer, Caroline

Der Einfluss schriftsprachbezogener Voraussetzungen auf die Textqualität in Klasse 1

Dem Erwerb von Schreibkompetenz kommt in der gegenwärtigen Diskussion im Kontext nationaler wie internationaler Schulleistungsstudien und der Überprüfung von Bildungsstandards eine besondere Bedeutung zu. Dabei herrscht Konsens darüber, dass die Fähigkeit von Kindern, eigene Texte zu schreiben, zwar wesentlich von der Grundschule vermittelt wird, aber auch auf den literarischen und medialen Erfahrungen im Vorschulalter aufbaut. Dementsprechend weisen Schulanfänger eine hohe Heterogenität hinsichtlich ihrer schriftsprachlichen Entwicklung auf (Martschinke, Kammermeyer, Frank & Mahrhofer, 2002). Neben allgemeinen Kompetenzen wie Merkfähigkeit und Aufmerksamkeit kommt der phonologischen Bewusstheit als Vorläuferfähigkeit für den Schriftspracherwerb eine hohe Bedeutung zu (z.B. Bradley & Bryant, 1985; Lundberg, Frost & Petersen, 1988; Schneider, Roth, Küspert & Ennemoser, 1998). Mit phonologischer Bewusstheit ist die metakognitive Fähigkeit gemeint, vom Wortsinn zu abstrahieren und sich auf die lautlichen Aspekte von Sprache zu konzentrieren. Neben der Abstraktionsfähigkeit beinhaltet die phonologische Bewusstheit auch die Fähigkeit zur Selbstbeobachtung und eine überlegte, explizite Kontrolle (Lundberg, 1994). Im engeren Sinn wird unter phonologischer Bewusstheit die Fähigkeit verstanden, nicht nur Anfangs- und Endlaute in einem gesprochenen Wort identifizieren, sondern das gesamte Wort in seine lautlichen Bestandteile zerlegen zu können. Bedeutsam ist die Prognosekraft der phonologischen Bewusstheit im Hinblick auf die Entwicklung der Lese- und Rechtschreibfähigkeit. Während es Kindern mit einer bereits im Vorschulalter gut entwickelten phonologischen Bewusstheit relativ leicht und zügig gelingt, sich die Schriftsprache anzueignen, gelten Kinder mit einer schwach ausgeprägten phonologischen Bewusstheit als Risikokinder des Schriftspracherwerbs (Schründer-Lenzen, 2007). Im Rahmen des Literacy-Konzepts kann zudem vermutet werden, dass Erfahrung mit Literatur in früher Kindheit positive Effekte auf das Lesen und Schreiben lernen besitzt. Bislang fehlen jedoch breit angelegte Studien der Schreib- und Deutschdidaktik, welche die familiären Kontextdaten zur Förderung in der Vorschulzeit bei Eingangsuntersuchungen berücksichtigen.

Trotz der gut dokumentierten Bedeutung der phonologischen Bewusstheit auf die Entwicklung der Lese- und Rechtschreibfähigkeit liegen keine Ergebnisse zu deren Einfluss auf die Qualität von Schülertexten in der Grundschule vor. Im Rahmen der DFG-Studie NaSch1 (Narrative Schreibkompetenz in Klasse 1) wird daher der Frage nachgegangen, inwiefern schriftsprachspezifische Vorläuferfähigkeiten die Textqualität in der ersten Jahrgangsstufe vorhersagen können. Das Projekt NaSch1 greift dabei auf Daten der Längsschnittstudie PERLE (Persönlichkeits- und Lernentwicklung von Grundschulkindern, vgl. Lipowsky, Faust & Kastens, 2013) zurück, in der für 540 Schüler aus 38 Klassen in der Eingangsuntersuchung zu Schulbeginn das schriftsprachliche Vorwissen der Schulanfänger erhoben wurde.

Die phonologische Bewusstheit wurde zu Beginn des ersten Schuljahres unter Einsatz der Untertests Anlaute hören und Laute sprechen des LEst 4-7 (Moser, Berweger & Lüchinger-Hutter, 2004) erfasst, während für die Buchstaben-Laut-Zuordnung der Untertest Buchstaben lesen herangezogen wird (Lotz et al., 2013). Ebenfalls zu Schulbeginn fand eine Elternbefragung statt, bei der die familiären Kontextdaten erhoben wurden. Die Qualität der Texte aus der ersten Jahrgangsstufe wurde im Projekt NaSch1 unter Berücksichtigung linguistischer, fachdidaktischer und empirischer Ansprüche holistisch und anhand eines Kriterienkatalogs ausgewertet. Diesen Analyseinstrumenten liegt ein Modell der Textqualität zugrunde, das sich an den Schreiberfahrungen der Schüler sowie an dem schreibdidaktischen Setting orientiert. Die bei der kriterialen Textqualitätsanalyse berücksichtigten 12 Items lassen sich in formal-strukturelle (z.B. Anrede- und Schlussformel), sprachliche (sprachsystematische und kohärenzstiftende Aspekte) sowie inhaltliche Kriterien (z.B. Perspektivenübernahme) einteilen. Sowohl die schriftsprachlichen Vorläuferfähigkeiten (Gresser, Pohl, Corvacho del Toro, Greb & Faust, 2009; vgl. auch Lotz et al., 2013) als auch die kriterial erfasste Textqualität (Pohlmann-Rother, Schoreit & Kürzinger, i.V.) liegen Raschskaliert vor, wobei WLE-Schätzer als Personenparameter für die Analysen verwendet wurden.

Die Untersuchung des Effekts der phonologischen Bewusstheit sowie der Buchstaben-Lautkenntnis wird mittels einer Pfadanalyse berechnet. Erste Ergebnisse belegen – unter Kontrolle der sozialen Herkunft (HISEI), der Intelligenz sowie der Klassenzusammensetzung – einen schwachen Effekt der Aspekte der phonologischen Bewusstheit und der Buchstaben-Laut-Zuordnung auf die Textqualität.

Anke, Schmitz

Sachtexte mit globaler Textkohäsion verständlich gestalten – Welche Schülerinnen und Schüler profitieren im Hinblick auf ihr Textverständnis?

Da die Wissensvermittlung in weiten Teilen auf dem Verstehen von Sachtexten gründet, gelten gute Leseleistungen als Dreh- und Angelpunkt für schulischen Erfolg. Dies betrifft nicht nur das Lernen im Deutschunterricht, sondern ebenfalls die Arbeit in Sachfächern, wie Politik, Geschichte oder Biologie. Vielen Schülerinnen und Schülern bereitet das Sachtextverstehen allerdings Schwierigkeiten (vgl. Naumann, Artelt, Schneider & Stanat, 2010). Diese äußern sich unter anderem darin, dass sich die Schülerinnen und Schüler auf die Entschlüsselung von Wortbedeutungen und auf lokale Textzusammenhänge konzentrieren. Solche punktuellen Leseleistungen erschweren die Erarbeitung neuer Lerngegenstände, da der Gesamtzusammenhang des Textes nicht erschlossen wird.

Der Verständlichkeit von Sachtexten kommt für ein adäquates Textverständnis ein zentraler Stellenwert zu. Ausgehend hiervon fokussiert der vorliegende Beitrag, ob sprachliche Textgestaltungsmerkmale das Sachtextverständnis von Schülerinnen und Schülern verbessern können. Die Studie basiert auf der Konstruktions-Integrations Theorie von Kintsch (1998), indem Interaktionen zwischen Text- und Lesermerkmalen untersucht werden. Analysiert wird, ob alle Schülerinnen und Schüler oder nur bestimmte Schülergruppen das Angebot für sich nutzen. Im Fokus der Textgestaltung steht die globale Textkohäsion, die den ‚roten Faden‘ eines Textes erzeugt. Unter globaler Textkohäsion werden linguistische Textverknüpfungen auf der syntaktischen, lexikalischen und grammatischen Textebene subsumiert, die über benachbarte Sätze hinweg Textelemente miteinander verbinden (vgl. de Beaugrande & Dressler, 1981). Beispiele globaler Textkohäsion sind Überschriften, Textverweise und die Explikation rhetorischer Relationen. Das Merkmal globale Textkohäsion wird untersucht, da es bislang undifferenziert in Kombination mit der lokalen Textkohäsion betrachtet wurde (vgl. Ozuru, Dempsey & McNamara, 2009). Es mangelt an empirischer Evidenz, ob die globale Textkohäsion das Sachtextverständnis von Schülerinnen und Schülern fördern kann, obgleich ihr aus theoretischer Sicht eine allgemeine verständnisfördernde Wirkung zugesprochen wird (vgl. Kintsch, 1998; Schnotz, 2006).

In einem experimentellen Design lesen 741 Schülerinnen und Schüler der neunten Jahrgangsstufe aus Gymnasien und Gesamtschulen entweder einen global kohäsiven oder einen global inkohäsiven Sachtext mit dem Thema ‚Afrikas Wirtschaftsboom‘ (Between Design). Das Textverständnis der Schülerinnen und Schüler wird mit einem Verständnistest nach dem Lesen des Textes offline erfasst (vgl. OECD, 2010). Die Itemkonstruktion des Verständnistests orientiert sich an den kognitionspsychologischen Subskalen der Lesekompetenz aus PISA 2009 und beinhaltet Fragen zum Suchen und Extrahieren von expliziten Informationen, Aufgaben des Typs Kombinieren und Interpretieren sowie Items, die ein Reflektieren und Bewerten des Gelesenen erfordern. Als vorhersagekräftige Prädiktoren des Textverständnisses, die die Wirkung der globalen Textkohäsion moderieren können, werden kognitive und motivationale Verständnisvoraussetzungen der Schülerinnen und Schüler erfasst. Das thematische Vorwissen zum Sachtext wird durch einen 10-minütigen schriftlichen Wissenstest abgefragt. Mit dem standardisierten Lesegeschwindigkeits- und -verständnistest (LGV 6-12) wird die Fähigkeit zum sinnentnehmenden Lesen auf der Satz- und auf der Textebene erhoben (vgl. Schneider, Schlagmüller & Ennemoser, 2007).

Die Interaktionsanalysen ergeben, dass die Schülerinnen und Schüler von der erhöhten globalen Kohäsionsdichte im Hinblick auf ihr Sachtextverständnis unterschiedlich stark profitieren. Schülerinnen und Schüler, die gut sinnverstehend lesen können, nutzen das Angebot der globalen Textkohäsion unabhängig ihres Umfanges an thematischem Vorwissen gleich stark. Jene Schülerinnen und Schüler, die mehr thematisches Vorwissen haben, jedoch gleichzeitig weniger gut sinnentnehmend lesen können, profitieren am wenigsten von der Textkohäsion. Darüber hinaus zeigt sich bei den Schülerinnen und Schülern, die wenig thematisches Vorwissen besitzen und gleichzeitig weniger gut lesen können, eine ebenfalls deutliche Wirkung der Textkohäsion. Signifikante Interaktionen der globalen Textkohäsion mit der Fähigkeit zum sinnentnehmenden Lesen liegen nicht vor. Das thematische Interesse beeinflusst die Nutzung der globalen Kohäsion nicht, sondern zeigt einen Haupteffekt auf das Sachtextverständnis.

Die differenziellen Befunde lassen sich hinsichtlich der Passung und der Gestaltung von mehr oder weniger kohäsive Texten in heterogenen Unterrichtskontexten diskutieren. Darüber hinaus können mögliche Konsequenzen für Schulbuchautoren aufgezeigt werden.

Fleckenstein, Johanna; Reble, Raja; Möller, Jens

Der Einfluss immersiven Unterrichts auf die Beherrschung von Rechtschreibstrategien in der Erstsprache

Im immersiven Unterricht, in dem Schülerinnen und Schüler primär in einer Fremdsprache unterrichtet werden, besteht häufig die Befürchtung, dass Defizite in grundlegenden erstsprachlichen Kompetenzen wie dem Schriftspracherwerb auftreten. Diese Befürchtungen sind jedoch meist unbegründet: Wenn die richtigen Lernvoraussetzungen gegeben sind, erbringen immersiv unterrichtete Schülerinnen in ihrer Erstsprache (L1) und den Sachfächern vergleichbare bis bessere Lernleistungen (Genesee, 2004). Diese Vorteile lassen sich theoretisch zum einen mit sogenannten Transfereffekten begründen, die nach Cummins (1981) aufgrund eines allgemeinen Sprachenspeichers zwischen einer erworbenen Zweitsprache (L2) und der L1 auftreten können. Zum anderen ergibt sich eine weitere mögliche Ursache aus den verbesserten exekutiven Funktionen immersiv unterrichteter Schülerinnen und Schüler.

In beiden Fällen beschränken sich die Vorteile auf komplexe (meta-)kognitive Fähigkeiten und Fertigkeiten. Nach Cummins (1981) stellt das Buchstabieren und Entschlüsseln von Schrift hingegen rein sprachspezifische Oberflächenmerkmale dar. Auch empirische Studien unterstützen die Annahme, dass Transfer zwischen Sprachen nur bei komplexeren Aufgaben stattfindet (Gebauer, Zaunbauer & Möller, 2012; Turnbull, Hart & Lapkin, 2003).

Rechtschreibung einfach als ein Oberflächenmerkmal abzutun, ist aus linguistischer Sicht jedoch zu trivial. Denn der Orthographie liegen verschiedene komplexe Konstruktionsprinzipien zugrunde, für deren Anwendung Schreiblerner anspruchsvolle kognitive Leistungen erbringen müssen. Diese der Schreibung zugrundeliegenden Prinzipien lassen sich operationalisierbaren Rechtschreibstrategien zuordnen.

Die Frage ist nun, ob es einen Transfer dieses Strategiewissens von immersiv unterrichteten Schülerinnen und Schülern gibt, welcher sich positiv auf die Verwendung von Rechtschreibstrategien in der L1 auswirkt. In diesem Fall müssten immersiv unterrichtete Schülerinnen und Schüler in ihrer L1 bessere orthographische Leistungen in den einzelnen Rechtschreibstrategien aufweisen, als monolingual unterrichtete Schülerinnen und Schüler. Obwohl ein Transfer bei sprachspezifischen Kompetenzen (z.B. der allgemeinen Rechtschreibleistung, verstanden als Oberflächenaspekt; Gebauer, Zaunbauer, & Möller, 2012) nicht möglich ist, sollte ein Transfer bei konkreten Rechtschreibstrategien (verstanden als komplexere kognitive Kompetenzen) von der L2 in die L1 stattfinden.

In dem vorliegenden Beitrag wurde die Beherrschung von Rechtschreibstrategien mit der Hamburger Schreibprobe (HSP; May, 2002) an N = 350 Grundschülerinnen und -schülern dritter Klassen an Schulen in Schleswig-Holstein und Hamburg erhoben. Die Stichprobe umfasste zwei Teilstichproben: monolingual unterrichtete Schülerinnen und Schüler (n = 162) sowie immersiv unterrichtete Schülerinnen und Schüler (n = 188).

Die HSP testet drei verschiedene Rechtschreibstrategien: die alphabetische Strategie (AS), die orthographische Strategie (OS) und die morphematische Strategie (MS). Die interne Konsistenz der drei Skalen lag bei $\alpha = .73$ (AS), $\alpha = .81$ (OS) bzw. $\alpha = .74$ (MS). Alle drei Skalen korrelierten hoch miteinander. Zur Untersuchung der Mittelwertunterschiede zwischen den Gruppen wurde eine einfaktorielle multivariate Varianzanalyse mit allgemeiner Rechtschreibleistung (Gramphemtreffer) als Kovariate gerechnet. Erwartungsgemäß ließ sich ein Haupteffekt der Gruppe feststellen, der selbst nach Kontrolle der allgemeinen Rechtschreibleistung bestehen blieb: Die Tests der Zwischensubjekteffekte zeigten statistisch signifikante Unterschiede in der AS ($F[1, 347] = 11.08$; $p = .001$; $d = .36$) und der MS ($F[1, 347] = 13.10$; $p < .001$; $d = .39$), nicht aber in der OS ($F[1, 347] = 1.71$; ns).

Die Ergebnisse zeigen erwartungsgemäß, dass immersiv unterrichtete Schülerinnen und Schüler die AS und die MS besser beherrschen als monolinguale Schülerinnen und Schüler. Allerdings blieb der erwartete Effekt bei der OS aus. Dies lässt sich damit erklären, dass die OS – im Gegensatz zur AS und MS – auf normierten Merk- und Regelementen beruht und somit nicht die Komplexitätsansprüche aufweist, die für einen erfolgreichen Transfer notwendig wären. Detaillierte Analysen der Fehlerschwerpunkte sollen weitere Anhaltspunkte für die Interpretation der Effekte liefern.

Kauper, Tabea

Struktur und Entwicklung professioneller Kompetenz angehender Lehrkräfte in der 1. und 2. Phase der Ausbildung

In Deutschland beschäftigten sich in den vergangenen Jahren mehrere Studien mit der Erfassung professioneller Kompetenz von Lehrkräften (z.B. COACTIV: Kunter et al., 2011; TEDS-M: Blömeke, Kaiser, & Lehmann, 2010; LEK: König & Seifert, 2012; BilWiss: Terhart et al., 2012). Professionelle Lehrkompetenz wird dabei als Zusammenspiel von kognitiven und motivationalen Aspekten sowie der Fähigkeit zur Selbstregulation aufgefasst (Kunter, Baumert, Blum, & Neubrand, 2011) und gilt als notwendige Voraussetzung für erfolgreiches Handeln im Unterricht (Baumert & Kunter, 2006, 2011). Dabei wird davon ausgegangen, dass diese professionelle Kompetenz grundsätzlich veränderbar und prinzipiell erlernbar ist (Bauer et al., 2010; Kunter, Kleickmann, Klusmann & Richter, 2011).

Die Beiträge des Symposiums widmen sich dieser Frage nach der Veränder- und Vermittelbarkeit während der 1. und 2. Phase der Lehramtsausbildung in Deutschland und der Schweiz. Thematisch stehen die Bereiche pädagogisches Wissen, subjektive Überzeugungen, motivationale Orientierungen und Unterrichtswahrnehmung in Anlehnung an das COACTIV-Modell von Baumert und Kunter (2006, 2011) im Mittelpunkt. Konkret wird den Fragen nachgegangen, welche Strukturen zur Beschreibung pädagogischen Wissens bei Lehramtsstudierenden der Naturwissenschaft und Mathematik angemessen sind (Beitrag 1, Pollmeier et al.), welche Rolle Mentoring im Vorbereitungsdienst für die Reflexion von Praxiserfahrungen und damit einhergehend für die professionelle Unterrichtswahrnehmung und -handlung hat (Beitrag 2, Linniger et al.). Ein weiterer Beitrag (3, Kauper et al.) beschäftigt sich mit der Veränderbarkeit von subjektiven Lehrüberzeugungen im Übergang von Studium und Vorbereitungsdienst, sowie Beitrag 4 (Huber et al.) mit der Frage nach der Rolle von individuellen Determinanten für die Entwicklung eines professionellen Selbstkonzepts und der Selbstwirksamkeit bei Lehramtsstudierenden in der Schweiz.

Bearbeitet werden diese Fragen aus den verschiedenen Perspektiven der Disziplincluster, denen die Referierenden angehören (Psychologie, Erziehungswissenschaften und Pädagogik der Naturwissenschaften). Die Zusammenstellung von Beiträgen aus Deutschland und der Schweiz soll zudem eine Perspektive aus verschiedenen Ausbildungssystemen in der Lehrprofessionalisierung ermöglichen.

Zu den in den unterschiedlichen Projekten angewendeten und in den Beiträgen berichteten Studiendesigns und Erhebungsmethoden zählen neben der Testentwicklung (pädagogisches Wissen) auch die längsschnittliche Erfassungen von Lehrerüberzeugungen und motivationalen Orientierungen, sowie videogestützte onlinebasierte Verfahren zur Erfassung professioneller Unterrichtswahrnehmung.

Ziel des Symposiums ist es, diese unterschiedlichen Perspektiven und methodischen Zugänge vor der Gemeinsamkeit der Forschungsfragen herauszustellen und zu diskutieren.

Freitag
13. März

9:00 Uhr

HGB 20

Chair(s)

Kauper, Tabea
Bernholt, Andrea**Struktur und Entwicklung professioneller Kompetenz angehender Lehrkräfte
in der 1. und 2. Phase der Ausbildung**

Pollmeier, Judith; Hohenstein, Friederike; Kleickmann, Thilo; Zimmermann, Friederike; Köller, Olaf; Möller, Jens

Zur Struktur pädagogischen Wissens von Lehramtsstudierenden mit mathematisch-naturwissenschaftlichen Fächern

Das Professionswissen als Teil der professionellen Kompetenz von Lehrkräften wird als ein zentraler Faktor für die Unterrichtsqualität und die Lernzuwächse von Schülerinnen und Schülern erachtet. Dabei wird hauptsächlich zwischen drei Aspekten des Wissens unterschieden: Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen (vgl. Baumert & Kunter, 2006; 2011). Während vor allem die fachbezogenen Aspekte Fachwissen und fachdidaktisches Wissen bereits ausführlicher bearbeitet wurden, ist die Forschung zum pädagogischen Wissen erst angelaufen (z.B. Voss, Kunter & Baumert, 2011).

Das pädagogische Wissen wird als „fachunabhängiges Wissen über die Gestaltung von Lehr-Lern-Prozessen“ (Baumert & Kunter, 2011, S.38) verstanden. Bisherige Forschung zu diesem Aspekt hat sich häufig auf eine indirekte Erfassung dieses Wissens gestützt oder auf einzelne Aspekte des Wissens fokussiert (vgl. Voss et al., 2011). Neuere Studien, die eine direkte und umfassende Erfassung des pädagogischen Wissens anstrebten, hatten in einem ersten Schritt die Entwicklung angemessener Messinstrumente für diesen Bereich zu bewältigen (vgl. König, Blömeke, Paine, Schmidt & Hsieh, 2011; Hohenstein, Kleickmann, Zimmermann, Köller & Möller, eingereicht; Voss et al., 2011).

Die Befunde zur Struktur des pädagogischen Wissens weisen auf eine Mehrdimensionalität des Konstruktes hin. Beispielsweise legten Voss und Kollegen (2011) für ihr Testinstrument zum pädagogischen Wissen Evidenz für die vier Dimensionen Wissen über Klassenführung, Wissen über das Unterrichtsmethoden, Wissen über Leistungsbeurteilung und Wissen über Schülerheterogenität und Lernen vor und konnten insbesondere ein hierarchisches Modell belegen, bei denen die gefundenen Wissensfacetten ihrerseits einen Faktor zweiter Ordnung bildeten, der das gesamte pädagogische Wissen repräsentierte.

Der vorliegende Beitrag beschäftigt sich mit der Frage, welche Struktur zur Beschreibung des pädagogischen Wissens im Lehramtsstudium der Naturwissenschaften und Mathematik angemessen ist. Dabei sollen insbesondere die Mehrdimensionalität sowie eine mögliche hierarchische Struktur des Konstruktes untersucht werden.

Zur Untersuchung der Fragestellung bearbeiteten insgesamt 1036 Lehramtsstudierende der Naturwissenschaften und Mathematik an 12 Hochschulstandorten einen schriftlichen Test zur Erfassung des pädagogischen Wissens (Hohenstein, Kleickmann, Zimmermann, Köller & Möller, eingereicht). Das Instrument enthielt 73, größtenteils geschlossene, Aufgaben zur Erfassung des Wissens in den Facetten Lehren (17 Aufgaben), Lernen (20 Aufgaben), Klassenführung (23 Aufgaben) und Leistungsbeurteilung (13 Aufgaben). Die Übereinstimmung der Kodierer für die Antworten auf die offenen Aufgaben war zufriedenstellend (Cohen's $\kappa > .60$). Nach dem Ausschluss von 6 Aufgaben aufgrund ungenügender Trennschärfe wurden die restlichen 67 Aufgaben den Analysen zur Untersuchung der Struktur unterzogen.

Explorative Faktorenanalysen auf Grundlage der einzelnen Aufgaben zeigten eine gute Passung eines vierdimensionalen Modells, das Modellen mit weniger Faktoren überlegen war. Die Inspektion der rotierten Faktorladungen der Aufgaben aus dem Modell ergab, dass die gefundenen Faktoren die der Aufgabenentwicklung zu Grunde gelegten Wissensfacetten widerspiegeln.

Für weitere konfirmatorische Faktorenanalysen wurden die Aufgaben zunächst sog. Item-Parcels zufällig zugeordnet (je drei pro Wissensfacette), die als manifeste Indikatoren in die Analysen gingen. Auch hier bestätigte sich die Überlegenheit eines vierdimensionalen Modells, welches einen signifikant besseren Fit an die Daten aufwies als Modelle mit niedrigerer Dimensionalität. Des Weiteren hatte auch ein hierarchisches Modell eine gute Passung an die Daten ($\chi^2(50)=156.66$, $p < .001$, RMSEA=.05, CFI=.96, TLI=.94). In diesem Modell repräsentierte ein Faktor zweiter Ordnung das gesamte pädagogische Wissen. Die Ladungen der vier Wissensfacetten auf dem Generalfaktor lagen zwischen .71 und .87. Die vier Subskalen wie auch die gesamte Skala pädagogischen Wissens zeigten eine ausreichende bis gute interne Konsistenz.

Insgesamt legen die Ergebnisse eine mehrdimensionale Struktur pädagogischen Wissens mit den Facetten Lehren, Lernen, Klassenführung und Leistungsbeurteilung nahe. Weitere Analysen auf der Basis längsschnittlicher Daten werden erforderlich sein, um eine mögliche Veränderung der Struktur pädagogischen Wissens im Verlaufe des Studiums zu untersuchen.

Freitag
13. März

9:00 Uhr

HGB 20

Chair(s)

Kauper, Tabea
Bernholt, Andrea

Struktur und Entwicklung professioneller Kompetenz angehender Lehrkräfte
in der 1. und 2. Phase der Ausbildung

Linninger, Christina; Kunina-Habenicht, Olga; Stürmer, Kathleen; Seidel, Tina; Kunter, Mareike
Wie lernen angehende Lehrkräfte? Zur Rolle von Reflexion und Mentoring in der Lehrerbildung

Professionelle Kompetenz gilt als erlernbare Grundlage qualitätvollen Lehrerhandelns (Baumert & Kunter, 2006). Damit ist die Lehrerbildung zunehmend in den Fokus des Forschungsinteresses gerückt und es stellt sich zum einen die Frage, welche Lernprozesse stattfinden und welche Wirkung sie im Professionalisierungsprozess haben, und zum anderen wie Lerngelegenheiten diese Lernprozesse bei angehenden Lehrkräften unterstützen können. Diesen beiden Fragen möchte der vorliegende Beitrag nachgehen.

Auf der Ebene der Lernprozesse wird die Reflexion von Erfahrungen als bedeutsamer Lernprozess bei angehenden Lehrkräften diskutiert, allerdings fehlt es noch an empirischer Evidenz der Wirkmechanismen von Reflexionsprozessen (Darling-Hammond & Richardson, 2009; Terhart, 2013). Es wird lediglich davon ausgegangen, dass eine reflektierende Auseinandersetzung mit Praxiserfahrungen zu zunehmend professionellerem Handeln führt. Auf der Ebene der Lerngelegenheiten ist Mentoring in der Berufseinstiegsphase ein typisches Element der Lehrerbildung (Beijaard et al., 2010), doch auch hier gilt es, die genauen Wirkmechanismen einzelner Ausbildungselemente noch besser zu erforschen (Ingersoll & Strong, 2011).

Der vorliegende Beitrag setzt an diesen Forschungslücken der Wirkmechanismen von Lernprozessen und Lerngelegenheiten in der Lehrerbildung an. Auf der Seite der Lernprozesse gehen wir als erstes der Frage nach, ob das allseits geforderte Reflektieren der eigenen Praxiserfahrungen auch tatsächlich mit professionellerer Wahrnehmung von Unterricht und qualitativollerem Unterrichtshandeln zusammenhängt. Reflexion wird dabei in Anlehnung an theoretische und empirische Vorarbeiten (Atkins & Murphy, 1993; Michalsky & Kramarski, 2014) als Lernprozess mit zwei Komponenten betrachtet: die kritische Analyse von Erfahrungen (Beurteilung) und das Generieren neuer Handlungsalternativen (Modifikation). Es wird angenommen, dass sich das Reflektieren von Erfahrungen in einer Zunahme von praxisrelevantem Handlungswissen manifestiert, sich also in professionellerer Wahrnehmung von Unterrichtssituationen sowie in professionellerem Unterrichtshandeln widerspiegeln sollte.

Als zweites soll auf der Seite der Lerngelegenheiten die Qualität des Mentoring näher betrachtet werden mit der Frage, inwieweit die Interaktion mit der Ausbildungslehrkraft Lernen durch Reflexion bei angehenden Lehrkräften unterstützen kann. Richter und Kollegen (2013) konnten bereits zeigen, dass sich eine konstruktivistisch orientierte (im Gegensatz zu transmissiv orientierter) Interaktion günstig auf die Entwicklung von Motivation und Wohlbefinden von angehenden Mathematiklehrkräften auswirkt. Angelehnt an diese Ergebnisse sowie den Annahmen konstruktivistischer Lerntheorien nehmen wir an, dass eine konstruktivistisch orientierte Interaktion mit der Ausbildungslehrkraft Reflexionsprozesse bei angehenden Lehrkräften anstoßen kann.

Die Forschungsfragen wurden an einer Stichprobe von 362 angehenden Lehrkräften aus Nordrhein-Westfalen untersucht, die im Rahmen des Forschungsprogramms BilWiss (Terhart et al., 2012) im Vorbereitungsdienst längsschnittlich begleitet wurden. Eingesetzte Instrumente waren eigenkonstruierte Fragebogenskalen zur Reflexion und der Selbsteinschätzung des eigenen Unterrichtsverhaltens (Aspekte konstruktive Unterstützung, kognitive Aktivierung und Klassenführung) am Ende des Vorbereitungsdiensts, ein videogestütztes onlinebasiertes Verfahren zur Erfassung der professionellen Unterrichtswahrnehmung (Seidel & Stürmer, 2014) am Ende des Vorbereitungsdiensts sowie Fragebogenskalen zur konstruktivistischen und transmissiven Mentor-Mentee-Interaktion im Vorbereitungsdienst (Richter et al., 2014).

In latenten Strukturgleichungsmodellen zeigten sich hypothesenkonforme Ergebnisse: Die Reflexion von Erfahrungen zeigte sowohl mit der objektiv erfassten professionellen Unterrichtswahrnehmung als auch mit der Selbsteinschätzung des Unterrichtshandelns am Ende des Vorbereitungsdiensts positive Zusammenhänge. Das Ausmaß der Reflexion wiederum wurde durch die Qualität der Lerngelegenheit des Mentoring beeinflusst: Dabei wirkt sich konstruktivistisch geprägte Interaktion im Gegensatz zu transmissiv geprägter Interaktion günstig auf die Anwendung von Reflexionsprozessen ein Jahr später aus.

Die Ergebnisse liefern Hinweise darauf, dass das Reflektieren von Erfahrungen tatsächlich in Zusammenhang mit professionellerer Wahrnehmung und Handeln steht. Des Weiteren konnten wir zeigen, dass die Gestaltung von Lerngelegenheiten, in diesem Fall die Art der Interaktion beim Mentoring, maßgeblich dafür sein kann, inwieweit solche Reflexionsprozesse angestoßen werden. Dies bringt Implikationen für die Gestaltung der Lehrerbildung sowie für die Professionalisierung von Mentorinnen und Mentoren im Vorbereitungsdienst mit sich.

Kauper, Tabea; Bernholt, Andrea; Bauer, Johannes; Zimmermann, Friederike

Verändern sich subjektive Lehrüberzeugungen im Übergang zwischen Studium und Vorbereitungsdienst?

Subjektive Überzeugungen von Lehrpersonen gelten als eine bedeutsame Facette professioneller Kompetenz, da sie die Wahrnehmung und Deutung von Unterrichtssituationen ebenso wie die an die Schülerinnen und Schüler und den Unterricht gerichteten Zielvorgaben und Erwartungen beeinflussen können (Baumert & Kunter, 2006). Der Einfluss der lehr- und lernbezogenen Überzeugungen auf das Lehrerhandeln, die Schülerwahrnehmung des Unterrichts und die Schülerleistung konnte bereits in mehreren Studien beschrieben werden (Schlichter, 2012, Dubberke et al., 2008, Staub & Stern, 2002). Die wissenschaftlich fundierte und reflektierte Vorstellung von Lehren und Lernen kann somit als notwendige Voraussetzung für professionelles Handeln im Lehrberuf verstanden werden (z.B. Dann, 1994; Baumert & Kunter, 2006).

Ergebnisse bisheriger empirischer Studien können sowohl Veränderungen der Überzeugungen während der universitären Lehrerbildung, als auch durch steigende Berufserfahrungen nachweisen (z.B. Voss, Kleickmann, Kunter & Hachfeld, 2011). Allerdings sind diese Ergebnisse aufgrund der Heterogenität der Untersuchungsanlagen zum Teil inkonsistent und kaum vergleichbar. Während einige Studien eine Entwicklung von eher transmissiven Vorstellungen hin zu konstruktivistischen Überzeugungen beschreiben, berichten andere Studien gegenteilige Effekte (Schlichter, 2012). Zudem lässt sich unter der Perspektive des „Praxisschocks“ beobachten, dass Lehrkräfte mit dem Zeitpunkt des Berufseinstieges vermehrt auf transmissive Überzeugungen zurückgreifen (Richter et al., 2013).

Anknüpfend an diese Befunde soll im vorliegenden Beitrag der Frage nach der Entwicklung subjektiver Lehrüberzeugungen zwischen dem Studium und den ersten Berufserfahrungen durch den Vorbereitungsdienst nachgegangen werden. Im Zentrum stehen dabei folgende Fragestellungen: 1.) Welche Lehrüberzeugungen existieren zum Ende des Studiums und zu Beginn des Vorbereitungsdienstes? 2.) Verändern sich die Überzeugungsprofile durch den Beginn des Vorbereitungsdienstes? 3.) Finden im Laufe des ersten Jahres im Vorbereitungsdienst weitere Veränderungsprozesse statt?

Den Analysen liegt eine drei Erhebungswellen umfassende Längsschnittstudie aus dem Forschungsprojekt PaLea: Panel zum Lehramtsstudium mit 421 Lehramtsstudierenden aus 11 Hochschulen zugrunde, die die Messzeitpunkte Studienabschluss (T1), Beginn (1-5 Monate) des Vorbereitungsdienstes (T2) sowie ein Jahr nach der ersten Befragung im Vorbereitungsdienst (T3) umfasst.

Die subjektiven Lehrüberzeugungen wurden über drei Fragebogenskalen erfasst, die transmissive und konstruktivistische Lehrüberzeugungen messen. Bei den konstruktivistischen Überzeugungen wurde zwischen lehrerbezogenen und schülerbezogenen Anschauungen unterschieden (Bauer & Drechsel, 2010).

Zur Beantwortung der ersten Fragestellung wurden in einem ersten Schritt latente Profilanalysen (LPA) über alle Messzeitpunkte berechnet. Hierbei zeigt sich, dass zu jedem der drei Messzeitpunkte zwei Profile beschrieben werden können. Ein Profil zeigt dabei eine weitgehend ausbalancierte Zustimmung zu allen drei Überzeugungen über das Lehren, das zweite Profil zeichnet sich durch eine überdurchschnittliche Zustimmung zu beiden konstruktivistischen Überzeugungen aus. In einem zweiten Analyseschritt wurde die Stabilität der Zugehörigkeit zu einem Profil von Messzeitpunkt 1 zu 2 und von Messzeitpunkt 2 zu 3 mittels Latent Profile Transition Analysis (LPTA) geprüft. Es zeigt sich eine hohe Wahrscheinlichkeit für den Verbleib im jeweiligen Profil über alle Messzeitpunkte (Profil 1: 85%; Profil 2: 69%), aber auch eine Wahrscheinlichkeit von 30% für den Wechsel vom eher konstruktivistischen in das ausbalancierte Profil von Messzeitpunkt 1 zu 2. Umgekehrt findet ein Wechsel nur zu einer Wahrscheinlichkeit von 15% statt. Zwischen den Messzeitpunkten 2 und 3 bleiben die Zugehörigkeiten zu einem Profil stabil (Profil 1: 95%; Profil 2: 92%).

Die Ergebnisse zeigen, dass sich latente Profile von subjektiven Lehrüberzeugungen sowohl bei Studierenden im Master als auch im Vorbereitungsdienst identifizieren lassen, die sich vor allem hinsichtlich ihrer Zustimmung zu konstruktivistischen Überzeugungen unterscheiden. Die Zugehörigkeiten zu einem der Profile sind erwartungskonform stabil, es zeigen sich daneben jedoch bedeutsame Veränderungen im Übergang von der ersten in die zweite Phase der Lehrerbildung. Die Befunde deuten darauf hin, dass sich die Lehrüberzeugungen bei einigen angehenden Lehrkräften mit dem Einstieg in das Berufsleben relativieren. In weiteren Analysen soll geklärt werden, ob und wie sich die verschiedenen Gruppen im Vorbereitungsdienst hinsichtlich weiterer Kompetenz- und Persönlichkeitsmerkmale charakterisieren lassen.

Chair(s)

Kauper, Tabea
Bernholt, Andrea

Struktur und Entwicklung professioneller Kompetenz angehender Lehrkräfte
in der 1. und 2. Phase der Ausbildung

Trempler, Kati

Umgang mit wissenschaftlicher Evidenz im Bildungsbereich

Hintergrund: In diesem Symposium kommen vier Beiträge zusammen, in denen empirische Studien vorgestellt werden, die sich mit dem Umgang von Studierenden mit bildungswissenschaftlichen Theorien und Befunden im Kontext von schulischen Alltagssituationen befassen. Die Beiträge befassen sich mit der Modellierung, Messung und Förderung von Kompetenzen, die im Zusammenhang mit evidenzbasierter Praxis im Bildungsbereich zunehmend an Bedeutung gewinnen (Bromme, Prenzel & Jäger, 2014).

Forschungsfrage: Wie können verschiedene Aspekte von Kompetenzen im Umgang mit bildungswissenschaftlicher Evidenz bei Studierenden im Bildungsbereich erfasst und gefördert werden?

Übersicht: Der Beitrag von Trempler und Kollegen beschäftigt sich mit der Modellierung und Messung von Kompetenzen im Umgang mit bildungswissenschaftlicher Evidenz: Der Informationssuche sowie der Bewertung von Studien kommt für den Umgang mit Evidenz eine besondere Bedeutung zu. In dem Beitrag werden theoretische und empirische Argumente für die Inhalts- und Konstruktvalidität der vorgestellten Instrumente zur Erfassung der Teilkompetenzen der Informationsauswahl und der Bewertung von Studien präsentiert.

Im zweiten Beitrag von Hetmanek und Kollegen wird untersucht, ob strukturierte Kurzfassungen im Gegensatz zu Originalartikeln über empirische Primärstudien dazu beitragen können, bildungswissenschaftliche Befunde effektiver an Praktiker im Bildungsbereich zu vermitteln. In diesem Beitrag wird eine experimentelle Studie vorgestellt, die zeigt, dass diese Form der Aufbereitung von Forschung das Verständnis fördern und die kognitive Belastung und wahrgenommene Aufgabenschwierigkeit reduzieren können.

Der dritte Beitrag von Wenglein und Kollegen beschäftigt sich mit der Messung und systematischen Förderung der Kompetenz zur Bewertung und Nutzung bildungswissenschaftlicher Evidenz auf der Basis von Vignetten zu schulbezogenen Problemen. Es werden Ergebnisse zu den Effekten eines kooperativen Trainingsansatzes zur Förderung des kompetenten Umgangs mit Evidenz von Lehramtsstudierenden vorgestellt.

Im vierten Beitrag von Klein und Kollegen steht die Anwendung bildungswissenschaftlicher Theorien und entsprechenden Wissens zur Erklärung komplexer schulischer Situationen bei Lehramtsstudierenden im Fokus: Es werden Ergebnisse aus einer Trainingsstudie vorgestellt, in der typische Fehler als Lerngelegenheiten für den Aufbau von negativem Wissen herangezogen werden, um damit den Umgang mit wissenschaftlichen Theorien in Anwendungssituationen zu schulen.

Abschließend werden die Beiträge von Stefan Fries aus der Perspektive der Psychologie des Lehrens und Lernens diskutiert und eingeordnet. Der im Symposium untersuchte Umgang von Studierenden im Bildungsbereich mit Evidenz stellt ein wichtiges Thema mit hoher Praxisrelevanz dar. Innerhalb des Symposiums wird das Thema mit einzelnen Beiträgen aus der empirischen Bildungsforschung und der Pädagogischen Psychologie, zum Teil unter Beteiligung von Fachdidaktikern, aus unterschiedlichen Perspektiven betrachtet.

Freitag
13. März

9:00 Uhr

HGB 30

Chair(s)

Trempler, Kati
Hetmanek, Andreas

Trempler, Kati; Hetmanek, Andreas; Wecker, Christof; Kiesewetter, Jan; Wermelt, Mia; Fischer, Martin R.; Fischer, Frank; Gräsel, Cornelia

Nutzung von Evidenz im Bildungsbereich – Validierung eines Instruments zur Erfassung von Kompetenzen der Informationsauswahl und Bewertung von Studien

Theoretischer Hintergrund: Evidenzbasierte Praxis ist ein Ansatz aus der Medizin, der nun zunehmend auch für den Bildungsbereich gefordert wird. Evidenzbasierte Praxis wird nach dem Vorbild der Medizin definiert als die Nutzung der besten verfügbaren und vor allem empirisch abgesicherten Evidenz für das Treffen und die Begründung beruflicher Entscheidungen (Sackett et al., 1996). In den KMK-Standards für die Lehrerbildung wird beispielsweise seit 2004 u.a. gefordert, dass Lehrkräfte in der Lage sein sollten, Ergebnisse der Bildungsforschung zu rezipieren und zu bewerten sowie für die eigene berufliche Tätigkeit nutzbar zu machen. Für die Erfüllung dieser Anforderung wird hier die Bezeichnung „kompetentes Argumentieren mit Evidenz“ eingeführt. Ein Ziel des diesem Beitrag zu Grunde liegenden Forschungsprojektes ist die Entwicklung eines Instruments zur Messung der dafür nötigen Kompetenzen.

Die kompetente Nutzung von Evidenz stellt eine komplexe kognitive Anforderung für die Praktiker dar. Dabei sind Fertigkeiten bei der Identifikation und Formulierung eines evidenzbasiert zu lösenden Problems, der systematische Suche, Auswahl und Bewertung von relevanten und aussagekräftigen wissenschaftlichen Untersuchungen sowie deren Einbeziehung in praktische Entscheidungsprozesse gefragt. Bei der Entwicklung einer theoretischen Charakterisierung für das kompetente Argumentieren mit Evidenz im Bildungsbereich kann ausgehend von Anstätzen aus der Medizin (Harden et al., 1999) auf Arbeiten aus dem Bereich des „information problem solving“ (Brand-Gruwel, Wopereis, & Walraven, 2009) und das „evidence-based reasoning framework“ (Brown et al., 2010) zurückgegriffen werden. Zudem ist davon auszugehen, dass an kompetenten Leistungen auch weitere Determinanten beteiligt sind. Daher werden in der vorliegenden Untersuchung auch allgemeine kognitive Fähigkeiten, spezifisches forschungsmethodisches wie inhaltliches Wissen sowie Fähigkeiten des wissenschaftlichen und kritischen Denkens und subjektive Überzeugungen in die Analyse einbezogen. Auf Grund ihrer grundsätzlichen Bedeutung konzentrieren wir uns in diesem Beitrag zunächst auf die Informationsauswahl und die Bewertung von Studien.

Fragestellung: Im Fokus des vorliegenden Beitrages steht die Validierung der Instrumente zur Erfassung der Teilkompetenzen der Informationssuche und der Bewertung von Studien. Die Fragestellung lautet: Welche empirischen Argumente für die Inhalts- und Konstruktvalidität können für das entwickelte Verfahren zur Messung der Informationsauswahl und der Bewertung von Studien angeführt werden?

Methode: Es wurden Daten von 341 Studierenden aus unterschiedlichen Studierendengruppen erhoben. Die Erhebung erfolgte in zwei Schritten: Zunächst wurden in einem online-Fragebogen u.a. Variablen zu kognitiven Fähigkeiten, Wissen und Überzeugungen erfasst. Die Kompetenzmessung wurde im Labor durchgeführt und war eingebettet in ein Szenario, in dem die Teilnehmenden ein Lehrkonzept vervollständigen mussten. Hierfür waren evidenzbasiert zwei instruktionale Entscheidungen zu treffen. Im Rahmen jeder der zwei Entscheidungen mussten zunächst aus zehn Treffern einer fiktiven Trefferliste auf Grundlage von Kurzinformationen (Titel, Autoren, Zeitschrift, Jahr) und dann aus sechs Treffern ergänzt durch Abstracts die entscheidungsrelevantesten Studien ausgewählt werden (Informationssuche). Im nächsten Schritt mussten strukturierte Kurzfassungen von vier Studien gelesen und anhand eines vorgegebenen Ratingschemas bewertet werden (Bewertung von Studien). Die Auswertung erfolgte jeweils auf Grundlage einer Musterlösung.

Ergebnisse: Die Ergebnisse weisen im Bereich der Inhaltsvalidität zufrieden stellende Übereinstimmungen zwischen befragten Experten und den im Projekt entwickelten Musterlösungen auf. Hinsichtlich der Konstruktvalidität weisen die Ergebnisse zudem darauf hin, dass die untersuchten Fertigkeiten Informationsauswahl und Bewertung von Studien voneinander sowie von weiteren Determinanten getrennt werden können, jedoch unter einander Zusammenhänge in erwartbarer Größenordnung aufweisen. Die Ergebnisse dieser Studie liefern damit erste Hinweise auf eine konstruktvalide Erfassung der Teilkompetenzen der Informationssuche und der Bewertung von Studien. Weitere Untersuchungen zur Absicherung der Validität der vorgestellten Instrumente werden angestrebt.

Freitag
13. März

9:00 Uhr

HGB 30

Chair(s)

Trempler, Kati
Hetmanek, Andreas

3

Hetmanek, Andreas; Wecker, Christof; Trempler, Kati; Kiesewetter, Jan; Gräsel, Cornelia; Fischer, Martin R.; Fischer, Frank
Effekte strukturierter Kurzfassungen empirischer Studien auf Textverständnis, kognitive Belastung und subjektive Aufgabenschwierigkeit

Theoretischer Hintergrund: Evidenzbasierte Praxis setzt qualitative hochwertige Forschung, die eine hohe Praxisrelevanz aufweist, im gleichen Maße voraus wie effektive Kommunikationskanäle, in denen entsprechende Ergebnisse in die Praxis kommuniziert werden können. Im medizinischen Bereich wird die Idee der Evidenzbasierung seit einiger Zeit erfolgreich verfolgt (Montori und Guyatt, 2008). Im Bildungsbereich wird die Idee aktuell verstärkt aufgegriffen (z.B. Bromme, Prenzel, und Jäger, 2014). Der Ansatz verlangt von Praktikern, dass sie über die Entwicklungen in der empirischen Forschung auf dem Laufenden sind. In der Medizin können sich die praktizierenden Ärzte aus systematischen Reviews und Leitlinien über aktuelle Erkenntnisse zu neuen diagnostischen Verfahren oder effektiven Therapien informieren. In komplizierteren Fällen gehört es allerdings zum professionellen Selbstverständnis, auch originale Forschungsartikel als Entscheidungshilfe heranzuziehen.

Originale Forschungsartikel zu lesen, zu verstehen und für die praktische Entscheidungsfindung heranzuziehen, ist eine kognitiv anspruchsvolle Anforderung. Für die Rezeption wissenschaftlicher Artikel sind Praktiker in der Regel nicht besonders gut ausgebildet. Zudem verfügen sie meist nicht über allzu große Zeitrressourcen. Außerdem ist ein Großteil der empirischen Bildungsforschung mittlerweile internationalisiert: Die meisten - insbesondere qualitativ hochwertige - Artikel erscheinen in englischer Sprache. Dies stellt eine zusätzliche Herausforderung dar, da Englisch für die meisten Praktiker im Bildungsbereich höchstens eine Zweitsprache ist. Diese Hürden zu überwinden, ist ein wichtiger Schritt, wenn evidenzbasierte Praktiken sich im Bildungsbereich weiter verbreiten sollen.

Strukturierte Kurzfassungen stellen ein mögliches Mittel dar, diese Hürden zu überwinden und Hauptfragestellungen, entscheidende methodische Eigenschaften, sowie Hauptergebnisse von Studien effektiver an Praktiker zu kommunizieren.

Fragestellung: Welchen Einfluss haben strukturierte Kurzfassungen im Vergleich zu Originalartikeln auf das Textverständnis, die kognitive Belastung und die subjektive Aufgabenschwierigkeit?

Methode: Zur Beantwortung dieser Frage wurden Daten von 48 Lehramtsstudierenden erhoben. Die Teilnehmenden arbeiteten unter Zeitbeschränkung mit ausgedruckten Versionen der Studien und bearbeiteten unmittelbar nach dem Lesen computerbasierte Tests zum Textverständnis (offenes Aufgabenformat) sowie zur kognitiven Belastung und subjektiven Aufgabenschwierigkeit (Rating-Items). In Teil 1 der Studie wurden vier deutsche empirische Zeitschriftenartikel, in Teil 2 vier englische empirische Zeitschriftenartikel bearbeitet. Die beiden Teile waren eingebettet in ein fiktives Szenario, in dem zur Fertigstellung eines Lehrkonzeptes zwei instruktionale Entscheidungen getroffen werden mussten: In Teil 1 musste entschieden werden, ob ein Gruppenpuzzle durchgeführt werden soll, in Teil 2, ob nur korrekte oder auch fehlerhafte Videobeispiele zum Einsatz kommen sollen. In beiden Teilen wurde ein within-subjects-Design realisiert. Dazu wurden die vier Studien in zwei Paare aufgeteilt. Ein Paar wurde in Form der Originalartikel, das jeweils andere Paar als strukturierte Kurzfassungen vorgelegt. Die Zuordnung des Präsentationsformats (Originalartikel vs. strukturierte Kurzfassung) zu den Textpaaren sowie die Reihenfolge der Studien wurden ausbalanciert. Alle 48 Reihenfolge-Textpaar-Kombinationen wurden umgesetzt.

Ergebnisse: In Teil 1 (deutschsprachigen Zeitschriftenartikel) ergaben Varianzanalysen mit Messwiederholung statistisch signifikante Haupteffekte der Darstellungsform auf das Textverständnis, $F(1; 46) = 9.61, p = .003, \eta_p^2 = .17$, die kognitive Belastung, $F(1; 46) = 114.87, p < .001, \eta_p^2 = .71$, und die subjektive Aufgabenschwierigkeit, $F(1; 46) = 136.71, p < .001, \eta_p^2 = .75$: Mit strukturierten Kurzfassungen waren das Textverständnis höher und die kognitive Belastung und die subjektive Aufgabenschwierigkeit niedriger. In Teil 2 (englischsprachige Zeitschriftenartikel) konnte dieses Effektmuster repliziert werden.

Diese Ergebnisse legen nahe, dass strukturierte Kurzfassungen von empirischen Forschungsartikeln das Verständnis von Studien erleichtern und dabei helfen können, die empfundene kognitive Belastung und Schwierigkeit der Aufgabe zu reduzieren. Dieses – oder ähnliche Darstellungsformen – könnten somit einen Beitrag zu einer effektiveren Kommunikation von Forschungsergebnissen an Praktiker leisten.

Wenglein, Sandra; Bauer, Johannes; Prenzel, Manfred

Training der Kompetenz angehender und aktiver Lehrkräfte zur Bewertung und argumentativen Nutzung von Evidenz

Theoretischer Hintergrund: Das Prinzip des evidenzbasierten Handelns, sowie das argumentative Nutzen von Evidenz, spielt im Bildungswesen eine immer größere Rolle (Fischer et al. 2014; Wiseman, 2010). Der Praxistransfer hat jedoch bisher in vielen Bereichen nicht, oder nur unzureichend, stattgefunden. Gerade Lehrkräfte begründen ihr Handeln kaum auf Grundlage wissenschaftlicher Evidenz (Williams & Coles, 2007), obwohl dies in den Standards für Lehrerbildung gefordert wird (KMK, 2004). Ein möglicher Grund dafür ist, dass es aktuell im Lehramtsstudium in der Regel keine systematische Förderung der Kompetenz zur Bewertung und Nutzung bildungswissenschaftlicher Evidenz gibt.

Im Anschluss an diese Entwicklungen wurde ein Training für Lehramtsstudierende zum kompetenten Umgang mit Evidenz entwickelt. Die Grundidee dieses Trainings besteht darin, innerhalb kurzer Zeit in kooperativen Lernsettings (Gräsel, 2006; Collins et al, 1989) Heuristiken zu zwei Teilkompetenzen zu vermitteln: (a) Bewertung und (b) argumentatives Nutzen von Evidenz.

Die Heuristiken zur kritischen Bewertung von Evidenz haben zum Ziel, den Studierenden kognitive Strategien zu vermitteln, mit denen Sie die komplexe Aufgabe der Bewertung von Evidenz auch mit geringem Vorwissen bewältigen können (Abelson, 1995; Elsevier, 2014). Die Heuristiken zum Argumentativen Nutzen, und damit dem berufspraktischen Einsatz von Evidenz, beruhen auf Schemata von Kuhn (1997) und Toulmin (1996).

Fragestellungen: Ziel der vorliegenden Studie ist es, erste Erkenntnisse zu Trainierbarkeit und Nützlichkeit der vermittelten Heuristiken im Kontext von evidenzbasiertem Argumentieren zu gewinnen, sowie zu eruieren, inwiefern sich das Lernen in kooperativen Situationen auf den individuellen Lernerfolg auswirkt. Dazu dienen sowohl videografierte Gruppendiskussionen als auch individuell verfasste, schriftliche Statements. Konkret fokussiert die Studie zwei Fragestellungen: (a) Wie manifestieren sich die erlernten Heuristiken in Gruppendiskussionen und in den schriftlichen Statements während des Trainings und führt ihre Nutzung zu einer gesteigerten Qualität der Argumentation? (b) Gelingt der Transfer der Heuristiken aus dem kooperativen Setting (Gruppendiskussion) während des Trainings auf eine nachfolgende Situation, in der Evidenz individuell bewertet und genutzt werden muss, und lässt sich hier eine Konsistenz in der Argumentationsqualität feststellen?

Methode: Die Analyse erfolgt qualitativ (N=12) durch Auswertung der videografierten Gruppendiskussionen und eines auf Fallvignetten basierenden Testinstruments (Wilson, 2005; Heininger, 2013). Zur Auswertung dient ein in Anlehnung an bestehende Instrumente entwickeltes Kategoriensystem (Kuhn, 1997; Wilson, 2005; Bell, 2010; Felton & Kuhn, 2011). Mit Hilfe dieses Kategoriensystems werten zwei Rater unabhängig voneinander sowohl die Gruppendiskussionen als auch die schriftlichen Statements aus.

Zudem werden über Fragebögen und Interviews im Sinn von „Cognitive Labs“ Prozess- und Akzeptanzdaten erfasst, die eine zusätzliche Einschätzung der ablaufenden Lernprozesse und Wirkungen des Trainings erlauben. Als relevante Hintergrundvariablen werden epistemologische Überzeugungen, Überzeugungen zur Nutzung von Evidenz und die Ambiguitätstoleranz erfasst.

Ergebnisse: Die Durchführung und Auswertung der Studie läuft aktuell, so dass die Ergebnisse auf der Konferenz vorgestellt werden können. Erste Analysen der schriftlichen Statements zeigen eine hohe Interrater-Reliabilität der Rater-Einschätzung der Argumentationsqualität (Intraklassenkorrelation über alle Items hinweg ICC = .95; Übereinstimmungskoeffizienten der einzelnen Items erreichten zwischen 0.90 und 0.97). Parallel dazu werden aktuell die videografierten Gruppendiskussionen ausgewertet, um einen Vergleich zu den schriftlichen Statements zu ermöglichen. Auch diese Ergebnisse werden auf der Konferenz präsentiert werden können.

Freitag
13. März

9:00 Uhr

HGB 30

Chair(s)

Trempler, Kati
Hetmanek, Andreas

Klein, Martin; Wagner, Kai; Klopp, Eric; Stark, Robin

Förderung anwendbaren bildungswissenschaftlichen Wissens bei Lehramtsstudierenden anhand fehlerbasierten kollaborativen Lernens: Replikation und Nachhaltigkeit bisheriger Befunde

Theoretischer Hintergrund: Die Anwendung bildungswissenschaftlichen Wissens zur Erklärung komplexer schulischer Situationen ist eine der Grundlagen effektiven Lehrerhandelns im Schulalltag (Bromme, 2008; Neuweg, 2007). Eine solche Theorienutzungscompetenz im Sinne der Fähigkeit zu wissenschaftlich reflektiertem Handeln wird als Schlüsselqualifikation von Lehrkräften betrachtet (Meier, 2006). Studierende weisen hier jedoch Defizite auf. Diese führen zu häufigen Fehlern, wie etwa der Erklärung komplexer schulischer Situationen anhand von Alltagswissen oder fehlerhaften Umgang mit empirischer Evidenz (Stark, 2005). Um diese Fehler als Lerngelegenheit nutzbar zu machen, wurde eine auf instruktionalen Fehlern (Oser, 2007) basierende Lernumgebung entwickelt. Die Lernumgebung wurde als Bestandteil von regulären Seminaren der Lehramtsausbildung eingesetzt und von den Studierenden in Dyaden kollaborativ anhand externaler Kollaborationsskripts (Fischer, Kollar, Stegmann & Wecker, 2013) bearbeitet. Ziel ist die Förderung der Kompetenz zur Erklärung schulischer Situationen auf Basis bildungswissenschaftlicher Theorien. Die Lernwirksamkeit der Lernumgebung und verschiedener instruktionaler Unterstützungsmaßnahmen wurden in mehreren Studien unserer Arbeitsgruppe nachgewiesen (Wagner, Klein, Klopp & Stark, 2014; Wagner et al., eingereicht).

Fragestellung: Die vorliegende Studie untersucht zusätzlich zur Replikation bisheriger Ergebnisse insbesondere die Nachhaltigkeit der erzielten Lerneffekte sowie die Effekte einer erweiterten Seminarkonzeption durch eine Follow-up-Testung am Semesterende.

Methode: N = 143 Studierende bearbeiteten die oben beschriebene Lernumgebung im Rahmen regulärer Seminare des Lehramtsstudiums. Variiert wurde anschließend die Konzeption der auf die Bearbeitung der Lernumgebung folgenden Seminarphase. Eine systematisch auf der Konzeption der Lernumgebung aufbauende erweiterte Seminarkonzeption wurde anhand von Blended Learning und adaptiver instruktionaler Unterstützung umgesetzt und mit einer regulären Seminarkonzeption verglichen. Erhoben wurden konzeptuelles Fehlerwissen, strategisches Fehlervermeidungswissen, prozedurales Fehleridentifikationswissen und Erklärungswissen (vgl. Wagner et al., 2014).

Ergebnisse: Die Effekte der Bearbeitung der Lernumgebung aus früheren Studien konnten repliziert werden und waren auch noch am Semesterende nachweisbar. Die erweiterte Seminarkonzeption erbrachte keinen zusätzlichen Lernfortschritt, verbesserte aber die Nachhaltigkeit der Lerneffekte in zwei Wissensarten gegenüber der Kontrollgruppe. Insgesamt fielen die Ergebnisse am Semesterende deskriptiv nur geringfügig schlechter aus als unmittelbar nach der Bearbeitung der Lernumgebung, so dass die Nachhaltigkeit der Lerneffekte als gesichert gelten kann.

Wolter, Ilka

Heterogenität als Herausforderung: Befunde zur Kompetenzmessung im Rahmen des Nationalen Bildungspanels.

Im Rahmen dieses interdisziplinären Symposiums werden vier Beiträge vorgestellt, die sich mit unterschiedlichen Aspekten der Kompetenzmessung im Nationalen Bildungspanel (Blossfeld, Roßbach & von Maurice, 2011) beschäftigen und dabei die besonderen Bedeutungen und Auswirkungen der Zugehörigkeit zu einer spezifischen Teilpopulation untersuchen. Heterogenität wird bezogen auf sprachlichen und sozioökonomischen Hintergrund, sonderpädagogischen Förderbedarf und Geschlecht in den Blick genommen. In diesem Symposium wird die Relevanz von Heterogenitätsaspekten, wie dem Migrationshintergrund und sozioökonomischen Status, zunächst am Beispiel der Entwicklung von naturwissenschaftlichen Kompetenzen am Übergang vom Kindergarten zur Grundschule verdeutlicht (Beitrag 1). Weiterhin beschäftigen wir uns mit Fragen der Validität und Fairness in der Kompetenzmessung für unterschiedliche Gruppen und thematisieren Möglichkeiten zur Erhöhung dieser über Veränderungen in der Instruktion der Testaufgaben (Beiträge 2 und 3) oder der Auswahl von Testmaterial (Beitrag 4).

Im ersten Beitrag von Hahn und Schöps wird die Entwicklung naturwissenschaftlicher Kompetenzen unter Berücksichtigung des Migrations-, sozioökonomischen und sprachlichen Hintergrundes der Kinder untersucht. Die Autorinnen untersuchten Kinder am Übergang vom Kindergarten in die Grundschule und fanden bedeutsame Disparitäten im Hinblick auf den Migrationshintergrund der Kinder. Allerdings weisen ihre Befunde darauf hin, dass die Sprache den stärksten Einfluss auf die naturwissenschaftlichen Kompetenzen ausübt.

Auf diese Befunde aufbauend wird im zweiten Beitrag von Lockl, Händel und Artelt untersucht, wie die deklarative Metakognition mit dem sprachlichen Hintergrund bei Kindern der 3. Jahrgangsstufe zusammenhängt. Die Autorinnen konnten zeigen, dass die Leistungen in einem metakognitiven Wissenstest mit dem Wortschatz der Kinder korrelieren. Darüber hinaus verglichen sie zwei unterschiedliche Testbedingungen und fanden heraus, dass die Unterschiede aufgrund der sprachlichen Kompetenzen der Kinder dann geringer ausfallen, wenn die Kinder die vorgelesenen Aufgabentexte selbst mitlesen können.

Der dritte Beitrag von Nusser und Weinert greift die Frage nach der Gestaltung von Instruktionen von Kompetenztestaufgaben wieder auf und fokussiert Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf. Es wurde untersucht, ob über eine intensive Instruktionsphase gegenüber einer Standard-Instruktion (d.h. vergleichbar der Instruktion in Regelschulen) die Qualität der Kompetenzmessung in der 3. Grundschulklasse verbessert werden kann. Die Autorinnen konnten zeigen, dass eine aufmerksamkeitsfördernde und motivierende Instruktion die Testleistungen der Kinder im Textverständnis positiv beeinflusst.

Im vierten Beitrag von Wolter, Lockl, Koller, Gehrer und Artelt wird die Auswahl von Testmaterial auf die Kompetenzmessung thematisiert. Hierbei gingen die Autorinnen der Frage nach, ob die Geschlechtstypisierung von Inhaltsbereichen in Lesetexten das Textverständnis und die Selbsteinschätzung des eigenen Textverstehens von männlichen und weiblichen Versuchspersonen bedingt.

Abschließend diskutiert Kessels die Beiträge sowie Aspekte der Heterogenität in der Bildung und Bildungsforschung.

Freitag
13. März

9:00 Uhr

HGB 40

Chair(s)

Wolter, Ilka
Lockl, Kathrin**Heterogenität als Herausforderung: Befunde zur Kompetenzmessung im Rahmen des Nationalen Bildungspanels.**

Hahn, Inga; Schöps, Katrin

Die Bedeutung von Sprache und Herkunft für die Entwicklung naturwissenschaftlicher Kompetenz (vom Kindergarten bis zur ersten Klasse).

Theoretischer Hintergrund: Sprache und Herkunft wirken sich bereits früh auf die Entwicklung von Kompetenzen aus. Bisher wurden im Rahmen von Studien in diesem Bereich meist kognitive Fähigkeiten sowie die sprachliche und mathematische Kompetenzentwicklung betrachtet. Hier ergeben sich bereits gravierende Disparitäten, die sich über die weitere Bildungslaufbahn als relativ stabil erweisen oder sich sogar in einer Art Schereneffekt noch vergrößern (Alexander, Entwisle, & Olson, 2007; Schneider & Stefanek, 2004). In der aktuellen Literatur werden diese Disparitäten neben individuellen Faktoren des Kindes vor allem auf das Elternhaus zurückgeführt. Hier liegen sogenannte „Risikofaktoren“ (Lehrl, Kuger, & Anders, 2014) im sozio-ökonomischen Status sowie im Migrationshintergrund der Kinder bzw. ihrer Eltern (Mistry, Biesanz, Chien, Howes, & Benner, 2008). Dubowy et al. (2008) fanden im Hinblick auf Sprache und vorwissensabhängige Fertigkeiten wie Zählen und Allgemeinwissen signifikante Unterschiede zwischen deutschsprachigen Kindern und Kindern, die aus Elternhäusern mit mindestens einem nicht-deutschsprachigen Elternteil stammen. Diese Befunde erwiesen sich in der Studie auch unter Kontrolle des sozioökonomischen Status (ISEI) als stabil. Dennoch zeigten sich hier ebenfalls signifikante Zusammenhänge zwischen dem ISEI und den kindlichen Leistungen im Alter von 3 bis 4 Jahren. Der sozioökonomische Status trägt auch vermittelt über den Spracherwerb zu domänenübergreifenden Disparitäten in der Kompetenzentwicklung bei (Weinert, 2006). Kinder weniger gebildeter Mütter oder aus Familien mit geringerem sozioökonomischem Status zeigen einen langsameren Aufbau ihres Wortschatzes (Roberts, Jurgens, & Burchinal, 2005).

Trotz weitreichender Studien zu unterschiedlichen Kompetenzdomänen fehlen bisher differentielle Befunde zum Status und zur Entwicklung naturwissenschaftlicher Kompetenzen von Kindern im Elementar- und Primarbereich, die die genannten Einflussfaktoren in den Blick nehmen. Die aktuelle Forschung des Nationalen Bildungspanels (National Educational Panel Study – NEPS) möchte einen Beitrag dazu leisten, diese Lücke zu schließen.

Fragestellungen: Der vorliegende Beitrag befasst sich mit der Frage, auf welche Weise der sozioökonomische Hintergrund, die Sprache und speziell der Migrationshintergrund differentielle Einflüsse auf die Entwicklung der naturwissenschaftlichen Kompetenz haben. Weiterhin soll die Frage beantwortet werden, inwiefern es in der Entwicklung der naturwissenschaftlichen Kompetenz für bestimmte Risikokinder (geringer sozioökonomischer Status, verzögerter Wortschatzerwerb und oder Migrationshintergrund) vom Kindergarten bis zur ersten Klasse zu einem Schereneffekt kommt.

Methode: Die Stichprobe der Studie besteht aus 1956 Kindergartenkindern (w=50%, m=50%, im Mittel 4;9 Jahre alt) und 6511 Erstklässlern (w=49%, m=51%, im Mittel 6;6 Jahr alt). Im Längsschnitt liegen Daten von 534 Kindern (w=49%, m=51%) mit einem Messzeitpunkt im Kindergarten und in der ersten Klasse vor.

Die naturwissenschaftliche Grundbildung der Kinder wurde ebenso wie ihr Wortschatz anhand der entsprechenden NEPS-Tests für den Kindergarten und für die erste Klasse erhoben. Die Hintergrunddaten der Kinder zur Sprache zu Hause bzw. zum Migrationshintergrund und zum sozioökonomischen Status wurden anhand der Elternfragebögen erhoben.

Die Fragestellungen werden anhand von latenten Regressionen und Strukturgleichungsmodellen mit Mplus (Muthén & Muthén, 2012) beantwortet.

Ergebnisse: Erste Ergebnisse des ersten Messzeitpunkts im Kindergarten zeigen, dass sich Kinder mit und ohne Migrationshintergrund bereits zu diesem frühen Zeitpunkt im Alter von knapp 5 Jahren signifikant in ihrer naturwissenschaftlichen Kompetenz unterscheiden (Cohen's $d = 1.2$). In der Regressionsanalyse erweist sich hier die Sprache (Wortschatz) als Faktor mit der größten prädiktiven Kraft ($\beta=0.64$, $p<.05$) für die Ausprägung der naturwissenschaftlichen Kompetenz. Kinder mit geringerem Deutsch-Wortschatz zeigen demnach geringere naturwissenschaftliche Kompetenz.

Ausgehend von den Ergebnissen im Kindergarten wird der Vortrag Antworten darauf geben, inwiefern sich die gleichen Disparitäten bei Kindern der ersten Klasse zeigen. Darüber hinaus wird anhand der Längsschnittergebnisse der Kinder mit zwei Messzeitpunkten betrachtet, inwiefern sich diese Disparitäten vom Kindergarten zur ersten Klasse noch vergrößern.

Freitag
13. März

9:00 Uhr

HGB 40

Chair(s)

Wolter, Ilka
Lockl, Kathrin

Heterogenität als Herausforderung: Befunde zur Kompetenzmessung im Rahmen des Nationalen Bildungspanels.

Lockl, Kathrin; Händel, Marion; Artelt, Cordula

Kompetenztestung bei Kindern mit unterschiedlichem Sprachhintergrund – Effekte zweier Testbedingungen am Beispiel eines Tests zum metakognitiven Wissen.

Ein wesentliches Ziel des Nationalen Bildungspanels (NEPS) ist es, Kompetenzen zu erfassen, die als bedeutsam für Bildungsverläufe und gesellschaftliche Teilhabe betrachtet werden (Weinert et al. 2011). Die Messung der Kompetenzen soll dabei möglichst objektiv, reliabel und valide erfolgen. Um dem Anspruch einer hohen Validität gerecht zu werden, sollen konstruktirrelevante Hindernisse reduziert werden, so dass sich die Testergebnisse idealerweise ausschließlich auf die Konstrukte beziehen, die gemessen werden sollen. Dies ist bei Tests, die sprachlich vorgegeben werden, nicht immer ohne Einschränkungen gegeben, denn es besteht die Möglichkeit, dass Sprachkenntnisse im Deutschen die eigentliche Testleistung mitbeeinflussen.

Vor diesem Hintergrund wurde in der vorliegenden Studie der Frage nachgegangen, wie sprachliche Anforderungen in einer Testsituation variiert werden können, um für Kinder mit unterschiedlichem Sprachhintergrund eine möglichst valide und faire Messung von Kompetenzen zu ermöglichen. Konkret wurden die Auswirkungen zweier Testbedingungen (vorlesen vs. mitlesen) auf die Leistung in einem Test zum metakognitiven Wissen untersucht.

An der Studie, die eine Entwicklungsstudie im Rahmen von NEPS darstellte, nahmen 319 Schülerinnen und Schüler der 3. Jahrgangsstufe teil. Der eingesetzte Test zum metakognitiven Wissen erfasst das konditionale metakognitive Strategiewissen, d. h. das Wissen darüber, welche Strategie in Abhängigkeit von der jeweiligen Anforderung/Situation besonders geeignet ist (Flavell 1979; Paris et al. 1983). Die einzelnen Aufgaben bestehen jeweils aus einem Szenario und drei möglichen Handlungsoptionen. Die Schülerinnen und Schüler wurden gebeten, die Handlungsoptionen hinsichtlich ihrer Nützlichkeit in dem jeweiligen Szenario auf einer 3-stufigen Skala zu bewerten. Die Inhalte der Aufgaben wurden in den Aufgabenheften mit Hilfe von Bildern veranschaulicht. Um die Effekte unterschiedlicher sprachlicher Anforderungen zu untersuchen, wurden die Schülerinnen und Schüler zufällig einer von zwei Gruppen zugewiesen: einer Gruppe, in der sie den von der Versuchsleiterin oder dem Versuchsleiter vorgelesenen Aufgabentext mitlesen konnten („Mitlese-Bedingung“) oder einer Gruppe, in der sie den Text lediglich vorgelesen bekamen („Vorlese-Bedingung“). Der Sprachhintergrund der Kinder wurde durch die Frage erfasst, ob „Deutsch die vorwiegend zu Hause gesprochene Sprache“ darstellt. Zudem wurde zur Erfassung der Sprachkompetenzen ein Test zum rezeptiven Wortschatz (PPVT) durchgeführt.

Die Auswertung des Tests zum metakognitiven Wissen erfolgte anhand von Paarvergleichen (Nützlichkeit der Strategie X im Vergleich zur Nützlichkeit von Strategie Y) basierend auf zuvor eingeholten Expertenratings (vgl. Händel, Artelt & Weinert, 2013, Lockl, 2012). In Bezug auf ihre Leistung im Wortschatztest wurden die Schülerinnen und Schülern per Mediansplit in eine Gruppe mit hohem und eine Gruppe mit niedrigem Wortschatz eingeteilt.

Erste Analysen zeigen, dass die erzielten Leistungen im metakognitiven Wissenstest signifikant, aber eher gering mit dem Wortschatz der Kinder korrelieren. Zudem weisen Kinder mit deutschem Sprachhintergrund höhere Werte im metakognitiven Wissen auf als Kinder ohne deutschen Sprachhintergrund. Interessanterweise ergab eine zweifaktorielle Varianzanalyse eine signifikante Interaktion zwischen Sprachkompetenzen (Mediansplit Wortschatz) und den Testbedingungen: Deutliche Unterschiede im metakognitiven Wissen zwischen Kindern mit höherem vs. niedrigem Wortschatz zeigten sich in der „Vorlese-Bedingung“, nicht aber in der „Mitlese-Bedingung“. Ähnliche Befunde ergaben sich auch, wenn anstatt des Wortschatzes der Sprachhintergrund berücksichtigt wird. Insgesamt deuten die Ergebnisse darauf hin, dass die Möglichkeit, die Aufgabentexte mitlesen zu können, die Unterschiede zwischen sprachlich guten und weniger guten Schülern verringert und somit die Validität und Fairness der Testung erhöht. Die Ergebnisse werden im Hinblick auf die unterschiedlichen sprachlichen Anforderungen in den beiden Testbedingungen diskutiert (vgl. Hale et al., 2005). Aus den dargestellten Befunden lassen sich Implikationen für andere Tests ableiten, die sprachlich vorgegeben werden, deren Ziel aber nicht die Messung der sprachlichen Kompetenzen selbst ist.

Freitag
13. März

9:00 Uhr

HGB 40

Nusser, Lena; Weinert, Sabine

Adäquate Testinstruktionen für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf im Bereich Lernen: eine Interventionsstudie im Rahmen des Nationalen Bildungspanels.

Standardisierte Bedingungen und Testinstruktionen sichern – nicht nur in large-scale Studien – eine objektive, reliable und vergleichbare Erhebung von Kompetenzen (Pitoniak & Royer, 2001). Anleitung und Ausfüllhinweise sollen die teilnehmenden Personen gut auf die Aufgaben vorbereiten und einen optimalen Testablauf gewährleisten (Sax, 1997). In Bezug auf die Testung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf Lernen (SPF-L) können sich Instruktionen, die mit Blick auf Regelschulgruppen konzipiert wurden, jedoch als nicht adäquat erweisen und eine gruppenspezifische Hürde darstellen. Ergebnisse von Machbarkeitsstudien zur Erfassung der Lesekompetenz bei Schülerinnen und Schülern an Förderschulen in 5. Klassen im Rahmen des Nationalen Bildungspanels (NEPS; Blossfeld, Roßbach & von Maurice, 2011) zeigen ein erhöhtes Auftreten ungültiger Antworten, wie z.B. Mehrfachzustimmungen bei Multiple-Choice-Aufgaben, im Vergleich zu Schülergruppen an Hauptschulen (Südkamp, Pohl & Weinert, 2014), was als Indiz eines unzureichenden Verständnisses der Ausfüllhinweise interpretiert werden kann. So nennt Zielinski (1996) neben dem aufgabenspezifischen Vorwissen und der Lernmotivation auch die Fähigkeit, Instruktionen zu verstehen, als einen internen Bedingungsfaktor von Lernschwierigkeiten (siehe auch Vernooij, 2007). Entsprechend konnte in Studien eine verbesserte Testperformanz von Förderschülerinnen und -schülern erreicht werden, wenn eine vorgeschaltete Gewöhnungsphase und zusätzliche Einarbeitungszeit vor der Administration eines Matrizen-Tests zur Erfassung des Lernpotentials implementiert wurde (Hessels, 2009; siehe auch Scruggs & Williams, 1984).

Neben eingeschränktem Sprachverständnis kann auch eine reduzierte Aufmerksamkeitsspanne, die bei der Schülergruppe mit SPF-L häufig zu finden ist (Grünke, 2004), dazu führen, dass die Aufnahme der Ausfüll- und Ablaufhinweise möglicherweise nur eingeschränkt erfolgt. Bereits kurze Koordinationsübungen mit einer Dauer von 10 Minuten tragen zur Förderung der Konzentration bei und haben positive Effekte auf die Konzentrationsleistung (Budde et al., 2008). Dynamische Spiele werden beispielsweise bei Konzentrationstrainings eingesetzt, um Spannungen abzubauen und Regellernen zu fördern (Krowatschek, Krowatschek & Reid, 2011).

Ziel der vorliegenden Studie war es zu testen, ob durch eine vor der Testung des Textverständnisses (ELFE 1-6; Lenhard & Schneider, 2006) durchgeführte intensivere Instruktionsphase bzw. durch eine andere Testvorbereitungsphase eine Bearbeitungs- und Leistungsverbesserung bei Schülerinnen und Schülern mit SPF-L erzielt werden kann. In einem experimentellen Design, das neben der regulären Instruktion zwei weitere, experimentell-variierte, Bedingungen vorsieht, wurden die Effekte von zwei Interventionen (siehe unten) geprüft.

369 Schülerinnen und Schüler mit SPF-L im 3. und 4. Jahrgang nahmen an der Studie teil (Altersdurchschnitt: M = 9,71 Jahre, 34% weiblich). Insgesamt wurden an 49 Förderschulen in vier deutschen Bundesländern 92 Testsitzungen realisiert. Die teilnehmenden Schülerinnen und Schüler wurden zufällig einer von drei Bedingungen des Testdesigns zugewiesen. Die Studie inklusive computergestützter Datenerfassung wurde von einem externen Erhebungsinstitut mit in der Standardisierung geschulten Testleiterinnen und -leitern durchgeführt.

Neben der regulären Standard-Instruktion (Kontrollgruppe) wurden zwei Interventionen implementiert. Schülerinnen und Schüler in der Interventionsgruppe 1 begannen die Testsitzung mit einem gemeinsamen Bewegungsspiel zur Förderung der Aufmerksamkeit und erhielten danach die reguläre Standard-Instruktion. Als Intervention 2 wurde eine intensivierte Instruktion entwickelt, die neben altersgerechten motivierenden Aspekten auch zusätzliche konkrete Hinweise zu den Aufgaben und deren Bearbeitung sowie interaktive Übungsphasen enthielt. Diese erweiterte Instruktion sollte das Verständnis der Anleitung sicherstellen sowie die Schülerinnen und Schüler an die Testmodalitäten gewöhnen und im Hinblick auf die Testbearbeitung motivieren.

Erste Ergebnisse zeigen, dass Schülerinnen und Schüler, die an einer der beiden Interventionen teilgenommen hatten, signifikant bessere Ergebnisse im nachfolgenden Textverständnis test zeigten als die Kontrollgruppe. Besonders deutlich ist der Unterschied zwischen der Kontrollgruppe mit Standardinstruktion und der Interventionsgruppe 2, die zudem auch einen reduzierten Anteil nicht instruktionsgemäßer Angaben aufweist. Schülerinnen und Schüler mit unterschiedlichen kognitiven Grundfähigkeiten reagierten dabei unterschiedlich auf die beiden Interventionen, indem sich differenzierte Effekte der beiden verschiedenen Testvorbereitungsphasen zeigten. Diese Interaktionseffekte sowie Implikationen für Kompetenzerhebungen bei dieser heterogenen Schülergruppe werden im Vortrag diskutiert.

Freitag
13. März

9:00 Uhr

HGB 40

Chair(s)

Wolter, Ilka
Lockl, Kathrin

Heterogenität als Herausforderung: Befunde zur Kompetenzmessung im Rahmen des Nationalen Bildungspanels.

Wolter, Ilka; Lockl, Kathrin; Koller, Ingrid; Gehrler, Karin; Artelt, Cordula

Science-Fiction oder Liebesgeschichte? Der Zusammenhang der Geschlechtstypisierung von Inhaltsbereichen in Lesetexten mit dem Textverständnis und der prozeduralen Metakognition.

Studien zeigen, dass Lesen grundsätzlich als eine eher weibliche Tätigkeit wahrgenommen wird (z. B. Steffens & Jelenec, 2011). Damit sind viele Erklärungsmechanismen für das Leseverhalten und die Lesekompetenz von Mädchen und Jungen (bzw. Frauen und Männern) verbunden. Vor dem Hintergrund von Befunden, die zudem zeigen, dass Personen je nach ihrem eigenen Geschlecht systematisch unterschiedliche Präferenzen in Bezug auf bestimmte Inhaltsbereiche sowie Textsorten (z. B. Worthy, Moorman & Turner, 1999) beim Lesen zeigen, möchten wir überprüfen, ob die Geschlechtstypisierung von Themenbereichen des angebotenen Lesematerials das Textverständnis sowie die Selbsteinschätzung des eigenen Textverstehens (als einen Indikator der prozeduralen Metakognition) erklären kann.

In dieser Studie gehen wir der Frage nach, ob die Geschlechtstypisierung von Inhaltsbereichen in der Lesekompetenzmessung relevant für die prozedurale Metakognition von Personen ist. Ausgehend von Befunden, die zeigen, dass Interessen und Freizeitaktivitäten von Personen geschlechtstypisiert ausfallen (z. B. Athenstaedt, Mikula, & Bredt, 2009; Hyde, 2014; Sue, Rounds, & Armstrong, 2009), nehmen wir an, dass Personen das Wissen um generalisierte geschlechtstypische Präferenzen nutzen, um ihre Selbsteinschätzung zu ihren Leistungen in den Lesetexten abzugeben. Deshalb nehmen wir an, dass sich Personen in Inhaltsbereichen, die als typischer für ihr Geschlecht wahrgenommen werden, kompetenter einschätzen als sie tatsächlich sind, also ihre tatsächliche Leseleistung für diese Leseinhalte überschätzen. Darüber hinaus untersuchen wir, ob in den Themeninhalten, die dem anderen Geschlecht zugeordnet werden, tendenziell eine Unterschätzung des eigenen Textverstehens vorliegt.

Als Datengrundlage zur Beantwortung dieser Forschungsfrage dienen uns die Daten von über 1300 Personen aus 4 Studien zur Messung von Lesekompetenzen und prozeduraler Metakognition im Lebenslauf, die im Rahmen des Nationalen Bildungspanels (NEPS, Blossfeld, Roßbach & von Maurice, 2011) erhoben wurden. Die Messung der Lesekompetenzen erfolgte anhand der Rahmenkonzeption des NEPS (vgl. Artelt, Weinert & Carstensen, 2013; Gehrler, Zimmermann, Artelt & Weinert, 2013). Die aufgabenbezogene, prozedurale Metakognition wurde über die Abfrage der selbsteingeschätzten richtig beantworteten Aufgaben zu jeder Leseinheit erfasst (siehe Händel, Artelt & Weinert, 2013; Lockl, 2013).

Zur Beantwortung unserer Fragestellungen wurden in einem ersten Schritt die Leseinheiten von externen Expertinnen und Experten dahingehend beurteilt, ob die Lesetexte als geschlechtstypisch (männlich oder weiblich) oder geschlechtsneutral bezüglich des Themenbereichs (d. h. Inhaltsbereich) eingeschätzt werden können. Auf Grundlage dieser Kategorisierung wurde dann die Genauigkeit der Einschätzung der eigenen Leistung (d. h. die Differenzwerte der aufgabenbezogenen Selbsteinschätzung zu den entsprechenden Leseleistungen) der Personen analysiert.

Erste vorläufige Ergebnisse zeigen, dass Inhaltsbereiche und Themen in Lesetexten als geschlechtstypisch klassifiziert werden können. Tendenziell lässt sich bereits zeigen, dass Personen ihre individuelle Selbsteinschätzung an Inhaltsbereichen, die als zu ihrem eigenen Geschlecht zugehörig wahrgenommen werden, erwartungsgemäß adaptieren. Weitere regressionsanalytische Auswertungen sollen dann zeigen, wie robust diese interindividuellen Effekte aufgrund der Geschlechtstypisierung der Themenbereiche der Leseinheiten ausfallen. In zusätzlichen Analysen möchten wir der Frage nachgehen, ob das individuelle Leseverhalten diese Befunde erweitern kann. Außerdem interessieren wir uns dafür, wie die geschlechtstypisierten Themenbereiche der Leseinheiten mit verschiedenen Textsorten (z. B. Sachtexte, literarische Texte) korrespondieren. Wir erwarten, dass weitere Erklärungsmechanismen zu Geschlechtsunterschieden in der aufgabenbezogenen Metakognition und Lesekompetenz sichtbar werden. Die Befunde werden in Hinblick auf Präferenzen von Lesematerial von Mädchen und Jungen (bzw. Frauen und Männern) und dessen Implikationen für die Kompetenzentwicklung in diesem geschlechtskonnotierten Bereich diskutiert. Darüber hinaus lassen sich aus unseren Befunden wertvolle Hinweise zur Berücksichtigung der Testfairness in der Kompetenzmessung des Bereichs Lesen ableiten. Die ersten Ansätze werden vorgestellt und diskutiert.

Freitag
13. März

9:00 Uhr

HGB 40

Stürmer, Kathleen

Differenzielle Effekte bei der Arbeit mit Unterrichtsvideos in der Lehrerbildung

Unterrichtsvideos dienen in den letzten Jahren als ein vielversprechendes Tool, um einen praxisnahen Kompetenzaufbau in der Lehrerbildung zu unterstützen. Sie werden in der Lehraus- und -fortbildung insbesondere genutzt, um fachliche und überfachliche Lernprozesse zu initiieren sowie Lernergebnisse und den Kompetenzaufbau von (angehenden) Lehrpersonen empirisch zu messen und zu beschreiben. Dabei ist zahlreichen Studien der letzten Jahre die Grundannahme gemein, dass Unterrichtsvideos grundlegend dazu dienen, die Unterrichtspraxis als solche zu repräsentieren. Es wird somit davon ausgegangen, dass Videos einen authentischen Anforderungsrahmen für die Lehrerprofessionalisierung darstellen. Mit Blick auf die Erfassung und Unterstützung von Lernprozessen bei (angehenden) Lehrpersonen gilt allerdings: „Video is but a tool“ (Seago, 2004). Das bedeutet, dass basierend auf den individuellen Lernzielen geeignetes Videomaterial als auch passgenaue, instruktionale Designs gewählt werden müssen (Blomberg et al., 2013). Bisher fehlt es allerdings insbesondere an differenziellen Perspektiven in Bezug auf die beiden genannten Aspekte und einer Darstellung ihrer Effekte auf Lernprozesse und Lernergebnisse. Dazu liefert das Symposium einen Beitrag, indem empirische Studien aus vier, interdisziplinären Wissenschaften (Erziehungswissenschaft, Psychologie, Lehrerbildung) für die Lehraus- und -fortbildung vorgestellt werden. In Beitrag 1 und 2 steht insbesondere die Rolle des Videomaterials für die Erfassung und Entwicklung Professioneller Unterrichtswahrnehmung im Vordergrund. Beitrag 3 und 4 fokussieren vor allem instruktionale Designs zur Einbettung von Videos zur Unterstützung von Wissen, motivationalen Orientierungen und performativen Prozessen.

Beitrag 1: Stürmer & Seidel

Im Beitrag wird gezeigt, dass die Beobachtung von Unterrichtsvideos mit unterschiedlichen, individuellen Bearbeitungsprozessen (Emotion/Interesse) bei Lehramtsstudierenden verschiedener Schularten einhergeht, was sich wiederum auf ihre Fähigkeit zur Professionellen Unterrichtswahrnehmung auswirkt.

Beitrag 2: Hellermann et al.

Im Beitrag wird gezeigt, dass in einem quasi-experimentell variierten Trainingsprogramm zur Professionellen Unterrichtswahrnehmung in der Grundschullehrerausbildung insgesamt die Kombination aus fremden und eigenen Videos die höchsten Lerneffekte zeigt.

Beitrag 3: Kumschick et al.

Im Beitrag wird auf der Grundlage videobasierter Fallarbeit gezeigt, dass variierte Instruktionen differenzielle Effekte u.a. hinsichtlich des Wissenserwerbs und der Motivation bei Lehramtsstudierenden hervorrufen.

Beitrag 4: Jerabek et al.

Im Beitrag wird im Rahmen einer videobasierten Intervention in der Lehrerfortbildung gezeigt, dass Lehrpersonen im Austausch mit einem Facilitator eine positive Diskussionskultur entwickeln und differenzielle Bezüge im Transfer der Fortbildungsinhalte für den eigenen Unterricht im Verlauf eines Jahres herstellen.

Diskussion: Fritz Oser

Chair(s)

Stürmer, Kathleen
Gröschner, Alexander

Stürmer, Kathleen; Seidel, Tina

Emotion und Interesse bei der Videobeobachtung: Effekte auf die wahrgenommene Repräsentativität der gesehenen Situation und die Professionelle Unterrichtswahrnehmung bei Lehramtsstudierenden

Unterrichtsvideos sind im Rahmen von Lehrerbildung ein vielversprechendes Tool geworden um anwendungsbezogene Aspekte professionellen Wissensaufbau zu unterstützen und zu beschreiben (König et al., 2014; Santagata & Angelici, 2010; Seidel & Stürmer, 2014). Gerade in der universitären Lehrerbildung wird ihr Potenzial darin gesehen, die Aneignung theorie-praxisintegrierender Wissensbestände zu untersuchen (Seidel, Blomberg, & Renkl, 2013; Stürmer, Königs, & Seidel, 2013, 2014), indem die Fähigkeit von Lehramtsstudierenden fokussiert wird Unterrichtssituationen adäquat zu beobachten und zu interpretieren. Mit dem Anspruch formativer Evaluation von Lernprozessen, sind in den letzten Jahren reliable, standardisierte Instrumente entwickelt wurden, in denen Unterrichtsvideos mit geschlossenen Rating Items kombiniert werden um Professionelle Unterrichtswahrnehmung zu erfassen (z.B. Seidel & Stürmer, 2014). Dabei wird angenommen, dass Unterrichtsvideos einen authentischen Anwendungskontext im Sinne der Kontextrepräsentativität als Voraussetzung für die Kontextgenauigkeit von Kompetenzerfassung darstellen (Shavelson, 2012). Allerdings zeigen erste Studien, dass die Bearbeitung von Unterrichtsvideos mit unterschiedlichen, emotionalen sowie motivationalen Prozessen einhergeht (z.B. Seidel, Stürmer, Blomberg, Kobarg, & Schwindt, 2011). In dieser Studie fokussieren wir deshalb auf die Frage, inwiefern individuelle Prozesse bei der Beobachtung von Unterrichtsvideos die Wahrnehmung der Repräsentativität der beobachteten Situation beeinflussen und inwieweit dies mit der Ausprägung in der Professionellen Unterrichtswahrnehmung im Rahmen videobasierter, standardisierter Kompetenzerfassung zusammenhängt.

Fragestellungen: (1) Wie hängen individuelle Bearbeitungsprozesse (Emotion/Interesse) bei der Beobachtung mit der wahrgenommenen Repräsentativität (Authentizität/Immersion) der Unterrichtsvideos zusammen? (2) Wie wirkt sich der Zusammenhang zwischen individuellen Bearbeitungsprozessen (Emotion/Interesse) und der wahrgenommenen Repräsentativität (Authentizität/Immersion) auf die Kontextgenauigkeit (Professionelle Unterrichtswahrnehmung/ Selbstwirksamkeitserwartung) aus?

Methode: An der Studie nahmen N = 247 Lehramtsstudierende unterschiedlichen Semesters (M = 4.14, SD = 1.30), Fächerkombinationen und Studienstandorten teil. Alle Teilnehmer bearbeiteten das online Tool Observer, welches die Professionelle Unterrichtswahrnehmung mit Blick auf pädagogisch-psychologische lernwirksame Unterrichtskomponenten erfasst ($\alpha = .95$). Im Tool werden den Teilnehmern sechs 2-4 minütige Unterrichtsvideos aus verschiedenen Fächern präsentiert. Die Unterrichtsvideos sind in Rating Items eingebettet, die auf die Interpretationsfähigkeit der Lehramtsstudierenden fokussieren. Die Antworten der Teilnehmer werden mit einer Expertennorm abgeglichen. Zusätzlich zur Professionellen Unterrichtswahrnehmung wird im Tool nach jedem Unterrichtsvideo die Selbstwirksamkeit die beobachtete Unterrichtssituationen adäquat einschätzen zu können, erfasst.

Die individuellen Bearbeitungsprozesse wurden durch drei Skalen nach jedem Unterrichtsvideo erfasst: negative Emotion (Beispielitem: „Das Gesehene hat mich verärgert“, 1' trifft nicht zu, 4' trifft zu, $\alpha = .84$), positive Emotion (Beispielitem: „Das Gesehene hat mir Freude bereitet“, 1' trifft nicht zu, 4' trifft zu, $\alpha = .78$) und Interesse (Beispielitem: „Das Gesehene war für mich interessant“, 1' trifft nicht zu, 4' trifft zu, $\alpha = .81$).

Die Kontextrepräsentativität wurde durch zwei Skalen nach jedem Unterrichtsvideos sowie nach der Gesamtbearbeitung des Instruments erfasst: Authentizität (Beispielitem: „Das Gesehene empfand ich als authentisch“, 1' trifft nicht zu, 4' trifft zu, $\alpha = .74$) und Immersion (Beispielitem: „Ich war voll und ganz bei der Sache“, 1' trifft nicht zu, 10' trifft zu, $\alpha = .86$).

Ergebnisse: Erste regressionsanalytische Auswertungen zeigen mit Blick auf die Fragestellungen, dass (1) die individuellen Bearbeitungsprozesse mittel bis stark mit der wahrgenommen Repräsentativität der Unterrichtsvideos zusammenhängen, und (2) sich die wahrgenommen Repräsentativität der Unterrichtsvideos auf die Ausprägung der Professionellen Unterrichtswahrnehmung sowie der Selbstwirksamkeitserwartung auswirkt. Dieser Effekt wird durch die individuellen Bearbeitungsprozesse moderiert.

Diskussion: Die vorliegende Studie spricht dafür eine differenzielle Perspektive bei der standardisierten Kompetenzerfassung Professioneller Unterrichtswahrnehmung einzunehmen. Sie zeigt die Bedeutung individueller Faktoren wie Emotion und Interesse ausgelöst durch die Arbeit mit Unterrichtsvideos auf und unterstützt die Diskussion um die Frage, wie ein professioneller Blick losgelöst von solchen Faktoren bei Lehramtsstudierenden geschult werden kann.

Freitag
13. März

9:00 Uhr

HGB 50

Chair(s)

Stürmer, Kathleen
Gröschner, Alexander

Differenzielle Effekte bei der Arbeit mit Unterrichtsvideos in der
Lehrerbildung

Hellermann, Christina; Gold, Bernadette; Holodynski, Manfred

Förderung der klassenführungsspezifischen Unterrichtswahrnehmung – Wie effektiv ist ein Training anhand eigener bzw. fremder Unterrichtsvideos?

Klassenführung stellt eine der bedeutsamsten Lehrerkompetenzen dar. Sie hat Effekte auf die Motivation und Schulleistungen der Schüler_innen (u.a. Hattie, 2009; Kunter, Baumert & Köller, 2007), sowie die Lehrer_innenbelastung (Klusmann, Kunter, Voss, & Baumert, 2012). Gleichzeitig nennen junge Lehrpersonen klassenführungsbezogene Aspekte als einflussreichstes Problemfeld (Veenman, 1984) und fühlen sich diesbezüglich schlecht vorbereitet (Fideler & Haskelhorn, 1999). Es stellt sich demzufolge die Frage, wie konzeptuelles Wissen über Klassenführung bereits im Verlauf des Studiums praktisch eingebettet und gefördert werden kann.

Als Vermittler zwischen Wissen und Handeln wird die professionelle Unterrichtswahrnehmung gesehen (Schwindt, 2008). Diese Fähigkeit, lernrelevante Situationen im Unterricht bemerken und theoriegeleitet interpretieren zu können (Sherin, 2007), kann - im Gegensatz zur Handlungskompetenz - bereits in der ersten Phase der Lehramtsausbildung gefördert werden. Ein prominenter Ansatz zur Theorie-Praxis-Verknüpfung in der Lehrerbildung stellt der Einsatz von Unterrichtsvideos dar (Blomberg, Renkl, Sherin, Borko & Seidel, 2013). Es gibt jedoch bisher kaum Erkenntnisse darüber, wie sich unterschiedliche Förderkonzepte auf die professionelle Wahrnehmung auswirken.

Zur Überprüfung dieser Frage wurden zwei videobasierte universitäre Seminare zur Förderung der professionellen Wahrnehmung von Klassenführung in einem quasi-experimentellen Prä-Post-Follow-Up-Kontrollgruppen-Design verglichen. Die erste Trainingsgruppe (N = 36) analysierte Videomaterial fremder Lehrpersonen. Die zweite Trainingsgruppe (N = 32) analysierte sowohl Videos aus dem eigenem als auch aus fremdem Unterricht. Die unbehandelte Kontrollgruppe (N = 28) erhielt keine Förderung zum Thema Klassenführung.

Der zweiten Trainingsgruppe wurde im Gegensatz zur ersten Trainingsgruppe ermöglicht, eine Unterrichtssituation sowohl aus der Innenperspektive (aus der Sicht des Lehrenden) als auch aus der Außenperspektive (aus der Sicht eines externen Beobachters) zu analysieren. Wir gehen davon aus, dass der Abgleich dieser Perspektiven zu einem höheren Lernzuwachs in der professionellen Wahrnehmung von Klassenführung führt.

Die Messung im Prä-Post und Follow-Up-Test erfolgte mit einem bereits validierten videobasierten Diagnoseinstrument, das aus vier kurzen Videoclips aus dem naturwissenschaftlichen Grundschulunterricht und insgesamt 47 Ratingitems zu den Klassenführungsfacetten Monitoring ($\alpha = .81$), Strukturierung ($\alpha = .77$) und Regeln / Routinen ($\alpha = .81$) besteht (Gold & Holodynski, eing.).

Die Ergebnisse der Kovarianzanalysen zeigten, dass das Training mit fremden und eigenen Unterrichtsvideos die professionelle Wahrnehmung am effektivsten fördert ($\eta_p^2 = .145$). Ergebnisse des Follow-Up-Tests einige Monate nach der Intervention werden im Rahmen des Vortrags präsentiert.

In der abschließenden Diskussion wird die Bedeutung der Ergebnisse mit Hinblick auf die Generalisierbarkeit erörtert. Die Dropouts, die sich nicht gleichmäßig auf die Interventionsgruppen verteilen, werden diskutiert.

Freitag
13. März

9:00 Uhr

HGB 50

Chair(s)

Stürmer, Kathleen
Gröschner, Alexander

Differenzielle Effekte bei der Arbeit mit Unterrichtsvideos in der
Lehrerbildung

Kumschick, Irina Rosa; Piwowar, Valentina; Barth, Victoria Luise; Krysmanski, Katharina; Ophardt, Diemut; Thiel, Felicitas

Differenzielle Effekte eines videobasierten Trainings zum Klassenmanagement

Der Umgang mit Unterrichtsstörungen ist eine zentrale Herausforderung für Novizen-Lehrkräfte (Admiraal, Korthagen, & Wubbels, 2000) und ein wichtiger Teilbereich des Klassenmanagements (Ophardt & Thiel, 2013). Da störungsspezifisches Fallwissen im Gegensatz zu Fach- und Fachdidaktischem Wissen schwer lehrbar ist und insbesondere im praktischen Handlungsvollzug angeeignet und erworben wird, brauchen Lehramtsstudierende eine gezielte Lernunterstützung in Form von Classroom Case Exemplars (Moreno & Valdez, 2007; Van Tartwijk & Driessen, 2009). Während bisherige Studien den Wert videobasierter Lerngelegenheiten zur Aneignung von Klassenmanagement bzw. dem Umgang mit Unterrichtsstörungen nachweisen konnten, liegen kaum Befunde vor, die das Lernarrangement, in das diese Fallarbeit eingebettet ist, untersuchen. Daher haben wir ein Training zum Aufbau von störungspräventivem und intervenierendem Wissen für Studierende entwickelt, in dem das Lernsetting auf der Grundlage videobasierter Fallarbeit variiert wird. Im Rahmen einer quasi-experimentellen Prä-Post Studie wurden 148 Lehramtsstudierende zwei Versuchsbedingungen zugeordnet: Dem klassischen problem-basierten Lernen (PBL), in dem die zentralen theoretischen Konzepte und empirischen Befunde zum wirksamen Umgang mit Störungen selbstständig in Gruppen erarbeitet wurden (N=73) oder dem instruierten problem-basierten Lernen (I-PBL), in dem die Studierenden in den ersten beiden Seminarsitzungen eine umfassende theoretische Instruktion erhielten (N=75). Beide Gruppen arbeiteten anschließend in Kleingruppen mit videobasierten, in Bezug auf den Umgang mit Störungen problematischen Fällen. Hierbei wurden dysfunktionale Strategien der Lehrkraft identifiziert und alternative, funktionale Strategien entwickelt.

Wir nahmen an, dass nach Teilnahme an den Interventionen (1) Studierende beider Gruppen ihre Lehrerselbstwirksamkeit und ihre Kompetenzen im Klassenmanagement verbessern (erfasst durch Selbsteinschätzung) (2) Studierende der I-PBL Gruppe jedoch mehr Wissen aufweisen als Studierende der PBL Gruppe (erfasst durch einen Multiple-Choice Wissenstest) (3) Studierende der PBL Gruppe hingegen eine höhere Motivation angeben als Studierende der I-PBL Gruppe (erfasst durch Selbsteinschätzung)

Die Hypothesen konnten bestätigt werden. Die differenziellen Befunde ergeben Implikationen für die Lehrerbildung und werden diskutiert.

Freitag
13. März

9:00 Uhr

HGB 50

Jerabek, Martina; Gröschner, Alexander; Seidel, Tina

Lernen von Lehrpersonen in einer videobasierten Fortbildung: Empirische Befunde zur Lernatmosphäre, Diskussionskultur und dem Transfer von Fortbildungsinhalten

Theoretischer Hintergrund: Lehrerfortbildungen dienen der Unterstützung von kontinuierlichen Lernprozessen über die Berufskarriere (Vermunt & Endedijk, 2011). Bisherige Forschung betont insbesondere die Berücksichtigung effektiver Komponenten von Fortbildungen (Borko, 2004; Desimone, 2009; Lipowsky, 2010; Wilson, 2013), die Relevanz der Initiierung einer professionellen Lerngemeinschaft (Bolam et al., 2005) sowie die Unterstützung des Transfers von Lerninhalten in die berufliche Praxis (z.B. durch praxisrelevante Videobeispiele) (vgl. Gröschner, Seidel, Pehmer et al., 2014; Tripp & Rich, 2012; van Es et al., 2014). Dabei stehen häufig vor allem Veränderungen im Wissen und unterrichtlichen Handeln bzw. Lernerfolge der Schüler/-innen im Mittelpunkt. Selten werden die eigentlichen Lernprozesse der Lehrpersonen in den Fortbildungen sowie die Gestaltung der Lehrerfortbildungen differenzierter auch im Kontext der „Facilitation“ empirisch untersucht (Lipowsky, 2010). Erste Studien, die die Lernprozesse der Lehrpersonen fokussieren, zeigen, dass u.a. eine positive Lernatmosphäre, eine wertschätzende Gesprächskultur sowie eine Verdeutlichung eigener Handlungsrountinen das Lernen von Lehrpersonen fördern (Borko et al., 2008; Gröschner et al., 2014). Die vorliegende Studie leistet dazu einen Beitrag, indem im Rahmen einer einjährigen, videobasierten Lehrerfortbildung zur Klassengesprächsführung (dem sog. „Dialogischen Videozirkel“, DVC) die Lernprozesse der Lehrpersonen auch in Bezug auf die lernwirksame Gestaltung der Fortbildung im Mittelpunkt standen (Gröschner, Seidel, Kiemer et al., 2014).

Fragestellungen: (1) Inwiefern besteht im DVC eine positive Lernatmosphäre zwischen den Lehrpersonen und mit dem Facilitator? (2) Inwieweit besteht eine konstruktive Diskussionskultur und inwiefern verändert sich diese im Verlauf der Fortbildung?(3) Inwiefern werden im DVC von den Lehrkräften Bezüge zum eigenen Unterricht hergestellt und inwieweit wird der Transfer der Fortbildungsinhalte in dem eigenen Unterricht thematisiert? Inwiefern zeigen sich Unterschiede zwischen den Lehrpersonen je nachdem, ob das eigene oder ein fremdes Unterrichtsvideo gezeigt wird?

Methode: Im Schuljahr 2011/12 nahmen sechs Lehrkräfte (Alter $M=39.5$, $SD=5.43$) mit mathematisch-naturwissenschaftlichen Unterrichtsfächern an zwei Durchläufen des DVC teil (Gröschner et al., 2014). Jeder Videozirkel (DVC 1 und DVC 2) bestand aus drei ineinandergreifenden Workshops (ca. zwei Stunden) und einer gefilmten Unterrichtsstunde jeder Lehrperson. Der erste Workshop zielte darauf ab, bestehende Unterrichtspläne mit Fokus auf das produktive Klassengespräch zu überarbeiten und anzupassen. Anschließend wurde die vorbereitete Unterrichtsstunde jeder Lehrperson gefilmt. In Workshop 2 und 3 wurden Sequenzen der gefilmten Unterrichtsstunde gezeigt und reflektiert.

Die Workshops ($N=6$) wurden videografiert und mittels eines hoch-inferenten Kodierschemas (Gröschner & Jerabek, 2014) ähnlich vorliegender Studien in 2-Minuten-Intervallen von zwei unabhängigen Ratern kodiert. Dabei wurden für den vorliegenden Vortrag drei Aspekte berücksichtigt: Lernatmosphäre (Beispielitem: „Die Atmosphäre in der Gruppe ist wertschätzend“; ICC = 1.00), Diskussionskultur (Beispielitem: „Es wird auf Regeln der Gesprächsführung hingewiesen“; ICC = .90) und Transfer der Fortbildungsinhalte (Beispielitem: „Die Lehrkräfte identifizieren sich mit dem gezeigtem Unterricht“; ICC = .78).

Ergebnisse: (1) Hinsichtlich der ersten Forschungsfrage zeigt sich, dass grundsätzlich eine positive Lernatmosphäre zwischen den Lehrpersonen untereinander und mit dem Facilitator über die beiden DVC hinweg besteht. (2) Bezüglich der zweiten Frage wird gezeigt, dass zwischen den Lehrpersonen und mit dem Facilitator eine konstruktive Diskussionskultur besteht, die sich im Verlauf der Fortbildung zugunsten einer stärkeren Initiierung der Lehrkräfte bei Abnahme durch den Facilitator verändert.(3) Schließlich wird deutlich, dass die Lehrpersonen verstärkt Bezüge zu ihrem Unterricht herstellen und somit den Transfer der Fortbildungsinhalte thematisieren. Dabei zeigt sich, dass die Reflexion des eigenen Unterrichtsvideos zu einer stärkeren Identifikation mit dem eigenen Unterricht gegenüber der Betrachtung von fremden Unterrichtsvideos führt.

Diskussion und Implikationen: Die vorliegende Studie zeigt in differenzieller Weise die Bedeutung eines lernwirksamen Austauschs zwischen Lehrpersonen mit einem Facilitator im Rahmen einer videobasierten Fortbildung. Es wird aufgezeigt, wie Lernprozesse von Lehrpersonen in Bezug auf die lernwirksame Gestaltung von Fortbildungen gefördert werden können. Zukünftige Forschung zu Lehrerfortbildung sollte den Blick noch stärker auf Lernprozesse von Lehrpersonen innerhalb des Angebots legen (Lipowsky, 2010).

Freitag
13. März

9:00 Uhr

HGB 50

Chair(s)

Stürmer, Kathleen
Gröschner, Alexander

3

Gebauer, Miriam Marleen

Sozial-kognitive Merkmale des Lernens: Determinanten, Bedeutung und Unterrichtsgestaltung

Um Wissenserwerbsprozesse von Schülerinnen und Schülern im Rahmen des schulischen Unterrichts zu optimieren, ist eine Vielzahl an Aspekten zu berücksichtigen. Von hoher Relevanz sind dabei sozial-kognitive Merkmale von Schülerinnen und Schülern, deren Bedeutung für schulischen Bildungserfolg in zahlreichen Studien festgestellt werden konnte (Usher & Pajares, 2008). Ziel des Symposiums ist es Beiträge zusammenzubringen, die sich dieser Thematik aus unterschiedlichen Perspektiven nähern und lern- und leistungsunterstützende Determinanten des schulischen Unterrichts untersuchen. Die gemeinsame Fragestellung des Symposiums lautet: Welche Determinanten sind von Bedeutung für sozial-kognitive Merkmale im Lernprozess?

Der erste Beitrag von Gebauer et al. wendet sich der Vorhersage der Selbstwirksamkeit durch die basic human needs zu. Die fundamentalen Voraussetzungen für motiviertes Handeln ähneln konzeptionell den Determinanten der Selbstwirksamkeitsüberzeugung, eine funktionale Bedeutung der Autonomieunterstützung kann jedoch theoretisch nicht für die Selbstwirksamkeit angenommen werden. Die Ergebnisse stützen die theoretischen Postulate - Nur das Erfüllen des Kompetenzerlebens und der sozialen Eingebundenheit sagt die Selbstwirksamkeit vorher.

Der zweite Beitrag von Zander et al. betrachtet die Selbstwirksamkeit, eine Facette der motivationalen Dynamik, als Moderator zwischen Lernumgebung und Lernverhalten. Es werden in diesem Beitrag differenzielle Muster für die affektive und kognitive Dimension des Umgangs mit Fehlern berichtet, die unter Einbezug sozial-kognitiver Theorien und aktueller Ansätze ko- und selbst-regulierten Lernens diskutiert werden.

Beitrag drei von Schöber et al. untersucht die Akkuratheit der Einschätzung der Lehrenden bezogen auf die Selbstwirksamkeit der Schülerinnen und Schüler, die sie unterrichten. Eine adäquate Einschätzung der Schülermerkmale gilt als Voraussetzung für individuellen Unterricht. Die Ergebnisse weisen darauf hin, dass die Fähigkeit der Lehrenden Selbstwirksamkeit der Kinder einzuschätzen mit der Einschätzungsleistung der schulische Leistungen korrespondiert.

Der vierte Beitrag von Dresel et al. wendet sich den unterrichtlichen Strukturen und dessen Bedeutung für einen konstruktiven Umgang mit Fehlern zu. Dabei stehen hier neben sozial-kognitiven Merkmalen auch Leistungsmerkmale von Schülerinnen und Schülern als abhängige Variable im Fokus. Die längsschnittliche Untersuchung zeigt, dass sich der konstruktive Umgang mit Fehlern tendenziell und ein konstruktives Fehlerklima mit größeren objektiven Lernzuwächsen assoziiert ist.

Die Heterogenität des Symposiums zeichnet sich dadurch aus, dass es sozial-kognitive Merkmale im Bildungskontext aus verschiedenen Perspektiven (Erziehungswissenschaft, Psychologie und Lehrerbildung in ihrer Bedeutung betrachtet). Beitrag eins ist für die pädagogisch-psychologische Forschung von Relevanz. Beiträge zwei und vier sind in der Erziehungswissenschaft (Relevanz für schulische Praxis) und Pädagogischen Psychologie zu verorten. Beitrag drei liefert Erkenntnisse, die für die Lehrerbildung nutzbar gemacht werden können.

Freitag
13. März

9:00 Uhr

HGC 20

Chair(s)

Gebauer, Miriam Marleen
McElvany, Nele**Sozial-kognitive Merkmale des Lernens: Determinanten, Bedeutung und Unterrichtsgestaltung**

Gebauer, Miriam; McElvany, Nele; Ferdinand, Hanna; Huelmann, Thorben; Bos, Wilfried; Köller, Olaf; Schöber, Christian

Zur Bedeutung der basic human needs für sozial-kognitive Merkmale

Theoretischer Hintergrund: Kompetenzüberzeugungen und Wirksamkeitserwartung sind für die motivationale Dynamik von zentraler Bedeutung (Schunk & Meece, 2005). Das Konzept der Selbstwirksamkeitsüberzeugung (SW) (Bandura, 1997) definiert die selbstbezogene Kompetenzüberzeugung bezogen auf zukünftige Handlungen und die Einschätzung des zu erwartenden Erfolgs oder Misserfolgs. Empirische Evidenzen zeigen die Bedeutung der Selbstwirksamkeit im schulischen Kontext in zahlreichen Zusammenhängen. So ist die bildungsbezogene und fachspezifische Selbstwirksamkeit von wesentlicher Relevanz für die Leistungen von Schülerinnen und Schülern in allen Altersstufen (Usher & Pajares, 2008). Deci und Ryan (2000) postulieren in ihrem motivationstheoretischen Modell eine organismic-dialectical Perspektive, die Motivation hinsichtlich ihrer Qualität und Ausrichtung bestimmen sowie das Erfüllen von basic human needs als die Voraussetzungen für motivierte Handlungen darstellt. Dabei sind diese Motivationskonzepte (extrinsisch/intrinsische Motivation) ebenfalls von Bedeutung für die schulischen Leistungen von Schülerinnen und Schülern (Wang, 2008). Die fundamentalen Voraussetzungen für motiviertes Handeln ähneln konzeptionell den Determinanten der Selbstwirksamkeitsüberzeugung. Die Erfüllung von Kompetenzerleben, sozialer Eingebundenheit und Autonomieerleben finden sich in den Ausführungen Banduras ebenfalls in vergleichbarer Form als Bedingungen für die Herausbildung von Selbstwirksamkeit (Bandura, 1997). Empirische Studien konnten Zusammenhänge zwischen intrinsischer Motivation und Selbstwirksamkeit zeigen (Yi-Guang, McKeachie & Yung, 2003). Die Bedeutung der Erfahrung von Autonomieunterstützung (Gillet, Vallerand & Lafrenière, 2011), sozialer Eingebundenheit (Kunter, 2004) und von Kompetenzerleben (Tsai, Kunter, Lüdtke, Trautwein & Ryan, 2008) für motiviertes Lernen von Schülerinnen und Schülern wurde in zahlreichen Studien untersucht. Studien, die beide Theorien integrieren und die Bedeutung des Erfüllens der basic human needs für die Selbstwirksamkeit untersuchen, fokussieren zumeist das Erleben von Autonomieunterstützung (Velez, 2008). Ziel der vorliegenden Untersuchung ist die Prüfung der Bedeutung des Erfahrens der basic human needs im schulischen Kontext für die bildungsbezogene und fach- und domänenspezifische Selbstwirksamkeit. Wobei davon auszugehen ist, dass Autonomieerleben weniger Relevanz für die Selbstwirksamkeit hat, hingegen die Erfüllung des Kompetenzerlebens und der sozialen Eingebundenheit von Bedeutung für bildungsbezogene und fach- und domänenspezifische Selbstwirksamkeit sind (Bandura, 1998, Niemiec & Ryan, 2009)

Fragestellung: (1) Sind die basic human needs im Rahmen des Deutsch- und Mathematikunterrichts der siebten Klassenstufe empirisch trennbar? H1: Es wird erwartet, dass jedes Konstrukt als separate, untereinander korrelierte Faktoren spezifizierbar ist. (2) Werden die basic human needs im Rahmen des Deutsch- und Mathematikunterrichts der siebten Klassenstufe erfüllt? H2: Die basic human needs können im schulischen Unterricht erfüllt werden. (3) Lassen sich die domänen- und bildungsbezogene Selbstwirksamkeit durch die basic human needs vorhersagen? H3: Die Erfüllung der basic human needs sind von positiver Bedeutung für die bildungsbezogene und fach- und domänenspezifische Selbstwirksamkeit von Schülerinnen und Schülern.

Methode: Die Stichprobe umfasst N = 1.350 Schülerinnen und Schüler der siebten Jahrgangsstufe (MAlter = 12.15; 49.5% weiblich). Die Faktorstruktur wurde mittels konfirmatorischer Faktorenanalyse (CFA) überprüft, die Zusammenhänge anhand eines Strukturgleichungsmodells mit latenten Variablen unter Berücksichtigung der Clusterung der Daten.

Ergebnisse: Anhand einer CFA konnte die theoriekonforme Drei-Faktorenstruktur bestätigt werden. Die Mittelwerte der Skalen zur Erfassung der basic human needs liegen im positiven Bereich der Antwortskala.

Die Ergebnisse zur Überprüfung der dritten Forschungsfrage zeigen akzeptable Modellfits ($\chi^2(df) = 1844.041(418)$; CFI = .941; TLI = .935; RMSEA = .050; SRMR = .034). Es konnten statistisch signifikante ($p < .05$) prädiktive Zusammenhänge zwischen Kompetenzerleben und bildungsbezogener ($\beta = .351$), mathematikbezogener ($\beta = .341$) und lesebezogener ($\beta = .249$) Selbstwirksamkeit festgestellt werden. Auch gelang die Vorhersage der bildungsbezogenen ($\beta = .308$), mathematikbezogenen ($\beta = .201$) und lesebezogenen ($\beta = .232$) Selbstwirksamkeit durch das Erleben der sozialen Eingebundenheit. Die Ergebnisse bestätigen die angenommene Bedeutung der basic human needs. Somit ist die Zusammenführung der Theorien gewinnbringend und Erkenntnisse können in die schulische Praxis umgesetzt werden, um den Wissenserwerbsprozess im Rahmen des Unterrichts zu optimieren.

Chair(s)

Gebauer, Miriam Marleen
McElvany, Nele

Sozial-kognitive Merkmale des Lernens: Determinanten, Bedeutung und Unterrichtsgestaltung

Zander, Lysann; Wolter, Ilka; Kreutzmann, Madeleine

Soziale Peerbeziehungen als Determinanten des konstruktiven Umgangs mit Fehlern im Unterricht

Konstruktiver Umgang mit Fehlern fördert erfolgreiche Lernprozesse und umfasst nach Spychiger, Kuster und Oser (2006) eine affektive und eine kognitive Dimension: (a) geringe Fehlerangst und (b) positive Orientierung, aus Fehlern zu lernen (vgl. Kreutzmann, Zander & Hannover, 2013). Jüngere Forschung zeigt Belege für die Bedeutung des Lernkontextes für den individuellen Umgang mit Fehlern (z.B. Steuer, Rosentritt-Brunn & Dresel, 2013). Bisher weitgehend unberücksichtigt ist die Rolle der Peers, also der Klassenkameradinnen und Klassenkameraden, im jeweiligen Lernkontext für den konstruktiven Umgang mit Fehlern. Dabei ist plausibel, dass insbesondere natürlich bestehende Zusammenarbeitsbeziehungen im Schulklassenkontext mit konstruktiven Einstellungen zu Fehlern in Lernsituationen, also geringerer Fehlerangst und höherer Fehlerlernorientierung zusammenhängen.

In diesem Beitrag (Zander, Kreutzmann & Wolter, 2014) haben wir untersucht, inwiefern die affektive und die kognitive Dimension, d.h. Fehlerangst und Fehlerlernorientierung mit der Struktur der Zusammenarbeitsnetzwerke in Schulklassen zusammenhängen. Dabei betrachten wir zwei Facetten kollaborativer Peernetzwerke genauer, nämlich (1) die Gesamtstruktur der Netzwerke, genauer die Dichte der Zusammenarbeitsbeziehungen zwischen verschiedenen Schülerinnen und Schülern und (2) die individuelle Einbindung von Schülerinnen und Schülern in reziproke, also gegenseitige, Zusammenarbeitsbeziehungen zu einzelnen anderen Mitschülern und Mitschülerinnen.

Im Sinne der sozial-kognitiven Lerntheorie von Bandura (2012) haben wir angenommen, dass der Selbstwirksamkeit der Schülerinnen und Schüler eine bedeutsame Rolle in der Moderation des Zusammenhangs zwischen Merkmalen des Lernkontextes und individuellem Fehlerumgang zukommt.

Hierfür haben wir ein Sample von 448 Berliner Schülerinnen und Schülern der 5. und 6. Jahrgangsstufe ($n = 239$ Mädchen, 53.3 %) in 21 Schulklassen aus 11 Grundschulen analysiert. Konstruktiver Umgang mit Fehlern wurde operationalisiert über die Skala von Spychiger et al. (2006). Für die Messung der allgemeinen Selbstwirksamkeit kam die Skala von Jerusalem und Schwarzer (2001) zum Einsatz. Über ein soziometrisches Verfahren erhoben wir die Kollaborationsnetzwerke der Schulklassen („Mit wem arbeitest du gern zusammen?“). Aus den Nennungen der Schülerinnen und Schüler wurden mit Hilfe der Software UCInet (Borgatti, Everett & Freeman, 2002) Indizes für die Netzwerkdichte (auf Klassenebene, Anteil realisierter Nennungen an tatsächlichen Nennungen) und die Einbindung in reziproke Beziehungen (auf Individualebene, Zentralitätsdegree) gebildet.

In einer Reihe von Mehrebenenanalysen in Mplus (Muthén & Muthén, 2002-2008) fanden wir differenzielle Muster für die affektive und kognitive Dimension des Umgangs mit Fehlern. Schüler/innen berichteten geringere Fehlerangst in Klassen mit dichteren Kollaborationsnetzwerken und wenn sie gut in diese Netzwerke eingebunden waren, jedoch nur bei hoher allgemeiner Selbstwirksamkeit (signifikante Crosslevel-Interaktion, $B = -5.56$, $SE = 0.97$, $t = -5.74$, $p < .000$).

Schüler/innen zeigten höhere Fehlerlernorientierung, wenn sie stark in wechselseitige kollaborative Beziehungen eingebunden waren und wenn sie eine hohe allgemeine Selbstwirksamkeit hatten ($B = -2.53$, $SE = 0.83$, $t = -3.06$, $p = .002$). Netzwerkdichte in der Gesamtklasse war kein Prädiktor höherer Fehlerlernorientierung (n.s.).

Die Ergebnisse korrespondieren mit Befunden von Altermatt und Broady (2009), die zeigten, dass Kinder stärker adaptive Reaktionen auf Fehler zeigen, in dem Maße wie enge Freunde und Freundinnen ihnen nach einer kognitiven Niederlage Hilfe anbieten. Insgesamt legen unsere Befunde nahe, dass sowohl die Gesamtstruktur des Zusammenarbeitsnetzwerks als auch die individuelle Einbindung in reziproke Hilfsnetzwerke bedeutsam mit konstruktivem Fehlerumgang zusammenhängen, wobei schulischer Selbstwirksamkeit zur Nutzung dieser Ressourcen eine bedeutsame Rolle zukommt. Die Ergebnisse werden weiterhin vor dem Hintergrund sozial-kognitiver Theorien und aktueller Ansätze ko- und selbst-regulierten Lernens (z.B. Järvelä, Volet & Järvenoja 2010; Zander, im Druck) diskutiert. Die Berücksichtigung der identifizierten Zusammenhänge, die zunächst der Replikation bedürfen, liefern nicht zuletzt sinnvolle Möglichkeiten für die Gestaltung von Interventionen, die auf die Unterstützung von Lernprozessen abzielen.

Chair(s)

Gebauer, Miriam Marleen
McElvany, Nele

Schöber, Christian; Huelmann, Thorben; Köller, Olaf; McElvany, Nele; Bos, Wilfried; Ferdinand, Hanna; Gebauer, Miriam
Können Lehrkräfte die Selbstwirksamkeitsüberzeugungen Ihrer Schulkinder im Fachbereich Lesen korrekt beurteilen

Die diagnostische Kompetenz oder auch Urteilsgenauigkeit von Lehrkräften gilt als eine Schlüsselqualifikation im Kontext des Lehrens und Lernens und ist eine Voraussetzung für die individuelle Förderung von Schulkindern durch Lehrkräfte (Artelt & Gräsel, 2009). Für Schulleistungen haben sich die fachspezifischen Selbstwirksamkeitsüberzeugungen von Schulkindern in zahlreichen Studien als starker Prädiktor erwiesen (Köller & Möller, 2010). Für eine optimale Förderung der Schulleistungen von Schulkindern erscheint es daher unumgänglich, dass Lehrkräfte deren fachspezifische Selbstwirksamkeitsüberzeugungen zutreffend einschätzen können. Bisherige Befunde zur Urteilsgenauigkeit von Lehrkräften bezüglich anderer motivationaler Konstrukte deuten jedoch auf eine unzureichende Genauigkeit der Lehrkrafturteile hin (Spinath, 2005, Dicke, Lüdtke, Trautwein, Nagy & Nagy, 2012).

Im Vergleich dazu können Leistungen der Schulkinder treffender durch ihre Lehrkräfte beurteilt werden (Südkamp, Kaiser & Möller, 2012). Zudem zeigte sich, dass Lehrkräfte z.B. bei der Beurteilung von Intelligenz die schulische Performanz als Grundlage ihres Urteils heranzogen (Rost und Hanses, 1997). Weitere Forschungsergebnisse zeigten, dass Stereotype, die unter anderem aufgrund familiärer Hintergrundmerkmale der Schulkinder entstehen, die Lehrkrafturteile moderierten (Glock & Krolak-Schwerdt, 2013).

Fragestellung: Unsere Studie soll daher die folgenden Forschungsfragen beantworten: (1) Wie genau können Lehrkräfte die lesespezifischen Selbstwirksamkeitsüberzeugungen ihrer Schulkinder beurteilen? (2) Wird das Lehrkrafturteil durch das Urteil der schulischen Performanz (Deutschnote) mediert? (3) Wird diese Mediation durch den Migrationshintergrund der Schulkinder moderiert?

Methode: Zur Beantwortung unserer Forschungsfragen verwendeten wir Fragebogendaten von 1.589 Schulkindern des siebten Jahrgangs und ihren Deutschlehrkräften. Nach Ausschluss von Klassengruppen mit weniger als sechs durch ihre Deutschlehrkraft bewerteten Schulkindern verblieben 1473 Schulkinder (50% weiblich; Alter (Jahre): $M = 12.1$, $SD = 0.6$) und 87 Lehrkräfte (84% weiblich; Alter (Jahre): $M = 43.3$, $SD = 11.9$; Lehrerfahrung (Jahre): $M = 14.0$, $SD = 11.7$) für unsere Analysen. Durch eine geschichtete Stichprobenziehung konnten überproportional viele Schulkinder mit Migrationshintergrund (24% türkisch, 14% ehemalige UdSSR und Polen, 20% sonstige Länder) befragt werden.

Die Schulkinder bearbeiteten sechs Items mit vierfach gestuftem Antwortformat zu ihrer lesespezifischen Selbstwirksamkeitsüberzeugung (Beispielitem: „Ich bin sicher, dass ich auch den Stoff in Unterrichtstexten, die sehr schwierig zu lesen sind, verstehen kann.“; Cronbachs Alpha = .88). Die Deutschlehrkräfte beantworteten leicht modifizierte Items, um die Selbstwirksamkeitsüberzeugungen ihrer Schulkinder einzuschätzen (Beispielitem: „Die Schülerin/der Schüler ist sich sicher, dass sie/er auch den Stoff in Unterrichtstexten, die sehr schwierig zu lesen sind, verstehen kann.“; Cronbachs Alpha = .94).

Es wurden Mehrebenenregressionsmodelle (Kriterium = Lehrkrafturteil) mit Mplus 7.11 (Muthén & Muthén, 2012) unter Verwendung des robusten Maximum-Likelihood-Schätzers berechnet, um die geschachtelte Datenstruktur (Schulkinder in Klassen) zu berücksichtigen. Dabei wurden die Prädiktoren (Prädiktor: Schulkindselbsteinschätzung, Mediator: Deutschnote, Moderatoren: Migrationshintergrund) auf Individualebene am Gruppenmittelwert zentriert und zusätzlich die entsprechenden Gruppenmittelwerte auf Gruppenebene als Prädiktoren aufgenommen, um die Effekte beider Ebenen trennen zu können (Preacher, Zhang & Zyphur, 2011).

Ergebnisse: Die Ergebnisse zur ersten Forschungsfrage zeigen, dass die Lehrkräfte die lesespezifischen Selbstwirksamkeitsüberzeugungen zu anderen motivationalen Konstrukten vergleichbar schwach einschätzen konnten ($\beta = 0.19$, $p < .001$, $R^2 = 0.04$). Zur zweiten Forschungsfrage zeigte sich auf der Individualebene ein signifikanter indirekter Effekt in Höhe von $0.087 (= \hat{a}_{within} \times \hat{b}_{within}$; $\hat{a}_{within} = -0.294$, 95% CI [-0.361, -0.227], $p < .001$; $\hat{b}_{within} = -0.295$, 95% CI [-0.339, -0.252], $p < .001$) mit einem 95% Konfidenzintervall von [0.064, 0.109], welches die Null nicht enthielt. Gleichzeitig ergab sich ein direkter Effekt von $\hat{c}_{within} = 0.123$; 95% CI [0.068, 0.178], $p < .001$. Das Urteil der Lehrkräfte bezüglich der lesespezifischen Selbstwirksamkeit ihrer Schulkinder wurde also nur teilweise über die Leistungseinschätzung vermittelt. Die Ergebnisse bedeuten, dass bei einem Anstieg der Schulkindselbsteinschätzung um einen Punkt die Fremdeinschätzung durch die Lehrkraft um 0.123 Punkte direkt zunimmt und eine weitere Zunahme um 0.087 Punkte über die Deutschnote vermittelt wird.

Eine Moderation dieser Effekte als Gegenstand unserer dritten Forschungsfrage zeigte sich hingegen nicht.

+

Dresel, Markus; Fasching, Michaela; Praetorius, Anna; Janke, Stefan; Nitsche, Sebastian; Dickhäuser, Oliver Dickhäuser

Freitag
13. März

9:00 Uhr

HGC 20

Chair(s)

Gebauer, Miriam Marleen
McElvany, Nele

Sozial-kognitive Merkmale des Lernens: Determinanten, Bedeutung und Unterrichtsgestaltung

Selbstreguliertes Lernen aus Fehlern im Fach Mathematik: Entwicklung über zwei Schuljahre und Abhängigkeiten von unterrichtlichen Bedingungen

Ein adaptiver Umgang mit Fehlern im Lernprozess birgt Chancen für Lernzuwächse und eine angemessene Wissensrepräsentation des Lerngegenstands (Hascher & Hagenauer, 2010; Oser & Spychiger, 2005). Die Aufrechterhaltung der Lernmotivation nach Fehlern und die Nutzung von Fehlern als Anlass für gezielte Lernhandlungen können damit als wichtige Aspekte des selbstregulierten Lernens gelten (Dresel, Schober, Ziegler, Grassinger & Steuer, 2013). Indes reagieren Schüler(innen) oftmals maladaptiv auf Fehler: Sie erleben selbstwertbedrohliche Emotionen, ihre Motivation sinkt und sie sehen Fehler vielfach nicht als Lernchance (z.B. Weinert, 1989). Von was es abhängt, dass Schüler(innen) das in Fehlern enthaltene Lernpotential nutzen oder es verstreichen lassen, ist eine für die pädagogische Praxis wichtige Frage. Mehrere Dekaden an Forschung zu Reaktionsmustern nach Misserfolg haben eine Reihe an individuellen Determinanten identifiziert (Überblick bei Schunk, Pintrich & Meece, 2008), die sich auch auf die Selbstregulation im Nachgang von spezifischen Fehlern übertragen lassen (Dresel et al., 2013; Tulis, Grassinger & Dresel, 2011). Im Gegensatz dazu ist bislang noch weniger gut verstanden, welche Aspekte des Unterrichts einen adaptiven Umgang mit Fehlern fördern. In der jüngeren Forschung wurde dazu das Konzept des Fehlerklimas vorgeschlagen, das als eigenständiges und multidimensionales Unterrichtsmerkmal aufgefasst wird (Oser & Spychiger, 2005; Steuer, Rosentritt-Brunn & Dresel, 2013; für eine ähnliches Konzept im betrieblichen Bereich siehe Van Dyck, Frese, Baer & Sonnentag, 2005). Bislang liegen allerdings ausschließlich querschnittliche Hinweise darauf vor, dass ein konstruktives Fehlerklima im Unterricht das Lernen aus Fehlern auf Individualebene der Schüler(innen) fördert.

Deshalb wird in dem vorliegenden Beitrag der Frage nachgegangen, wie sich der Umgang von Schüler(inne)n mit Fehlern über die Zeit verändert, welche Unterschiede zwischen unterschiedlichen Schulklassen in dieser Entwicklung bestehen und ob diese Entwicklungen vom Fehlerklima im Unterricht abhängen. Ergänzend wird geprüft, ob auch objektive Lernzuwächse umfangreicher ausfallen, wenn im Unterricht ein konstruktives Fehlerklima herrscht.

Im Beitrag präsentieren wir Ergebnisse einer Längsschnittstudie, die über die komplette 5. und 6. Jahrgangsstufe angelegt war und fünf Messzeitpunkten umfasste. Sie war im Mathematikunterricht des Gymnasiums angelegt. Der Stichprobe gehörten 1.440 Schüler(innen) in 69 Schulklassen an (erweiterte Stichprobe für die ersten drei Messzeitpunkte: 3.104 Schülerinnen und Schüler in 158 Klassen). Das Fehlerklima und der individuelle Umgang mit Fehlern wurden mit validierten Instrumenten erfasst (Dresel et al., 2013; Steuer et al., 2013). Die Schulleistung im Fach Mathematik wurde mittels standardisierter Schulleistungstests operationalisiert (HRT 1–4, DEMAT 5+, DEMAT 6+; Götz, Lingel & Schneider, 2013a, b; Haffner, Baro, Parzer & Resch, 2005).

Konfirmatorische Zwei-Ebenen-Faktorenanalysen bestätigten zunächst die angenommene Faktorenstruktur des Fehlerklimas mit acht Subfacetten und einem einheitlichen Faktor zweiter Ordnung (vgl. Steuer et al., 2013). Wachstumskurvenmodelle erbrachten eine negative Entwicklung, wonach sich der individuelle Umgang der Schüler(innen) mit Fehlern im Fach Mathematik über die Zeit verschlechterte. Die Stärke dieser Verschlechterung variierte zwischen Schulklassen und hing systematisch vom Fehlerklima im Mathematikunterricht ab: Je konstruktiver es war, desto milder verlief die negative Entwicklung. Darüber hinaus indizierten die Befunde, dass ein konstruktives Fehlerklima mit größeren objektiven Lernzuwächsen assoziiert ist.

In der Gesamtschau implizieren die vorliegenden Ergebnisse, dass – zumindest zu Beginn der Gymnasialzeit – Bedarf besteht, den Umgang von Schüler(inne)n mit Fehlern zu fördern und dass dies durch instruktionales Handeln im Unterricht gelingen kann. Praktische Implikationen für eine in dieser Hinsicht förderliche Unterrichtsgestaltung lassen sich insbesondere aus den einzelnen, auf das Lehrkräfthandeln bezogenen Subfacetten des Fehlerklimas ableiten.

Südkamp, Anna

Erfolgsfaktoren schulischer Inklusion

Seit Verabschiedung der UN-Behindertenrechtskonvention (UN-BRK, 2008) und dem Beschluss der Kultusministerkonferenz (KMK, 2011) zur Umsetzung von Inklusion werden Möglichkeiten und Grenzen der Umsetzung in der Öffentlichkeit vielfältig diskutiert. Das zentrale Ziel schulischer Inklusion besteht darin, die gleichberechtigte soziale Teilhabe aller Schülerinnen und Schüler, insbesondere der Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf, im Kontext Schule zu gewährleisten. Dies stellt vor dem Hintergrund zunehmender Heterogenität der Schülerschaft eine Herausforderung für die Akteure im Bildungssystem dar. Wie Studien zeigen, wird eine gleichberechtigte soziale Teilhabe nicht automatisch durch das gemeinsame Unterrichten aller Schülerinnen und Schüler erreicht (Huber & Wilbert, 2012). Der Erfolg schulischer Inklusion hängt von zahlreichen Faktoren ab, wie z. B. den Einstellungen und Selbstwirksamkeitserwartungen der beteiligten Akteure, den Kompetenzen der Schülerinnen und Schüler und der Gestaltung des Unterrichts.

In der Empirischen Bildungsforschung werden derzeit zunehmend Bemühungen unternommen Erfolgsfaktoren gelingender Inklusion zu identifizieren. In diesem Symposium werden vier Studien aus unterschiedlichen wissenschaftlichen Disziplinen vorgestellt, die diese Erfolgsfaktoren empirisch untersuchen.

Der erste Beitrag von Michel Knigge (Pädagogische Psychologie, Universität Halle-Wittenberg) beschäftigt sich mit den Einstellungen und Selbstwirksamkeitserwartungen von Lehramtsstudierenden zur Inklusion. Es wird die Frage geprüft, inwiefern diese verhaltensrelevanten Überzeugungen der Lehramtsstudierenden durch praktische Erfahrungen sowie Seminarbesuche im Bereich Inklusion beeinflusst werden.

Im zweiten Beitrag von Timo Lüke (Inklusionspädagogik, Universität Potsdam) wird experimentell untersucht, ob Befragungen von Lehrkräften zur Einstellung zur Inklusion durch sozial erwünschtes Antwortverhalten verfälscht sein könnten. Die Studie geht insbesondere der Frage nach, ob die antizipierte Einstellung der befragenden Organisation die berichtete Einstellung von Lehrkräften beeinflusst.

Im dritten Beitrag von Stefanie Brimmers (Rehabilitationspsychologie, TU Dortmund) wird die soziale Integration von Schülerinnen und Schülern als Kriterium für eine gelingende Inklusion untersucht. Mit Hilfe soziometrischer Verfahren wird geprüft, wie gut Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf und mit Migrationshintergrund in Klassen gemeinsamen Lernens integriert sind und welche Schülermerkmale die soziale Integration beeinflussen.

In dem Beitrag von Stefan Voß (Institut für Sonderpädagogische Entwicklungsförderung und Rehabilitation, Universität Rostock,) werden aktuelle Ergebnisse einer Interventionsstudie zum Rügener Inklusionsmodell vorgestellt. Es wurde untersucht, wie sich Grundschülerinnen und -schüler im inklusiven Unterricht auf Basis des „Response to Intervention“-Ansatzes entwickeln im Vergleich zu Grundschülerinnen und -schülern, die nach dem herkömmlichen Unterrichts- und Fördersystem in Mecklenburg-Vorpommern unterrichtet wurden.

Für die Diskussion konnte Prof. Dr. Franz Wember (Rehabilitation und Pädagogik bei Lernbehinderungen, TU Dortmund) gewonnen werden. Er wird die Beiträge kritisch würdigen und zur Diskussion über die Herausforderungen von Inklusion (Wember, 2013) anregen.

Freitag
13. März

9:00 Uhr

HGC 30

Knigge, Michel; Rotter, Carolin

Zusammenhänge von inklusionsbezogenen Einstellungen und Selbstwirksamkeitserwartungen in Abhängigkeit von Seminarbesuchen und praktischen Erfahrungen vor und während eines Semesters bei Lehramtsstudierenden in Halle und Hamburg

Einstellungen und Selbstwirksamkeitserwartungen sind bedeutsame Antezedenzen für das Verhalten einer Person. Dies lässt sich neben dem Verweis auf zahlreiche empirische Untersuchungen zu beiden Konstrukten mit dem Verweis auf die Theorieklasse der Wert-Erwartungs-Modelle veranschaulichen (Eccles, Barber, Updegraff, & O'Brien, 1998). Diese Modelle besagen, dass die Wahl zwischen verschiedenen Handlungszielen und damit auch die Handlungen zur Zielerreichung maßgeblich dadurch bestimmt werden, welchen Wert eine Person dem angenommenen Zielzustand und der Zielerreichungshandlung selbst beimisst sowie wie hoch die Erwartungen sind, bei der Zielerreichung Erfolg zu haben. Während Einstellungen eine zentrale Dimension der Wert-Komponente darstellen, sind Selbstwirksamkeitserwartungen ein zentraler Aspekt der Erwartungskomponente.

Entsprechend darf angenommen werden, dass bei der gegenwärtigen Umsetzung der schulpolitischen Vorgabe, ein möglichst inklusives Schulsystem zu entwickeln, inklusionsbezogene Einstellungen und Selbstwirksamkeitserwartungen von Lehrkräften eine wichtige Rolle spielen. In Modellen zur Beschreibung der professionellen Kompetenz von Lehrkräften lassen sich diese als Teil der relevanten Überzeugungen der Lehrkräfte begreifen (Baumert & Kunter, 2013). Um sicherzustellen, dass Lehrkräfte hier über günstige Voraussetzungen zum Umgang mit Inklusion verfügen, sollten bereits Lehramtsstudierende dabei unterstützt werden, Überzeugungen zu entwickeln, die produktive Lehr-Lern-Situationen für alle Schülerinnen und Schüler wahrscheinlicher machen.

Die vorliegende Untersuchung beschäftigt sich entsprechend mit der Frage, welchen Einfluss praktische Erfahrungen mit sowie Seminarbesuche zu Inklusion Einfluss auf inklusionsbezogene Einstellungen und Selbstwirksamkeitserwartungen bei Lehramtsstudierenden haben. Der Zusammenhang zwischen praktischen Erfahrungen sowie theoretischem Wissen und Einstellungen zu Inklusion konnte wiederholt gezeigt werden. In einem Review von Avramidis und Norwich (2002) wurden insgesamt drei zentrale Variablencluster identifiziert, die einen Zusammenhang mit der Einstellung zu Inklusion haben. Unterschieden wurden hier in Variablen auf Seiten der Kinder (z. B. Art der Behinderung), Umweltvariablen (z. B. Vorhandensein von Fortbildungen) sowie Eigenschaften der Lehrkräfte. Bei den zuletzt genannten zeigen, neben nur moderat einflussreichen Variablen wie dem Geschlecht oder der religiösen Überzeugung, vor allem Erfahrungsvariablen einen bedeutsamen Zusammenhang mit der Einstellung zu Inklusion. Während für Zusammenhänge zwischen Erfahrungen und Einstellungen ein umfangreicher Forschungsstand vorliegt, gibt es bisher nur wenige Studien, die sich mit der Veränderung über die Zeit in Abhängigkeit von gegenwärtig gemachten Erfahrungen unter der Berücksichtigung von Vorerfahrungen beschäftigen. Hier liegt ein Forschungsdesiderat vor, dem sich die vorliegende Studie widmet.

Angelehnt an Antonak und Larrivee (1995) sowie Avramidis, Bayliss und Burden (2000) wurde im Rahmen der EiLInk-Studie ein Erhebungsinstrument konzipiert, welches eine multidimensionale Struktur von kognitiven und affektiven Einstellungen zu Inklusion berücksichtigt. Zusätzlich wurden verschiedene Facetten von inklusionsbezogenen Selbstwirksamkeitserwartungen nach Kopp (2009) erfasst. Die Datenerhebung erfolgte im Längsschnitt mit zwei Messzeitpunkten am Anfang und am Ende des Wintersemesters 2013/2014 bei Lehramtsstudierenden in der Mitte ihres Studiums an der Martin-Luther-Universität Halle Wittenberg und an der Universität Hamburg, wobei Seminarbesuche und praktische Erfahrungen im Bereich Inklusion vor und während des laufenden Semesters erfasst wurden.

Anhand von ersten querschnittlichen varianzanalytischen Auswertungen der Rohdaten zeigten sich Befunde, dass vor allem bei gleichzeitigen theoretischen und praktischen Erfahrungen die Einstellungen zu Inklusion positiver ausfielen. Dabei zeigten sich Unterschiede zwischen den Standorten und den angestrebten Schulformen, an denen unterrichtet werden soll. Gegenwärtig werden vertiefende Analysen mit multipel imputierten Daten im Längsschnitt zu Veränderungen der inklusionsbezogenen Einstellungen und Selbstwirksamkeitserwartungen über ein Semester durchgeführt. Dabei wird das Vorwissen aus früheren Semestern als zusätzlicher Faktor einbezogen. Entsprechend kann gezeigt werden, ob theoretische und/oder praktische Erfahrungen während des Studiums Einfluss auf zentrale Aspekte von Lehrkräfteüberzeugungen im Zusammenhang mit Inklusion nehmen. Die Befunde werden mit Bezug auf ihre grundlagenwissenschaftlichen und ihre praktischen Implikationen diskutiert, wobei ein besonderer Fokus auf die Lehrkräfteausbildung gelegt wird.

Lüke, Timo; Grosche, Michael

Experimentelle Befunde zum Einfluss sozialer Erwünschtheit bei der Messung der Einstellung zur Inklusion

Die Einstellungen zur Inklusion sind in der aktuellen Schul- und Bildungsforschung ein hochaktuelles Thema. In zahlreichen Arbeiten werden neue Instrumente vorgestellt oder bereits veröffentlichte Instrumente zur Untersuchung übergeordneter Fragestellungen genutzt. Meist wird dabei die Einstellung von Lehrkräften und Schulleitungen zur Inklusion als wichtiger Einflussfaktor für deren Gelingen betrachtet. Dabei werden die Ergebnisse der sozialpsychologischen Einstellungsforschung häufig noch zu wenig berücksichtigt: Beispielsweise ist unklar, ob und in welchem Ausmaß die von Befragten berichtete Einstellung zur Inklusion durch soziale Erwünschtheit verzerrt sein könnte. Da Inklusion wahrscheinlich als politisch korrekte Idee wahrgenommen wird (Avramidis & Norwich, 2002; Leyser et al., 1994), wäre eine bewusste oder unbewusste Beeinflussung in Richtung einer positiveren Einstellung zu vermuten. Ohne Kenntnis dieser Problematik können Studien – beispielsweise Befragungen von Universitäten – kaum sinnvoll interpretiert werden.

In zwei Experimenten haben wir daher folgende Forschungsfrage bearbeitet: Beeinflussen soziale Kontextinformationen, wie die (antizipierte) Einstellung der befragenden Organisation die berichtete Einstellung der Befragten?

Im ersten Experiment wurden die TeilnehmerInnen einer Onlinebefragung randomisiert vier Bedingungen zugeordnet, in denen wir Hinweise auf die Einstellung der befragenden Organisation manipulierten. Dabei haben wir fiktive Organisationen, die für oder gegen schulische Inklusion argumentierten sowie eine Universitätsbedingung, in der wir keine Argumente nannten, genutzt. Anschließend wurden die Einstellung zur Inklusion der Befragten und empirisch belegte Prädiktoren dieser Einstellung erfasst.

Zur Messung der Einstellung wurde die Skala „Einstellung zur Integration“ (Kunz et al., 2010) mit den Subskalen „Schulische Förderung und Unterstützung“ (SFU) und „Soziale Integration“ (SI) in einer Onlineversion eingesetzt. Insgesamt nahmen 231 Personen (70% weiblich) im Alter von durchschnittlich 32 Jahren (SD = 12) an der Untersuchung teil. Die Rekrutierung erfolgte über Internetforen und an öffentlichen Plätzen, um nicht übermäßig interessierte Lehrkräfte und Studierende zu berücksichtigen.

Die Ergebnisse bestätigen hypothesenkonforme Zusammenhänge zwischen der bei der befragenden Organisation vermuteten und der für sich selbst berichteten expliziten Einstellung zur Inklusion. In hierarchischen Regressionen auf die Einstellungsskalen der TeilnehmerInnen erreichen die aus dem Forschungsstand abgeleiteten Prädiktoren (Block 1) zusammen eine Varianzaufklärung von 8.6% (SFU) bzw. 9.7% (SI). Signifikante Zusammenhänge finden sich für die generellen politischen Ansichten der Befragten ($\beta_{SFU} = -.219^{***}$; $\beta_{SI} = -.126^*$), den regelmäßigen Kontakt zu Menschen mit Behinderung ($\beta_{SFU} = -.164^*$; $\beta_{SI} = -.140^*$) und die Zugehörigkeit zur pädagogischen Profession ($\beta_{SFU} = .138^*$; $\beta_{SI} = .182^*$). Durch die Aufnahme der von den TeilnehmerInnen vermuteten Einstellung auf Seiten der befragenden Organisation (Block 2), werden beide Modelle deutlich verbessert ($R^2_{SFU} = 20.9\%$, $\Delta R^2_{SFU} = .122^{***}$; $R^2_{SI} = 19.4\%$; $\Delta R^2_{SI} = .098^{***}$). Während sich die übrigen Betagewichte nicht relevant verändern, wird die vermutete Einstellung der befragenden Organisation in beiden Modellen zum stärksten Prädiktor ($\beta_{SFU} = .355^{***}$; $\beta_{SI} = .318^{***}$).

Dass die TeilnehmerInnen in der Universitätsbedingung ebenso positiv antworteten wie die TeilnehmerInnen der Pro-Bedingung, aber signifikant positiver als die anderen Gruppen, verdeutlicht die dem politischen Mainstream entsprechende sozial erwünschte Antworttendenz bei der berichteten Einstellung zur Inklusion ($\eta^2_{SFU} = .066$; $\eta^2_{SI} = .045$). Damit ist nachgewiesen, dass der soziale Kontext der Befragungssituation die explizite Einstellung von Personen stärker beeinflusst als relevante individuelle Merkmale.

In einem zweiten Experiment wurde in einem analogen Design untersucht, ob sich der Effekt auch ohne inhaltliche Argumente – nur über die Identität der befragenden Organisation – replizieren lässt. Die Universitätsbedingung wurde weiter ausdifferenziert: In der einen Gruppe wird lediglich die Universität genannt, in der anderen auch die Zugehörigkeit zur „Inklusionspädagogik“. Die Ergebnisse dieses zweiten Experimentes werden zum Zeitpunkt der Tagung vorliegen.

Bei der momentan sehr ausgiebigen Verwendung von Fragebogen zur Einstellung zur Inklusion in der Forschung wird der Einfluss sozialer Kontextinformationen nicht ausreichend berücksichtigt. Die damit verbundenen Validitätsprobleme werden bisher zu wenig reflektiert. Mit der vorliegenden Studie sollen auf diese Problematik hingewiesen und Desiderata aufgezeigt werden.

Freitag
13. März

9:00 Uhr

HGC 30

Chair(s)
Südkamp, Anna

Brimmers, Stefanie; Südkamp, Anna; Lange, Sarah; Wolf, Sylvia Mira; Tröster, Heinrich

Welche Faktoren beeinflussen die soziale Integration im Gemeinsamen Unterricht?

Gemeinsamer Unterricht (GU) zeichnet sich dadurch aus, dass alle Kinder zusammen unterrichtet werden. Die Heterogenität in Klassen gemeinsamen Lernens kann entlang verschiedener Dimensionen beschrieben werden. So werden beispielsweise Kinder mit und ohne sonderpädagogischen Förderbedarf (SFB) sowie Kinder mit und ohne Migrationshintergrund zusammen unterrichtet (Sander, 2008). Politik, Eltern und Lehrkräfte sehen die soziale Integration aller Schülerinnen und Schüler als wichtiges Ziel des Gemeinsamen Unterrichts an (Kultusministerkonferenz, 2011; Preuss-Lausitz, 2002). Aktuelle Untersuchungen zeigen allerdings, dass Kinder mit SFB häufiger von sozialer Ausgrenzung betroffen sind als ihre Mitschüler (z.B. Frostad & Pijl, 2007; Huber & Wilbert, 2012). Die soziale Position von Schülerinnen und Schülern mit Migrationshintergrund wurde bisher noch nicht explizit in Klassen gemeinsamen Lernens untersucht. Studien in Regelklassen zeigen aber, dass auch Kinder mit Migrationshintergrund häufiger ungünstige soziale Positionen in der Klasse einnehmen (z.B. von Grünigen, Kochenderfer-Ladd, Perren & Alsaker, 2012). Auf der Suche nach Ursachen für ungünstige soziometrische Positionen, wurden bisher hauptsächlich Schülermerkmale untersucht. Soziale Kompetenz und Schulleistungen scheinen demnach mögliche Prädiktoren für die soziale Position in der Klasse zu sein. Newcomb, Bukowski und Pattee (1993) untersuchten in einer Metaanalyse den Zusammenhang zwischen der soziometrischen Position und dem Sozialverhalten bzw. kognitiven Fähigkeiten. Beliebte Kinder zeichnen sich durch mehr geselliges Verhalten ($d=.383$), weniger Aggressionen ($d=-.258$) und Rückzugsverhalten ($d=-.109$) sowie höhere kognitive Fähigkeiten ($d=.247$) im Vergleich zu durchschnittlich beliebten Kindern aus. Diese Metaanalyse berücksichtigt allerdings keine Studien des Gemeinsamen Unterrichts und geht nicht explizit auf Kinder mit SFB oder Migrationshintergrund ein. Im Gegensatz zu Newcomb et al. (1993), die insgesamt sowohl kognitive als auch soziale Variablen als bedeutsam für die soziale Position in der Klasse ansehen, findet Huber (2006) in einer Stichprobe im GU, dass Schulleistungen in einem stärkeren Zusammenhang ($r=.45$) zur soziometrischen Position stehen als Sozialkompetenz ($r=.18$) oder Aggressivität ($r=-.18$). Huber (2006) untersuchte dabei den sozialen Status von Kindern mit SFB, bezieht jedoch keine Kinder mit Migrationshintergrund in seine Analysen ein.

Die vorliegende Studie geht zwei Fragestellungen nach. Zum einen wurde die soziale Position von Kindern mit SFB und von Kindern mit Migrationshintergrund in Klassen mit GU untersucht. Des Weiteren wurde der Zusammenhang von Schulleistungen und sozialer Kompetenz mit der soziometrischen Position im GU überprüft.

Um diese Fragen zu untersuchen, wurden 741 Schülerinnen und Schüler (mittleres Alter 10.6 Jahre, $SD=1.9$) aus 33 Klassen des Gemeinsamen Unterrichts befragt. 102 teilnehmende Schülerinnen und Schüler (14%) haben einen SFB (davon 70% im Bereich Lernen) und 199 Kinder (27%) einen Migrationshintergrund (davon 42% türkischer Abstammung). Die soziale Integration wurde über sechs soziometrische Wahl- und Ablehnungsfragen erfasst (vgl. Moreno, 1967; Huber & Wilbert, 2012) und zu einem Integrationsstatus verrechnet (Petillon, 1980). Als Indikator für die Schulleistungen liegen die Mathematik- und Deutschnoten des vergangenen Zeugnisses vor. Darüber hinaus wurden die Lehrkräfte gebeten, die soziale Kompetenz der Kinder auf einer sechsstufigen Skala einzuschätzen.

Die Ergebnisse zur ersten Fragestellung zeigen, dass Kinder mit SFB ($M=-1.05$, $SD=0.90$) einen deutlich niedrigeren Integrationsstatus einnehmen als ihre Mitschülerinnen und Mitschüler ohne SFB ($M=0.17$, $SD=0.91$, $t(734)=12.7$, $p<.001$). Kinder mit und ohne Migrationshintergrund unterscheiden sich hingegen nicht hinsichtlich der soziometrischen Position ($MMH=-0.02$, $SDMH=1.01$; $MohneMH=0.02$, $SDohneMH=0.98$, $t(729)=0.45$, $p>.05$). Zur Überprüfung der zweiten Fragestellung, wurde eine Regressionsanalyse unter Berücksichtigung der Mehrebenenstruktur der Daten mit R (R Core Team, 2014; Paket „lme4“) durchgeführt. Es zeigte sich, dass die soziale Kompetenz ($\beta=0.48$, $p<.001$) und die Mathematiknote ($\beta=0.17$, $p<.001$) in einem bedeutsamen Zusammenhang mit der soziometrischen Position in der Klasse stehen, die Deutschnote allerdings nicht ($\beta=0.08$, $p<.05$).

Die Ergebnisse machen deutlich, dass die soziale Integration von Kindern mit Migrationshintergrund in heterogenen Lerngruppen gut gelingt, für die soziale Integration von Kindern mit SFB aber noch mehr getan werden muss. Unterstützung bei der Entwicklung sozialer Kompetenzen scheint hier ein guter Ansatzpunkt zu sein.

Voß, Stefan; Blumenthal, Yvonne; Hartke, Bodo

Der Response to Intervention-Ansatz (RTI) – ein konkretes Rahmenmodell zur erfolgreichen Umsetzung von Inklusion?

Fragen nach der Umsetzung von schulischer Inklusion beziehen sich derzeit nicht mehr auf das ob, sondern vielmehr steht das wie im Mittelpunkt. In diesem Zusammenhang rücken die Gelingensbedingungen von Inklusion zunehmend in den Fokus der Forschung. Als eine zentrale Voraussetzung auf schulischer Ebene konnte die Einstellung der am Umsetzungsprozess beteiligten Akteure herausgearbeitet werden (u. a. Luiselli & Diament, 2002). Werden Lehrkräfte befragt, wie Sie zur schulischen Inklusion stehen, ergibt sich ein uneindeutiges Bild. Die Mehrheit befürwortet schulische Inklusion im Allgemeinen, kann sich jedoch die konkrete Umsetzung in der eigenen Klasse schwer bis kaum vorstellen (z. B. Scruggs, Mastropieri, & Leins, 2011; von de Boer, Pijl & Minnaert, 2011). Somit scheint die konkrete strukturelle (Um)Gestaltung von Unterricht und Förderung, angefangen von den innerschulischen Routinen, Rollen und Verantwortungen bis hin zur Einbeziehung außerschulischer Unterstützungssysteme neben der Einstellung der Lehrpersonen, eine weitere wesentliche Gelingensbedingung für die Umsetzung von Inklusion in der Schule darzustellen. Ein vielversprechendes konzeptuelles Rahmenmodell zur strukturellen (Um)Gestaltung von Unterricht und Förderung bildet der Response to Intervention-Ansatz (RTI; u. a. NCRTI, 2010). Als Präventions- und Inklusionsansatz von Lern- und Entwicklungsproblematiken gewinnt der RTI-Ansatz auch im deutschsprachigen Raum zunehmend an Bedeutung (Blumenthal, Kuhlmann & Hartke, 2014; Walter, 2008). Auf Basis einer Kooperation zwischen der Rostocker Universität und den Rügener Grundschulen gelingt es seit dem Schuljahr 2010/2011 erstmalig flächendeckend alle Grundschüler einer Region nach dem RTI-Ansatz im Rahmen des Rügener Inklusionsmodells (RIM) zu unterrichten (Mahlau, Blumenthal, Diehl, Schöning, Sikora, Voß & Hartke, 2014; Mahlau, Diehl, Voß Hartke, 2011).

Der Vortrag beleuchtet zunächst im Überblick die zentralen Kernelemente des RTI-Ansatzes sowie deren konkrete Umsetzung im Beschulungskonzeptes RIM und informiert anschließend über die Befunde zur Wirksamkeit des Ansatzes. Im Rahmen der Studie wird der Frage nachgegangen, inwieweit der inklusive Unterricht nach dem RIM eine tragfähige Alternative zum herkömmlichen Unterrichts- und Fördersystem in Mecklenburg-Vorpommern darstellt. In einem Zweigruppenversuchsplan werden dazu die Effekte von Unterricht und Förderung in den Grundschulen auf Rügen (Experimentalgruppe: 441 Kinder aus 23 Klassen/ 12 Schulen) und in der Hansestadt Stralsund (Kontrollgruppe: 385 Kinder aus 20 Klassen/ 8 Schulen) zum Ende der Grundschulzeit im Schuljahr 2013/14 vergleichend untersucht. Kriterien für die Berechnungen stellen dabei einerseits die schulischen Leistungen in den Domänen Mathematik, Lesen und Rechtschreiben sowie andererseits die emotional-soziale Entwicklung der Kinder sowie ihre soziale Stellung in den Klassen dar. Zur Analyse differentieller Effekte werden Kindergruppen mit unterschiedlichen schulischen Lernvoraussetzungen zum Zeitpunkt der Einschulung gebildet.

In Zusammenhang mit den Vorjahresergebnissen werden vergleichbare schulische Leistungen der beiden Untersuchungsgruppen bei höheren Kennwerten bezüglich der emotional-sozialen Entwicklung sowie der sozialen Integration der Schülerinnen und Schüler der Experimentalgruppe erwartet. Zudem wird davon ausgegangen, dass sich positive Effekte durch das RIM vor allem bei schulisch gefährdeten Kindern zeigen.

Freitag
13. März

9:00 Uhr

HGC 30

Chair(s)
Südkamp, Anna

Morgenroth, Stefanie

Social Support in schulischen Kontexten

Der Arbeitsplatz und Lernort Schule stellt für die Lehrkräfte und die Schülerinnen und Schüler gleichermaßen einen sozial interaktiven Raum dar. Neben den Lehrkräften und den Schülerinnen und Schülern lassen sich die Eltern und Freunde als weitere Interaktionspartner nennen. Schulische Qualifizierungs- und Prüfungsphasen sind eine besondere Herausforderung für die sozialen Beziehungen aller genannten Akteure. Hierbei spielt die soziale Unterstützung für das physische und psychische Wohlbefinden eine wichtige, wenngleich auch ambivalente Rolle. Die Ambivalenz der Bedeutung sozialer Unterstützung drückt sich darin aus, dass sie auf der einen Seite als belastend bzw. störend und auf der anderen Seite als hilfreich wahrgenommen werden kann. Das Symposium rückt den letztgenannten Aspekt in den Fokus. Während zahlreiche Studien den positiven Effekt der sozialen Unterstützung im Allgemeinen belegen, fehlt es an Untersuchungen, die die Bedingungen der Wirkungskraft konkretisieren. An dieser Stelle knüpfen die vier Beiträge des Symposiums an und tragen mit einem interdisziplinären Blick aus der Psychologie und der Lehrerbildung sowie der Kombination aus quantitativen und qualitativen Forschungsdesigns zur Illustration einer effektiven sozialen Unterstützung bei.

Im Vortrag von Käser und Neugebauer mit dem Titel Social Support in der zweiten Phase der Lehramtsausbildung gelingt es, die zentrale Bedeutung der sozialen Unterstützung für Referendare aus psychologischer Sicht zu untermauern. Neben der Bedeutung verschiedener Support-Quellen während der schulpraktischen Ausbildung wird der Zusammenhang zwischen Persönlichkeitsmerkmalen und der Suche nach bzw. der Zufriedenheit mit sozialer Unterstützung diskutiert. Im qualitativen Beitrag von Morgenroth werden erstmals unterschiedliche Interaktionsmuster von Lehrkräften dargestellt, die zur Suche nach sozialer Unterstützung führen und damit als Wegbereiter für eine intensive Lehrerkooperation bzw. eine erfolgreiche Lehrerprofessionalisierung dienen. Braasch und Buchwald konzentrieren sich in ihrem Beitrag auf die Bedeutung von Facebook als Quelle sozialer Unterstützung in Lern- und Prüfungskontexten. Hierbei geben qualitative Analysen erste wichtige Hinweise hinsichtlich der Mobilisierung von Social Support im Rahmen von sozialen Online Netzwerken. Ringeisen, Raufelder und Bürgermeister schließen das Symposium mit ihrem Beitrag Unterstützung und Druck der Eltern als Prädiktoren von Prüfungsangst: Geschlechtsspezifische Effekte bei Jugendlichen ab. Ergebnisse eines Multigruppenvergleichs auf Strukturgleichungsebene tragen u.a. zur differenzierten Darstellung geschlechtsspezifischer Wirkungspfade zwischen der elterlichen Unterstützung und der Prüfungsangst ihrer Kinder bei.

Freitag
13. März

9:00 Uhr

HGC 40

Chair(s)

Morgenroth, Stefanie
Braasch, Martina

Käser, Udo; Neugebauer, Hanna

Social Support in der zweiten Phase der Lehramtsausbildung

Theoretischer Hintergrund: Die Bedeutung sozialer Unterstützung für den erfolgreichen Umgang mit Belastungen im Alltag von Lehrerinnen und Lehrern ist umfangreich dokumentiert (Weddig, 1993; Barth, 1997; Käser & Wasch, 2009; vbw, 2014). Im Sinne eines ressourcentheoretischen Erklärungsansatzes (Buchwald, Schorn & Morgenroth, 2011) stellt sich der Zusammenhang zwischen sozialer Unterstützung und dem Umgang mit beruflicher Belastung so dar, dass Lehrkräfte, die sozial unterstützt werden, Belastungen eher positiv als Eustress erleben und leichter bewältigen. Bei fehlendem Social Support sind sie hingegen eher in einer Verlustspirale verhaftet (Buchwald & Hobfoll, 2004), die bei chronischer Belastung in Burnout münden kann.

Allerdings gibt es nur wenige Studien, die die Wechselwirkung von sozialer Unterstützung und beruflicher Belastung für die Ausbildungsphase analysieren, obwohl bereits Rauin (2007) auf den engen Zusammenhang zwischen dem Belastungserleben während des Lehramtsstudiums, des Referendariats und im späteren Beruf hinweist. Auch Käser, Leuchs, Neugebauer und Wasch (2012) zeigen, dass die Prädiktorenstruktur für Burnout bei Lehramtsstudierenden, Referendaren und Lehrkräften vergleichbar ist.

Fragestellung: Vor diesem Hintergrund wird die Bedeutung von Social Support in der zweiten Phase der Lehramtsausbildung näher untersucht. Folgenden Fragen wird nachgegangen: (1) Inwieweit suchen Referendare berufsbezogene soziale Unterstützung? (2) Wie zufrieden sind Referendare mit der berufsbezogenen soziale Unterstützung durch private / kollegiale Kontakte? (3) Gibt es Unterschiede hinsichtlich der Suche nach sowie der Zufriedenheit mit berufsbezogener sozialer Unterstützung in Abhängigkeit von biographischen und beruflichen Kontextvariablen? (4) Inwiefern ist soziale Unterstützung bedeutsam für das Erleben von beruflicher Belastung in der zweiten Phase der Lehrerausbildung?

Methode: Die Studie wurde als Befragung online durchgeführt. Im Fragebogen wurden biographische und berufsbezogene Personenmerkmale erfasst. Neben dem Belastungserleben, das durch den MBI nach Maslach, Jackson und Leiter (1996) gemessen wurde, wurden Neurotizismus, Extraversion, Verträglichkeit und Gewissenhaftigkeit als Persönlichkeitsmerkmale durch das NEO-FFI (Borkenau & Ostendorf, 1993) sowie arbeitsbezogenes Verhalten und Erleben durch den AVEM (Schaarschmidt & Fischer, 2008) erfasst. Außerdem wurden Lehrerselbstwirksamkeit, berufliche Zufriedenheit, Suche nach sozialer Unterstützung und die Zufriedenheit mit der sozialen Unterstützung erhoben (Käser & Wasch, 2009). Für alle Skalen liegen zufrieden stellende Reliabilitätskoeffizienten vor ($\alpha > .6$).

Die Stichprobe besteht aus 592 Referendare: 100 Männer und 492 Frauen im Alter zwischen 23 und 51 Jahren ($M = 27.5$ Jahre, $SD = 3.7$ Jahre), die zum Zeitpunkt der Erhebung durchschnittlich seit 13.5 Monaten ($SD = 5.7$ Monate) an Grundschulen ($n = 197$), Hauptschulen ($n = 34$), Realschulen ($n = 71$), Gesamtschulen ($n = 76$) oder Gymnasien ($n = 214$) ausgebildet wurden. Die meisten Referendare unterrichteten Deutsch (42%), Mathematik (28%) oder Englisch (21%). 70% unterrichteten ein sprachliches Fach, 60% ein Fach aus dem Bereich Mathematik-Naturwissenschaft-Technik, 49% eine Geisteswissenschaft und 21% künstlerisch-pragmatisches Fach.

Ergebnisse: Im Ergebnis wird deutlich, dass Referendare im Mittel stark nach sozialer Unterstützung im Beruf suchen und mit der sozialen Unterstützung, die sie in der zweiten Ausbildungsphase erhalten, im Allgemeinen zufrieden sind. Während sich dies für Referendare der verschiedenen Unterrichtsfächer in vergleichbarer Weise zeigt, suchen Frauen stärker nach Social Support als ihre männlichen Kollegen. Hinsichtlich der Zufriedenheit mit der sozialen Unterstützung liegen keine Geschlechtsunterschiede vor. Des Weiteren zeigen sich ebenfalls Unterschiede hinsichtlich der Quellen von Social Support: Während die Zufriedenheit mit der Unterstützung durch das private Umfeld und durch Kolleginnen und Kollegen überaus hoch ausfällt, liegen für die Hauptseminarleiter, Fachseminarleiter, Ausbildungskoordinatoren und die Schulleitung signifikant niedrigere Werte vor. Des Weiteren zeigen sich erwartungsgemäß signifikante Zusammenhänge mit Persönlichkeitsmerkmalen. So geht hohe Extraversion mit einer größeren Zufriedenheit mit und einer stärkerer Suche nach Social Support durch Kolleginnen und Kollegen einher, während Referendare mit hohen Werten in Neurotizismus seltener nach sozialer Unterstützung suchen und zugleich mit der sozialen Unterstützung durch das Kollegium eher unzufrieden sind. Schließlich bestätigt sich die hohe Bedeutung von Social Support für den erfolgreichen Umgang mit beruflicher Belastung im Referendariat der Lehramtsausbildung.

Freitag
13. März

9:00 Uhr

HGC 40

Chair(s)

Morgenroth, Stefanie
Braasch, Martina

Morgenroth, Stefanie

Social Support und Autonomie als Wegbereiter intensiver Lehrerkooperation

Theoretischer Hintergrund: Pädagogisch psychologische Forschungsansätze zur Lehrerkooperation postulieren, dass gemeinsame Ziele und Aufgaben, Vertrauen und ein gewisses Maß an Autonomie wesentliche Bedingungen für eine intensive Lehrerkooperation sind (Gräsel, Fussangel & Pröbstel, 2006). Die Autonomie gilt in diesem Prozess als ambivalentes Merkmal; zu viel Autonomie schränke die Gruppenkohäsion ein, wohingegen sich zu wenig Autonomie negativ auf die Motivation auswirke (Gräsel, Fussangel & Pröbstel, 2006, S. 208). Während sich der Mangel an intensiver Lehrerkooperation in zahlreichen Forschungsarbeiten bestätigen lässt (vgl. Fussangel, 2008, Pröbstel, 2008, Soltau, 2012), fehlt es an Studien, die dazu beitragen das Wirkungsgefüge zwischen der Autonomie und dem notwendigen Social Support als wesentliche Bedingung intensiver Lehrerkooperation zu konkretisieren.

Fragestellung: Dem genannten Forschungsdesiderat wird mit folgender Fragestellung nachgegangen: Unter welchen Umständen gelingt es Lehrkräften ihre Autonomie bzw. Eigenständigkeit trotz des interaktiven Charakters der sozialen Unterstützung zu bewahren, um intensive Lehrerkooperation auszuüben?

Methode: Es wurden problemzentrierte Einzelinterviews mit 14 Lehrkräften an zehn Grundschulen in Nordrhein-Westfalen durchgeführt. Alle Interviewpartner waren weiblichen Geschlechts, da sich die männlichen Lehrkräfte zu keinem Interview bereit erklärten. Begründet wurde dies über den Mangel an Anonymität, die aufgrund der überwiegend weiblichen Zusammensetzung der Kollegien an deutschen Grundschulen nicht gewährleistet werden konnte. Die Stichprobe beschränkt sich zudem ausschließlich auf Lehrkräfte, die sich im Rahmen einer telefonischer Anfrage freiwillig zu einer Interview-Teilnahme bereit erklärt haben. Insgesamt konnte die Interviewlänge aufgrund der relevanten Themenbereiche auf maximal 45 Minuten begrenzt werden. Die Interviews wurden mit einem Aufnahmegerät gesichert und im Anschluss transkribiert. Mittels der Interview-Transkripte wurde in einem ersten Auswertungsschritt ein Codierleitfaden nach der qualitativen Inhaltsanalyse (Mayring, 2008) erstellt. In einem zweiten Schritt wurden mittels der Methode der empirisch begründeten Typenbildung (Kluge, 2000) vier unterschiedliche Handlungsmuster entwickelt, die verdeutlichen, wann es Lehrkräften gelingt ihre Autonomie trotz der interaktiven Bedingungen einer intensiven Lehrerkooperation zu bewahren.

Ergebnisse: Teils mittels eines theoretisch-deduktiven Zugangs auf Basis der pädagogisch-psychologischen Forschungsansätze, teils induktiv in der Aufarbeitung der qualitativen Daten konnten vier verschiedene Autonomie-Muster abgebildet werden, die Wege zu einer intensiven Lehrerkooperation verdeutlichen: Ein Bedarfsweckendes-, ein Erkenntnisreiches-, ein Synergie- und ein Unsicherheits-Muster. Diese Muster variieren je nach intrinsischer oder extrinsischer Handlungsmotivation und je nach der Bewahrung der beruflichen Selbstwirksamkeit in Abhängigkeit von vertrauten und fremden Handlungsräumen. Somit gelingt es den Lehrkräften bei dem Bedarfsweckenden-Muster, die Initiierung zur Lehrerkooperation durch Dritte, als Mehrwert und nicht als Einschränkung der Autonomie zu deuten, indem sie selbstkritisch und -reflexiv mit den Anmerkungen der Kolleginnen und Kollegen umgehen. Bei dem Erkenntnisreichen-Muster hingegen bleibt die Autonomie bewahrt, da die Ausübung einer intensiven Lehrerkooperation intrinsisch motiviert ist; hierbei suchen die Lehrkräfte also aktiv nach Unterstützung. Es zeigt sich, dass dieses Muster in der Regel angewendet wird, wenn bislang fremde Handlungsräume, wie bspw. der Umgang mit Eltern oder die Übernahme einer neuen Klasse, erschlossen werden. Das Synergie-Muster findet Anwendung, wenn Lehrkräfte die gegenseitige Unterstützung als notwendige Voraussetzung und Mehrwert für ihre persönliche Weiterentwicklung verstehen. Die Autonomie der Lehrkräfte, die nach diesem Muster handeln ist durch die Suche nach vertrauten und sympathischen Kooperationspartner gesichert. Im Rahmen des Unsicherheitsmusters bleibt die Autonomie bewahrt, indem die Lehrkräfte antizipatorisch handeln und frühzeitig nach Unterstützung suchen, sobald sie mit Aufgaben eines unbekanntes Handlungsfeldes konfrontiert werden.

Inwieweit sich diese Muster in ihrer Effektivität unterscheiden und sich an anderen Stichproben replizieren lassen, ist in weiteren Analysen zu prüfen. Implikationen für die weitere Forschung zur Förderung intensiver Lehrerkooperation und die schulpraktische Umsetzung werden diskutiert.

Freitag
13. März

9:00 Uhr

HGC 40

Chair(s)

Morgenroth, Stefanie
Braasch, Martina

Braasch, Martina; Buchwald, Petra

Social Support in sozialen Online Netzwerken – Facebook als Quelle sozialer Unterstützung in Lern- und Prüfungskontexten

Theoretischer Hintergrund: Wie die repräsentative Untersuchung des Bundesverbandes für Informationswirtschaft, Telekommunikation und neue Medien e.V. (2013) ergab, integrieren Jugendliche in Deutschland ihr soziales Online Netzwerk zu einem bedeutenden Anteil in ihre Alltagswelt. 89% der Nutzer im Alter bis 30 Jahren sind auf ihren Lieblings-Netzwerken täglich aktiv (BITKOM 2013). Facebook gilt in Deutschland als das beliebteste soziale Online Netzwerk, bei dem 56% aller Internetnutzer angemeldet sind (BITKOM 2013) und dessen Nutzungshäufigkeit auch in 2014 noch weiterhin ansteigt (Facebook Statista Dossier 2013, S. 23). Die Funktionen der Plattform Facebook erleichtern Prozesse der Aufrechterhaltung und Verwendung sozialen Kapitals (vgl. Steinfield, Ellison & Lampe 2008), wodurch ein Zugriff auf Ressourcen, insbesondere sozialer Natur (Schümer & Buchwald, 2012), ermöglicht wird. Eine Besonderheit dieser Ressourcen besteht in der Loslösung traditioneller Barrieren wie Ort und Zeit, um auf sie zugreifen zu können, da die neuen Medien, insbesondere Smartphones, neue Möglichkeiten hierzu schaffen (Perotti & Hair, 2011). Dieser Zugriff auf Ressourcen innerhalb des sozialen Online Netzwerkes spielt für die Stressforschung dann eine Rolle, wenn untersucht wird, welche Formen von Ressourcen im Online Netzwerk für die Stressbewältigung nutzbar gemacht werden können (vgl. Braasch, Buchwald & Morgenroth 2014). Ein Faktor, der den Stressbewältigungsprozess wesentlich beeinflusst, ist Social Support (Buchwald, Schwarzer & Hobfoll 2004, Schwarzer & Buchwald, 2004). Um in stresshaften Zeiten auf Social Support zugreifen zu können, müssen Menschen diese soziale Ressource bereits präventiv in stressfreien Zeiten aufrechterhalten und erweitert haben (Buchwald 2002).

Fragestellung: Da in sozialen Online Netzwerken insbesondere Nutzer im Alter zwischen 14 und 29 Jahren (BITKOM 2013) aktiv sind, gehören Schülerinnen und Schüler sowie Studierende zu den größten Nutzergruppen. Ihnen gemeinsam ist, dass sie regelmäßig durch Prüfungen unter besonderem Stress stehen und auf Copingstrategien angewiesen sind, um diesem Stress zu begegnen (vgl. Buchwald 2002). In der vorliegenden Untersuchung wird der Frage nachgegangen, ob die Facebook-Plattform als soziale Ressource bei Stress eingesetzt und der Zugang zu Social Support dort aktiv gesucht wird.

Methode: In einer qualitativen Untersuchung wurden teilstandardisierte Interviews mit Hilfe eines Leitfadens mit Personen im Alter zwischen 12 und 32 Jahren geführt (N = 36), die Prüfungsstress in der Schule oder im Studium erlebten. Die Stichprobe wurde durch das inhaltliche Interesse an der Gruppe (Helferich 2010) eingangs so eingegrenzt, dass nur Schülerinnen und Schüler, Studierende sowie junge Lehrkräfte und Referendare, die ebenfalls mit Prüfungsstress konfrontiert sind oder kurzfristig waren, interviewt wurden. Nach der Transkription der Interviews wurde zur Auswertung ein Codierleitfaden nach der qualitativen Inhaltsanalyse (Mayring 2008) erstellt und auf die Interviews angewendet.

Ergebnisse: Die Ergebnisse verweisen auf ein tendenziell dichotomes Verhalten der Nutzer hinsichtlich der Verwendung von Facebook in Prüfungsphasen, die sich in zwei Gruppen beschreiben lassen. Für die eine Gruppe spielte Facebook zur Mobilisierung von Social Support während dieser Phasen eine wichtige Rolle, um den Kontakt zur sozialen Außenwelt auch dann noch aufrecht zu erhalten, wenn der durch die Prüfungsvorbereitung entstehende Stress dies eigentlich nicht mehr zulässt. Die oftmals selbst auferlegte soziale Isolation konnte in dieser Gruppe durch das Online Netzwerk durchbrochen werden und soziale Unterstützungsleistungen gewährleisten, auf die diese Gruppe sonst hätte verzichten müssen. Hingegen zeigt die andere Gruppe einen so disziplinierten Umgang mit Facebook während der Prüfungsphase, dass sie ihre sonst hochfrequente Nutzung zu diesen Zeiten stark einschränkt, um negative Effekte auf das Lernverhalten zu vermeiden. Interessant ist hierbei, dass es sich bei den Befragten aus der zweiten Gruppe um die jüngere Gruppe, hauptsächlich Schülerinnen und Schüler, handelt, die so genannten Digital Natives. Die Gründe für dieses divergente Verhalten stehen zur Diskussion und gilt es, in Folgestudien zu untersuchen.

Chair(s)

Morgenroth, Stefanie
Braasch, Martina

Ringeisen, Tobias; Raufelder, Diana; Buergermeister, Anika

Unterstützung und Druck der Eltern als Prädiktoren von Prüfungsangst: Geschlechtsspezifische Effekte bei Jugendlichen

Theoretischer Hintergrund: Das sozio-kontextuelle Modell der Prüfungsangstentwicklung (SoKoMoP; Eccles, 2007; Gutman & Eccles, 2007; Frome & Eccles, 1998) postuliert geschlechtsspezifische Wirkungspfade zwischen schulbezogenem Elternverhalten und dem Auftreten von Prüfungsangst bei Jugendlichen. Eccles und Kollegen nehmen an, dass sich geschlechtsbezogene Kompetenz- und Leistungserwartungen von Mutter und Vater in deren (Rollen-)Verhalten widerspiegeln, welches die kognitiv-affektiven Manifestationen von Prüfungsangst bei Jungen und Mädchen unterschiedlich beeinflussen soll. Empirische Belege fehlen weitgehend, da bisherige Studien den kombinierten Einfluss von Mutter und Vater gemessen und/oder die Facetten von Prüfungsangst nicht differenziert haben (Brand & Klimes-Dougan, 2010; Leung, Yeung, & Wong, 2010).

Fragestellung: Im Einklang mit den Annahmen des SoKoMoP wurde in der vorliegenden Studie untersucht, ob geschlechtsspezifische Wirkungspfade zwischen schulbezogenem Elternverhalten (Unterstützung und Druck) und den vier Facetten von Prüfungsangst (Besorgtheit, Interferenz, Mangel an Zuversicht und Emotionalität; Hodapp, Rohrman & Ringeisen, 2011) bei jugendlichen Schülerinnen und Schülern bestehen: Es wurde angenommen, dass ausgeprägter mütterlicher (bzw. väterlicher) Druck mit erhöhter Prüfungsangst bei Mädchen (bzw. Jungen) einhergeht, während mütterliche Unterstützung beide Geschlechter vor einem Mangel an Zuversicht schützt.

Methode: Eine Schülerstichprobe der Jahrgangsstufen 9 und 10 (N = 845; 465 Mädchen, 380 Jungen; Mage = 15.32; SD = 0.49) füllten Fragebögen zur wahrgenommenen elterlichen Unterstützung und zum elterlichen Druck (separat für Mutter und Vater; Reitzle, Metzke & Steinhausen, 2001) sowie zu den vier Facetten von Prüfungsangst (Besorgtheit, Interferenz, Mangel an Zuversicht und Emotionalität; Hodapp et al., 2011) aus. Die Schüler stammten aus 22 Sekundarschulen (10 Oberschulen und 12 Gymnasien), die zufällig aus den insgesamt 124 Sekundarschulen in Brandenburg ausgewählt worden waren. Auf Basis eines Multigruppenvergleichs wurden anhand von Strukturgleichungsmodellen mögliche Geschlechtsunterschiede im Zusammenhangsmuster der Variablen untersucht.

Ergebnisse: Wie erwartet zeigten sich geschlechtsspezifische Wirkungspfade zwischen Elternverhalten und Prüfungsangst: Mütterlicher Druck ging mit erhöhter Werten für Emotionalität und Interferenz bei Mädchen einher, während väterlicher Druck (bzw. väterliche Unterstützung) positiv mit Interferenz (bzw. Besorgtheit) bei Jungen korrelierte. Unabhängig vom Geschlecht der Schüler wurde Mangel an Zuversicht positiv durch mütterlichen Druck (bzw. negativ durch mütterliche Unterstützung) vorhergesagt. Die Ergebnisse legen nahe, dass primär das schulbezogene Verhalten der Mutter für den Aufbau von leistungsbezogener Zuversicht bei Mädchen und Jungen verantwortlich ist, während die anderen Facetten der Prüfungsangst durch gleichgeschlechtliche Wirkungspfade zwischen Eltern und Jugendlichen beeinflusst werden. Implikationen für die weitere Forschung zu elterlichen Angstprädiktoren und die Nutzbarmachung der Ergebnisse für den schulischen Kontext werden diskutiert.

Freitag
13. März

9:00 Uhr

HGC 40

Chair(s)

Morgenroth, Stefanie
Braasch, Martina

Schwabe, Ulrike

„Neue Ungleichheiten in einer neuen Hochschullandschaft?“

Die empirisch gut belegten Bildungsungleichheiten nach sozialer Herkunft zeigen sich nach wie vor auch im tertiären Bildungssystem, obwohl im Zuge der Bildungsexpansion die Beteiligung an Hochschulbildung insgesamt beträchtlich zugenommen hat. Wenngleich akademische Abschlüsse damit ihre Exklusivität eingebüßt haben, sind sie immer noch eine entscheidende Voraussetzung zur Erlangung vorteilhafter Erwerbspositionen und höheren Einkommens. Im Zuge dieser Expansion hat sich das deutsche Hochschulsystem beträchtlich ausdifferenziert. Das geplante Symposium thematisiert die Konsequenzen dieser Expansions- und Differenzierungsdynamiken der letzten Jahre für soziale Ungleichheitsmuster im Hochschulbereich: Reduktion oder Vergrößerung? Diese Frage steht im Zentrum einer internationalen Debatte (Shavit et al. 2007), allerdings gibt es bislang kaum aktuelle Forschungsergebnisse aus Deutschland dazu. Einerseits wurden durch die institutionelle Ausdifferenzierung neue Möglichkeiten der verstärkten Beteiligung von Kindern aus weniger privilegierten Familien geschaffen (Inklusion). Dies hat eine stärkere Heterogenität in der Zusammensetzung der Studierendenschaft zur Folge. Andererseits besteht Grund zur Annahme, dass Schüler aus sozial benachteiligten Familien in weniger angesehene oder weniger ertragreiche Bildungsgänge des Tertiärsystems kanalisiert werden, während die privilegierteren Schichten die neuen Differenzierungsmöglichkeiten zu ihren Gunsten nutzen (Ablenkung). Diese Differenzierung nach Ansehen und Ruf bringt somit eine heterogenere Hochschullandschaft hervor.

Das geplante Symposium beleuchtet die Frage nach einer Reduktion oder Verstärkung sozialer Ungleichheit anhand theoretisch fundierter, empirischer Analysen. Zunächst werden die Auswirkungen der Bologna-Reform untersucht. Dazu zeigt Martin Neugebauer, inwiefern die Verkürzung des ersten Studienzyklus zu einer Inklusion von Studienberechtigten aus niedrigeren Sozialschichten geführt hat und inwieweit, umgekehrt, die neu geschaffene Übergangsschwelle vom Bachelor- zum Masterstudium sozial selektiv wirkt. Markus Lörz, Nicolai Netz und Heiko Quast stellen sich anschließend die Frage, ob sich trotz der Ausweitung der Mobilitätsopportunitäten im Zuge dieser Reform soziale Ungleichheiten als persistent erweisen und sozial selektive Studienverläufe hervorbringen. Steffen Jakstats Beitrag fokussiert ferner die Promotionsphase als dritten Zyklus des Bologna-Prozesses und untersucht, inwieweit bei diesem Übergang herkunftsspezifische Unterschiede bestehen. Schließlich untersuchen Felix Weiss, Steffen Schindler und Maria Gerth die Einführung von nationalen Hochschulrankings und fragen, ob diese neue Stratifizierung des zuvor als insgesamt homogen zu bezeichnenden Hochschulsystems eine neue Dimension sozial ungleicher Beteiligung nach sich zieht.

Ein abschließender Diskussionsbeitrag von Marita Jacob ordnet diese Befunde in einen übergeordneten Zusammenhang ein und gibt eine Antwort auf die Frage, inwiefern sich durch neue Differenzierungslinien neue soziale Ungleichheitsmuster manifestieren. Zusammenfassend liefern die Beiträge erste empirische Befunde zu den Konsequenzen aktueller Veränderungen der deutschen Hochschullandschaft für soziale Ungleichheit. Gleichsam wird die Belastbarkeit der bisherigen Ergebnisse thematisiert. Damit leistet dieses Symposium einen wichtigen Beitrag zur evidenzbasierten Reformfolgebewertung.

Freitag
13. März

9:00 Uhr

HGC 50

Chair(s)

Schwabe, Ulrike
Neugebauer, Martin

Neugebauer, Martin

Welchen Einfluss hat die Bachelor-Master Struktur auf soziale Ungleichheiten in der Hochschule?

Der Bologna-Prozess hat durch die Umstellung auf eine Bachelor-Master Struktur die Hochschullandschaft nachhaltig verändert. Einerseits führte die Umstellung zu einer Verkürzung der Zeit bis zum ersten „berufsbefähigenden“ Studienabschluss. Andererseits wurde eine neue Übergangsschwelle zwischen Bachelor- und Masterstudium eingeführt. Der Beitrag untersucht die Konsequenzen dieser Reform für soziale Ungleichheiten vor dem Hintergrund des selbstgesteckten Ziels der Bologna-Reformer, „dass Studierende ihr Studium ungehindert durch ihre sozialen oder wirtschaftlichen Voraussetzungen abschließen können“.

Die Länge der Studiendauer wird immer wieder als eine Ursache genannt, warum Studienberechtigte aus niedrigeren Sozialschichten deutlich seltener studieren als Studienberechtigte aus höheren Schichten. Nach der Reform dauert ein Bachelorstudium nun genauso lange wie eine Ausbildung, weshalb angenommen werden kann, dass sich der Anteil an Studierenden aus niedrigeren Sozialschichten erhöht hat. Der Beitrag nutzt die zeitliche Variation der Umstellung einzelner Studiengänge, um auf Basis von deutschlandweiten (Pseudo)-Paneldaten den Effekt der Verkürzung auf den Anteil von Studierenden aus niedrigeren Sozialschichten zu schätzen. Ergebnisse aus fixed-effects Regressionen deuten auf persistente Ungleichheiten im ersten Studienzyklus hin.

In einem zweiten Schritt untersucht der Beitrag, ob am neu geschaffenen Übergang in das Masterstudium soziale Disparitäten zum Tragen kommen. Die Analysen zeigen, dass Bachelorabsolventen aus niedrigeren Bildungsschichten seltener ein Masterstudium aufnehmen, als solche aus höheren Bildungsschichten. Die beobachteten Unterschiede lassen sich weitgehend auf Unterschiede in der vorangegangenen Bildungskarriere zurückführen: Bildungsfernere Schichten schließen ihr Bachelorstudium häufiger an Fachhochschulen und in solchen Studienfächern ab, die die Aufnahme eines Masterstudiums allein schon wegen des mangelnden Angebots an Masterstudiengängen erschweren (z.B. soziale Arbeit). Zudem erzielen sie schlechtere Studienleistungen, die für den Zugang zum Masterstudium bedeutsam sind.

Da die Verkürzung zu keinem sozialen Ausgleich führt, während die Übergangsschwelle neue soziale Selektion provoziert, trägt die neue Studienstruktur in der Summe zu einer Ausweitung sozialer Ungleichheiten bei.

Freitag
13. März

9:00 Uhr

HGC 50

Chair(s)

Schwabe, Ulrike
Neugebauer, Martin

Lörz, Markus; Netz, Nikolai; Quast, Heiko

Warum ziehen untere Sozialgruppen ein Auslandsstudium seltener in Betracht?

Mit der Bildungsexpansion und der Internationalisierung der Arbeitsmärkte ist die Auslandsmobilität der deutschen Studierenden in den vergangenen Jahrzehnten deutlich angestiegen. Mittlerweile sammelt etwa jeder Dritte im Rahmen seines Studiums (erste) Auslandserfahrungen und diese wirken sich positiv auf die späteren Einkommens- und Karriereaussichten aus. Differenzierte Analysen stellen in diesem Zusammenhang jedoch fest, dass in der studienbezogenen Auslandsmobilität deutliche herkunftsspezifische Unterschiede bestehen und sich diese Unterschiede in den vergangenen Jahren kaum reduziert haben. Aus makro- und mikrotheoretischer Perspektive liegen verschiedene Erklärungsansätze für die Mobilitätsentscheidungen vor, empirisch ist es bislang jedoch nicht hinreichend gelungen die Ursachen und Prozesse zu identifizieren, die den herkunftsspezifischen Unterschieden zugrunde liegen.

Auf Basis der DZHW-Studienberechtigendaten 2010 beschäftigt sich der vorliegende Beitrag in theoretischer und empirischer Hinsicht mit der Frage aus welchen Gründen bildungsferne Gruppen seltener ein Auslandsstudium in Betracht ziehen. Es zeigt sich, dass die Schülerinnen und Schüler bereits während der Schulzeit erste Auslandserfahrungen sammeln und diese die (geplanten) nachschulischen Bildungswege maßgeblich beeinflussen. Da Kinder aus bildungsfernen Familien seltener über solche Erfahrungen verfügen, geringere Unterstützungsleistungen vom Elternhaus zu erwarten haben und sich von einem Auslandsaufenthalt geringere Erträge versprechen, ziehen sie seltener ein Auslandsstudium in Betracht.

Freitag
13. März

9:00 Uhr

HGC 50

Jaksztat, Steffen

Bildungsherkunft und Zugang zur Promotion

Der Beitrag behandelt den Zusammenhang zwischen der Bildungsherkunft von Universitätsabsolvent(inn)en und der Wahrscheinlichkeit einer Promotionsaufnahme nach dem Studium. Vor dem Hintergrund habitus- und entscheidungstheoretischer Überlegungen wird erstens überprüft, inwieweit beim Übergang in die Promotionsphase herkunftsspezifische Unterschiede bestehen, und zweitens, welche Prozesse und Mechanismen zu diesen Unterschieden führen. Darüber hinaus soll in dem Beitrag die Frage gestellt werden, ob und in welchem Umfang sich herkunftsspezifische Übergangswahrscheinlichkeiten in die Promotionsphase im Zeitverlauf geändert haben. Als Datengrundlage für die Analysen dienen die bundesweit repräsentativen Absolvent(inn)enstudien des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW).

Die gängige These, dass aufgrund zunehmend homogener Leistungspotenziale und Bildungsaspirationen bei höheren Bildungsstufen keine Herkunftseffekte auftreten (Mare 1980, 1981), scheint sich am Beispiel der Promotion für Deutschland nicht zu bestätigen. Die multivariaten statistischen Analysen zeigen, dass die Wahrscheinlichkeit der Aufnahme einer Promotion deutlich mit der Bildungsherkunft variiert. Fairlie-Effektdekompositionen verdeutlichen, dass sich dieser Herkunftseffekt zu großen Teilen auf die herkunftsspezifische Fachrichtungswahl, Leistungsunterschiede sowie auf Hilfskrafttätigkeiten während des Studiums zurückführen lässt.

Vorhandene Studien zum Berufserfolg von Promovierten geben Hinweise darauf, dass ein Dokortitel mit fach- und branchenspezifischen monetären sowie nichtmonetären beruflichen Erträgen, wie einem höheren Einkommen, einer höheren Beschäftigungsadäquanz oder einer höheren beruflichen Zufriedenheit, verbunden ist (Enders & Bornmann 2001; Engelage & Schubert 2009; Heineck & Matthes 2012; Falk & Küpper 2013). Vor diesem Hintergrund erhält die Untersuchung sozial ungleicher Zugangschancen zur Promotion und der zugrunde liegenden Entscheidungs- und Rekrutierungsprozesse eine hohe gesellschaftliche Relevanz.

Freitag
13. März

9:00 Uhr

HGC 50

Chair(s)

Schwabe, Ulrike
Neugebauer, Martin

„Neue Ungleichheiten in einer neuen Hochschullandschaft?“

Weiss, Felix; Schindler, Steffen; Gerth, Maria

Hochschulrankings als Kriterium für neue soziale Selektivität in der Studienentscheidung

Durch die interne Ausdifferenzierung des deutschen Hochschulsystems im Zuge seiner Expansion wurden neue Möglichkeiten der verstärkten Beteiligung von Kindern aus weniger privilegierten Haushalten geschaffen. Andererseits stellt sich die Frage, ob mit der Ausdifferenzierung eine neue Dimension sozial ungleicher Beteiligung einhergeht. In unserem Beitrag stellen wir die empirische Frage, ob ein solcher Prozess durch die Veröffentlichung von Hochschulrankings in den 1990er Jahren angestoßen wurde. Insbesondere auf Initiative des CHE (Centrum für Hochschulentwicklung), einer Einrichtung die eng mit der deutschen Hochschulpolitik verweben ist, wurden Fachbereiche nach unterschiedlichen Qualitätskriterien, wie ihrer Publikationstätigkeit oder der Zufriedenheit der aktuellen Studierenden, verglichen und die Ergebnisse in breit rezipierten Medien veröffentlicht.

Aus theoretischer Sicht können widersprechende Hypothesen entwickelt werden. Zum einen könnten Rankings durch einen öffentlichen Vergleich bewirken, dass bisher nur in akademischen Kreisen bekannte Informationen über Hochschulqualität allen zugänglich gemacht wird, sozusagen eine Demokratisierung von Information stattfindet. Dies könnten gerade leistungsstarke junge Erwachsene aus bildungsfernen Haushalten dazu nutzen, bessere Hochschulen zu wählen und den Informationsvorsprung von Akademikerkindern auszugleichen. Andererseits könnte aber auch eine neue Dimension sozial ungleicher Teilnahme an Hochschulbildung entstehen. So ist zu erwarten, dass die steigende Nachfrage nach den gut gerankten Programmen zu erhöhten Zulassungsvoraussetzungen führt. Im Lichte eines immer geringeren Vorteils durch Hochschulbildung per se kann außerdem erwartet werden, dass junge Erwachsene aus Akademikerfamilien, das „Exzellenz-Label“ als eine neue Möglichkeiten zur sozialen Differenzierung sehen und verstärkt Anstrengungen unternehmen, um an diesen Hochschulen zu studieren. Darüber hinaus könnte ein gutes Ranking-Ergebnis als Signal für hohe Anforderungen im Studium gesehen werden, das Kinder aus bildungsfernen Familien als Risiko in Bezug auf den Studienerfolg interpretieren könnten. Sie blieben dann möglicherweise sogar verstärkt solchen Hochschulen fern.

Empirisch untersucht unser Beitrag den Einfluss der sozialen Herkunft auf die Wahl einer (später) hoch gerankten Hochschule vor und nach der Veröffentlichung der Rankingergebnisse. Wir beschränken uns auf die zwei großen und häufig gerankten Fachbereiche der Wirtschafts- und der Ingenieurwissenschaften. Als zentrales Ranking konzentrieren wir uns auf das CHE-Ranking. Empirisch vergleichen wir anhand der HIS/DZHW-Studienanfängerbefragung die Eintrittskohorten 1993, 1996, 2000 und 2005 im Hinblick auf die Unterschiede nach sozialer Herkunft beim Eintritt in eine Hochschule, die nach dem CHE-Ranking 1998 ein Top-Ranking in der Professorenempfehlung oder der Bewertung der Forschungsleistung erhielten. Hierbei beachten wir sowohl Universitäten als auch Fachhochschulen. Unsere Befunde zeigen, dass in den beiden letzten Eintrittskohorten verstärkt junge Erwachsene aus privilegierten Haushalten in den Jahrgängen 2000 und 2005 in hoch gerankte Hochschulen eintreten. Allerdings konnte ein Anstieg der Ungleichheit beim Eintritt in diese Hochschulen bereits beim Eintrittsjahrgang 1996 festgestellt werden, was der These widerspricht, dass alleine das CHE-Ranking für die Unterschiede verantwortlich zu machen ist. Stattdessen schlussfolgern wir, dass die Rankings Teil einer generelleren Entwicklung hin zu einer stärkeren Ausdifferenzierung des Bildungssystems sind, die mit neuen sozialen Ungleichheiten einhergeht.

Freitag
13. März

9:00 Uhr

HGC 50

Chair(s)

Schwabe, Ulrike
Neugebauer, Martin

Scharenberg, Katja; Rudin, Melania; Müller, Barbara; Meyer, Thomas; Hupka-Brunner, Sandra
Die Bedeutung von schulischen Basiskompetenzen und Schulformzugehörigkeit am Ende der Pflichtschulzeit für den Bildungsstand im Erwachsenenalter in der Schweiz

Am Ende der Pflichtschulzeit sollten Schülerinnen und Schüler „für das Leben gerüstet“ (Zahner et al., 2002) sein und grundlegende Basiskompetenzen erworben haben, die ihnen einen erfolgreichen Weg ins Erwachsenenleben ermöglichen. PISA erhebt dabei den Anspruch, Kompetenzen zu erfassen, die die Fähigkeit junger Menschen zur aktiven Teilhabe am gesellschaftlichen Leben abbilden und anschlussfähig für kontinuierliches Weiterlernen sind (für PISA 2000 vgl. bereits Baumert, Stanat & Demmrich, 2001; für PISA 2012 vgl. Sälzer & Prenzel, 2013). In diesem Sinne sind die erzielten Basiskompetenzen am Ende der Pflichtschulzeit Ergebnisse (Output) schulischer Bildungsprozesse, deren langfristige Wirkungen (Outcomes) sich dagegen erst an Bildungserträgen z.B. in Form von weiterführenden Bildungsabschlüssen messen lassen (Ditton, 2000).

Dem Bildungssystem kommt dabei in mindestens zweierlei Hinsicht eine immense Bedeutung zu (Fend, 1981): Einerseits soll es Jugendliche unabhängig von ihrer Herkunft befähigen und sie auf dem Weg ins Erwachsenenleben begleiten. Andererseits bewertet Schule die Leistungen der Jugendlichen und stellt Leistungseinschätzungen (Schulnoten, besuchte Schulform, Bildungsabschlüsse) bereit, die weiterführenden Schulen und Lehrbetrieben als „Signal“ (Spence, 1973) für die Leistungsfähigkeit und Produktivität von Schulabgängern dienen (Kronig, 2010). Auch aus lebenslaufsoziologischer Perspektive lässt sich vermuten, dass die institutionelle Ausgestaltung des Bildungssystems und der Zeitpunkt individueller Übergänge den weiteren Bildungsverlauf und Bildungsstand beeinflussen (Blossfeld & Mills, 2009; Maaz, Watermann & Baumert, 2007).

Unserem Beitrag liegen folgende Forschungsfragen zugrunde: Inwieweit vermögen schulische Leistungssignale und Basiskompetenzen den späteren Bildungserfolg im Erwachsenenalter und die Weiterqualifikation über die Pflichtschulzeit hinaus vorherzusagen? Welchen Einfluss haben die von PISA gemessenen Basiskompetenzen sowie schulische Leistungsindikatoren (z.B. Schulnoten, Schulformzugehörigkeit) im Verhältnis zu individuellen Herkunftsmerkmalen (z.B. soziale Herkunft, Migrationshintergrund) und Laufbahnfaktoren (z.B. verzögerter Übergang in postobligatorische Ausbildungen der Sekundarstufe II) auf den erreichten Bildungsstand im Erwachsenenalter?

Datengrundlage ist die Schweizer Studie „Transitionen von der Erstausbildung ins Erwerbsleben“ (TREE). TREE ist in der Schweiz die erste repräsentative Längsschnitterhebung zum Übergang Jugendlicher von der Schule ins Erwerbs- und Erwachsenenleben. Die TREE-Stichprobe umfasst rund 6.000 Jugendliche, die an PISA 2000 (BFS & EDK, 2002) teilnahmen, im selben Jahr aus der Schulpflicht entlassen wurden und seitdem über nunmehr neun Messzeitpunkte längsschnittlich begleitet werden.

Abhängige Variable ist der höchste erreichte Bildungsabschluss zehn Jahre nach Schulaustritt (kein Abschluss: 10 %, allgemeinbildender [11 %] oder berufsbildender Abschluss [50 %] auf der Sekundarstufe II, Abschlüsse auf der Tertiärstufe [29 %]). Multinomial logistische Regressionsanalysen (Backhaus, Erichson, Plinke & Weiber, 2000) berücksichtigen bei der Schätzung durchschnittlicher Marginaleffekte (1) individuelle Herkunftsmerkmale (z.B. Migrationshintergrund, soziale Herkunft), (2) schul- und ausbildungsverlaufsbezogene Merkmale (z.B. PISA-Lesekompetenz, Schulformzugehörigkeit, Schulnoten, Ausbildungsstatus in den ersten beiden Jahren nach Ende der Pflichtschulzeit) sowie (3) sozio geografische Besonderheiten in der Schweiz (z.B. Sprachregion).

Die Ergebnisse (Scharenberg et al., 2014) lassen sich wie folgt zusammenfassen: (1) Die am Ende der Pflichtschulzeit erreichte PISA-Lesekompetenz erweist sich als relevanter Prädiktor für den höchsten erreichten Bildungsabschluss im jungen Erwachsenenalter: Personen mit mittlerer oder hoher PISA-Lesekompetenz haben eine signifikant höhere Chance, einen Hochschulabschluss zu erwerben, als Personen mit niedriger PISA-Lesekompetenz. (2) Die Schulformzugehörigkeit am Ende der Sekundarstufe I ist dagegen für alle Abschlussarten entscheidend: Für Personen, die einen Schultyp mit Grundanforderungen besucht haben, ist im Vergleich zum Besuch anderer Schultypen zwar die Chance auf einen Abschluss der Sekundarstufe II signifikant höher, sie haben aber signifikant geringere Chancen auf einen Hochschulabschluss. Für sie besteht zudem ein erhöhtes Risiko, ohne weiterführenden Abschluss zu bleiben. (3) Auch die Schulnoten sind für die Selektionsprozesse nach der Pflichtschulzeit zentral: Gute Schulnoten in der Unterrichtssprache oder in Mathematik erhöhen die Wahrscheinlichkeit auf einen Hochschulabschluss.

Die Befunde verdeutlichen somit, dass sich schulische Basiskompetenzen, Leistungsindikatoren und die Schulformzugehörigkeit in der Sekundarstufe I über individuelle Herkunftsmerkmale und Laufbahnfaktoren hinaus selbst zehn Jahre nach Beendigung der Pflichtschulzeit noch nachhaltig auf den Bildungsstand im jungen Erwachsenenalter auswirken. Darüber hinaus zeigen sich sprachregionale Disparitäten im Bildungsstand.

Freitag
13. März

9:00 Uhr

GC 03/142

Chair(s)
Scharenberg, Katja

Schoemann, Klaus; Weiss, Christina; Knauber, Carolin

Entwicklung der Basiskompetenzen über den Lebensverlauf – Analysen mit PIAAC und CiLL Daten

Die Basiskompetenzen wie Literalität, alltagsmathematische Kompetenzen und technologiebasiertes Problemlösen sind bereits nach Vollendung der schulischen Vollzeitausbildung sehr heterogen über die Bevölkerung verteilt (Rammstedt 2013). Das ist zunächst nicht überraschend unter Berücksichtigung der unterschiedlichen Ausbildungsdauern von Personen bis zum Alter von 25 Jahren. Ab diesem Zeitpunkt im Lebensverlauf gestalten sich die Berufsverläufe und Kompetenzverläufe jedoch zusätzlich sehr unterschiedlich. Arbeitsmarkttheorien wie die Humankapitaltheorie oder die Segmentationstheorie lassen deutlich unterschiedliche weitere Berufsverläufe erwarten. Die Ungleichheiten beim Einstieg in den Arbeitsmarkt setzen sich meistens über den Lebensverlauf hinweg fort. Höhere Ausbildung ermöglicht eher eine Erweiterung der Kompetenzen im Gegensatz zu niedrig qualifizierten Personen, die meistens weniger Möglichkeiten für Weiterbildung und beruflichen Aufstieg realisieren.

Der Beitrag untersucht mit Daten von PIAAC für Personen im Alter zwischen 25-65 und für Personen zwischen 66-80 Jahren mit Daten der Studie „Competences in later life CiLL“ CiLL die Auswirkungen des Werdegangs im allgemeinen und beruflichen Bildungssystem auf die Basiskompetenzen zu unterschiedlichen Zeitpunkten des Lebensverlaufs (Gebrande et al 2014; Tippelt et al. 2014). Die forschungsleitende Hypothese besteht in der Untersuchung der Heterogenität in den Basiskompetenzen über die Lebensspanne von 25 bis 80 Jahren. Aufbauend auf den deskriptiven Befunden zur Heterogenität des Kompetenzerhalts bzw. –abbaus nach Altersgruppen werden Regressionsmodelle vorgestellt, die wichtige Faktoren für die beobachtbaren Prozesse aufzeigen.

Erste Untersuchungen weisen auf die Bedeutung von Aktivitäten wie Weiterbildung und den beruflichen Kontext hin. Für den Kompetenzerhalt bis in die Altersgruppe 66-80 erweist sich neben der Gesundheit auch die Bedeutung der Dauer des Arbeitslebens, d.h. ein möglichst später Übergang in den Ruhestand als wichtig. Der selektive Prozess der Beteiligung an Weiterbildung (Schömann und O'Connell 2002) wirft somit einen langen Schatten für nicht im Erwerbsleben stehende Personen und nicht in der Weiterbildung aktive Personen. Zusätzlich bietet ehrenamtliche Tätigkeit einen Beitrag zum Kompetenzerhalt insbesondere im höheren Alter.

Im Vortrag wird thematisiert, wie sich die Heterogenität über die Altersgruppen hinweg darstellt und ob es angemessen ist von einer zu erwartenden anhaltenden oder gar zunehmenden Ungleichheit über den Lebensverlauf zu sprechen. Da bisher nur Querschnittsdaten vorliegen wird ein Ausblick auf mögliche Fragestellungen für die in Vorbereitung befindliche PIAAC-Längsschnittstudie angesprochen.

Freitag
13. März

9:00 Uhr

GC 03/142

Chair(s)

Scharenberg, Katja

Kaak, Silvio; Heinrichs, Karin; Lipowski, Katrin; Wuttke, Eveline; Kracke, Bärbel

Der Fragebogen zur Berufswahlkompetenz – Ein Instrument zur Individualisierung der Berufsorientierung?

Der Übergang von der Schule in eine nachschulische Ausbildung ist eine zentrale Entwicklungsaufgabe im Jugendalter, die von der Schule durch gezielte pädagogische Angebote zur beruflichen Orientierung unterstützt werden soll. Um die Heranwachsenden im beruflichen Explorationsprozess individualisiert begleiten zu können, bedarf es der Messung und Förderung spezifischer Berufswahlkompetenzen. Hierfür wurde im Forschungsprojekt Thüringer Berufsorientierungsmodell (ThüBOM) ein Kompetenzmodell zur Berufswahl entwickelt. Dieses beschreibt vier Phasen, die Heranwachsende durchlaufen, um am Ende der Schulzeit eine begründete Berufswahlentscheidung treffen zu können. Drei Kompetenzdimensionen liegen dem Modell zugrunde - Wissen, Motivation und Handlung. Es dient als Rahmenmodell für die Förderung von Berufswahlkompetenzen und die Konstruktion von diagnostischen Instrumenten. Eine gezielte Diagnostik ist die Grundlage, um aktuellen Anforderungen an Individualisierung und den Umgang mit Heterogenität im Kontext der Berufsorientierung gerecht zu werden. Mithilfe des Fragebogens zur Berufswahlkompetenz lässt sich der Entwicklungsstand von SchülerInnen im Berufswahlprozess feststellen und auf Grundlage der Ergebnisse können individualisierte Fördermaßnahmen und Entwicklungsangebote im Rahmen der schulischen Berufsorientierung bereitgestellt werden.

Fragestellungen: Im Rahmen des Beitrags wird geprüft, ob das Fragebogeninstrument in verschiedenen Schulformen und Altersstufen gleichermaßen eingesetzt werden kann und Entwicklungsunterschiede sensibel erfasst werden.

Methode: Zur Beantwortung der Fragestellung wurden Querschnittsdaten von 700 SchülerInnen unterschiedlicher Schulformen (Gymnasien, Gesamtschulen und Berufsbildende Schulen) und Altersstufen aus den Bundesländern Hessen und Nordrhein-Westfalen erhoben und mithilfe multivariater statistischer Verfahren analysiert.

Ergebnisse: Erste Varianz- und Regressionsanalysen weisen auf systematische Unterschiede der Berufswahlkompetenz zwischen verschiedenen Schulformen und Altersstufen hin. Unter den SchülerInnen an beruflichen Schulen zeigt sich, dass insbesondere Gymnasiasten eine geringere Berufswahlkompetenz aufweisen. Zudem ist bei SchülerInnen, die noch keine Vorstellungen dazu entwickelt haben, welchen Beruf Sie anstreben möchten, die Berufswahlkompetenz niedriger ausgeprägt, als bei denjenigen, die bereits wissen, was sie nach dem Schulabschluss tun werden. Die Berufswahlkompetenz erweist sich darüber hinaus als Prädiktor des individuellen Fortschritts im Berufswahlprozess. Die Ergebnisse stützen somit die Annahme, dass das Instrument zur Individualisierung der schulischen Berufsorientierung beiträgt.

Volodina, Anna; Nagy, Gabriel

Prädiktoren des Übergangs in die berufliche Erstausbildung: Welche Konstruktbereiche eignen sich besonders für die Vorhersage individueller Ausbildungswahlentscheidungen

Der schulische und berufliche Ausbildungsweg eines Individuums ist von vielen Wahlentscheidungen gekennzeichnet. Die Entscheidung für eine bestimmte berufliche Ausbildung hat besonders weitreichende Konsequenzen für das weitere Berufsleben, da sie den Rahmen der späteren beruflichen Möglichkeiten maßgeblich determiniert. Die Untersuchung personenseitiger Determinanten der Berufswahl hat eine lange Tradition. Jedoch fehlt es an empirischen Untersuchungen, in denen die eine gemeinsame Betrachtung der Vorhersageleistung verschiedener Konstruktbereiche vorgenommen wird. Das Ziel der vorliegenden Arbeit ist die Identifikation von Konstruktbereichen, die sich besonders zur Vorhersage der Berufswahlentscheidungen von Realschülern, die den Übergang in die berufliche Erstausbildung anstreben, eignen. Konkret werden folgende Konstruktbereiche untersucht: Kognitive und fachgebundene Kompetenzen, Schulnoten, akademische Selbstkonzepte und berufliche Interessen.

Die bisherige Forschung liefert eine Fülle von Hinweisen, dass sich die Mitglieder unterschiedlicher Berufsgruppen in ihren kognitiven und fachgebundenen Kompetenzen systematisch voneinander unterscheiden (z.B. Gottfredson, 1986). Da die mit objektiven Testverfahren erfassten Kompetenzen den Individuen jedoch nicht bekannt sind, liegt es nahe anzunehmen, dass deren Effekte auf die Übergangentscheidungen durch die in der Schule in Form von Zensuren rückgemeldeten Leistungen zurückgeführt werden können (z.B. Trost, 2005). Eine Vielzahl der in der pädagogischen Psychologie verorteten theoretischen Modelle legen dabei nahe, dass die Effekte von Schulnoten weitgehend durch die daraus abgeleiteten subjektiven Fähigkeitsselbsteinschätzungen (d.h. akademische Selbstkonzepte; Marsh & Yeung, 1997) und damit assoziierten Interessen vermittelt werden (z.B. Eccles, 1983). Im Bereich der Interessenforschung haben sich dabei die von Holland (1997) entwickelte Interessentaxonomie als besonders leistungsstarker Prädiktor von Berufswahlentscheidungen herausgestellt (Päßler & Hell, 2012).

Die Datengrundlage des vorliegenden Beitrags liefert die Studie „Transformation des Sekundarschulsystems und akademische Karrieren: Klassenstufe 10“ (TOSCA-10). Im vorliegenden Beitrag konzentrieren wir uns auf diejenigen Schülerinnen und Schüler, die in der 10. Klasse eine Realschule besucht haben und angegeben haben, dass sie sich für eine berufliche Ausbildung beworben haben (N = 972, Durchschnittsalter 16.78 (SD = .64), 57.0% weiblich). Schülerkompetenzen in den Bereichen Mathematik, Biologie, Deutsch, Englisch, Wirtschaft und Technik wurden durch etablierte Instrumente erhoben. Die individuelle kognitive Grundfähigkeit wurde durch den Kognitiven Fähigkeitstest (Heller & Perleth, 2000) erfasst. Halbjahresnoten in den Fächern Mathematik, Englisch, Deutsch, Physik, Biologie, Erd- und Gemeinschaftskunde wurden den Schülerakten entnommen. Akademische Selbstkonzeptskalen lagen für die Bereiche Mathematik, Sprache, Wirtschaft, Technik/Handwerk und Englisch vor (Schwanzer, Trautwein, Lüdtke & Sydow, 2005). Zur Messung der beruflichen Interessen wurde der revidierte Allgemeine Interessen-Struktur-Test (Bergmann & Eder, 2005) eingesetzt. Das Geschlecht fungierte als Kontrollvariable und wurde den Schülerakten entnommen. Die abhängige Variable bildete der von den Schülerinnen und Schülern angegebene Wunschberuf, der anhand des von der Bundesagentur für Arbeit (2010) entwickelten Berufssystems kategorisiert wurde (fünf homogene Berufsfeldkategorien).

Die Hauptergebnisse multinomialer Regressionsanalysen lassen sich wie folgt zusammenfassen: Gemessen an der Pseudo-R²-Statistik von Nagelkerke erwiesen sich die beruflichen Interessen als die leistungsstärksten Prädiktoren der Berufspräferenzen (R² = .60), während objektive Testleistungen (R² = .24), Schulnoten (R² = .22), akademische Selbstkonzepte (R² = .42) und das Geschlecht (R² = .23) schwächere Zusammenhänge aufwiesen. Die Vorhersageleistung der beruflichen Interessen konnte durch die simultane Verwendung aller Prädiktoren nur unwesentlich erhöht werden ($\Delta R^2 = .06$).

Die ermittelten Befunde legen nahe, dass berufliche Interessen nach Holland (1997) als proximale Determinanten individueller Berufspräferenzen von Realschulabgängern, die am Übergang in die berufliche Erstausbildung stehen, verstanden werden können. Insofern der Übergang in die berufliche Erstausbildung interessenbasiert stattfindet, stellen sich in Folge des Zusammenhangs zwischen beruflichen Interessen und objektiven Testleistungen, Schulnoten sowie akademischen Selbstkonzepten bereits beim Eintritt in die berufliche Erstausbildung systematische Unterschiede zwischen den Berufsgruppen in allen betrachteten Personenmerkmalen ein.

Freitag
13. März

9:00 Uhr

GC 03/142

Chair(s)

Scharenberg, Katja

Dag Tjaden, Jasper

Unterschiede in den Bildungsentscheidungen von Hauptschülern am Ende der Pflichtschulzeit

Der Beitrag widmet sich der Frage, welchen Einfluss gruppenspezifische Bildungspräferenzen auf den Verbleib von Hauptschülern nach der Pflichtschulzeit haben. Der Schwerpunkt liegt auf Unterschieden zwischen Hauptschülern mit und jenen ohne Migrationshintergrund. Gerade Jugendliche mit Migrationshintergrund weisen geringere Einmündungschancen in eine (betriebliche) Berufsausbildung (u.a. Diehl et al. 2009; Hunkler 2010) und ein höheres Verbleibsrisiko im sogenannten ‚Übergangssystem‘ auf (Ulrich 2008). In den vorhandenen Studien spielen Bildungspräferenzen zur Erklärung ethnischer Unterschiede eine Nebenrolle. Das überrascht insbesondere deshalb, weil an anderer Stelle gezeigt wurde, dass sich Bildungspräferenzen bzw. Aspirationen erheblich in Bezug auf den Migrationshintergrund unterscheiden (Stanat et al. 2010; Gresch 2012; Relikowski et al. 2012). Weitere Studien deuten z.B. für den Übergang in die Sekundarstufe I und die Hochschule darauf hin, dass die relativ hohen Übergangschancen auf die hohen Bildungsaspirationen von Jugendlichen mit Migrationshintergrund zurückzuführen sind (Kristen et al. 2008; Kristen und Dollmann 2010).

Auch Hauptschüler stehen in Deutschland nach Ablauf der 9. Klasse vor einer Entscheidung: ‚Aufwärtsqualifizierung‘ oder Berufsausbildung (Schuchart 2010). In den meisten Bundesländern ist es mittlerweile möglich den mittleren Abschluss an der gleichen Schule zu erwerben. Darüber hinaus besteht in regionaler Abhängigkeit ein vielseitiges Angebot für berufsbildende und allgemeine Bildungsabschlüsse. Die Analyse der Situation von Hauptschülern ist aufgrund der (institutionellen)Komplexität der Bildungslandschaft weitaus komplexer als an anderen Bildungsetappen.

Nach Boudon (1974) wird der Übergang von Hauptschülern nach der 9. Klasse als sekundärer Effekt (choice effect) modelliert. Sekundäre Effekte beschreiben leistungsunabhängige Einflüsse auf einen Bildungsübergang. Bildungsentscheidungen werden in der soziologisch orientierten Bildungsforschung als subjektive Kosten-Nutzen-Kalkulation verstanden (Erikson und Jonsson 1996; Breen und Goldthorpe 1997; Esser 1999). Unterschiede in den Übergängen zwischen Jugendlichen mit und jenen ohne Migrationshintergrund werden somit auf Unterschiede in den zieloptimierenden, nutzenmaximierenden Abwägungen über die Bildungsalternativen zurückgeführt.

Die empirische Analyse basiert auf Daten des ‚Nationalen Bildungspanels‘ (NEPS 2010, Startkohorte 4) (Blossfeld et al. 2011). Für die Untersuchung werden die erste und die zweite Wellte (2010-2011) der Neuntklässler-Kohorte verwendet. Es handelt sich somit um prospektive Längsschnittdaten, die es ermöglichen, die kausale Richtung der Effekte zu berücksichtigen. Die NEPS-Daten bieten darüber hinaus einige Vorteile: eine repräsentative Stichprobe von Hauptschulen in Deutschland; Informationen zum Humankapital in Form von Schulnoten und standardisierten Leistungstests; weitgehende Informationen zum sozialen Hintergrund und zu den Bildungspräferenzen der Schüler, sowie eine differenzierte Erhebung des Migrationshintergrunds.

Die vorläufigen Analysen in Form von logistischen und multinominalen Regressionen deuten auf erhebliche sekundäre Effekte in Bezug auf den Migrationshintergrund hin. In den meisten Fällen können Bildungspräferenzen die ethnischen Unterschiede im Übergang an der ersten Schwelle aufklären. Hauptschüler mit Migrationshintergrund tendieren an dieser Schwelle häufiger zu schulischen Bildungsangeboten, die den Erwerb eines mittleren Abschlusses bzw. der Fachhochschulreife ermöglichen, als zum Direkteinstieg in die Berufsausbildung. Dieser Befund lässt eine Vielzahl interessanter Interpretationen zu. Aus der Perspektive der Bildungssoziologie werden die Ergebnisse in Bezug auf drei verbreitete Thesen diskutiert: Selbstselektion, Zuwanderungsoptimismus und ‚blocked opportunity‘-Annahmen. Aus der Perspektive der Praxis verweist der Befund (der starke Einfluss der Bildungspräferenz) auf das Potential von Berufsorientierungsmaßnahmen.

Lösch, Thomas; Kelava, Augustin; Nagengast, Benjamin; Trautwein, Ulrich

Wer überschätzt sich? Zur Erklärung von verzerrten Selbsteinschätzung von Abiturleistungen

Theoretischer Hintergrund: In Schulen werden häufig Prüfungsleistungen erbracht. Noch bevor Schülerinnen und Schüler (SuS) eine externe Rückmeldung in Form von Noten erhalten, mutmaßen sie über ihre erzielten Leistungen. Diese Einschätzungen sind häufig verzerrt (Dunning, Heath & Suls, 2004), was bei habituellem Vorkommen möglicherweise negative Folgen haben kann. Beispielsweise führten verzerrte Selbsteinschätzungen bereits in experimentellen Studien zu schlechteren Leistungen (Thiede, Anderson & Theriault, 2003). Dies galt insbesondere für Überschätzungen (Dunlosky & Rawson, 2012). Dennoch ist wenig über die personenbezogenen und institutionsbezogenen Determinanten bekannt, die im realen Schulkontext zu Überschätzung führen.

Auf Individualebene (d.h. personenbezogen) zeigte sich in früheren Studien, dass sich männliche Schüler stärker als Schülerinnen überschätzten (Lundeberg, Fox & Puncoha, 1994). Weiterhin konnte gezeigt werden, dass eine Überschätzung mit einer geringeren Kompetenz einherging (Kruger & Dunning, 1999). Darüber hinaus schien das (allgemeinere) akademische Selbstkonzept konkretere Selbsteinschätzungen (z.B. die Einschätzung der Note in einem gegebenen Test) zu beeinflussen (Zhao & Linderholm, 2008). Die Persönlichkeitsvariable Narzissmus zeigte ganz allgemein einen Zusammenhang mit Selbstüberschätzung (Back et al., 2013).

Auf Schulebene (d.h. institutionsbezogen) beschreibt der Big-Fish-Little-Pond-Effekt (BFLPE; Marsh, 1987) einen negativen Zusammenhang zwischen der durchschnittlichen Kompetenz auf Schulebene und dem akademischen Selbstkonzept. Ein ähnlicher Effekt könnte ebenfalls bei der Einschätzung von Abiturleistungen wirken und zu Unterschätzung führen.

Um die Überschätzung von SuS zu erklären, werden in dieser korrelativen Studie verschiedene Prädiktoren aus den Bereichen Persönlichkeit und Motivation herangezogen und um den Effekt von Kontextvariablen im Bereich Schule ergänzt.

Fragestellung: Das Ziel dieser Studie ist die Identifikation von Prädiktoren der Überschätzung von Abiturleistungen. Auf Schülerebene werden die Hypothesen untersucht, dass männliches Geschlecht, niedrige Kompetenz, ein hohes akademisches Selbstkonzept, sowie hoher Narzissmus Überschätzung vorhersagen (unter Kontrolle der jeweils anderen Prädiktoren). Auf Schulebene wird die Hypothese untersucht, dass es einen BFLPE der Überschätzung gibt.

Methode: Die Stichprobe umfasste N = 5,566 SuS (53.16 % weiblich), die 2005 in Hamburg ihr Abitur machten (siehe Trautwein, Köller, Lehmann & Lüdtke, 2007). Das Durchschnittsalter betrug M = 19.87 Jahre (SD = 1.02). Nach den schriftlichen Abiturprüfungen schätzten die SuS ihre jeweiligen Prüfungsnoten ein. Die Differenzen in Englisch und Mathematik aus dieser Einschätzung und der tatsächlichen Abiturleistung wurden als abhängige Variablen verwendet. Als Prädiktoren wurden standardisierte Leistungstests in Mathematik und Englisch (Köller & Trautwein, 2004), domänenspezifische Selbstkonzeptskalen ($\alpha = .86$ in Mathematik und $\alpha = .84$ in Englisch), sowie eine adaptierte Skala des Narcissic Personality Inventory ($\alpha = .86$; Raskin & Hall, 1979) verwendet.

Statistische Analysen wurden mit Mplus 7.12 (Muthen & Muthen, 2012) durchgeführt. Zur Überprüfung der Hypothesen wurden mehrere Mehrebenen-Regressionsmodelle geschätzt. Um die Modelle korrekt zu spezifizieren wurden individuelle Prädiktoren parallel auch auf Schulebene aufgenommen. Zusätzlich wurde die Schulform als Kovariate verwendet. Die Prädiktoren wurden in folgender Reihenfolge aufgenommen: Geschlecht, Kompetenz, Schulform, Selbstkonzept, Narzissmus. Zur Schätzung der Effekte auf Schulebene wurde ein manifester Kovariaten-Ansatz gewählt (vgl. Lüdtke et al., 2008). Es wurden separate Modelle für Englisch und Mathematik geschätzt.

Ergebnisse: Insgesamt zeigte sich, dass individuelle Fähigkeiten aber auch Merkmale der Persönlichkeit und Schule eine Überschätzung vorhersagen. Auf Individualebene überschätzten sich männliche Schüler im Durchschnitt stärker als Schülerinnen. Weiterhin stand eine geringere Kompetenz mit einer stärkeren Überschätzung in Zusammenhang. Das akademische Selbstkonzept ging in Mathematik mit Überschätzung einher. Narzissmus dagegen sagte eine stärkere Überschätzung in Englisch vorher. Auf Schulebene trat der BFLPE in Mathematik nicht und in Englisch in umgekehrter Form auf, d.h. SuS überschätzten sich mehr an Schulen mit höherer durchschnittlicher Kompetenz. Mit der vorliegenden Studie wurde somit ein erster Schritt unternommen, Überschätzung einerseits mit Variablen auf Individualebene und gleichzeitig mit Variablen auf Schulebene in einem angewandten Kontext zu verknüpfen.

Freitag
13. März

9:00 Uhr

GC 03/149

Chair(s)

Lösch, Thomas

Flunger, Barbara; Trautwein, Ulrich; Nagengast, Benjamin; Lüdtke, Oliver; Niggli, Alois; Schnyder, Inge
Langes Sitzen an den Hausaufgaben bedeutet nicht unbedingt höheren Lernerfolg: Lerntypen während der Hausaufgabenbearbeitung und ihr Zusammenhang mit Leistungsentwicklung

Hausaufgaben (HA) gehören insbesondere in den Fächern Mathematik, Deutsch und Fremdsprache zu den zentralen Bestandteilen des Schulunterrichts und sollen u.a. den Lernerfolg steigern. Das Engagement von Schülerinnen und Schülern bei der Hausaufgabenbearbeitung kann durch den Aufwand an Zeit und an Compliance beschrieben werden (vgl. u.a. Schmitz & Skinner, 1993). Die Ergebnisse von Studien zum Zusammenhang zwischen Hausaufgabenzeit und Schülerleistung sind inkonsistent (z.B., Cooper, Robinson, & Patall, 2006). Bezüglich der HA-Compliance wurde weitgehend konsistent ein positiver Zusammenhang mit Schülerleistung gefunden (Trautwein & Köller, 2003; Trautwein, 2007; Trautwein, Lüdtke, Schnyder & Niggli, 2006). Wenn HA-Zeitaufwand und HA-Compliance verglichen werden, scheint HA-Compliance demnach einen stärkeren Zusammenhang zur Schülerleistung zu haben.

Allerdings wurden bisher die Wirkungsweisen von HA-Zeit und Compliance lediglich getrennt voneinander betrachtet. Man könnte jedoch annehmen, dass sich Unterschiede zwischen Schülerinnen und Schülern in spezifischen Ausprägungen bezüglich beider Variablen finden lassen. Eine Gruppe von Schülerinnen und Schülern könnte durch großes HA-Engagement charakterisiert sein, was durch ein hohes Maß an Zeitaufwand und Compliance gekennzeichnet sein sollte. Eine andere Gruppe von Schülerinnen und Schülern verbringt hingegen eventuell viel Zeit mit Hausaufgaben, arbeitet dabei jedoch wenig sorgfältig. HA-Zeitaufwand könnte eine wichtige Informationsquelle für das Hausaufgabenverhalten darstellen, wenn HA-Compliance simultan berücksichtigt und das Zusammenspiel der Variablen untersucht wird. Zusätzlich könnten bei Schülerinnen und Schülern weitere Konfigurationen der HA-Bearbeitung (in anderen Worten Lerntypen) identifiziert werden, die sich wiederum durch ein charakteristisches Wirkungsgefüge von HA-Zeitaufwand und Compliance von den weiteren Lerntypen unterscheiden lassen.

Der vorliegende Beitrag untersucht drei Fragestellungen, die innerhalb eines DFG-Projekts zur Re-Analyse eines großen Datensatzes zum HA-Verhalten erforscht wurden.

Die Studie basiert auf einer Stichprobe von 1915 Schweizer Schülerinnen und Schülern der 8. Jahrgangsstufe, die zu ihrem HA-Verhalten im Fach Französisch zu zwei Zeitpunkten (September/Oktober und Juni 2004) befragt worden waren. Erstens wurde erforscht, inwieweit spezifische Lerntypen bestimmt werden können, auf Grundlage von HA-Zeit und HA-Compliance. Es wurde erwartet, dass sich mindestens vier Lerntypen mit folgenden Charakteristika finden lassen würden: (1) Hoher Zeitaufwand/hohe Compliance („hoch engagierte Lerner“), (2) niedriger Zeitaufwand/hohe Compliance („schnelle Lerner“), (3) hoher Zeitaufwand/niedrige Compliance („sich abmühende Lerner“), (4) niedriger Zeitaufwand/niedrige Compliance („Minimalist“). Zweitens wurde angenommen, dass für Lerntypen mit hoher HA-Compliance im Längsschnitt bessere Leistungsergebnisse gefunden werden können, da in bisherigen Studien konsistent positive Assoziationen von HA-Compliance und Schülerleistung nachgewiesen wurden. Drittens wurde untersucht, inwieweit verschiedene Merkmale der Schülerinnen und Schüler (IQ, Wertüberzeugungen, Kompetenzeinschätzung und Gewissenhaftigkeit) mit ihrer Klassifikation als spezifische Lerntypen assoziiert sind.

Empirisch können Lerntypen mittels personen-zentrierter Verfahren, wie der Latenten Profilanalyse (LPA, Vermunt & Magidson, 2002), identifiziert werden. Die LPA kann ermitteln, inwieweit sich verschiedene Gruppen von Schülerinnen und Schülern finden lassen, die ein ähnliches Profil an spezifischen Ausprägungen von HA-Zeit und HA-Compliance zeigen, das sich von den Profilen der anderen Gruppen unterscheidet. Dabei können auch Assoziationen zu den genannten Schülermerkmalen untersucht werden.

Mithilfe der LPA wurden fünf unterschiedliche Lerntypen gefunden. Anhand ihrer Mittelwertsprofile wurden sie bezeichnet als „schnelle Lerner“, „hoch engagierte Lerner“, „Durchschnittsschüler“, „sich abmühende Lerner“ und „Minimalisten“. Das Resultat der Untersuchung der Leistungsunterschiede zwischen Lerntypen mittels multipler Regressionsanalyse war, dass unter Kontrolle der Effekte der Ausgangsleistung zum ersten Messzeitpunkt „schnelle Lerner“ und „hoch-engagierte Lerner“ eine bessere schulische Leistungsentwicklung aufwiesen als die anderen Lerntypen. Die schlechtesten Lernergebnisse wiesen „sich abmühende Lerner“ auf. Bezüglich der dritten Fragestellung zeigte sich, dass Lerntypen, die durch ein hohes Maß an HA-Compliance charakterisiert sind, höhere Gewissenhaftigkeit und Wertüberzeugungen aufweisen als die anderen Lerntypen.

Insgesamt erwies sich HA-Zeitaufwand als bedeutsam für die adäquate Beschreibung von HA-Verhalten, wenn simultan die HA-Compliance berücksichtigt wurde. Es konnte gezeigt werden, dass durch Einsatz des personen-zentrierten Ansatzes neue Erkenntnisse erzielt werden können, die das HA-Verhalten von Schülerinnen und Schülern differenzierter beschreiben als bisherige Ansätze.

Freitag
13. März

9:00 Uhr

GC 03/149

Chair(s)
Lösch, Thomas

Lintorf, Katrin; Buch, Susanne R.; Sparfeldt, Jörn R.; Rost, Detlef H.

Anstrengungsvermeider – Lustlos oder hilflos?

Theoretischer Hintergrund: Der Anstrengungsvermeidungstest (AVT) von Rollett und Bartram (1977, 1998) soll die „Neigung von Schülern, schulischen Anstrengungen aus dem Weg zu gehen“ erfassen (1977, S. 6). Zur Genese dieser bereichsspezifischen Neigung zur Anstrengungsvermeidung sollen wiederholte negativ erlebte (schulische) Anstrengungen beitragen (Rollett & Rollett, 2010). Dafür sprechen Befunde z.B. zu Überforderungserfahrungen (Rollett, 2004; Rollett & Hanfstingl, 2005). Unabhängig von der Diskussion um die Konzeption eines Anstrengungsvermeidungsmotivs (Jopt, 1982, 1984; Rollett, 1983) hat sich der AVT in der Praxis (z.B. schulpyschologische Beratung) bewährt und ist immer noch in Verwendung.

Dabei stellt sich die Frage, ob sich unterschiedliche Subgruppen anstrengungsvermeidender Schüler/innen differenzieren lassen. Helmke und Rheinberg (1996) fanden bei Grundschulkindern vier Cluster, wobei zwei eine hohe Anstrengungsvermeidungstendenz aufwiesen: Ähnlich wie der desorganisierte vs. apathische Arbeitsstil nach Rollett und Bartram (1998) zeigten diese Cluster Unterschiede im Profil kognitiver und motivational-affektiver Merkmale. Helmke und Rheinberg (1996) sahen darin differenzielle Bedingungsfaktoren der Anstrengungsvermeidung, etwa aktuelle Überforderung vs. motivationale Defizite, und wiesen auf die Notwendigkeit einer differenziellen Diagnostik hin. Unklar ist, ob und inwieweit sich die Befunde auf ältere Schüler/innen generalisieren lassen.

Die insbesondere für Diagnostik und Intervention relevante Frage der Identifikation verschiedener Anstrengungsvermeidungstypen ist jenseits der Befunde von Helmke und Rheinberg (1996) bislang kaum empirisch untersucht. Für den desorganisierten vs. apathischen Arbeitsstil existieren nur Einzelfallberichte aus der lerntherapeutischen Praxis (Rollett & Bartram, 1998). Eine Clusteranalyse von Wagner, Spiel und Tranker (2003), in die neben der Anstrengungsvermeidungstendenz allerdings ausschließlich kognitive Merkmale eingingen, ergab bei einer Stichprobe von Sechst- und Siebtklässlern keine differenziellen Profile, sondern nur ein einzelnes Cluster mit überdurchschnittlicher Anstrengungsvermeidung.

Fragestellung: Lassen sich unter Berücksichtigung kognitiver, motivationaler und affektiver Merkmale Gruppen von Anstrengungsvermeidern („Typen“) mit unterschiedlichen Profilausprägungen differenzieren und validieren?

Methode: An der Studie nahmen 1079 Acht- bzw. Neuntklässler aus 44 Gymnasialklassen teil. Das Geschlechterverhältnis war ausgeglichen (53% Mädchen, 47% Jungen).

Zur Typisierung wurden in Mplus latente Profilanalysen berechnet. In Anlehnung an das Vorgehen von Helmke und Rheinberg (1996) gingen neben der Anstrengungsvermeidungsskala (AVT, Rollett & Bartram, 1998) folgende kognitive und motivational-affektive Variablen in die Analysen ein: die Mathematik- und Deutschnote, die Subtests Analogien und Zahlenreihen des I-S-T 70 (Amthauer, 1970), eine Skala zur Erhebung des allgemeinen schulischen Interesses (angelehnt an Schiefele, Krapp & Wild, 1993 sowie Sparfeldt, Rost & Schilling, 2004), die Skala zur Erfassung des Selbstkonzepts schulischer Leistungen und Fähigkeiten (SKSLF, Rost & Lamsfuß, 1992; Rost & Sparfeldt, 2002) und drei Skalen zur Erfassung der Prüfungsängstlichkeit (Aufgeregtheit, Besorgnis, TAI-G, Hodapp, 1991, Kognitive Angstmanifestation, DAI, Rost & Schermer, 1997).

Zur Validierung dienten folgende Variablen: das Geschlecht, eine Skala zur Erfassung des schulischen Hilflosigkeitserlebens (angelehnt an Trautwein & Köller, 2003), die Skalen zur Erfassung der Lern- und Leistungsmotivation (SELLMO, Spinath, Stiensmeier-Pelster, Schöne & Dickhäuser, 2002) und die Self-Esteem-Skala von Rosenberg (1965).

Ergebnisse: Eine Auswertung verschiedener Kriterien zur Bewertung der Modellgeltung bzw. Festlegung der Klassenzahl (vgl. z.B. Geiser, 2011; Nylund, Asparouhov & Muthén, 2007; Rost, 2006) spricht für eine Lösung mit drei bis vier Klassen. Die Drei-Klassen-Lösung zeigt im Wesentlichen niveaueverschiedene Profile. In der Vier-Klassen-Lösung werden dagegen zwei Gruppen von Anstrengungsvermeidern mit jeweils ungünstigen Profilen weiter ausdifferenziert. Diese unterscheiden sich weniger in der Ausprägung kognitiver, sondern eher in der Ausprägung motivational-affektiver Merkmale. Erwartungsgemäß befinden sich in der Gruppe mit insgesamt unterdurchschnittlicher Prüfungsangst und mit, im Vergleich zur anderen Gruppe, geringerem allgemeinen schulischen Interesse mehr Jungen. In der Gruppe mit hoch überdurchschnittlicher Angstaussprägung und unterdurchschnittlichem Selbstkonzept finden sich dagegen mehr Mädchen. Darüber hinaus ergeben sich differenzielle Bezüge zu den weiteren Validierungsvariablen: So zeigt die letztgenannte Gruppe einen deutlich geringeren Selbstwert ($d = 0.53$) und ein höheres Hilflosigkeitserleben ($d = 1.09$) als die erstgenannte Gruppe. Die Befunde werden vor dem Hintergrund differenzieller Interventionsmöglichkeiten diskutiert.

Freitag
13. März

9:00 Uhr

GC 03/149

Chair(s)

Lösch, Thomas

Gaspard, Hanna; Dicke, Anna-Lena; Flunger, Barbara; Schreier, Brigitte; Häfner, Isabelle; Trautwein, Ulrich; Nagengast, Benjamin
Nebenwirkungen von fachspezifischen Motivationsinterventionen? Effekte einer Intervention im Fach Mathematik auf die Motivation im Fach Deutsch

Mit Verweis auf den Fachkräftemangel wird häufig festgestellt, dass sich nicht genügend Jugendliche für MINT-Fächer (Mathematik, Informatik, Naturwissenschaften, Technik) interessieren (z.B. Akademie der Wissenschaften Schweiz, 2014). In den letzten Jahren wurde eine Reihe von Interventionen zur Förderung der Motivation im MINT-Bereich entwickelt (z.B. Blackwell, Trzesniewski & Dweck, 2007; Harackiewicz, Rozek, Hulleman & Hyde, 2012; Hong & Lin-Siegler, 2012; Hulleman & Harackiewicz, 2009). Einige dieser Interventionen setzten bei der Bedeutung von mathematisch-naturwissenschaftlichen Kompetenzen für das Leben der Jugendlichen an, um so deren Wertüberzeugungen zu fördern (Harackiewicz et al., 2012; Hulleman & Harackiewicz, 2009). In bisherigen Studien konnte gezeigt werden, dass die Wertüberzeugungen und das Interesse sowie die Leistung in Mathematik und Naturwissenschaften und die Wahl entsprechender Kurse mit Hilfe solcher Interventionen tatsächlich gefördert werden können.

Offen bleibt die Frage, ob Interventionen in MINT-Fächern auch einen Einfluss auf die Motivation in anderen Fächern haben. Sowohl die Erwartungs-Wert-Theorie von Eccles und Kollegen (Eccles et al., 1983) als auch das Internal/External-Frame-of-Reference-Modell von Marsh (1986) gehen davon aus, dass die Motivation von Schülerinnen und Schülern auch von Vergleichen über Schulfächer hinweg beeinflusst wird. Typischerweise lassen sich Kontrasteffekte zwischen sprachlichen und nicht-sprachlichen Schulfächern finden (Möller & Marsh, 2013). Theoretisch wird in Bezug auf dimensionale Vergleichsprozesse implizit die Annahme getroffen, dass die Motivation in einem Schulfach die Motivation in einem anderen Schulfach negativ beeinflussen kann. Interventionsstudien, die auf Schülermotivation im MINT-Bereich abzielen, bieten die Möglichkeit, diese Annahme experimentell zu untersuchen. Bisherige Studien dazu fehlen bislang allerdings.

Die vorliegende Studie widmet sich daher der Frage, ob eine Intervention zur Förderung der Motivation im Fach Mathematik negative Effekte auf die Motivation im Fach Deutsch hat. Hierzu werden Daten einer Interventionsstudie genutzt, in der die Wertüberzeugungen von Schülerinnen und Schülern für Mathematik in der neunten Klassenstufe gefördert wurden. Es werden Effekte dieser Intervention auf die fachspezifischen Wertüberzeugungen sowie das fachspezifische Selbstkonzept als zentrale motivationale Faktoren im Sinne der Erwartungs-Wert-Theorie (Eccles et al., 1983) betrachtet.

An der Studie nahmen insgesamt 82 neunte Klassen aus 25 Gymnasien in Baden-Württemberg teil. Die teilnehmenden Klassen wurden randomisiert einer von zwei Interventionsbedingungen oder einer Wartekontrollgruppe zugewiesen. Die Intervention bestand aus einer doppelstündigen Unterrichtseinheit zur Nützlichkeit der Mathematik. In der ersten Interventionsbedingung verfassten die Schülerinnen und Schüler einen Text zur persönlichen Relevanz der Mathematik. In der zweiten Interventionsbedingung wurden Interviewzitate zur Nützlichkeit von Mathematik bewertet. Vor der Intervention (Prätest) sowie sechs Wochen (Posttest) und fünf Monate nach der Intervention (Follow-Up) füllten insgesamt 1916 Schülerinnen und Schüler Fragebögen aus. Darin wurden sowohl Wertüberzeugungen als auch das Selbstkonzept für Mathematik und Deutsch mit parallelen Skalen erfasst. Die Interventionseffekte wurden in Mehrebenen-Regressionsanalysen unter Kontrolle der Ausgangswerte geprüft.

Für das Fach Mathematik fanden sich zum Posttest sowie zum Follow-Up positive Effekte der Zitate-Bedingung im Vergleich zur Wartekontrollgruppe auf die Wertüberzeugungen der Schülerinnen und Schüler ($\beta = .21, p < .001$). Außerdem fand sich zum Posttest ein positiver Effekt der Zitate-Bedingung auf das Selbstkonzept der Schülerinnen und Schüler ($\beta = .10, p < .05$). Für die Text-Bedingung wurden keine signifikanten Effekte auf Wertüberzeugungen und das Selbstkonzept für Mathematik gefunden.

Für das Fach Deutsch fanden sich keine Effekte der Interventionsbedingungen zum Posttest. Allerdings zeigte sich zum Follow-Up ein negativer Effekt der Zitate-Bedingung auf die Wertüberzeugungen ($\beta = -.18, p < .01$) sowie ein tendenziell negativer Effekt ($\beta = -.10, p < .10$) auf das Selbstkonzept der Schülerinnen und Schüler.

Die Ergebnisse deuten darauf hin, dass erfolgreiche Interventionen im MINT-Bereich tatsächlich negative Effekte auf die Motivation in sprachlichen Fächern haben können. Diese Effekte können im Sinne ipsativer Vergleiche als eine Verschiebung des motivationalen Profils von sprachlichen Fächern hin zu nicht-sprachlichen Fächern verstanden werden. Solche Effekte sollten daher in Interventionsstudien regelmäßig überprüft werden.

Freitag
13. März

9:00 Uhr

GC 03/149

Chair(s)
Lösch, Thomas

Lazarides, Rebecca; Rohowski, Susanne; Ittel, Angela

Unterrichtsgestaltung, Motivation und berufliche Selbstexploration

Theoretischer Hintergrund: Berufs- und Studienwahlprozesse werden als zentrale Entwicklungsaufgabe im Jugendalter beschrieben (Havighurst, 1948). Die in dieser Studie untersuchte berufliche Selbstexploration, also die Reflektion eigener Interessen und Fähigkeiten, ist eine wichtige Voraussetzung für den erfolgreichen Einstieg ins Berufsleben (Stumpf, Austin, & Hartman, 1984). Inwieweit der Einstieg ins Berufsleben gelingt, ist abhängig von Heterogenitätsdimensionen wie Migrationshintergrund und Geschlecht. Jugendliche mit Migrationshintergrund gelten auch bei gleichen Leistungen beim Einstieg ins Berufsleben als benachteiligt gegenüber Jugendlichen ohne Migrationshintergrund. Die starke Geschlechtersegregation auf dem Arbeitsmarkt zeigt, dass auch das Geschlecht ein Faktor ist, der die Berufswahl beeinflusst. Es ist daher von Interesse, zu untersuchen, wie die berufliche Selbstexploration bei Jugendlichen verschiedener Gruppen ausgeprägt ist und inwiefern die von Lehrkräften gestaltete Unterrichtsgestaltung die berufliche Selbstexploration Jugendlicher begünstigt. Die Selbstbestimmungstheorie der Motivation (SDT; Deci & Ryan, 1985) beschreibt, dass Lehrkräfte, die im Unterricht Autonomie, Kompetenz und Verbundenheit unterstützen, selbstinitiiertes Explorationsverhalten fördern. Eccles (2005) verdeutlicht in ihrem Erwartungs-Wert-Modell, dass subjektive fachbezogene Werte (Interesse, persönliche Wichtigkeit, Nützlichkeit) von Schülerinnen und Schülern jenen Formen motivationaler Verhaltensregulation ähneln, die hohes Explorationsverhalten bewirken. Die vorgestellte Studie untersucht an den Schnittstellen von Psychologie, Erziehungswissenschaft und Lehrkräftebildung die Effekte wahrgenommener motivationsrelevanter Unterrichtsmerkmale (Unterstützung von Autonomie, Kompetenz, sozialer Eingebundenheit), des Geschlechts, des Migrationshintergrundes und subjektiver fachbezogener Werte auf die berufliche Selbstexploration.

Methode: In der Fragebogenstudie wurden Lernende der Sekundarstufe I zum Unterricht im Schulfach „Wirtschaft, Arbeit, Technik“ (WAT) befragt. Für die Analysen wurden Daten von 1780 Lernenden (45.8% weiblich) der 7. bis 10. Jahrgangsstufen aus 14 Berliner Integrierten Sekundarschulen ausgewertet. Das Durchschnittsalter lag bei $M = 14.47$ Jahren ($SD = 1.38$). Die Daten wurden von wissenschaftlichem Personal in dafür zur Verfügung gestellten Unterrichtsstunden (45 Minuten) erhoben. Die Testung der angenommenen Zusammenhänge erfolgte mittels linearer Strukturgleichungsmodellierung in Mplus unter Berücksichtigung der geclusterten Datenstruktur. Durch multipler Gruppenanalyse wurde untersucht, ob die Zusammenhänge zwischen den Geschlechtergruppen variierten.

Ergebnisse: Mädchen ($\beta = .089$, $SE = .024$, $p < .001$) berichteten signifikant höhere berufliche Selbstexploration als Jungen. Die im Unterricht wahrgenommene Unterstützung von Autonomie ($\beta = .109$, $SE = .053$, $p < .05$), Kompetenz ($\beta = .114$, $SE = .053$, $p < .05$), sozialer Eingebundenheit ($\beta = .106$, $SE = .040$, $p < .01$) und die von den Lernenden eingeschätzte Nützlichkeit des Fachunterrichts ($\beta = .162$, $SE = .034$, $p < .001$) begünstigten die berufliche Selbstexploration. Wahrgenommene Autonomieunterstützung im Unterricht wirkte positiv auf die berufliche Selbstexploration über die von den Lernenden wahrgenommene Nützlichkeit des Fachunterrichts ($\beta = .079$, $SE = .017$, $p < .001$). Jugendliche, die Deutsch nicht als Muttersprache angaben, empfanden das Fach WAT als weniger persönlich wichtig ($\beta = -.077$, $SE = .020$, $p < .001$) als Jugendliche deutscher Muttersprache. Es zeigte sich eine gute Passung zwischen den Daten und dem Modell, $\chi^2(433, N = 1701) = 1149.536$, $p < .001$, CFI = .97, TLI = 0.96; RSMA = .031 (90% CI = .029 - .033). Das Modell erklärte 16,1% der Varianz der beruflichen Selbstexploration. Das Modell war gleichermaßen für Mädchen und Jungen gültig.

Interpretation und Relevanz: Die Resultate sind relevant für Bildungspraxis und Lehrkräftebildung. Sie zeigen, dass Unterricht, der das Erleben von Autonomie, Kompetenz und Eingebundenheit begünstigt, nicht nur motivationsfördernden wirkt, sondern auch einen Beitrag im Berufswahlprozess leistet. WAT Unterricht, der von Jugendlichen als nützlich für ihre Zukunft betrachtet wird, fördert die Auseinandersetzung mit beruflichen Perspektiven. Interesse am WAT Unterricht oder seine persönliche Bedeutung sind dabei von geringer Wichtigkeit. Für die Unterrichtsgestaltung zeigt sich hier die Notwendigkeit, den Nutzen der vermittelten Fachinhalte zu verdeutlichen. Dass Jugendliche mit Migrationshintergrund den persönlichen Wert des Faches geringer einschätzen als Jugendliche ohne Migrationshintergrund wirft weitere Fragen zur gruppenspezifisch orientierten Unterrichtsgestaltung auf.

Lehmann, Malte; Rösken-Winter, Bettina

Problemlösekompetenzen von Ingenieursstudierenden in mathematisch-physikalischen Kontexten

Das BMBF-geförderte Projekt KoM@ING hat sich die Kompetenzmodellierung und -messung in Ingenieursstudiengängen, insbesondere für die Studieninhalte Höhere Mathematik (HM) und Technische Mechanik (TM), zum Ziel gesetzt. Im Rahmen eines Teil-Projektes wurden IRT-skalierte Tests zur Technischen Mechanik entwickelt (Nickolaus, Behrendt, Dammann, Ștefănică, & Heinze, 2013). Solche IRT-basierten Untersuchungen können den Lösungsprozess jedoch nicht in seiner Gesamtheit erfassen. D. h. auf welchen Wegen und mit welchen Mitteln Studierende Probleme lösen, kann nicht ausreichend erklärt werden, sodass insbesondere nicht modellgültige Aufgaben in der hier präsentierten Studie weiteren, qualitativen Analysen unterzogen wurden. Als theoretischer Rahmen dienten dabei Arbeiten zum Problemlösen. Zum einen können so innere Strukturen wie verwendete Heuristiken (Bruder, & Collet, 2011) untersucht werden. Hier stehen die Hilfsmittel und Strategien (zum Beispiel: Informative Figuren oder Rückwärtsarbeiten) im Fokus. Zum anderen können äußere Strukturen, die Phasen des Lösungsprozesses (Polya, 1957; Schoenfeld, 1985), Gegenstand der Untersuchung sein. Während Polya und Schoenfeld eher lineare Abläufe vorschlagen, weist Rott (2013) darauf hin, dass solche linearen Modelle nicht immer ausreichend sind, um Problemlöseprozesse adäquat abzubilden. Die theoretischen Sichtweisen erlauben nicht nur personenbezogene Aussagen, sondern können auch Charakteristiken von Aufgaben und ihre Wirkung auf den Problemlöseprozess beschreiben. Die Forschungsfrage für diesen Beitrag lautet daher:

Können sowohl über die Klassifizierung von Heuristiken und Problemlösephasen als auch über ihre Interaktionen schwierigkeitsgenerierende Merkmale von Aufgaben erfasst werden?

In der vorliegenden Studie wurden Folgen von zwei bis vier Aufgaben aus der o.g. Pilotierung untersucht, bei der mindestens eine Aufgabe nicht modellkonform war. Die Aufgaben einer Folge behandelten denselben inhaltlichen Kontext (Auflagerkräfte, Freikörperbild, Schwerpunkt und Allgemeine Kraftsysteme) und waren in aufsteigender Schwierigkeit angeordnet. Da nicht zu jeder Aufgabe Schätzer vorlagen, wurden die Aufgaben nach der Lösungsquote in der Pilotierung geordnet. Dabei waren die ersten Aufgabefolgen eher einfach (Lösungsquote Auflagerkräfte: 0,82/0,42/0,1 und Schwerpunkt: 0,6/0,47) und die zweiten eher schwierig (Kraftsysteme: 0,35/0,14/0 und Freikörperbild: 0,28/0,09/0,04/0). Insgesamt wurden vier Aufgabenfolgen mit elf Aufgaben eingesetzt, sodass n=16 Aufgabenbearbeitungen vorliegen. Jede Aufgabenfolge wurde von vier Teilnehmenden gelöst, dabei trat jede Kombination von zwei Aufgabenfolgen zweimal auf. Die Studierenden lösten die Aufgaben und setzten die Methode Lautes Denken (Weidle, & Wagner, 1982) ein, um die kognitiven Prozesse während des Lösungsprozesses beobachtbar zu machen. Anschließend wurden die Studierenden zu ihren Bearbeitungsprozessen auf der Basis eines Leitfadens interviewt. Die Daten wurden über Videografie erhoben und anschließend transkribiert. Die Datenanalyse wurde anhand eines Kategoriensystems zum Problemlösen auf Basis der Heuristiken nach Bruder und Collet (2011) und der Phasen nach Polya (1957) vorgenommen.

Alle Aufgaben forderten von den Studierenden die Arbeit an und mit Abbildungen. Die beiden einfacheren Aufgabenfolgen konnten jeweils mit einem bestimmten Schema, wie dem Aufstellen eines Gleichungssystems oder dem Anfertigen einer Tabelle, gelöst werden. Dieses Vorgehen wurde auch von nahezu allen Studierenden angewendet. Bei den beiden schwierigeren Aufgabenfolgen standen keine Standardverfahren zur Verfügung. Während es bei der Gesamtzahl der eingesetzten Heuristiken und der Anzahl der unterschiedlichen eingesetzten Heuristiken keine signifikanten Unterschiede zwischen den Teilnehmenden gab (Beispiel für Aufgabenfolge „Auflagerkräfte“: $Ft(3, 4,39)=1,22$, $p=0,40$), zeigte sich, dass erfolgreiche Studierende ihre Heuristiken besonders im Zusammenspiel mit den gegebenen Abbildungen miteinander kombinierten und diese nicht isoliert verwendeten. Dies war besonders bei den nicht modellkonformen Aufgaben zu beobachten. Außerdem beschränkten die Studierenden den Einsatz des Evaluierungsschrittes auf die schwierigen Aufgaben.

Die Ergebnisse zeigen deutlich, dass Aufgaben mit einer höheren Schwierigkeit ein vielseitiges Kombinieren von Heuristiken erfordern. Dies trifft vor allem auf die nicht modellkonformen Aufgaben zu. Dieses ist gemein, dass sie durch eine komplexere Aufgabenstellung und –skizze einen häufigen Wechsel zwischen verschiedenen Heuristiken erfordern, was die Ergebnisse von Rott (2013) bestätigt. Währenddessen sind einfachere Aufgaben durch ein lineares Vorgehen (theoriekonform zu Polya, 1957, und Schoenfeld, 1985) lösbar.

Altieri, Mike

Einflussfaktoren auf den Prüfungserfolg in der Mathematik-Vorlesung für Ingenieure

Theoretischer Hintergrund: Nach Untersuchungen zum Zusammenhang zwischen Schul- und Studiennoten gelangen Steyer et al. (2005) zu dem Schluss, dass es insgesamt sehr schwierig erscheint, Studienerfolg vorherzusagen. Eine Begründung dafür sehen sie in dem fehlenden Wissen über fachspezifische Anforderungen und empfehlen, fachspezifische Anforderungsanalysen durchzuführen. Diese werden in der vorliegenden Studie zum Prüfungserfolg in Lehrveranstaltungen zur Mathematik für Informatik, Naturwissenschaften und Technik (INT) einigen häufig genannten psychologischen Faktoren (nämlich Prüfungsangst, Aufschiebeverhalten und Selbstwirksamkeitserwartung) gegenübergestellt, um zu bestimmen, welche in ein Anforderungsprofil aufgenommen und ggf. in oder vor den Lehrveranstaltungen gefördert werden sollten. Die fachspezifischen Anforderungen werden dazu unterschieden nach den zentralen kognitiven Aktivitäten im mathematischen Modellierungskreislauf (z.B. Blum und Leiß (2005)), dem "Mathematisieren" (vom Aufgabentext zum mathematischen Modell) und/oder "Verarbeiten" (vom mathematischen Modell per Algorithmus zur Lösung). Dabei wird das Verarbeiten operationalisiert über Kalkülkompetenzen auf Hochschulniveau (Kalküle der Hochschulmathematik in angemessener Zeit fehlerfrei anwenden) und Schulniveau (Bruchrechentest), das Mathematisieren auf Schulniveau durch einen Leistungstest zu Grundvorstellungen und konzeptuellem Denken.

Zentrale Fragestellungen: (1) Welche fachspezifischen und psychologischen Faktoren haben auf das Ergebnis einer Mathematik-Klausur in den INT-Fächern den grössten Einfluss? (2) In welchem Verhältnis steht die Relevanz für den Prüfungserfolg von Faktoren, die eher dem "Mathematisieren" zuzuordnen sind gegenüber Faktoren, die eher dem "Verarbeiten" bzw. keiner dieser beiden Stationen des Modellierungskreislaufes zuzurechnen sind?

Methoden der Datenerhebung: Als Instrumente eingesetzt wurde ein aus acht Subtests bestehendes Testrepertoire: Die fachspezifischen Fähigkeiten im Bereich Verarbeiten wurden erfasst durch einen raschskalierten Speed- und Leistungstest zur Kalkülkompetenz sowie einen Test zur Bruchrechnung (Eigenentwicklung bzw. in Anlehnung an Padberg (1978)) sowie einem Konzentrationstest nach Satow (2011). Die Fähigkeiten im Bereich Mathematisieren wurden in einem selbstentwickelten Leistungstest zu Grundvorstellungen und konzeptuellem Denken erfasst. Als oft diskutierte mögliche psychologische Einflussfaktoren wurden Prüfungsangst mit dem TAI-G, vgl. Hodapp (1991), Aufschiebeverhalten (Schwarzer (1999)) sowie die Selbstwirksamkeitserwartung bei der Kalkülanwendung (Eigenentwicklung) einbezogen. Als Kontrollvariablen wurden Schlussfolgerndes Denken, Schulform, Grund-/Leistungskurs, G8/G9 erfasst.

Die Stichprobe bestand aus 1552 Teilnehmern der Erstsemester-Klausur "Höhere Mathematik I" für Ingenieur-Studiengänge aus dem Wintersemester 2013/14, von denen über 1000 Studierende an allen während der Vorlesungszeit durchgeführten Tests sowie an der genannten Klausur nach Ende der Vorlesungszeit teilgenommen haben.

Methoden der Datenauswertung: In Anlehnung an Ree und Earles (1991) wurden hierarchische lineare und logistische Regression zur Prognose der Klausurleistung bzw. des -ergebnisses "bestanden" oder "nicht bestanden" eingesetzt. Im letzteren Fall steht insbesondere das Prognosepotenzial (die korrekte Vorhersage des Klausurergebnisses) der einzelnen Subtests im Fokus. Durch die hohe Teilnehmerzahl konnte die Stichprobe in eine Lernstichprobe zur Modellentwicklung und eine Teststichprobe zur Modellprüfung geteilt werden. Bootstrapping ergab eine Normalverteilung des Prognosepotenzials, so dass dieses inklusive Konfidenzintervall für jeden Subtest sowie für jede beliebige Kombination bestimmt werden kann. Die Ergebnisse der Regressionsanalysen hinsichtlich erklärter Varianz und Relevanz der Subtests wurden anschließend durch Analyse der - insbesondere partiellen - Interkorrelationen sowie Hauptkomponenten betrachtet.

Ausgewählte Ergebnisse: Das Prognosepotenzial des saturierten Modells bezüglich "bestanden" oder "nicht bestanden" liegt bei knapp 80% (Erwartungswert des Prognosepotenzials bei zufälliger Prognose der Ergebnisse: 52%). Es zeigt sich, dass bei der gestellten "Standardklausur" (annähernd normalverteilte Ergebnisse mit $M = 34.4$ von 80 erreichbaren Punkten und $SD = 13.4$), die einbezogenen psychologischen Faktoren keinen signifikanten Einfluss haben, während die dem Bereich "Verarbeiten" zuzuordnenden Faktoren einen deutlich höheren Einfluss ausüben als die dem "Mathematisieren" zuzurechnenden Faktoren. Als Einzeltest hat der Bruchrechentest, der als Stellvertreter für die Überprüfung basaler Kompetenzen steht, bereits ein Prognosepotenzial von 70% gegenüber 76% des auf Hochschulmathematik bezogenen Kalkülkompetenztests. Bei der linearen Regression zeigt sich, dass der veranstaltungsnaher Kalkülkompetenztest mit Abstand den größten Anteil an der insgesamt erklärten Varianz von 45% hat. Die Ergebnisse werden durch die Analyse der Korrelationen bestätigt.

Nakoinz, Simone; Sumfelth, Elke

Analyse der Fähigkeit Studierender, Phänomene und submikroskopische Konzepte zu verknüpfen

Theoretischer Hintergrund: Lernende verschiedenen Alters zeigen Schwierigkeiten bei der Verknüpfung der Repräsentationsebenen von chemischen Inhalten (submikroskopisch, makroskopisch, symbolisch; Johnstone, 1991) (z.B. Dori & Kaberman, 2012). Sie sind beispielsweise nicht in der Lage, Phänomene auf Grundlage der Teilchenebene zu interpretieren (z.B. Meijer, Bulte & Pilot, 2009). Häufig kommt es deswegen zu Fehlvorstellungen, die auf dem Transfer makroskopischer Eigenschaften auf submikroskopische Konzepte basieren (z.B. Barke, 2006; Busker, 2010).

Die Verknüpfung der verschiedenen Repräsentationsebenen führt jedoch zu einem vertieften Verständnis chemischer Konzepte (z.B. Jaber & BouJaoude, 2012). Die Verknüpfung des Wissens der verschiedenen Ebenen setzt eine Reproduktion dieses Wissens voraus. Die submikroskopische und die symbolische Ebene stehen dabei eng miteinander in Zusammenhang, da die nicht wahrnehmbaren submikroskopischen Konzepte oft durch Symbole und Modelle visualisiert und auf diese Weise reproduziert werden (z.B. Treagust, Chittleborough & Mamiala, 2003). Das Wissen auf allen drei Repräsentationsebenen muss schließlich organisiert, integriert und zu einem mentalen Modell verknüpft werden.

Ziel dieses Projekts ist es, ein Testinstrument zu entwickeln, das Wissen auf submikroskopischer Ebene, Wissen auf makroskopischer Ebene und die Fähigkeit, die beiden Ebenen zu verknüpfen getrennt voneinander erfasst. Dieses Testinstrument soll dann für weitere Analysen bezüglich der VPs-Fähigkeit der Studierenden genutzt werden.

Fragestellungen: Ist es möglich, ein Testinstrument zu entwickeln, das Wissen auf submikroskopischer Ebene, Wissen auf makroskopischer Ebene und die Fähigkeit beide Ebenen miteinander zu verknüpfen getrennt voneinander misst?

In wie weit sind Erstsemester-Studierende im Fach Chemie im Modul „Allgemeine Chemie“ in der Lage, ihr Wissen über Phänomene und submikroskopische Zusammenhänge zu verknüpfen?

Methode: Im Rahmen dieses Projekts wurden Items für ein Testinstrument entwickelt, das die Fähigkeit zur Verknüpfung von Phänomenen und submikroskopischen Zusammenhängen messen soll (VPs). Der Test besteht dazu aus drei Itemtypen: Einerseits Items, welche die Reproduktion von Wissen auf makroskopischer (erster Itemtyp) oder auf submikroskopischer (zweiter Itemtyp) erfordern, und Items, bei denen eine Verknüpfung des Wissens der beiden Ebenen erforderlich ist (dritter Itemtyp). Zu einem chemischen Konzept wurde jeweils ein Item der drei Itemtypen entwickelt, sodass der Test insgesamt aus 30 Items zu zehn chemischen Konzepten besteht.

Bei der Hauptstudie im Wintersemester 2013/14 wurde der neu entwickelte und zuvor pilotierte Test jeweils zu Beginn und zum Ende des Semesters eingesetzt. Desweiteren wurden verschiedene Kontrollvariablen eingesetzt, wie zum Beispiel ein weiterer Fachwissenstest, ein Test zum schlussfolgernden Denken, Fragebögen zur Motivation, zum Interesse am Fach Chemie und zur chemiebezogenen Selbsteinschätzung. Es konnten insgesamt 101 Erstsemesterstudierende der Fächer Chemie und Water Science, des Lehramts Chemie und des Lehramts Biotechnik an der Universität Duisburg-Essen getestet werden.

Ergebnisse: Die Kennwerte für den neu entwickelten VPs-Test lagen bei der Hauptstudie im Wintersemester 2013/14 im zufriedenstellenden Bereich (EAP/PV-Reliabilität Pre: .86; Post: .80). Durch Dimensionsanalysen konnte gezeigt werden, dass es nicht möglich war ein Testinstrument zu entwickeln, bei welchem die drei Itemtypen des Tests empirisch voneinander getrennt werden können. Es lassen sich jedoch deutliche Tendenzen bezüglich der Unterschiede in der Schwierigkeit der verschiedenen Itemtypen erkennen. Die makroskopischen Items eines Itemblocks (submikroskopisch, makroskopisch, VPs) sind in über 80% der Fälle leichter als Items der beiden anderen Itemtypen.

Weitere Dimensionsanalysen ergaben, dass es möglich ist, das neu entwickelte Testinstrument von einem bereits bestehenden Fachwissenstest (Freyer, 2013) empirisch abzugrenzen. Weitere Analysen bezüglich des neu entwickelten Testinstruments und der erhobenen Kontrollvariablen folgen und werden auf der Tagung vorgestellt.

Freitag
13. März

11:15 Uhr

GA 03/49

Gigl, Florian; Spoden, Christian; Zander, Simon; Borowski, Andreas; Fischer, Hans E.

Entwicklung von physikalischem Fachwissen im ersten universitären Semester

Hintergrund: Professionswissen wird meist in die drei Wissensbereiche Fachwissen, Fachdidaktisches Wissen und Pädagogisches Wissen untergliedert (Shulman, 1987, Abell, 2007, Fischer, Borowski & Tepner, 2012). Um den besonderen Anforderungen der Lehramtsausbildung in Bezug auf das Fachwissen nachzukommen ist die Einführung von Subdimensionen angebracht. In der Mathematikdidaktik konnte im Rahmen der COACTIV-Studie bereits gezeigt werden, dass sich mathematisches Fachwissen von Lehrkräften inhaltlich und psychometrisch in die Dimensionen Schulwissen, vertieftes Schulwissen und universitäres Wissen trennen lässt (z.B. Baumert, et al., 2010). Die Dimension vertieftes Schulwissen stellt unter diesen als „Schulwissen vom höheren Standpunkt“ (Baumert et. al., 2010) eine Schlüsselkompetenz für Lehrkräfte dar. Eine empirisch abgesicherte Übertragung in die Physik steht aus.

Hierfür wurde ein dreidimensionales Kompetenzmodell von Fachwissen entwickelt. Die Dimension Schulwissen (SW) stellt Wissen auf Niveau des Grundkursabiturs dar, vertieftes Schulwissen (VSW) notwendiges Wissen zur Transformation universitärer Inhalte auf Fragestellungen der Schule. Universitäres Wissen (UW) wird im Rahmen der Fachausbildung an der Universität vermittelt, ist jedoch nicht Bestandteil des vertieften Schulwissens.

Im Rahmen eines Längsschnittdesigns wurden angehende Physiklehrkräfte während ihrer ersten zwei Studienjahre wissenschaftlich begleitet und das vorliegende dreidimensionale Modell zu zwei Messzeitpunkten validiert. Insbesondere während der kritischen Studieneingangsphase ergibt sich die Frage, ob und gegebenenfalls welche Indikatoren von Schulleistungen eine erfolgreiche Aneignung von Fachwissen im Studium vorhersagen können (vgl. z.B. Schuler, 2010).

Fragestellungen: Hieraus ergeben sich folgende Forschungsfragen: (1) Ist die Differenzierung von Fachwissen in drei Dimensionen (Schulwissen, vertieftes Schulwissen, universitäres Wissen) empirisch tragfähig?(2) Kann man mit Hilfe von Schulleistungsindikatoren Leistungszuwächse im Laufe des relevanten ersten Studiensemesters statistisch erklären?

Methode: Zu Beginn und Ende des Wintersemesters 2013/ 14 wurden N(1)=273 bzw. N(2)=216 Physikstudierende bestehend aus einer Zielgruppe (Lehramt BA Physik) und einer Vergleichsgruppe (Fach BA Physik) getestet. Dafür kam ein modellkonform konstruiertes Testheft mit 40 pilotierten Aufgaben im Themengebiet Mechanik zum Einsatz.

Die Untersuchung der Dimensionalität des Testinstrumentes wurde mittels konfirmatorischer Faktorenanalyse mit dem Rasch-Modell durchgeführt. Zu beiden Testzeitpunkten wurden die Modellannahme und ein eindimensionales Modell hinsichtlich der Modellpassung anhand von Informationskriterien verglichen. Ergänzend hierzu wurden Skalenreliabilitäten als Sekundärkriterien berücksichtigt.

Auf Basis einer im Längsschnitt identifizierbaren Stichprobe (N = 87) wurden zudem Zusammenhänge zwischen den Prädiktoren Rechenfertigkeiten, Motivation sowie verschiedenen Abiturnoten (Physik, Mathematik, Gesamtdurchschnitt) und dem Leistungszuwachs zwischen den beiden Testzeitpunkten in Ziel- und Kontrollgruppe berechnet.

Ergebnisse: Der Modellvergleich zwischen ein- und dreidimensionalem Modell mithilfe des AIC weist jeweils eine bessere Modellpassung des dreidimensionalen Modells aus ($\Delta AIC = 32$ beim Prätest bzw. $\Delta AIC = 27$ beim Posttest), die EAP/PV-Reliabilitäten geben keinen klaren Anhaltspunkt für eine der beiden Lösungen (eindimensional: .79 beim Prä- bzw. .83 beim Posttest; dreidimensional: .79/.73/.74 (SW/VSW/UW) beim Prä- bzw. .83/.79/.79 beim Posttest).

Die Zusammenhänge zwischen Leistungszuwächsen in den drei Dimensionen des Fachwissens und Rechenfertigkeiten, Motivation sowie Schulnoten fallen insgesamt eher gering aus. Der stärkste Zusammenhang in der Gruppe der Lehramtsstudierenden besteht zwischen Lernzuwächsen im UW und den Rechenfertigkeiten ($r = .30$, $p = .12$). In der Kontrollgruppe der Physik BA-Studierenden hängen Lernzuwächse im UW ebenfalls verhältnismäßig stark mit der Rechenfertigkeit ($r = .31$, $p = .05$) und zusätzlich mit der Motivation ($r = .35$, $p < .05$) zusammen. Alle übrigen Korrelationen sind nicht signifikant.

Diskussion: Das neu entwickelte Testinstrument ermöglicht es, Fachwissen von angehenden Lehrkräften in drei Dimensionen zu differenzieren. Damit kann es diagnostisch eingesetzt werden, um universitäre Lehrveranstaltungen für Lehramtsstudierende hinsichtlich späterer Berufsanforderungen anzupassen. Die hier vorgestellten Ergebnisse zur Prädiktion von Leistungszuwächsen stellen zudem relevante Informationen für die Studienwahl und -planung dar.

Freitag
13. März

11:15 Uhr

GA 03/49

Chair(s)
Lehmann, Malte

Kompetenzdiagnostik und Kompetenzentwicklung von Studierenden der
Naturwissenschaften

Stäbler, Franziska; Dumont, Hanna; Becker, Michael

What about the fish's achievement in a big pond? Eine simultane Analyse von Effekten der mittleren Leistung auf Selbstkonzept und Individualleistung im Längsschnitt

Inwiefern die Zusammensetzung der Schülerschaft einer Klasse und/oder Schule die Lern- und Leistungsentwicklung von Schülerinnen und Schülern beeinflusst, stellt seit langem einen bedeutenden Gegenstand der Bildungsforschung dar (Coleman et al., 1966). Empirische Untersuchungen, die die Auswirkungen der Lerngruppenkomposition auf die Leistungsentwicklung fokussierten, konnten zeigen, dass sich in Lerngruppen mit höherer mittlerer Leistung individuelle Leistungen besser entwickeln (im Überblick siehe Dumont, Neumann, Maaz, & Trautwein, 2013). Neben diesem positiven Effekt für die Leistungsentwicklung – in der Literatur unter Kompositionseffekt bekannt – konnte ein anderer Forschungsstrang belegen, dass der Besuch einer leistungsstarken Klasse mit Nachteilen im psychosozialen Bereich verbunden ist. Leistungsstarke Lerngruppen eröffnen Schülerinnen und Schülern viele Gelegenheiten für soziale Aufwärtsvergleiche, sodass die mittlere Leistung einer Lerngruppe negativ auf das Selbstkonzept wirkt – in der Literatur als Big Fish Little Pond-Effekt (BFLPE) bekannt (Marsh, 1987). Da sich das Selbstkonzept wiederum selbst auf die Leistungsentwicklung auswirkt (Marsh & Yeung, 1997), stellt sich die Frage, wie Kompositionseffekte und BFLP-Effekte über die Zeit miteinander interagieren. Bisherige Studien betrachteten die Wechselwirkungen zwischen Selbstkonzept und Leistung entweder ausschließlich auf der Individualebene (z. B. Pinxten, de Fraine, van Damme, & D'Haenens, 2013; Seaton, Parker, Marsh, Craven, & Yeung, 2014) oder untersuchten den BFLPE und Kompositionseffekte im Längsschnitt getrennt voneinander (z. B. Köller & Baumert, 2001; Köller, Trautwein, Lüdtke, & Baumert, 2006). Daher sollen im vorliegenden Beitrag Kompositionseffekte und BFLP-Effekte simultan über mehrere Zeitpunkte hinweg betrachtet werden, um der Frage nachzugehen, wie sich die Effekte der mittleren Leistung der Lerngruppe weitertransportieren und auf nachfolgende Individualleistungen der Schülerinnen und Schüler wirken.

Als Datenbasis diente die BIJU-Studie (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter), da sowohl das fachspezifische Selbstkonzept, als auch Leistungsmaße über drei Messzeitpunkte hinweg erhoben wurden. Zu Beginn, in der Mitte und am Ende der 7. Klassenstufe wurden 6207 Schülerinnen und Schüler aus 271 Klassen hinsichtlich ihres mathematischen Selbstkonzepts befragt sowie in ihren Fähigkeiten in Mathematik getestet. Das fachspezifische Selbstkonzept wurde über fünf Items ermittelt ($\alpha T1: .86$, $\alpha T2: .9$, $\alpha T3: .91$). Mathematische Fähigkeiten wurden über einen rasch-skalierten standardisierten Leistungstest gemessen. Die statistische Auswertung erfolgte mittels multi-level cross-lagged models in Mplus, wobei die unabhängigen Variablen auf der Individualebene am Gesamtmittelwert zentriert wurden.

Es zeigte sich, dass die mittlere Leistung der Klasse einen größeren Einfluss auf die Leistungsentwicklung als auf die Entwicklung des Selbstkonzepts der Schülerinnen und Schüler hatte, d.h. der Kompositionseffekt größer ausfiel als der BFLPE. Darüber hinaus war der BFLPE nur in der ersten Hälfte des Schuljahres sichtbar, während sich der Kompositionseffekt auch noch am Ende der siebten Jahrgangsstufe zeigte. Des Weiteren zeigte sich nur ein geringer Einfluss des Selbstkonzepts auf die Leistungen, so dass die durch die mittlere Klassenleistung hervorgerufenen Einbußen im Selbstkonzept keine größeren Leistungseinbußen nach sich zogen.

Insgesamt können die Befunde dahingehend interpretiert werden, dass das Selbstkonzept von Schülerinnen und Schülern zwar nach dem Wechsel in eine neue leistungsstarke Bezugsgruppe zu Beginn des Schuljahres Verluste erfährt, es aber im Laufe des Schuljahres zu keinen weiteren Einbußen im Selbstkonzept kommt, da die relative Position in der Lerngruppe weitestgehend konstant bleiben dürfte. Im Gegensatz dazu scheinen leistungsstarke Mitschülerinnen und Mitschüler fortwährend kognitive Anregungen zu ermöglichen, was letztlich den über das gesamte Schuljahr anhaltenden positiven Effekt der mittleren Klassenleistung auf die Leistungsentwicklung von Schülerinnen und Schülern erklären könnte.

Freitag
13. März

11:15 Uhr

GA 03/142

Chair(s)

Stäbler, Franziska

Schmidt, Isabelle; Scherrer, Vsevolod; Keller, Lena; Preckel, Franzis

Die Ausbildung akademischer Selbstkonzeptprofile bei Grundschulkindern: Eine Überprüfung des genesteten Marsh/Shavelson Modells

Theoretischer Hintergrund: Akademische Selbstkonzepte sind mentale Repräsentationen der eigenen akademischen Fähigkeiten (Moschner & Dickhäuser, 2010). Bezogen auf die Struktur des akademischen Selbstkonzeptes und dessen Entwicklungsverlauf gibt es bislang wenige Studien, die sich auf das gesamte Grundalter (d. h. Klasse 1-4) beziehen (z. B. Poloczek, Karst, Praetorius & Lipowsky, 2012). Nach Brunner et al. (2010) beschreibt für Sekundarstufenschüler/innen ein genestetes Modell das akademische Selbstkonzept optimal (genestetes Marsh/Shavelson Modell; kurz: NMS Modell). Im NMS-Modell sind die fachspezifischen akademischen Selbstkonzepte (z. B. mathematisches und verbales Selbstkonzept) unter einem generellen akademischen Selbstkonzept genestet. Damit bilden die fachspezifischen akademischen Selbstkonzepte, bereinigt um Anteile des allgemeinen akademischen Selbstkonzeptes, das fachspezifische Profil des akademischen Selbstkonzeptes ab. Innerhalb des NMS-Modells wurden negative Beziehungen zwischen fachspezifischen Selbstkonzepten gefunden, die eine Binnendifferenzierung des akademischen Selbstkonzeptes für ältere Schüler belegen (vgl. Brunner et al., 2010). Für Grundschulkindern steht eine solche Untersuchung noch aus (Ehm, 2014). Theoretisch sollte das akademische Selbstkonzept mit zunehmendem Alter differenzierter werden, was u.a. mit der kognitiven Entwicklung, dem Einfluss von Schulnoten oder einer realistischeren Selbsteinschätzung zusammenhängt (vgl. Harter, 1999). Hierfür gibt es erste empirische Hinweise (Marsh & Ayotte, 2003).

Fragestellung: (1) Lässt sich für die Struktur des akademischen Selbstkonzeptes im Grundschulalter ebenfalls eine sehr gute Passung des NMS-Modells nachweisen? (2). Findet eine Ausdifferenzierung des akademischen Selbstkonzeptes statt, werden also die Zusammenhänge zwischen verbalem und mathematischem Selbstkonzept mit zunehmender Klassenstufe geringer?

Methode.

Stichprobe: Untersucht wurde eine Stichprobe von 903 Erstklässler/innen (353 weiblich), 890 Zweitklässler/innen (385 weiblich), 1562 Drittklässler/innen (734 weiblich) und 1674 Viertklässler/innen (689 weiblich) aus 298 unterschiedlichen Klassen aus sechs Bundesländern (N = 5327).

Material: Das generelle akademische Selbstkonzept, das mathematische und das verbale Selbstkonzept (getrennt nach Lesen und Schreiben) wurden mit je drei Items erhoben (Bsp.: „Ich bin gut... in der Schule/in Mathe/im Schreiben/im Lesen“). Die Beantwortung erfolgte auf einer dreistufigen Smiley-Skala. Die Items entstammen der Kurzversion FEES-K (Baudson & Preckel, 2011) des FEES (Rauer & Schuck, 2003, 2004).

Design und Datenauswertung: Das NMS-Modell wurde mit einer konfirmatorischen Faktorenanalyse für geordnete kategoriale Daten (IFA mit WLSMV Schätzer) mit Mplus 7.2 geprüft. Zunächst wurde die Messinvarianz der Items über die Klassenstufen mittels Multigruppenvergleich aufeinanderfolgender Klassenstufen überprüft, da dies die Voraussetzung zum Vergleich der Zusammenhänge der fachspezifischen akademischen Selbstkonzepte ist (Brown, 2006). Im nächsten Schritt wurde die Korrelation des verbalen und mathematischen Selbstkonzeptes zwischen zwei aufeinanderfolgenden Klassenstufen gleichgesetzt und mit einem Modell mit in den Klassenstufen frei schätzbarer Korrelation über einen Chi²Diff.-Test verglichen.

Ergebnisse: Unsere Ergebnisse sprechen für eine Erweiterung des Geltungsbereiches des NMS-Modells von Brunner et al. (2010) auf das Grundschulalter. Das NMS-Modell wies in allen vier Klassenstufen einen sehr guten Fit auf (Klasse 1: CFI = .98; Klasse 2: CFI = .98; Klasse 3: CFI = .95; Klasse 4: CFI = .96). Von skalarer Messinvarianz über die Klassenstufen konnte ausgegangen werden. Das Modell mit frei schätzbarer Korrelation von verbalem und mathematischem Selbstkonzept wies eine bessere Passung auf die Daten auf als das Modell mit gleichgesetzter Korrelation (signifikanter Chi²Diff.-Test). Zwischen dem verbalen und mathematischen Selbstkonzept zeigte sich ein nicht signifikanter positiver Zusammenhang in der ersten Klasse und ein bis zur dritten Klassenstufe ansteigender negativer Zusammenhang (Klasse 1: $r = .06$ ns; Klasse 2: $r = -.254$, $p = .035$; Klasse 3: $r = -.44$, $p \leq .001$ Klasse 4: $r = -.41$, $p \leq .001$)).

Schlussfolgerungen: Die Beziehung zwischen dem verbalen und mathematischen Selbstkonzept im NMS Modell fiel mit zunehmender Klassenstufe negativer aus. Die Befunde sprechen für eine Ausdifferenzierung des akademischen Selbstkonzeptes während der Grundschulzeit. Interessanterweise zeigte sich diese auch ohne Notenrückmeldungen bereits in Klasse 1 und 2.

Freitag
13. März

11:15 Uhr

GA 03/142

Chair(s)
Stäbler, Franziska

Weber, Kira Elena; Freund, Alexander

Erfassung des Selbstkonzepts von Jugendlichen mit Förderbedarf im Schwerpunkt Lernen: Entwicklung und Validierung eines deutschsprachigen Messinstruments

Theoretischer Hintergrund: Vor dem Hintergrund der aktuellen bildungspolitischen Inklusionsdebatte ergeben sich für die pädagogische Psychologie eine Vielfalt spezifischer Forschungsfragen, wie z.B. die Frage nach den Auswirkungen dieser heterogenen Beschulungsform auf das Selbstkonzept von Jugendlichen mit Förderbedarf im Schwerpunkt Lernen. Betrachtet man den aktuellen Forschungsstand, wird ersichtlich, dass sich eine Vielzahl internationaler Studien mit dem Selbstkonzept von Schülerinnen und Schülern mit Lernbehinderung und dem Einfluss unterschiedlicher Beschulungsformen beschäftigen (eine Übersicht hierzu bieten z.B. Chapman, 1988 und Zeleke, 2004), dem jedoch nur wenige empirische Studien aus dem deutschsprachigen Raum gegenüber stehen (z.B. Sauer, Ide & Borchert, 2007). Die bisherigen deutschsprachigen Studien konzentrierten sich dabei zumeist auf Grundschul Kinder, sodass ein akuter Forschungsbedarf zum Selbstkonzept von Jugendlichen besteht. Die Erfassung diverser Selbstkonzeptfacetten mit Hilfe von ökonomischen, zielgruppengerechten, reliablen und validen Messinstrumenten stellt dabei eine wichtige Forschungsgrundlage dar. Im Fokus der vorliegenden Studie stand daher die theoriegeleitete Entwicklung und empirische Überprüfung eines deutschsprachigen Instruments zur Erfassung des Selbstkonzepts von Jugendlichen mit Förderbedarf im Schwerpunkt Lernen.

Fragestellungen: Erfasst der entwickelte Fragebogen multiple Facetten des Selbstkonzepts von Jugendlichen mit Förderbedarf im Schwerpunkt Lernen zielgruppengerecht, ökonomisch, reliabel und valide? Kann die a priori postulierte Struktur konfirmatorisch belegt werden? Sind alle Einzelskalen Rasch-skalierbar im Sinne des eindimensionalen Rasch-Modells? Ist der Fragebogen auch bei Schülerinnen und Schülern anderer Schulformen einsetzbar und eignet sich somit für schulformübergreifende Vergleiche?

Methode.

Entwicklung des Fragebogens: In Anlehnung an das Selbstkonzeptmodell von Shavelson, Hubner und Stanton (1976) und seinen Modifikationen (siehe hierzu Marsh, 1990) wurden acht Skalen mit insgesamt 38 Items entwickelt, die multidimensionale Facetten des Selbstkonzepts erfassen sollten. Vor der Pilotierung wurden Experteninterviews (N = 8) sowie zwei Vorstudien (N = 13; N = 27) durchgeführt.

Stichproben: Gesamtstichprobe (N = 660; MAlter = 14.72, SD = 1.49): Stichprobe I = 294 Schüler (39% weiblich) mit Förderbedarf im Schwerpunkt Lernen (MAlter = 15.04, SD = 1.27); Stichprobe II = 366 Schüler (41% weiblich) verschiedener Schulformen (MAlter = 14.09, SD = 1.37).

Statistische Analysen: Als statistische Kennwerte wurden Item- und Skalenmittelwerte, Standardabweichungen, Trennschärfen nach der KTT sowie interne Konsistenzen (Cronbachs Alpha) ermittelt. Im Rahmen konfirmatorischer Faktorenanalysen (KFA) wurden verschiedene Faktormodelle spezifiziert und auf ihre Passungsgüte verglichen. Als Indizes für einen guten Modellfit wurden das Verhältnis aus Chi-Quadrat (χ^2) und Freiheitsgraden (df), der Comparative-Fit-Index (CFI), der Tucker-Lewis-Index (TLI) sowie der Root Mean Square Error of Approximation (RMSEA) herangezogen. Darüber hinaus wurden die latenten Interkorrelationen der Selbstkonzeptskalen ermittelt. Für die Acht Einzelskalen wurden Mixed-Rasch-Modelle (MRM) berechnet, um eine mögliche Mehrdimensionalität der Skalen, unskalierbare Personen oder Antworttendenzen (Response Sets) zu identifizieren (Rost, 2004). Ein- bis Vier-Klassen-Lösungen wurden miteinander verglichen sowie die Ordnung der Schwellenparameter überprüft. Zur Bewertung der besten Modellgüte wurden das Bayesian Information Criterion (BIC) sowie das Consistent Akaike Information Criterion (CAIC) herangezogen.

Ergebnisse: Die Ergebnisse zeigten für beide Stichproben eine gute Reliabilität (Gesamtstichprobe: $M\alpha = .89$). KFA und das Interkorrelationsmuster der Selbstkonzeptskalen belegten die inhaltliche Ausdifferenzierung der Selbstkonzeptfacetten und bestätigten die faktorielle Struktur (Studie I: Modell-Fit für die Acht-Faktoren-Lösung: χ^2/df (1095.96/637) = 1.72, $p < .001$, CFI = .97, TLI = .97, RMSEA = .05). Wesentliche Befunde internationaler Studien zur Struktur des Selbstkonzepts (z.B. Marsh & Shavelson, 1985; Renick & Harter, 1988/ 2012; Schwanzer, 2002) konnten mit dem Fragebogen repliziert werden. So wies das generelle Selbstwertgefühl die höchste Korrelation zum Selbstkonzept des Aussehens auf und auch die vielfach empirisch nachgewiesene Unabhängigkeit zwischen dem verbalen und dem mathematischen Selbstkonzept (z.B. Marsh, 1990) konnte analytisch bestätigt werden. MRM-Analysen belegten für Stichprobe I und II weitestgehend die Eindimensionalität der Einzelskalen. Für die Gesamtstichprobe konnten diese Ergebnisse für vier der acht Skalen nicht repliziert werden, sodass sich hier Diskussionsbedarf zeigte.

Freitag
13. März

11:15 Uhr

GA 03/142

Chair(s)
Stäbler, Franziska

Jansen, Malte; Arens, Katrin

Mehrdimensionalität und Hierarchie akademischer Selbstkonzepte im sprachlichen Bereich

Theoretischer Hintergrund: Das akademische Selbstkonzept wird in der aktuellen Forschung als mehrdimensionales, hierarchisches Konstrukt aufgefasst (Marsh, 1990, 2007). Die bisherige Forschung fokussierte sich stark auf den Aspekt der Mehrdimensionalität und konnte vielfältige Belege dafür liefern, dass Schülerinnen und Schüler zwischen ihren selbstwahrgenommenen Fähigkeiten in verschiedenen Bereichen differenzieren. Diese Differenzierung tritt besonders stark zwischen der mathematischen und der sprachlichen Domäne auf (Marsh, 1986; Möller & Marsh, 2013). Aber auch innerhalb der sprachlichen Domäne konnte gezeigt werden, dass Schülerinnen und Schüler unabhängige Selbstkonzepte für verschiedene Sprachen aufweisen, also z. B. zwischen ihren Fähigkeiten in der Muttersprache und in Fremdsprachen unterscheiden. Neben dieser Differenzierung des akademischen Selbstkonzepts zwischen Schulfächern wurde bisher allerdings wenig betrachtet, ob sprachliche Selbstkonzepte auch innerhalb eines Faches eine Mehrdimensionalität aufweisen, also inwiefern Schüler für einzelne Fähigkeitsfacetten wie etwa Lesen und Schreiben trennbare Selbstkonzepte entwickeln (Lau et al., 1999; Yeung et al., 2000). Ebenfalls blieb die Frage offen, ob sich solche spezifischen Selbstkonzeptfacetten zu einem allgemeinen fachbezogenen Selbstkonzept aggregieren ließen, was als Beleg für die Hierarchie des Konstruktes gelten könnte.

Fragestellung: Ziel des Beitrags ist die Beschreibung der Struktur des akademischen Selbstkonzepts in drei Sprachen (Deutsch, Englisch, Französisch) und vier Fähigkeitsbereichen (Lesen, Schreiben, Zuhören, Sprechen). In Bezug auf die Mehrdimensionalität des akademischen Selbstkonzepts soll untersucht werden, ob sich innerhalb eines sprachlichen Fachs fähigkeitsbezogene Selbstkonzeptfacetten trennen lassen. Sollte sich diese Ausdifferenzierung zeigen, soll in Bezug seine Hierarchie geprüft werden, ob sich die fähigkeitsbezogenen Facetten zu einem Faktor höherer Ordnung aggregieren lassen und ob dieser Faktor von einem direkt erhobenen globalen Fachselbstkonzept trennbar ist. Dabei wird nicht nur die messtheoretische Trennbarkeit der Facetten und Faktoren untersucht, sondern auch potentielle differentielle Zusammenhänge zu Leistungsmaßen und zum Geschlecht in den Blick genommen.

Methode: Die Studie basiert auf einer Stichprobe von 4257 Schülerinnen und Schülern der 9. Jahrgangsstufe aller Schulformen. Alle Schülerinnen und Schüler erhielten Items zu ihren akademischen Selbstkonzepten in den Fächern Deutsch und Englisch, wobei einige Items nach ihren globalen Selbsteinschätzungen in diesen Fächern fragten (z. B. „in Deutsch bin ich gut“), andere hingegen nach Selbsteinschätzungen bezogen auf spezifische Fähigkeiten in diesen Fächern (z. B. „das Lesen von Texten fällt mir leicht“; in jedem Fach erhoben). Eine Teilstichprobe erhielt zusätzlich analoge Items zum Fach Französisch. Außerdem wurden die Fachnoten erhoben und Leistungstests in den Kompetenzbereichen Lesen und Zuhören eingesetzt. Die Selbstkonzeptstruktur wurde mit konfirmatorischen Faktorenanalysen modelliert.

Ergebnisse: Mehrdimensionale Modelle, in denen die fähigkeitsbezogenen Selbstkonzeptfacetten innerhalb der Fächer unterschieden wurden, zeigten in allen drei Fächern eine bessere Passung zu den Daten als unidimensionale Modelle. Differentielle Zusammenhänge der fähigkeitsbezogenen Selbstkonzeptfacetten zu Leistungstests im Lesen und Zuhören konnten allerdings nicht gezeigt werden. Die Selbstkonzeptfacetten konnten zu einem Faktor höherer Ordnung aggregiert werden, der hoch mit dem globalen fachbezogenen Selbstkonzept korrelierte ($\rho(\text{Deutsch}) = .85$; $\rho(\text{Englisch}) = .95$, $\rho(\text{Französisch}) = .95$). Die getesteten Faktormodelle waren messinvariant über die Geschlechtergruppen. Sowohl für den Faktor höherer Ordnung als auch für das globale fachbezogene Selbstkonzept zeigten sich Geschlechtseffekte zugunsten der Mädchen. Auch die Zusammenhänge der beiden Faktoren mit Noten und Testleistungen waren in ihrer Höhe vergleichbar. In den drei Fächern zeigten sich ähnliche Ergebnismuster.

Diskussion: Die Ergebnisse liefern neue Evidenz für die Mehrdimensionalität und Hierarchie akademischer Selbstkonzepte innerhalb des sprachlichen Bereichs. So konnte gezeigt werden, dass auch unterhalb der Fachebene trennbare Selbstkonzeptfacetten nachweisbar sind, was die Befunde zur Mehrdimensionalität akademischer Selbstkonzepte auf einer fähigkeitsspezifischen Ebene erweitert. Ein aus diesen Facetten aggregierter Faktor verhält sich empirisch sehr ähnlich zu einem direkt erhobenen fachbezogenen Selbstkonzept, was für eine hierarchische Struktur innerhalb des sprachlichen Selbstkonzepts spricht und somit Implikationen für die Diagnostik und Messung sprachlicher Selbstkonzepte hat. Eine sparsame Messung mit wenigen, global formulierten Items scheint die Breite der selbstwahrgenommenen fachlichen Fähigkeiten gut abzubilden.

Rieger, Sven; Göllner, Richard; Zettler, Ingo; Trautwein, Ulrich

Die Bedeutung sozialer und dimensionaler Vergleiche für das Selbstkonzept von Drittklässlern

Theoretischer Hintergrund: Schulische Selbstkonzepte gelten in der pädagogisch-psychologischen Forschung als wichtige Determinanten des Leistungsverhaltens von Schülerinnen und Schülern. Beispielsweise fördern Selbstkonzepte – vermittelt über motivationale Variablen – Lernprozesse in der Schule, erklären geschlechtsspezifische Leistungsunterschiede und beeinflussen das akademische Wahlverhalten von Schülern während des Übergangs in die universitäre Ausbildung (siehe z.B. Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005; Nagy et al., 2008). Hinsichtlich der Genese und Entwicklung schulischer Selbstkonzepte haben sich in den letzten Jahren zwei theoretische Ansichten etabliert (Marsh, 2007; Parker, Marsh, Lüdtke, & Trautwein, 2013; Möller, Pohlmann, Köller, & Marsh, 2009). Hierzu gehören (a) soziale Vergleiche mit den Mitschülerinnen und Mitschülern (Big-Fish-Little-Pond Effekt; kurz BFLPE; Marsh, 1987) und (b) internale Vergleiche, bei denen die Schülerinnen und Schüler die Leistungen in einem Fach mit den Leistungen in einem anderen Fach vergleichen (Internal/External frame of Reference Model; kurz I/E-Modell; Marsh, 1986).

Fragestellung: Beide Erklärungsansätze konnten anhand empirischer Arbeiten bestätigt werden, wenngleich im Rahmen der bisher zur Verfügung stehenden Studien beiden Modelle überwiegend getrennt voneinander und vornehmlich im Sekundarschulbereich untersucht wurden. Ziel des vorliegenden Beitrages ist es daher die Genese fachspezifischer Selbstkonzepte (Mathematik und Deutsch) bei Schülerinnen und Schülern der Grundschule zu untersuchen.

Methode: Die Untersuchung wurde im Rahmen der wissenschaftlichen Begleitung des Hector-Projektes durch das Hector-Institut für Empirische Bildungsforschung sowie des Deutschen Institut für Internationale Pädagogische Forschung (DIPF), Frankfurt a.M., durchgeführt. Datengrundlage für die vorliegende Arbeit sind Daten, die im Rahmen der Grundschul-Talente-Studie 2011 erhoben wurden. Die Datenerhebung erfolgte im Klassenverband und wurde durch geschulte Testleiter durchgeführt.

Stichprobe: Die Stichprobengröße beträgt $N = 2\,403$ Grundschülerinnen und Grundschüler (51.7% weiblich). Das durchschnittliche Alter betrug $M = 9.28$ Jahre ($SD = 0.56$ Jahre) mit einer Spanne von 8 bis 13 Jahren.

Instrumente: Als Leistungsindikatoren dienten der Deutsche Mathematiktest für zweite Klassen (DEMAT 2+; Krajewski, Liehm & Schneider, 2004) sowie die Würzburger Leise Leseprobe (WLLP; Küspert & Schneider, 1998). Das Selbstkonzept wurde durch den Self-Description Questionnaire I (SDQ I; Arens, Trautwein & Hasselhorn, 2011) erfasst.

Statistische Analyse: Für die vorliegende Untersuchung wurden Mehrebenenmodelle verwendet (Raudenbush & Bryk, 2002). Mehrebenenmodelle berücksichtigen einerseits, die bei Untersuchung mit mehreren Schulen bzw. Klassen entstehende hierarchische Struktur der Daten und bieten andererseits die Möglichkeit, inhaltliche Fragestellungen auf mehreren Ebenen zu untersuchen.

Ergebnisse: Es fanden sich sowohl für den BFLPE (sozialer Vergleich) als auch für das I/E-Modell (interner Vergleich) die typischen Ergebnismuster. Entsprechend hatte die jeweilige fachspezifische individuelle Leistung positive Effekte auf die jeweiligen Selbstkonzepte (in Deutsch und Mathematik), die mittleren Klassenleistungen hingegen wirkten sich negativ auf die fachgebundenen Selbstkonzepte aus (BFLPE). Gleichzeitig wirkte sich die Leistung in Mathematik negativ auf das Selbstkonzept im Fach Deutsch aus. Keine Effekte fanden sich für die Leistung in Deutsch auf das mathematische Selbstkonzept. Darüber hinaus fanden sich hohe Korrelationen zwischen den Kompetenzen und nur geringe Korrelationen zwischen den jeweiligen Selbstkonzepten (I/E-Modell). Die Ergebnisse werden diskutiert und in den Forschungsstand zum BFLPE und dem I/E-Modell eingearbeitet.

Lorenz, Jennifer; Stubbe, Tobias C.; Nonte, Sonja

Die Bedeutung der sozialen Herkunft für die Inanspruchnahme privaten Nachhilfeunterrichts

Im Verlauf der Sekundarstufe erhalten 21 Prozent der deutschen Schülerinnen und Schüler Nachhilfeunterricht (Autorengruppe Bildungsberichterstattung, 2012). Dennoch erfuhr dieses Thema in der Bildungsforschung über Jahrzehnte kaum Beachtung (Hass & Jäger, 2011). Erst in den vergangenen Jahren zeichnet sich eine Zunahme von Studien ab. Diese beschäftigen sich unter anderem mit der Frage, welche Rolle die soziale Herkunft von Schülerinnen und Schülern für die Teilnahme an Nachhilfeunterricht spielt. Während laut Schneider (2004) lediglich das Einkommen der Eltern mit der Inanspruchnahme von Nachhilfeunterricht zusammenhängt, findet Guill (2012) Hinweise darauf, dass für Hamburger Schülerinnen und Schüler auch der Bildungsabschluss und die berufliche Stellung der Eltern von Bedeutung sind. Gleichzeitig zeichnet sich in dieser Studie eine schulformabhängige Nachfrage nach Nachhilfeunterricht ab.

Bildungsentscheidungen, zu denen auch die Entscheidung für Nachhilfeunterricht zu zählen ist, werden in der empirischen Bildungsforschung üblicherweise mithilfe der rational choice Theorie modelliert. Anhand dieser Theorie konnte wiederholt gezeigt werden, dass Bildungsentscheidungen in Abhängigkeit der sozialen Herkunft getroffen werden und, dass dabei insbesondere das Motiv des Stuserhalts eine Rolle spielt. Besonders deutlich kam dieses Motiv bislang bei statushöheren Eltern bei der Entscheidung ihre Kinder nach der Grundschule auf ein Gymnasium zu schicken zum Tragen (Maaz, Hausen, McElvany & Baumert, 2006). Angesichts der Tatsache, dass das Gymnasium sich zur meistbesuchten Schulform in der Sekundarstufe I entwickelt hat (Autorengruppe Bildungsberichterstattung, 2014), ist davon auszugehen, dass es der Funktion statushöherer Familien den Erhalt ihres sozialen Status zu sichern allein nicht mehr nachkommen kann. Nachhilfeunterricht könnte von diesen Familien genutzt werden, um ihren Kindern einen Leistungsvorsprung zu verschaffen, der den Stuserhalt ermöglicht.

Ausgehend von der heterogenen Befundlage bestehender Studien wird im vorliegenden Beitrag zunächst anhand bundesweit repräsentativer Daten untersucht, inwiefern verschiedene Indikatoren der sozialen Herkunft mit der Entscheidung für Nachhilfeunterricht zusammenhängen. Weiterhin wird der Frage nachgegangen, ob das Motiv des Stuserhalts für Eltern von Gymnasiastinnen und Gymnasiasten bei der Entscheidung für Nachhilfeunterricht mehr ins Gewicht fällt als für Eltern von Schülerinnen und Schülern anderer Schulformen.

Für die empirischen Analysen werden Daten von Schülerinnen und Schülern der fünften und sechsten Jahrgangsstufe sowie von deren Eltern aus dem nationalen Bildungspanel (NEPS; Blossfeld, Roßbach & von Maurice, 2011) ($n = 5525$) genutzt. Die Daten wurden mithilfe des Pakets mice (van Buren & Groothuis-Oudshoorn, 2011) in R (R Development Core Team, 2014) multipel imputiert ($m=10$). Zur Überprüfung der Forschungsfragen wurden logistische Regressionsanalysen unter Berücksichtigung der Mehrebenenstruktur der Daten durchgeführt. Als abhängige Variable wurde die Entscheidung von Eltern für ihre Kinder in der sechsten Klasse Nachhilfeunterricht in Anspruch zu nehmen, nachdem diese in der fünften Klasse keinen erhalten hatten, betrachtet. Die soziale Herkunft wurde nach Bourdieu (1983) als ökonomisches (Einkommen, höchster Berufsabschluss der Eltern), kulturelles (Anzahl der Bücher im Haushalt, höchster Bildungsabschluss der Eltern) und soziales Kapital (Unterstützungsverhalten der Eltern) einbezogen. Als Stuserhaltmotiv wurden die Angaben der Eltern auf die Frage verwendet, wie sehr sie sich für ihr Kind einen ähnlichen oder besseren Bildungsabschluss oder beruflichen Status wünschen wie sie selbst. In allen Analysen wurde für die Schulleistungen, sowie für weitere Merkmale der Schülerinnen und Schüler (u.a. Migrationshintergrund und Geschlecht) kontrolliert.

Die Ergebnisse zeigen, dass kein eindeutiger Zusammenhang zwischen der sozialen Herkunft und der Inanspruchnahme von Nachhilfeunterricht besteht: Mehr Unterstützung und ein höherer Bildungsabschluss der Eltern zeigen einen negativen Zusammenhang mit der Chance auf Nachhilfeunterricht, während eine höhere Anzahl Bücher im Haushalt einen positiven Effekt aufweist. Einkommen und Berufsabschluss der Eltern hängen nicht mit der Entscheidung für Nachhilfeunterricht zusammen. Hinsichtlich des Stuserhaltmotivs zeigt sich, dass eine stärkere Ausprägung dieses Motivs die Chance auf Nachhilfeunterricht erhöht, es jedoch bei der Entscheidung für eine solche Förderung für Eltern von Gymnasiastinnen und Gymnasiasten nicht wichtiger ist als für Eltern von Schülerinnen und Schülern anderer Schulformen.

Freitag
13. März

11:15 Uhr

GA 03/149

Chair(s)

Lorenz, Jennifer

3

Hollmann, Jelena; Otterpohl, Nantje; Wild, Elke

Argumentationskompetenz fördern im Elternhaus? Erste Befunde zur Evaluation eines Elterntrainings zur Unterstützung informeller Erwerbsprozesse in Familien

Obwohl unumstritten ist, dass für den Bildungserfolg Heranwachsender maßgebliche Diskursfähigkeiten wesentlich in (z.T. schichtspezifisch geprägten) Routinen der Eltern-Kind-Kommunikation begründet sind, mangelt es an Familienbildungsprogrammen zum Ausgleich entsprechender primärer Bildungsdisparitäten. Der vorliegende Beitrag setzt an diesem Desiderat an und stellt erste Ergebnisse zur Akzeptanz und Wirksamkeit eines Elterntrainings vor, das sich an Eltern von SchülerInnen der fünften Klassenstufe richtet. Das Training wurde im Rahmen des Projekts Fundus* entwickelt und fokussiert auf familiäre Erwerbsbedingungen von Argumentationskompetenz (AK), die auf den, in interdisziplinärer Anstrengung gewonnenen, Erkenntnissen aufbaut und darauf abzielt, Eltern (1) die Relevanz alltäglicher sprachlicher Interaktionen für die Entwicklung der kindlichen AK zu verdeutlichen, (2) ihren Blick auf die langfristigen Vorteile einer in den Familien gepflegten "Verhandlungskultur" zu lenken, (3) ihnen Prinzipien positiver und autoritativer Erziehung so zu vermitteln, dass sie adaptiv, an den Erfordernissen des Einzelfalls angewendet werden können und (4) Eltern Hilfen zur Bewältigung schwieriger, hoch konfliktärer argumentativer Interaktionen an die Hand zu geben. Die Sitzungen wurden von Tandems (eigens geschulte Trainer mit linguistischer und psychologischer Expertise) durchgeführt und sahen eine Kombination von kurzen Input-Phasen und ausgedehnteren interaktiven Phasen in moderierten Kleingruppen oder im Plenum vor, in deren Rahmen mithilfe eigens ausgearbeiteter Videos, Rollenspiele und "Hausaufgaben" die zentralen Informationen vermittelt, eingeübt und unter Anwendungsgesichtspunkten reflektiert wurden.

Der vorliegende Bericht basiert auf den in einem Prä-Post-Kontrollgruppendesign erhobenen Daten von insgesamt 37 Eltern. Die Eltern wurden über die Schulen angesprochen, die vorab anhand soziodemographischer Merkmale parallelisiert und per Zufall der Experimental- oder der Kontrollbedingung zugewiesen wurden. Insgesamt 17 Eltern absolvierten und bewerteten das sieben Sitzungen umfassende und im Abstand von einer Woche stattfindende Training, die 20 Eltern der "Kontrollgruppe" bearbeiteten lediglich zum Prä- und Post-Zeitpunkt vorgelegte Fragebögen. Die Ergebnisse zeigen eine sehr hohe subjektive Zufriedenheit der TrainingsteilnehmerInnen: Praktisch alle würden das Training weiterempfehlen und 94 Prozent gaben an, dass das Training ziemlich bzw. sehr hilfreich für sie selber war ($M=4.47$; $SD=.62$; $Min=3$; $Max=5$) und dass ihre Teilnahme an dem Training ihren Kindern ziemlich bzw. sehr nutzte ($M=4.41$; $SD=.70$; $Min=3$; $Max=5$). Nicht minder zufriedenstellend sind die Bewertungen der Prozessqualität. So beurteilten die meisten Teilnehmer die Sitzungsinhalte als hilfreich und gaben an, dass sie die vermittelten Inhalte in ihr eigenes Verhaltensrepertoire aufnehmen konnten. Auch die Urteile zur Strukturqualität fielen durchweg positiv aus. Beispielsweise äußerte sich die Mehrheit der Teilnehmer als sehr zufrieden mit Blick auf die gestellten Aufgaben und Übungen ($M=3.53$; $SD=.51$; $Min=3$; $Max=4$) und die Möglichkeiten, eigene Erfahrungen und Probleme in den Treffen einzubringen ($M=3.82$; $SD=.39$; $Min=3$; $Max=4$). Um die Wirksamkeit des Trainings auch jenseits der subjektiven Zufriedenheit und Erfolgsbilanz zu überprüfen, wurde eine zweifaktorielle multivariate Varianzanalyse mit dem Messwiederholungsfaktor Messzeitpunkt und dem Faktor Gruppe gerechnet. (1) Für die Wichtigkeit einer hohen AK im Alltag konnte kein signifikanter Interventionseffekt gefunden werden. Die Eltern aus der Experimentalgruppe hingegen erkannten nach dem Training mehr Potential des Elternhauses beim Erwerb von AK als die Eltern der Kontrollgruppe ($F(1,35)=9.89$, $p<.05$, $\eta^2=.22$). (2) In der Experimentalgruppe konnten die Eltern nach dem Training die Konsequenzen unterschiedlicher dialogischer Verfahren besser erkennen und auf die Fähigkeiten des Kindes abstimmen ($F(1,35)=4.82$, $p<.05$, $\eta^2=.12$). Außerdem beschrieben die Eltern der Experimentalgruppe nach dem Training einen höheren sprachlichen Anregungsgehalt in der familialen Kommunikation ($F(1,35)=5.13$, $p<.05$, $\eta^2=.13$) im Vergleich zur Kontrollgruppe. Bei der Anwendung erwerbssupportiver Muster zeigte sich kein Interventionseffekt. (3) Im Vergleich zur Kontrollgruppe realisierten trainierte Eltern nach dem Training häufiger autoritatives Erziehungsverhalten ($F(1,35)=8.46$, $p<.05$, $\eta^2=.20$) und zeigten weniger dysfunktionales Erziehungsverhalten ($F(1,34)=8.89$, $p<.05$, $\eta^2=.21$). (4) Ferner konnte für unterstützende Reaktionen auf kindliche Emotionen zwar kein signifikanter Interventionseffekt nachgewiesen werden, die Eltern der Experimentalgruppe zeigten aber nach dem Training weniger unterdrückend-bestrafende Reaktionen auf die kindlichen Emotionen ($F(1,35)=6.68$, $p<.05$, $\eta^2=.16$). Insgesamt zeigt die Studie, dass die Teilnahme an dem Elterntraining die Eltern dazu befähigte, die sprachlich-kommunikative Weiterentwicklung sowie die psychosoziale Entwicklung ihrer Kinder besser unterstützen zu können. Abschließend wird der Mehrwert unter Berücksichtigung der Beschränkungen der Studie diskutiert.

Freitag
13. März

11:15 Uhr

GA 03/149

Chair(s)

Lorenz, Jennifer

Schmid, Christine

Hausaufgabenhilfe durch Eltern und Geschwister. Ein Mediationsmodell zur Erklärung von Geschwisterpositionseffekten auf der Grundlage von PISA 2000

Der Einfluss von Geschwistern auf die schulischen Leistungen von Kindern stellt ein bisher unterbelichtetes Thema in der empirischen Bildungsforschung dar. Ein Prozessmerkmal, über das ein solcher Einfluss stattfinden könnte, bildet die geschwisterliche Hausaufgabenhilfe.

Für die Hausaufgabenhilfe von Eltern hat sich in einer Metaanalyse gezeigt, dass in der Regel ein negativer Zusammenhang mit den schulischen Leistungen von Kindern besteht (Hill & Tyson 2009). Dieser negative Zusammenhang ist wahrscheinlich darauf zurückzuführen, dass vor allem diejenigen Kinder Hausaufgabenhilfe erhalten, die Probleme haben, den schulischen Leistungsanforderungen zu genügen. Letzteres spiegelt sich dann in niedrigeren Leistungstestwerten wieder.

Für die geschwisterliche Hausaufgabenhilfe kann von daher angenommen werden, dass sie ebenfalls einen negativen Zusammenhang mit Leistungstestwerten aufweist. Im vorliegenden Beitrag soll darüber hinaus geprüft werden, ob sich mittels eines Mediationsmodells nicht außerdem positive Effekte für die geschwisterliche Hausaufgabenhilfe nachweisen lassen.

In einem Beitrag, in dem auf der Grundlage der PISA 2000 Studie (N = 34 561) Effekte der Geschwisterkonstellation auf die Lesekompetenz von Schülerinnen und Schülern untersucht wurden (Schmid, eingereicht), hat sich gezeigt, dass Kinder mit einem älteren Geschwister eine niedrigere Lesekompetenz aufweisen als Kinder ohne ein älteres Geschwister. In den Analysen wurde für eine Reihe von Variablen kontrolliert, die mit dem Vorhandensein eines älteren Geschwisters konfundiert sein könnten (Zahl der Kinder, SES, Migrationshintergrund, Herkunft aus einem alleinerziehenden Haushalt oder einer Stieffamilie, Geschlecht). Außerdem hat sich gezeigt, dass geschwisterliche Hausaufgabenhilfe vor allem dann berichtet wurde, wenn ein älteres Geschwister vorhanden war.

In Fortführung dieser Analysen soll im vorliegenden Beitrag zunächst die Annahme überprüft werden, dass die elterliche Hausaufgabenhilfe abnimmt, wenn ein älteres Geschwister vorhanden ist. Diese Annahme speist sich aus der Vermutung, dass Eltern bei Vorhandensein eines älteren Geschwisters aufgrund eigener begrenzter Ressourcen (Downey, 2001) diesem die Aufgabe der Hausaufgabenhilfe häufiger überlassen.

Anschließend soll anhand eines Mediationsmodells (schrittweise multiple Regression) ermittelt werden, inwieweit der Unterschied in den Testleistungen zwischen Kindern mit und ohne einem älteren Geschwister auf die Häufigkeit der Hausaufgabenhilfe durch Geschwister oder durch Eltern zurückzuführen ist.

Im Ergebnis zeigte sich auch unter Einbezug aller genannten Kontrollvariablen ein deutlicher Anstieg der Hausaufgabenhilfe durch Geschwister bei Vorhandensein eines älteren Geschwisters (um 1.14 Skalenpunkte, Betagewicht: .45, $p < .000$). Ein Abfall der elterlichen Hausaufgabenhilfe bei Vorhandensein eines älteren Geschwisters war dagegen nur für die Väter statistisch signifikant (Mütter: 0.05 Skalenpunkte, Betagewicht: .02, n.s.; Väter: 0.08 Skalenpunkte, Betagewicht: .03, $p < .001$).

Im Regressionsmodell zur Überprüfung der Mediationseffekte verringerte sich wie erwartet im zweiten Schritt durch die Aufnahme der geschwisterlichen Hausaufgabenhilfe ins Modell der Unterschied in den Testleistungen zwischen Kindern mit und ohne ältere Geschwister (von 15 auf 6 Testpunkte). Dagegen ergab sich durch die Aufnahme der elterlichen Hausaufgabenhilfe ins Modell keine bedeutsame Veränderung. Sowohl die Hausaufgabenhilfe durch Geschwister wie auch die der Eltern wies einen negativen Zusammenhang mit der Lesekompetenz der Schülerinnen und Schüler auf.

Offenbar wird elterliche Hausaufgabenhilfe relativ unabhängig von der Frage erteilt, ob ein älteres Geschwister vorhanden ist oder nicht. Dementsprechend leistete die elterliche Hausaufgabenhilfe auch keinen Beitrag zur Erklärung des Unterschiedes in der Lesekompetenz zwischen Kindern mit und ohne ältere Geschwister. Dagegen verringerte sich der Unterschied in der Lesekompetenz zwischen Kindern mit und ohne ältere Geschwister deutlich durch die Aufnahme der geschwisterlichen Hausaufgabenhilfe ins Modell. Dieser Mediationseffekt weist auf eine positive Wirkung der geschwisterlichen Hausaufgabenhilfe auf die Testleistungen der jüngeren Geschwister hin.

Freitag
13. März

11:15 Uhr

GA 03/149

Chair(s)

Lorenz, Jennifer

Kollmayer, Marlene; Hodosi, Tanja; Schober, Barbara; Spiel, Christiane

Geschlechterstereotype jüngerer und älterer Eltern

Kulturell geprägte Vorstellungen darüber, was typisch Mann und typisch Frau ist, werden uns von frühester Kindheit an mitgegeben (Asendorpf & Neyer, 2007). Der Einfluss von Eltern auf ihre Kinder ist dabei wesentlich: Sie fungieren als Modelle, teilen ihr Wissen und ihre Erwartungen mit ihren Kindern und belohnen erwünschtes Verhalten (Carli & Bukkato, 2000). Altersunterschiede in Geschlechterstereotypen wurden bis jetzt vor allem im Kindes und Jugendalter untersucht. Nach einem Höhepunkt an Rigidität bezüglich Geschlechterrollen im Grundschulalter kommt es in der Jugend zu mehr Flexibilität und weniger Geschlechterstereotypen (Kulik, 2002). Im Erwachsenenalter wird wieder mehr an Geschlechterrollen festgehalten (Shepard & Hess, 1975). Besonders die Elternschaft und die Erziehungsaufgaben können im Erwachsenenalter zu einer Erhöhung geschlechterstereotyper Einstellungen führen (Calvo-Salguero, García-Martínez & Monteoliva, 2008). Inwieweit es innerhalb der Elternschaft in Abhängigkeit vom Alter Unterschiede in Geschlechterstereotypen gibt, ist kaum erforscht. Im Fokus der vorliegenden Studie stehen daher Geschlechterstereotype von Eltern in Abhängigkeit von ihrem Alter. Das Interesse gilt Eltern von Kindern zwischen 3 und 6 Jahren, da der Einfluss der Eltern auf ihre Kinder bezüglich Geschlechtsstereotypen in diesem Alter besonders groß ist (Gelman, Taylor & Nguyen, 2004). Da Spielsachen im Leben von Kleinkindern eine große Rolle spielen (Cherney & Dempsey, 2010) und Eltern über Spielzeugauswahl und Reaktionen auf das Spiel Kindern schon früh Werte und Einstellungen vermitteln (Wood, Desmarais & Gugula, 2002), wurde Spielzeug als zentrales Instrument für diese Studie ausgewählt.

Vor diesem Hintergrund wurde folgende Fragestellung untersucht: Unterscheiden sich jüngere von älteren Eltern hinsichtlich ihrer (1) Einstellungen zu Geschlechterrollen, (2) ihrer Einschätzung der Geschlechtstypisierung und (3) Erwünschtheit sowie (4) ihrer Bereitstellung verschiedener Spielzeuge?

Zur Beantwortung der Fragestellung wurde 324 Personen (79% Frauen) ein Fragebogen vorgegeben. Der Median des Alters der Befragten betrug 34 Jahre (Q25 = 30, Q75 = 38) und wurde als Trennwert zwischen jüngeren und älteren Eltern herangezogen. Die Kinder, auf sich die Eltern im Rahmen der Befragung bezogen, waren zwischen 2.49 und 6.98 Jahre alt.

Die Einstellung zu Geschlechterrollen wurde mittels Fragebogen zur normativen Geschlechtsrollenorientierung (NGRO, Athenstaedt, 2000) erhoben. Zur Erfassung der Geschlechtstypisierung von Spielsachen wurden den Eltern Bilder von Spielzeugen (z.B. Puppenhaus, Knetmasse, Matchboxautos) vorgegeben, die sie mittels visueller Analogskala danach beurteilen sollten, ob diese Spielzeuge eher „für sie“, „für ihn“ (Pole der Skala) oder „für beide“ (neutrale Mitte der Skala) geeignet sind. Zudem wurden die Eltern gebeten, die abgebildeten Spielzeuge auf Ratingskalen danach zu beurteilen, wie wünschenswert sie für Ihr Kind sind. Zur Erfassung der Bereitstellung verschiedener Spielzeuge wurden die Eltern gebeten anzugeben ob sie das Spielzeug tatsächlich zu Hause haben.

Die Ergebnisse zeigen, dass ältere Eltern ($M = 2.78$, $SD = 0.97$) auch unter Kontrolle des Bildungsstandes egalitärer eingestellt sind als jüngere Eltern ($M = 3.17$, $SD = 0.92$), $F(1, 317) = 9.85$, $p = .002$, $\eta_p^2 = .030$. Bezüglich der Einschätzung der Geschlechtstypisierung der einzelnen Spielzeuge wurden keine signifikanten Unterschiede zwischen den beiden Altersgruppen der Eltern festgestellt und sowohl ältere als auch jüngere Eltern finden „männliches“ Spielzeug für Jungen erwünschter ($M = 4.97$, $SD = 1.22$) als für Mädchen ($M = 2.95$, $SD = 1.40$) und „weibliches“ Spielzeug für Mädchen erwünschter ($M = 5.31$, $SD = 1.08$) als für Jungen ($M = 3.35$, $SD = 1.48$), $F(1, 317) = 308.46$, $p < .001$, $\eta_p^2 = .368$. Die Analyse der Bereitstellung von verschiedenen Spielzeugen zeigte, dass sowohl Kinder jüngerer als auch Kinder älterer Eltern mehr geschlechterstereotype Spielsachen zu Hause haben als Spielsachen, die für das andere Geschlecht als typisch angesehen werden.

Die Ergebnisse zeigen, dass Kinder schon in jungen Jahren geschlechterstereotyp sozialisiert werden und lassen auf einen Trend in Richtung Rückkehr zu traditionellen Geschlechterrolleneinstellungen bei Eltern schließen.

List, Marit Kristine; Nagy, Gabriel; Köller, Olaf

Wer ist von Test Speededness betroffen?

Test Speededness bezeichnet das Ausmaß, in dem Teilnehmende während der Testbearbeitung von Zeitdruck betroffen sind. Testteilnehmende gelten als von Speededness betroffen, wenn sie nicht alle Aufgaben eines Tests bearbeiten, wobei sich die Ausprägung des Merkmals durch den Anteil nicht erreichter Items bestimmen lässt (Glas & Pimentel, 2008). Test Speededness kann somit als ein auf Personenseite verortetes Merkmal verstanden werden, dessen Ausprägung interindividuell variiert.

Im Kontext von Schulleistungsstudien stellt sich die Frage, wie als Folge von Test Speededness auftretende fehlende Werte im Rahmen von IRT-Skalierungen behandelt werden sollen (Pohl & Carstensen, 2013). Nicht erreichte Items werden dabei zumeist entweder als fehlende Werte oder als Falschantworten kodiert. Der Kodierung als fehlende Werte liegt die Annahme zugrunde, dass Nichterreichen des Testendes unabhängig von Testleistung ist (Lord, 1980). Bei der Behandlung als Falschantworten wird die Zahl nicht erreichter Items neben dem Muster korrekter und falscher Antworten zur Messung des interessierenden Leistungsaspekts einbezogen, wie etwa in PISA zur Schätzung der Personenparameter (z.B. OECD, 2009). Beide Varianten legen eine unterschiedliche Bedeutung von Test Speededness nahe: So impliziert letztere Perspektive, dass Test Speededness systematisch mit der Leistung in den bearbeiteten Items sowie anderen Determinanten der zu messenden Fachleistung assoziiert ist.

Gegenstand des vorliegenden Beitrags ist die Untersuchung von (1) Korrelaten individueller Unterschiede in Test Speededness und (2) Konsequenzen von Speededness für die Evaluation der Zusammenhänge von Testleistung mit Hintergrundvariablen, die in Schulleistungsstudien typischerweise mit Fachleistung assoziiert werden (Schulform, Geschlecht, kognitive Grundfähigkeiten und Fähigkeitsselbstkonzept).

Wir konzentrieren uns auf einen Mathematiktest, der im Rahmen der BIJU-Studie (z.B. Köller, Baumert & Schnabel, 1999) in der siebten Jahrgangsstufe eingesetzt wurde (N=5609). Dieser Test wurde ausgewählt, da er vom Umfang (31 Items) her typisch für Leistungstests in Large-Scale-Studien und zudem deutlich gespeded ist (33 % der Testteilnehmenden erreichten nicht das letzte Testviertel).

Die Auswertungen erfolgen auf Grundlage des von Glas & Pimentel (2008) vorgeschlagenen zweidimensionalen IRT-Modells, das neben Personenfähigkeit den Bearbeitungsumfang, d.h. wie weit der Test bearbeitet wurde, als zweite Dimension beinhaltet. Höhere Werte auf der zweiten Dimension indizieren eine geringere Zahl nicht erreichter Items (einen geringeren Grad an Test Speededness). Das Modell wird nachfolgend um Prädiktoren für Personenfähigkeit und Bearbeitungsumfang erweitert. Dies ermöglicht die Abschätzung der Zusammenhänge von Bearbeitungsumfang mit Testleistung und Hintergrundvariablen und erlaubt zudem, die über den Bearbeitungsumfang vermittelten Effekte auf die Personenfähigkeiten abzuschätzen.

Die Befunde lassen sich wie folgt zusammenfassen: Die Zusammenhänge der Fähigkeitsvariable mit den Hintergrundmerkmalen fielen insgesamt erwartungskonform aus. Mathematikleistung war positiv mit kognitiven Grundfähigkeiten und dem Fähigkeitsselbstkonzept assoziiert; Gymnasialschülerinnen und -schüler sowie Jungen hatten im Mittel höhere Werte auf der Fähigkeitsdimension. Der Bearbeitungsumfang war negativ mit der Personenfähigkeit korreliert ($r = -.24$). Ferner waren Zugehörigkeit zum Gymnasium ($d = -.34$) und Höhe der kognitiven Grundfähigkeiten ($r = -.16$) negativ mit dem Bearbeitungsumfang assoziiert. Der Bearbeitungsumfang der Jungen überschritt denjenigen der Mädchen ($d = .20$), während Fähigkeitsselbstkonzept nicht mit Bearbeitungsumfang korrelierte ($r = .01$, $p > .10$). Trotz der signifikanten Zusammenhänge fielen die über den Bearbeitungsumfang vermittelten Effekte auf die Fähigkeitsvariable gering aus. Eine Konsequenz dieses Ergebnismusters ist, dass das Ausmaß von Test Speededness nur geringfügige Auswirkungen auf die ermittelten Zusammenhänge von Personenfähigkeit und Leistungsdeterminanten hatte.

Die Befunde indizieren, dass Test Speededness in der vorliegenden Studie nicht als Aspekt einer unitären Fähigkeitsdimension verstanden werden sollte, sondern eher Gewissenhaftigkeit individueller Testbearbeitung zu reflektieren scheint. Jedoch ergaben sich im vorliegenden Fall keine nennenswerten Auswirkungen von Test Speededness auf die Schlussfolgerungen hinsichtlich Korrelaten der Personenfähigkeiten. Inwieweit sich die hier berichteten Befunde auch in anderen Studien, Testkonfigurationen und Leistungsdomänen wiederfinden, bleibt zu untersuchen.

Freitag
13. März

11:15 Uhr

GB 03/49

Chair(s)

List, Marit Kristine

Pham, Giang

Unsicherheitsvermeidung als Prädiktor des Missingsanteils von Schülerantworten in internationalen Vergleichsstudien

Theoretischer Hintergrund: Bei internationalen Vergleichsstudien wie PISA, TIMSS, PIRLS etc. variiert der Anteil der ausgelassenen bzw. unzulässigen Antworten (omitted/invalid) stark nach Land. In der PISA Studie 2012 beispielsweise beträgt der durchschnittliche Anteil von fehlenden/unzulässigen Antworten (Missingsanteil) von Schülerinnen und Schülern in Albanien 24,2%, in Österreich 9,7%, in Deutschland 8,4%, während in Singapur nur 3,6%, und in USA nur 2,9%. Mit welchen theoretischen Modellen lassen sich diese Unterschiede erklären? In Frage kommt hier das Kulturdimensionen-Modell von G. Hofstede (1991), jüngstens von Geert Hofstede & Hofstede (2011). Auf Basis umfangreicher Befragungsdaten zu Werteinstellungen in über 70 Ländern hat Hofstede die Basisdimensionen verschiedener Kulturen identifiziert und konzipiert. Eine dieser Dimensionen ist die Unsicherheitsvermeidung (UAI), die sich auf die Toleranz einer Gesellschaft und deren Mitglieder gegenüber Unsicherheit und Unklarheit bezieht. Ausgehend vom Hofstede-Modell lässt sich die Hypothese aufstellen, dass der Missingsanteil in den Ländern mit hoher UAI-Ausprägung aufgrund der Angst vor Fehlern höher, und in den Ländern mit niedriger UAI-Ausprägung (Übergewicht von Hoffnung auf Erfolg gegenüber Angst vor Misserfolg) von Schülerinnen und Schülern geringer ausfällt.

Fragestellung: Es geht um drei Forschungsfragen: (1) Ist der Missingsanteil von Schülerantworten in internationalen Studien ein reliables Maß? Ist also der Missingsanteil von Schülerantworten ein konsistentes Merkmal dieser Länder? (2) Können die unterschiedlichen Missingsanteile durch die unterschiedliche Ausprägungen der Unsicherheitsvermeidung der Länder nach Hofstede vorhergesagt werden? Lassen sich (leistungsunabhängige) Gründe für das international unterschiedliche Schülerverhalten identifizieren? (3) Ist dieser Zusammenhang fächerübergreifend?

Methode: Zur Erfassung der Missingsanteile der Schülerantworten nach Land werden die veröffentlichten statistischen Kennwerten der Items von der internationalen Studie PISA in den Jahren 2000, 2003, 2006, 2009 und 2012 verwendet. Die Anzahl der Items pro Studie variiert zwischen 165 und 219. Für die fachspezifischen Analysen werden die Items nach Fach getrennt betrachtet. Die UAI-Länderwerte stammen aus der IBM-Indextabelle für die Unsicherheitsvermeidung für 66 Länder und Regionen von Geert Hofstede & Hofstede (2011, S. 221-222). In die Auswertungen werden die Daten von allen Ländern einbezogen, für die sowohl Schülerstestdaten als auch UAI-Werte vorhanden sind. Die Analysen finden separat nach Jahr statt. Die Anzahl der Länder in der jeweiligen Analyse pro Jahr bewegt sich zwischen 31 und 41. Als Maß der internen Konsistenz und Reliabilität des Missingsanteils werden der ICC und ICC2 (Bliese, 2000; Raudenbush & Bryk, 2002) berechnet. Um den Zusammenhang zwischen den UAI-Werten und den Missingsanteilen zu untersuchen, werden sowohl die bivariate Korrelation als auch die partielle Korrelation unter Berücksichtigung der mittleren Lösungshäufigkeit der Länder berechnet. Anschließend werden die Ergebnisse aller Jahre zusammengefasst.

Ergebnisse: Die Ergebnisse zeigen, dass mindestens 6% der Gesamtvariation der Missingsanteile der Items auf die Varianz zwischen den Landesmittelwerten zurückführbar sind ($ICC_{max} = 0,15$). Insgesamt erweist sich der mittlere Missingsanteil als ein reliables landesspezifisches Maß ($ICC2 \geq 0,92$). Der hypostasierte Zusammenhang zwischen dem UAI-Wert und dem Missingsanteil wird bestätigt, auch nach Kontrolle der mittleren Lösungshäufigkeit. Je höher der UAI-Wert eines Landes ist, desto höher fällt der mittlere Missingsanteil bei der PISA Studie aus ($r \geq 0,3$). Dieser Zusammenhang gilt für alle Jahre und alle Fächer (Mathematik, Lesen und Naturwissenschaft). Demzufolge können alle drei formulierten Hypothesen als vorläufig bestätigt angesehen werden. Die Ergebnisse deuten darauf hin, dass Länder mit hohen UAI-Werten in den Vergleichsstudien wie PISA benachteiligt werden, wenn die Missings bei der Skalierung ausschließlich als falsch bewertet werden (vgl. Robitzsch, in Druck). Diese Befunde liefern somit ein weiteres Argument für die aktuelle Diskussion in der scientific community bezüglich der Kodierung der Missings beim Large-Scale-Assessment (siehe Pohl, Gräfe, & Rose, 2014, Rohwer, 2013 und Robitzsch, in Druck). Zur Kreuzvalidierung der hier berichteten Ergebnisse werden demnächst auch TIMSS und PIRLS Daten miteinbezogen sowie separate Analysen für unterschiedliche Itemformate.

Freitag
13. März

11:15 Uhr

GB 03/49

Chair(s)

List, Marit Kristine

Frey, Andreas; Spoden, Christian

Eine neue Methode zum Umgang mit fehlenden Werten bei computerbasierten Tests

Theoretischer Hintergrund: Empirisch erhobene Testdaten enthalten zumeist einen substantiellen Anteil fehlender Werte. Der Umgang mit diesen fehlenden Werten ist nicht trivial. Sowohl Ignorieren als auch Umkodieren in falsche Antworten führt oft zu verzerrten Ergebnissen. Deshalb werden komplexere statistische Methoden wie full-information-maximum-likelihood (Arbuckle, 1996), multiple Imputation (Rubin, 1987) oder modellbasierte Ansätze (z.B. Rose, von Davier, & Xu, 2010) zum Umgang mit fehlenden Werten genutzt. Ein Nachteil dieser Verfahren ist, dass sie aufgrund ihrer Komplexität für Anwender undurchsichtig sind und/oder Probleme bei der praktischen Umsetzung resultieren. Mit diesem Beitrag schlagen wir eine leichter zugängliche Alternative vor, die bei computerbasierten Erhebungen verwendet werden kann. Die Methode nutzt Itembearbeitungszeiten. Sie basiert auf der Annahme, dass eine Schwelle der Bearbeitungszeit existiert, die Antworten, die den Prinzipien missing at random (MAR) beziehungsweise missing completely at random (MCAR) folgen, von Antworten trennt, die dem Prinzip missing not at random (MNAR) folgen. Dabei ist das irrtümliche Überspringen eines Items (MAR/MCAR) durch kurze Antwortzeiten charakterisiert, wohingegen das Auslassen eines Items aufgrund mangelnder Fähigkeit (MNAR) mit längeren Antwortzeiten einhergeht, die für das Lesen, Verstehen etc. nötig sind.

Fragestellung: Die neue Methode wird im Beitrag erläutert und anhand von Daten illustriert, die mit einem multidimensionalen computerbasierten Tests erhoben wurden. Konkret wird der Forschungsfrage nachgegangen, welchen Effekt die Anwendung der neuen Methode auf die Schätzung von (a) Itemparametern, (b) die multidimensionale Fähigkeitsverteilung und (c) die Reliabilität hat.

Methode.

Stichprobe und Instrumente: Als Datenbasis wurden die Testdaten von N = 1632 BerufsschülerInnen (46% weiblich; Durchschnittsalter: 21 Jahre) in den Domänen Mathematik (133 Items), Lesen (73 Items) und Naturwissenschaften (133 Items) analysiert. Jede Testperson bearbeitete ein Testheft mit 33 Items. Die Anzahl ausgelassener Items variierte von 0 bis 30 (Md = 3 Items). Nicht erreichte Items am Testende wurden nicht berücksichtigt. Statistisches Vorgehen: Die vorgeschlagene Methode basiert auf einem rekursiven Algorithmus zur Identifikation itemspezifischer Schwellen auf Basis der Punkt-biserialen Korrelation zwischen einer Indikatorvariable für fehlende Itemantworten und der zu messenden Fähigkeit. Bei zufällig ausgelassenen Items auf Basis von MAR/MCAR liegt diese Korrelation nahe dem Wert Null, während bei aufgrund unzureichender Fähigkeit nicht bearbeiteten Items (MNAR) die Korrelation mit der zu messenden Fähigkeit negativ ist und substantiell von Null abweicht. Der Algorithmus identifiziert die Antwortzeit, ab der die Punkt-biserialen Korrelation den Wert Null erreicht. Im Anschluss an die Identifikation dieser Schwelle werden fehlende Itemantworten als zufällig fehlender Wert (MAR/MCAR) beziehungsweise als „falsch“ (MNAR) kodiert. Zur Beantwortung der Forschungsfrage, wurden die Ergebnisse der neuen Methode mit jenen unter Kodierung aller fehlenden Werte als „falsch“ verglichen. Anschließend wurden die Verteilung der geschätzten Itemparameter, die dreidimensionale Fähigkeitsverteilung, sowie die Reliabilitäten analysiert.

Ergebnisse: Die neue Methode identifizierte Schwellen für die Antwortzeiten zwischen 1 und 48 Sekunden, wenn auch – aufgrund der geringen Anzahl fehlender Werte im Datensatz – nur bei 18 Items Schwellen bestimmt wurden. 42 ausgelassene Item-Antworten wurden als MAR/MCAR identifiziert und als zufällig fehlender Wert kodiert. Dies ging mit einer Veränderung der Itemparameter um bis zu 0,18 Logits einher. Varianzen, Kovarianzen und Reliabilitäten unterlagen marginalen Veränderungen. Obwohl die Effekte auf Stichprobenkennwerte gering ausfielen, wurden in Einzelfällen Fähigkeitsunterschiede von bis zu 0,85 Logits beobachtet.

Diskussion: Die vorgeschlagene Methode unterscheidet zufällig fehlende von nicht-zufällig fehlenden Item-Antworten effektiv. Sie ist trotz des rekursiven Algorithmus schnell, einfach anwendbar und auch für Praktiker mit grundlegenden Methodenkenntnissen leicht nachzuvollziehen und anzuwenden. Obwohl im vorliegenden Datensatz nur geringfügige Unterschiede zwischen der neuen und der konventionellen Methode vorlagen, konnten in Einzelfällen deutliche Unterschiede bei der Schätzung von Fähigkeiten identifiziert werden. Dies deutet darauf hin, dass die neue Methode – neben der Berücksichtigung bei der Berechnung von Stichprobenkennwerten – in der Individualdiagnostik eine nützliche Alternative zum undifferenzierten Rekodieren fehlender Werte darstellt.

Freitag
13. März

11:15 Uhr

GB 03/49

Chair(s)

List, Marit Kristine

Zupanic, Michaela; Bauer, Daniel; Ehlers, Jan P.; Fischer, Martin R.; Hofmann, Marzellus
Fehlervermeidung beim Progress Test Medizin durch Nutzung der Antwortoption „Weiss nicht“

Theoretischer Hintergrund: Mit dem formativen Progress Test Medizin (PTM) wird zu Beginn jeden Semesters an 15 deutschsprachigen Universitäten das kumulierte Wissen von Medizinstudierenden mit 200 MC-Fragen auf Absolvierendenniveau erfasst (Nouns & Georg, 2010). Mit der zusätzlich bereitgestellten Antwortoption „weiss nicht“ wird dabei ein qualitativer Aspekt des Lernens, die Konfidenz in das eigene medizinische Wissen berücksichtigt (Barenberg & Dutke, 2011). Dieser Ansatz ermöglicht eine frühe Förderung von Diagnosekompetenz insofern, dass somit Fehler 2. Art vermieden werden können. Die ärztliche Fehlervermeidungskultur gilt insbesondere für Hausärzt/innen, die eine Gatekeeper-Funktion im Gesundheitssystem erfüllen müssen, um die Patientensicherheit zu gewährleisten. Die Inzidenz von Diagnose- und Behandlungsfehlern hängt zuallererst vom ärztlichen Wissen ab, das nach Ende des Studiums im weiteren Berufsverlauf aber abzunehmen scheint (Ramsey et al., 1991).

Fragestellung: Welche Konfidenz haben Hausärzt/innen in das eigene medizinische Wissen, gemessen mit dem Progress Test Medizin?

Methode: Bei Fortbildungen oder Kongressen in 2011 nahmen 161 Hausärzt/innen (52 weiblich, 109 männlich) im durchschnittlichen Alter von 51.0 ± 8.4 Jahren (Median 52 J., Minimum 31 J., Maximum 67 J.) an der Studie teil. Das medizinische Wissen wurde mit 60 MC-Fragen erfasst, die unter Berücksichtigung von Itemschwierigkeit und Trennschärfe (Referenz: PJ-Studierende) sowie des Test-Blueprints aus dem Progress Test Medizin 21 ausgewählt wurden. Nach Altersmediandichotomisierung der Stichprobe wurden die Leistungen in den vier Subgruppen weiblich / jünger (N = 28), männlich / jünger (N = 51), weiblich / älter (N = 24) und männlich / älter (N = 58) analysiert, unter Verwendung der vier PTM Scores „richtige“, „falsche“, „total“ (richtige minus falsche) und „weiss nicht“. Mit einem Fragebogen wurden soziodemografische, ausbildungs- und berufsbezogene Daten einverständlich erfasst.

Ergebnisse: Die Hausärzt/innen erzielten ein gutes Gesamtergebnis mit einem PTM-Score „total“ von 43.5 ± 8.0 (Min. 11, Max. 60) bei durchschnittlich 50.0 ± 5.1 „richtigen“ Antworten. Frauen gaben insgesamt deutlich häufiger „weiss nicht“ an (4.7 ± 3.6) als Männer (3.0 ± 3.6 ; Mann-Whitney-U-Test: $U = 1864$, $p = .000$), bei einer vergleichbaren Fehlerzahl von 6.5 ± 3.7 ($U = 2644$, n. s.). Jüngere und ältere Hausärzt/innen gaben insgesamt gleich häufig „weiss nicht“ an (3.6 ± 3.6), unterschieden sich aber im PTM-Score „falsche“ mit 5.5 ± 3.2 bei den jüngeren und 7.4 ± 3.9 bei den älteren signifikant ($U = 2171$, $p = .000$). Deutliche Unterschiede zeigten sich für die vier Subgruppen in allen PTM-Scores und bei vertiefenden Analysen mit dem PTM-Score „falsche“ in den klinischen Fächern Innere Medizin (Kruskall-Wallis-Test: $\chi^2 = 10.6$, $p = .014$), mit der höchsten Fehlerzahl der Subgruppe weiblich / älter (2.3 ± 1.6) und Neurologie ($\chi^2 = 10.4$, $p = .015$), mit den meisten Fehlern der Subgruppe männlich / älter (1.1 ± 0.9). Zugleich fanden sich jedoch die wenigsten Angaben im PTM-Score „weiss nicht“ für das Fach Neurologie in der Subgruppe männlich / älter (0.4 ± 0.6 ; $\chi^2 = 9.5$, $p = .023$).

Diskussion: Bei der hier untersuchten Stichprobe zeigt sich der zu erwartende Rückgang des medizinischen Wissens im Berufsverlauf mit größerem Abstand zum Examen und fortschreitendem Alter. Gleichwohl muss von einem selection bias ausgegangen werden, da die Hausärzt/innen im Durchschnitt ohne weitere Testvorbereitung einen PTM-Score „total“ erreichen, der einer Leistung von 72 % entspricht. Die Konfidenz in das eigene medizinische Wissen scheint mit dem Geschlecht zu variieren und ist bei Männern unabhängig vom Alter stärker ausgeprägt. Die Nutzung der „weiss nicht“ Antwortoption im Progress Test Medizin wird in folgenden Untersuchungen unter Berücksichtigung der Aufgabencharakteristika, wie z. B. die medizinischen Fächer, analysiert.

Freitag
13. März

11:15 Uhr

GB 03/49

Chair(s)

List, Marit Kristine

Krampen, Dorothea; Hartig, Johannes; Gold, Andreas

Aufgabenstamm- und Antwortabhängigkeiten im C-Test

Insbesondere in der Bildungsforschung werden häufig Tests eingesetzt, die sich aus sogenannten Testlets zusammensetzen. Als Testlet werden mehrere Items bezeichnet, die sich auf einen gemeinsamen Aufgabenstamm beziehen und gemeinsam administriert werden (Wainer & Kiely, 1987). Das Testletformat stellt eine Herausforderung für moderne psychometrische Modelle dar, da es deren Grundannahme der lokalen stochastischen Unabhängigkeit verletzen kann. Abhängigkeiten zwischen Antworten auf Items innerhalb eines Testlets, sogenannte lokale Abhängigkeiten, sind ein regelmäßiges Phänomen. Zwei Formen lokaler Abhängigkeiten können dabei unterschieden werden (Marais & Andrich, 2008a, b): 1) Abhängigkeiten, die auf den gemeinsamen Aufgabenstamm zurückgehen (Aufgabenstammabhängigkeiten) und 2) spezifische Effekte des Lösens einzelner Items auf das Lösen darauf folgender Items (Antwortabhängigkeiten). Werden vorhandene Abhängigkeiten bei der Testauswertung ignoriert, kann dies unter anderem zu verzerrten Parameterschätzungen und Teststatistiken führen. Es wird ein neuartiger Ansatz zur Modellierung der beiden Formen lokaler Abhängigkeiten sowie ihrer Mischform im Rahmenwerk der Item-Response-Theorie vorgestellt (vgl. auch Hartig, 2013). Das Testletmodell zur Berücksichtigung von Aufgabenstammabhängigkeiten definiert einen Generalfaktor sowie zusätzlich testletspezifische Faktoren, die von dem Generalfaktor sowie allen übrigen testletspezifischen Faktoren unabhängig sind. Das Response-Dependence-Modell zur Berücksichtigung von Antwortabhängigkeiten definiert neben der allgemeinen latenten Personenfähigkeit autoregressive Zusammenhänge zwischen je zwei aufeinander folgenden Items innerhalb eines Testlets. Das Mischmodell kombiniert die Annahmen des Testletmodells mit denen des Response-Dependence-Modells. Die drei Modelle werden im Zuge einer systematischen empirischen Anwendung auf Daten des testletbasierten, lückentextähnlichen C-Tests (z. B. Klein-Braley, 1985) zur Erfassung der allgemeinen sprachlichen Kompetenz miteinander verglichen. Die Untersuchungsgrundlage bilden Daten aus der nationalen Schulleistungsstudie Deutsch Englisch Schülerleistungen International (DESI; DESI-Konsortium, 2008) von $n = 9\,816$ Schülerinnen und Schülern der neunten Jahrgangsstufe im allgemeinbildenden Schulwesen. Für alle Modelle werden im Rahmen ihrer empirischen Anwendung bayesianische Methoden der Parameterschätzung genutzt, mit einem Probitlink zur Verknüpfung der Modellvariablen. Die übergeordnete Zielsetzung des Beitrags ist demnach die Unterscheidung und angemessene psychometrische Modellierung der verschiedenen Formen lokaler Abhängigkeiten in testletbasierten Tests. Eine Analyse der Daten des C-Tests mittels der vorgestellten Modelle soll die Fragen beantworten, ob eine Entscheidung zwischen Aufgabenstamm- und Antwortabhängigkeiten möglich ist und ob beide Formen von Abhängigkeiten auch bei simultanem Vorliegen noch voneinander getrennt werden können. Es wird davon ausgegangen, dass im C-Test – aufgrund seiner spezifischen Beschaffenheit – vornehmlich Antwortabhängigkeiten oder Mischformen aus Aufgabenstamm- und Antwortabhängigkeiten auftreten und damit das Response-Dependence-Modell beziehungsweise das Mischmodell im Hinblick auf die Daten eine vergleichsweise bessere Passung aufweisen sollte. Deskriptive Ergebnisse auf Basis von Residualkorrelationen (Yen, 1993) zeigen zunächst, dass sowohl Aufgabenstamm- als auch Antwortabhängigkeiten in geringer bis moderater Höhe im C-Test vorhanden sind, wobei Antwortabhängigkeiten etwas stärker ausgeprägt scheinen. Die Betrachtung der durchweg substantiellen Modellparameter sowie Modellvergleiche anhand eines reduzierten Datensatzes legen dar, dass beide Formen lokaler Abhängigkeiten – auch bei simultanem Vorliegen – empirisch voneinander getrennt werden können. Das Mischmodell weist, gefolgt vom Response-Dependence-Modell, nach Beurteilung der 95 %-Konfidenzintervalle der Differenzen zwischen beobachteten und replizierten χ^2 -Werten die beste Passung auf die Daten auf und spiegelt die (zuweilen heterogene) Struktur der existierenden Abhängigkeiten adäquat wider. Ausgehend von den Ergebnissen werden Implikationen für die Testentwicklung und Testauswertung diskutiert sowie ausgewählte Abhängigkeiten einer tiefergehenden inhaltlichen Analyse unterzogen. Ein besonderes Augenmerk gilt zudem technischen Einschränkungen.

Saalbach, Henrik

Unterstützung individueller Lernprozess in der frühen naturwissenschaftlichen Bildung

Dieses Symposium geht der Frage nach, wie naturwissenschaftliches Lernen von Kindern im Vorschul- und Grundschulalter effektiv unterstützt werden kann. Als theoretischer Ausgangspunkt dienen Erkenntnisse zur Entwicklung einer Scientific Literacy, deren Einflussfaktoren und Einschränkungen bei jungen Kindern. Pädagogischen Fachkräften und Lehrpersonen sollten die Möglichkeiten und Einschränkungen des Lernens junger Kinder vertraut sein, und sie sollten über Kompetenzen verfügen, diese in ihrer Unterrichtsplanung zu berücksichtigen. Dies sollte sich vor allem in angemessenen Unterstützungsmaßnahmen innerhalb naturwissenschaftlicher Lernumgebungen spiegeln. Das interdisziplinär zusammengesetzte Symposium beleuchtet verschiedene Aspekte der Unterstützung individueller Lernprozesse aus entsprechend unterschiedlichen Perspektiven:

- (1) Aus einer psychologischen Perspektive nähern sich die Autoren des ersten Beitrags der Frage, welche kognitiven Voraussetzungen junge Kinder mitbringen und welche Möglichkeiten der Unterstützung sich darauf beziehend anbieten. Untersucht werden Ressourcen und mögliche Nachteile bilingualer Kinder beim komplexen Problemlösen. Die Ergebnisse zeigen einen Vorteil bilingualer Kinder in exekutiven Prozessen, wie sie in komplexen Problemlöseaufgaben benötigt werden. Dieser Vorteil wird aber verdeckt durch schlechtere Sprachkompetenzen. Dies wirft wichtige Fragen auf bzgl. gezielter Förderung naturwissenschaftlicher Kompetenzen bei bilingualen Kindern.
- (2) Aus einer fachdidaktischen Perspektive betrachten die Autoren der zweiten Studie das Handeln und die Kompetenzen von Kindergartenlehrpersonen. Sie untersuchen, mit welchen Äußerungen Kindergartenlehrpersonen naturwissenschaftliches Lernen von Kindern unterstützen und in welcher Beziehung die Äußerungen mit professionellem Wissen und Vorstellungen stehen. Die Befunde zeigen, dass die Fachsprache der Kindergartenlehrpersonen mit ihrem Fachwissen sowie die von ihnen vorgenommene Anregung des Konzeptwandels mit konstruktivistischen Vorstellungen zusammenhängen.
- (3) Ebenfalls eine vor allem fachdidaktische Perspektive vertritt der dritte Beitrag mit einer Studie zum Schlussfolgern durch Kindergarten- und Grundschulkindern im naturwissenschaftlichen Kontext. Die Autoren untersuchten darin, inwiefern prozessbegleitende Scaffolds mit unterschiedlichem Unterstützungsgrad angemessene Schlussfolgerungen bei den Kindern fördern. Die Autoren fanden, dass eine Förderung hypothesenbezogener Schlussfolgerungen bereits bei beiden Gruppen von Kindern möglich ist. Jedoch ergeben sich altersabhängig unterschiedliche Implikationen hinsichtlich des Unterstützungsgrades. Ferner zeigte sich, dass selbst bei einer Rücknahme der Unterstützung ein Fördereffekt erhalten bleibt.
- (4) Eine Erziehungswissenschaftliche Perspektive nimmt schließlich der vierte Beitrag ein, der auf die lernwirksame Gestaltung von naturwissenschaftlichem Grundschulunterricht fokussiert. Dabei geht es um die quantitative Evaluation eines Unterrichts, der auf Grundlage empirischer Befunde zu Gelingensbedingungen naturwissenschaftlichen Lernens entwickelt wurde. Die Autoren konnten eine deutlich bessere Lernentwicklung von Kindern im Modellunterricht im Vergleich zu Kindern im herkömmlichen Unterricht im Hinblick auf die Kriterien Lernerfolg, Interesse und Selbstkonzept feststellen.

Freitag
13. März

11:15 Uhr

HGB 20

Chair(s)

Saalbach, Henrik

Unterstützung individueller Lernprozess in der frühen naturwissenschaftlichen Bildung

Gunzenhauser, Catherine; Saaalbach, Henrik; Karbach, Julia

Ressourcen und Ansätze zur Unterstützung des komplexen Problemlösens bei bilingualen Grundschulkindern

Theoretischer Hintergrund: Zweisprachige Schülerinnen und Schüler mit Migrationshintergrund zeigen bereits während der Grundschulzeit schwächere Leistungen als monolinguale Gleichaltrige im naturwissenschaftlichen Lernbereich (Bos, Wendt, Köller, & Selter, 2012). Voraussetzung für eine effiziente Unterstützung der individuellen Lernprozesse dieser Kinder ist ein umfassendes Verständnis ihres komplexen Problemlösens, da diese Kompetenz als zentral für das naturwissenschaftliche Lernen betrachtet wird (Bos et al., 2012). Dabei ist zu beachten, dass die Leistung in komplexen Problemlöseaufgaben sowohl von exekutiven Funktionen als auch von sprachlichen Kompetenzen deutlich beeinflusst wird (Cheetham, Rahm, Kaller, & Unterrainer, 2012). Eine mögliche Ressource zur Unterstützung findet sich in den exekutiven Kontrollfunktionen: In mehreren Studien hat sich gezeigt, dass bilinguale Kinder hier Vorteile haben könnten (Greenberg, Bellana, & Bialystok, 2013). Sprachliche Prozesse könnten insbesondere durch (subvokale) verbale Selbstinstruktionen Einfluss auf das Problemlösen ausüben (Lidstone, Meins, & Fernyhough, 2010). Da bilinguale Kinder jedoch häufig defizitäre sprachliche Kompetenzen aufweisen, sind ihre möglichen Vorteile beim Lösen komplexer Problemlöseaufgaben aufgrund von besserer exekutiver Kontrolle schwer von ihren möglichen Nachteilen aufgrund von schwächeren Sprachkompetenzen zu entflechten (siehe Bialystok, Luk, Peets, & Yang, 2010).

Fragestellung: Ziel der vorliegenden Studie war daher die Untersuchung der Leistung bilingualer Grundschul Kinder in komplexen Problemlöseaufgaben unter Kontrolle sprachlicher Einflüsse. Dabei wurden zwei Ansätze verfolgt: Zum einen wurden die Deutschkenntnisse der Kinder in einem korrelativen Design kontrolliert. Zum anderen wurden die Kinder durch auditorische Interferenz daran gehindert, Selbstinstruktionen zu benutzen.

Unsere Annahme war zunächst, dass sich monolinguale Kinder und bilinguale Kinder in ihren Leistungen in komplexen Problemlöseaufgaben ohne Interferenz nicht unterscheiden würden. Bei statistischer Kontrolle der Deutschkenntnisse und der Verhinderung von Selbstinstruktionen erwarteten wir jedoch einen Vorteil der bilingualen Kinder.

Methode: Die Stichprobe bestand aus N = 53 monolingualen und N = 44 bilingualen Grundschulkindern aus dem Saarland, die sich hinsichtlich ihres Alters, ihrer durchschnittlichen fluiden Intelligenz und des sozioökonomischen Status ihrer Familien nicht unterschieden. In der bilingualen Gruppe befanden sich jedoch signifikant mehr Mädchen. Die Kinder waren im Durchschnitt 8.90 Jahre alt (SD = 0.77 Jahre). Als komplexe Problemlöseaufgabe wurde die Tower of London-Aufgabe (TOL) benutzt. Die Leistung in der TOL-Aufgabe ist von Sprachkompetenzen abhängig, und Grundschul Kinder verwenden während der Vorausplanung ihrer Spielzüge spontan verbale Selbstinstruktionen (Lidstone et al., 2010). Die Kinder lösten jeweils vier TOL-Items unter zwei Bedingungen (vgl. Lidstone et al., 2010): In der regulären Bedingung sagten die Kinder zuerst, wie viele Spielzüge sie benötigen würden und bauten dann ihre Türme. In der Interferenz-Bedingung sollten die Kinder jede Sekunde das Wort „Montag“ aussprechen, während sie ihre Spielzüge planten. Ein Item wurde als korrekt bewertet, wenn die Kinder nach der Planung die richtige Anzahl von Spielzügen nannten und diese Anzahl von Zügen dann während des Bauens des Turmes auch benötigten.

Die Deutschkenntnisse der Kinder wurden mit Kurzversionen des Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) und des Test for the Reception of Grammar (Bishop, 2003) gemessen.

Ergebnisse: Wie erwartet zeigten die monolingualen Kinder bessere Deutschkenntnisse als die bilingualen Kinder. Unter Kontrolle von Geschlecht und Alter unterschieden sich monolinguale und bilinguale Kinder nicht in ihrer Leistung in der regulären TOL-Bedingung. In einer hierarchischen Regression, in der Geschlecht, Alter und Deutschkenntnisse kontrolliert wurden, sagte Zweisprachigkeit jedoch signifikant und positiv die Leistung der regulären TOL-Bedingung vorher. Eine ANCOVA mit den Kovariaten Geschlecht und Alter zeigte, dass die bilingualen Kinder die monolingualen Kinder in ihrer Leistung in der Interferenz-Bedingung übertrafen.

Die Ergebnisse stützen die These, dass bilinguale Kinder einen Vorteil in exekutiven Prozessen haben, wie sie in komplexen Problemlöseaufgaben benötigt werden. Dieser Vorteil wird aber verdeckt durch Ihre schlechteren Sprachkompetenzen. Implikationen für die Förderung naturwissenschaftlicher Kompetenzen bei bilingualen Kindern werden diskutiert.

Studhalter, Ueli; Elmer, Anneliese; Tettenborn, Annette; Leuchter, Miriam; Saalbach, Henrik

Strategien der verbalen Lernunterstützung im frühen naturwissenschaftlichen Unterricht und deren Zusammenhang mit professionellen Kompetenzen der Kindergartenlehrpersonen

Im Rahmen eines sozial-konstruktivistischen Lernverständnisses gelten verbale Unterstützungsmassnahmen zur kognitiven Aktivierung als wichtiger Faktor für gelingende Lernprozesse. Solche Strategien der Lernunterstützung beinhalten beispielweise die Klärung von Ziel- und Aufgabenzusammenhang, Vereinfachungen, gezielte Fragen und das Anregen von Vergleichen (Eshach, Dor-Ziderman, & Arbel, 2011; Gentner & Namy, 1999; Saalbach & Schalk, 2011). Das Anbieten adäquater verbaler Lernunterstützung bedingt einen Rückgriff auf fachliches und fachdidaktisches Wissen, was in Modellen der professionellen Kompetenz von Lehrpersonen zum Ausdruck kommt (Baumert & Kunter, 2006; Magnusson, Krajcik, & Borko, 1999; Shulman, 1987). Das Vorhandensein einer fundierten fachlichen Wissensbasis ist jedoch, gerade für Kindergartenlehrpersonen, betrachtet man deren Ausbildungsstrukturen und -inhalte, keinesfalls gegeben (vgl. Appleton, 2003, 2008).

In der vorliegenden Studie wurde verbale Lernunterstützung von Kindergartenlehrpersonen in einer Lernumgebung zum Thema „Schwimmen / Sinken“ untersucht. Es wurde folgenden Fragen nachgegangen: (1) Wie oft verwenden Kindergartenlehrpersonen Fachsprache? Es wird vermutet, dass Kindergartenlehrpersonen eher selten explizite Fachsprache verwenden. (2) Wie oft wenden Kindergartenlehrpersonen spezifische Strategien der verbalen Lernunterstützung an? Es wird vermutet, dass Kindergartenlehrpersonen vor allem diejenigen Strategien der verbalen Lernunterstützung oft zeigen, welche auf Anleitung und Erklärung der Aufgabe abzielen. (3) Wie hängen die fachspezifisch-pädagogischen Vorstellungen, das Fachwissen, das fachspezifische Interesse und die Berufserfahrung mit der verbalen Lernunterstützung zusammen? Es wird vermutet, dass sozial-konstruktivistisch basierte Vorstellungen der Lehrpersonen positiv mit denjenigen Strategien der verbalen Lernunterstützung korrelieren, welche die Kinder zum Überlegen anregen und den Konzeptwandel herausfordern. Zudem wird erwartet, dass höheres Fachwissen auch mit einer deutlicher ausgeprägten Fachsprache einhergeht.

Eine Stichprobe von N = 33 weiblichen Kindergartenlehrpersonen erhielt in einer Weiterbildung eine Einführung in eine strukturierte Lernumgebung im Themenbereich ‚Schwimmen/Sinken‘. Die Lehrpersonen wurden anschliessend bei der Arbeit mit den Kindern an dieser Lernumgebung jeweils während einer 40-minütigen Sequenz videografiert. Zur Messung der verbalen Lernunterstützung wurde ein elaboriertes Kodierschema im Time-Sampling-Verfahren (total 4403 Slots; jeweils 15 Sekunden) angewendet (IRR: Cohen's Kappa = .78). Das Kodierschema beinhaltet sowohl die inhaltsspezifische Fachsprache, wie auch die inhaltsspezifische verbale Lernunterstützung, welche durch folgende vier Bereiche abgebildet wird: (A) Phänomen/Aufgabe/Vorgehen klären, (B) Aufmerksamkeit fokussieren, (C) Vorwissen aktivieren und (D) Überlegungen und Konzeptwandel herausfordern. Für die Fachsprache und die Lernunterstützung wurden prozentuale Anteile (zutreffende Slots/ total analysierte Slots) bestimmt und Mittelwerte über die gesamte Stichprobe berechnet.

Die fachspezifisch-pädagogischen Vorstellungen wurden mittels einem in einer Vorstudie entwickelten Fragebogen anhand von folgenden sechs Skalen erhoben (5-Punkte-Likert: 1=stimmt gar nicht, 5=stimmt völlig): (1) Lernen als Konzeptwandel, (2) Entwicklung eigener Deutungen, (3) Transmission, (4) Offene Handlungsorientierung, (5) Lernen durch Spiel und (6) Motiviertes Lernen. In diesem Fragebogen wurde u.a. auch das Fachwissen erfasst.

Die Kindergartenlehrpersonen verwenden erwartungskonform eher selten inhaltsspezifische Fachsprache (M= 11.9, SD= 9.1). Die verbale Lernunterstützung zeigt sich wie folgt: (A) Phänomen/Aufgabe/Vorgehen klären: 39.3 (8.4), (B) Aufmerksamkeit fokussieren: 14.2 (6.8), (C) Vorwissen aktivieren: 12.3 (5.4), und (D) Überlegungen und Konzeptwandel herausfordern: 14.5 (9.0). Folglich kann bestätigt werden, dass die Anleitung und Klärung der Aufgabe einen grossen Stellenwert einnimmt.

Bei zwei von sieben Skalen der fachspezifischen-pädagogischen Vorstellungen wurden signifikante Korrelationen mit einzelnen Bereichen der verbalen Lernunterstützung identifiziert: Positive Korrelation zwischen ‚Entwicklung eigener Deutungen‘ und ‚Überlegungen und Konzeptwandel herausfordern‘ ($r = .45, p = 0.01$), negative Korrelation zwischen ‚Transmission‘ und ‚Vorwissen aktivieren‘ ($r = -.36, p = 0.04$). Die Hypothese zur Verbindung von sozial-konstruktivistischen Vorstellungen mit spezifischeren Formen der Lernunterstützung kann teilweise bestätigt werden.

Weiter wurde erwartungskonform eine positive Korrelation zwischen Fachwissen und der im Unterricht gezeigten inhaltsspezifischen Fachsprache gefunden ($r = .39, p = 0.03$). Jedoch konnte keine Korrelation zwischen der Fachsprache und der verbalen Lernunterstützung gefunden werden.

In dieser Studie konnten Verbindungen zwischen Strategien der verbalen Lernunterstützung und den professionellen Kompetenzen aufgezeigt werden. Die Bedeutung des Fachwissens für den frühen naturwissenschaftlichen Unterricht konnte bestätigt werden (Leuchter & Saalbach, 2014).

Freitag
13. März

11:15 Uhr

HGB 20

Chair(s)
Saalbach, Henrik

Unterstützung individueller Lernprozess in der frühen naturwissenschaftlichen Bildung

Robisch, Christin; Stephan-Gramberg, Simone; Tröbst, Steffen; Möller, Kornelia; Hardy, Ilonca

Förderung hypothesenbezogener Schlussfolgerungen im naturwissenschaftlichen Kontext bei Kindergarten- und Grundschulkindern

Kognitionspsychologisch gilt das schlussfolgernde Denken als bedeutsam für die Prüfung von Hypothesen (Moshman, 1979). Das hypothesenbezogene Schlussfolgern ist definiert als die Evaluation gegebener Zustände in Bezug auf die Aussagekraft über den Wahrheitsgehalt einer Hypothese (Vergauwe, Gauffroy, Mosanyi, Dagry & Barrouillet, 2013). Schlussfolgerungen sind dann adäquat, wenn bekräftigende, widerlegende und für die Prüfung irrelevante Ereignisse korrekt unterschieden werden. Ein theoretisches Erklärungsmodell bietet die dual-process mental models theory des konditionalen Schlussfolgerns, welche das Schlussfolgern mit der Bildung entsprechender mentaler Modelle im Arbeitsgedächtnis erklärt (Gauffroy & Barrouillet, 2009). Insbesondere die adäquate Evaluation irrelevanter Ereignisse stellt dabei eine Herausforderung (nicht nur) für Kinder dar, wie sowohl kognitionspsychologische (Barrouillet, Gauffroy & Lecas, 2008; Gauffroy & Barrouillet, 2011; English, 1997) als auch fachdidaktische (Tröbst, Hardy & Möller, 2011) Befunde belegen.

Dennoch werden adäquate Schlüsse bei der Prüfung von Hypothesen von Kindern bereits in Kontexten des frühen naturwissenschaftlichen Lernens gefordert (OECD, 2007). Bisherige Studien zur Förderung der Schlussfolgerungsfähigkeit beschränkten sich allerdings primär auf Erwachsene und vernachlässigten naturwissenschaftliche Kontexte (z. B. English, 1997; Klauer, Stegmaier & Meiser, 1997). Unter Berücksichtigung der Bedingungsfaktoren adäquater Schlüsse – Arbeitsgedächtnis (Klauer et al., 1997), Inhibitionsfähigkeit (Handley, Capon, Beveridge, Dennis & Evans, 2004), Inhalt und Kontext (Thompson, 1994) sowie Wissenschaftsverständnis (Hardy, Stephan-Gramberg, Robisch, Tröbst, Möller, 2013) – wurden daher zwei Fördermaßnahmen entwickelt, um hypothesenbezogene Schlussfolgerungen bei Kindern in einem naturwissenschaftlichen Kontext zu fördern.

In einer experimentellen Laborstudie im mikrogenetischen Design (Siegler & Crowley, 1991) mit $N = 101$ DrittklässlerInnen wurde unter der exemplarischen Frage „Wie kommt es, dass ein Ball springt?“ (Thiel, 1987) untersucht, inwiefern (1) prozessbegleitende Scaffolds mit unterschiedlichem Unterstützungsgrad (KG = keine Scaffolds, EG 1 = aufmerksamkeitslenkende Scaffolds, EG 2 = aufmerksamkeitslenkende und modellierende Scaffolds) angemessene Schlussfolgerungen fördern und (2) ob bei einer Rücknahme der Unterstützung der etwaige Fördereffekt erhalten bleibt. Analog hierzu wurde eine Trainingsstudie mit $N = 63$ Kindergartenkindern ($M = 5,9$) durchgeführt.

Varianzanalytisch wurde für die Grundschul Kinder ein (1) signifikanter Haupteffekt des Faktors Bedingung [$F(2, 98) = 49.27, p < .001, \eta_p^2 = .50$] identifiziert. Kontrastanalysen spezifizierten den Befund, indem eine signifikante Überlegenheit der EG 1 [$p < .001; r = .73$] und der EG 2 [$p < .001; r = .77$] gegenüber der KG hinsichtlich der angemessenen Evaluation von Ereignissen nachgewiesen wurde, während sich die EG 1 und die EG 2 nicht überzufällig [$p = .352$] unterschieden. Zudem wurde (2) der erwartete Haupteffekt des Faktors Bedingung [$F(2, 98) = 8.33, p < .001, \eta_p^2 = .15$], bei vollkommener Rücknahme der Unterstützung, nachgewiesen. Kontrastanalysen ergaben, dass die zuvor geförderten EG 1 [$p < .001, r = .46$] und EG 2 [$p < .001, r = .38$] der nicht geförderten KG auch ohne Unterstützung hinsichtlich angemessener Schlussfolgerungen signifikant überlegen waren. Signifikante Unterschiede zwischen den Experimentalgruppen wurden nicht gefunden [$p = .463$]. Die Ergebnisse für die Kindergarten Kinder zeigen hingegen eine signifikante Überlegenheit der EG 2 gegenüber der EG 1 und KG (MANOVA $F(4, 55) = 2.02; p > .05; \epsilon_2 = .02$), insbesondere für die Evaluation von irrelevanten Ereignissen.

Eine Förderung angemessener Schlussfolgerungen im Kontext „Wie kommt es, dass ein Ball springt?“ ist folglich bereits bei Vor- und Grundschulkindern möglich; jedoch ergeben sich altersabhängig unterschiedliche Implikationen hinsichtlich der zusätzlichen Wirksamkeit modellierender zusätzlich zu aufmerksamkeitslenkenden Scaffolds. In weiteren Studien sollen die Fördermaßnahmen in einen Unterricht für ganze Schulklassen bzw. Kindergartengruppen implementiert werden.

Hartinger, Andreas; Grygier, Patricia; Kullmann, Harry; Tretter, Tobias; Ziegler, Florian

Lernen auf individuellen Lernwegen im Sachunterricht – Befunde aus dem Modellversuch GribS

Im Vortrag sollen die zentralen Befunde der Evaluation des Modellversuchs GribS („Grundschulen zur individuellen Förderung bayerischer Schülerinnen und Schüler“) berichtet werden. Dieser Modellversuch fand zwischen 2006 und 2011 statt; er wurde von der Stiftung Bildungspakt Bayern initiiert und an der Universität Augsburg wissenschaftlich begleitet.

In der Auseinandersetzung mit empirischen Befunden zu Gelingensbedingungen naturwissenschaftlichen Lernens (vgl. zusammenfassend z.B. Labudde & Möller, 2012; Minner, Levy & Century, 2011) und mit den zu dieser Zeit vorliegenden Ergebnissen zur individuellen Förderung im Sachunterricht (z.B. Blumberg, 2008) wurde ausgehend von den unterschiedlichen Vorerfahrungen der Kinder Unterricht zu ausgewählten Themen entwickelt, um individuelle Lernwege zu ermöglichen (vgl. z.B. Ziegler, Hartinger, Grygier & Lange, 2014). Er beinhaltet die folgenden drei Schritte:

- Explizierung der Vorerfahrungen und Präkonzepte der Kinder (z.B. durch die Begegnung mit erstaunlichen Phänomenen)
- Selbstständiges und eigenaktives Arbeiten an einer Fragestellung (z.B. durch die Bearbeitung von Forscheraufträgen)
- Bewährung oder Veränderung der subjektiv stimmigen Erklärungen im sozialen Austausch (z.B. durch Kreisgespräche, in denen die Erkenntnisse aus der eigenständigen Arbeit besprochen werden)

In einem Prä- Post- Follow-Up-Design, in die die Daten von über 1000 Schüler(inne)n aus GribS-Schulen und (von den jeweiligen Schulämtern als vergleichbar ausgewählten) Kontrollschulen einfließen, wurde der Modellversuch nach den Kriterien Lernerfolg, Interesse und Selbstkonzept evaluiert. Ziel war neben der Feststellung von Haupteffekten auch die Frage, inwieweit differenzielle Effekte bei unterschiedlichen Schüler(innen)gruppen (z.B. Mädchen/ Jungen; Kinder mit unterschiedlichem Vorwissen oder einer anderen Erstsprache als Deutsch...) auftreten.

Im Vortrag werden Befunde zu folgenden Fragen berichtet:

- Unterscheiden sich die Lernergebnisse der Kinder in den GribS-Schulen von denen in den Vergleichsschulen?
- Gibt es Unterschiede bezüglich der Entwicklung des Interesses an Naturwissenschaften und des Selbstkonzepts in den beiden Gruppen?
- Werden im GribS-Unterricht Kinder mit unterschiedlichen Lernvoraussetzungen gleichermaßen gefördert und/oder gelingt eine Reduzierung von Unterschieden?

Die Befunde, die durch die Berechnungen hierarchischer Regressionsanalysen gewonnen wurden, zeigen, a) dass es substantielle Unterschiede in der Lernentwicklung zwischen den verschiedenen Gruppen zu Gunsten der GribS-Klassen gibt, b) dass diese Unterschiede im Verlauf des Modellversuchs ständig größer wurden sowie c) dass keine differenziellen Effekte zu finden waren. Es profitierte keine Schülerteilgruppe innerhalb der Modellschulen mehr (oder weniger) als die gleiche Teilgruppe innerhalb der Vergleichsschulen.

Zu diskutieren sind die Befunde insbesondere auf der Grundlage neuerer Ergebnisse zur individuellen Förderung im naturwissenschaftlichen Lernen, z.B. bezogen auf die Frage, inwieweit bestimmte Teilprozesse des Unterrichts für die individuelle Förderung und die auf sie zurückzuführenden Ergebnisse in spezifischer Weise von Bedeutung sind.

Ohle, Annika

"Lehrkräfte, Unterrichtsqualität & Lernerfolge im Kontext von Videostudien – Perspektiven aus Didaktik, Erziehungswissenschaften und Psychologie"

Zur empirischen Beschreibung von Lehr-Lernprozessen werden häufig Modelle herangezogen, die neben individuellen Voraussetzungen der Lernenden auch Lehrkraftkompetenzen und Unterrichtsqualität in den Blick nehmen. Dabei haben sich insbesondere drei Merkmalsbereiche von Unterrichtsqualitätsmerkmalen als bedeutsam für Lernerfolge von Schülerinnen und Schülern (SuS) erwiesen: a) kognitive Aktivierung, b) Klassenführung, Regelklarheit und Struktur und c) unterstützendes, schülerorientiertes Unterrichtsklima (Klieme & Rakoczy, 2008). Die Erfassung dieser Merkmale im Unterricht ist jedoch perspektivenabhängig, da Lehrende und Lernende das Geschehen im Klassenraum unterschiedlich wahrnehmen können. Besonders für schwer beobachtbare Prozesse auf der Tiefenstruktur des Unterrichts hat sich die objektive Perspektive der Videobeobachtung als zielführend erwiesen (Clausen, 2002).

Im Rahmen dieses Symposiums werden vier Studien aus den Bereichen Mathematik und Sachunterricht vorgestellt, in denen Merkmale tiefenstruktureller Unterrichtsqualität mittels Videobeobachtung erfasst werden. Drei der Beiträge beschreiben darüber hinaus Zusammenhänge dieser Merkmale zu Lehrkraftkompetenzen und/ oder Lernerfolgen von SuS in Grundschule und Sekundarstufe I.

Im ersten Beitrag wird eine quasi-experimentelle Videostudie zur Implementation von Bildungsstandards bezüglich prozessbezogener Kompetenzen im Mathematikunterricht der 4. Jahrgangsstufe im Rahmen des SINUS-Programms vorgestellt. Dabei soll geklärt werden, welche Bereiche der prozessbezogenen Kompetenzen im Unterricht gefördert werden und ob sich die SINUS-Klassen hinsichtlich der Förderung von einer Kontrollgruppe unterscheiden. Die Ergebnisse weisen darauf hin, dass nicht alle Bereiche gleichermaßen gefördert werden. Der Gruppenvergleich zeigt Unterschiede im Ausmaß der Förderung prozessbezogener Kompetenzen.

Die zweite Studie untersucht Zusammenhänge zwischen der Qualität von Unterrichtsgesprächen und lerntheoretischen Überzeugungen von Lehrkräften. Es wurden in Videos aus dem Mathematikunterricht der Sekundarstufe I Schüler- und Lehrkraftäußerungen hinsichtlich ihrer Art und ihres kognitiven Anspruchs kodiert. Die Ergebnisse zeigen, dass sich drei Skalen von Unterrichtsgesprächsqualität identifizieren lassen, wobei die Skala „ko-konstruktives Gespräch“ positive Zusammenhänge zu konstruktivistischen Überzeugungen der Lehrkräfte ausweist.

Der dritte Beitrag befasst sich mit der Strukturierung von Sachunterricht der 4. Jahrgangsstufe und Zusammenhängen zur Schülerleistung. Dabei wurde eine Unterrichtsreihe nach der Struktur der Oser'schen Basismodelle, die – im Grundschulunterricht umstrittene - kognitiv anspruchsvolle Reflexionsphasen beinhalten, konzipiert, durchgeführt und videographiert. Die Videos wurden anschließend hinsichtlich des Lernangebotes und dessen Nutzung kodiert. Ein Prä-Post Vergleich zeigt einen hohen Lernzuwachs der SuS.

Im vierten Beitrag wird die Wirkungskette von Lehrkraftkompetenzen - Unterrichtsqualität - Schülerkompetenzentwicklung im Bereich Bild-Text-Integration im Sachunterricht der Jahrgangsstufe 4 untersucht. Hierfür wurden Unterrichtsvideos bezüglich vier tiefenstruktureller Qualitätsmerkmale analysiert. Erste Ergebnisse weisen auf unterschiedliche geartete Zusammenhänge zwischen Lehrkraftwissen, -einstellungen und -überzeugungen und den vier Unterrichtsqualitätsmerkmalen hin. Methodische Verfahren bei der Videoanalyse sowie Implikationen für Forschung und Schulpraxis werden im Rahmen des Symposiums diskutiert.

Freitag
13. März

11:15 Uhr

HGB 30

Menk, Marleen; Dalehefte, Inger Marie; Köller, Olaf

Instruktion von Aufgaben im Mathematikunterricht mit Fokus auf Prozessbezogene Kompetenzen - Eine Videostudie im Rahmen von SINUS an Grundschulen

Theoretischer Hintergrund: Die Bildungsstandards im Fach Mathematik für den Primarbereich wurden im Oktober 2004 als verbindlich für alle Bundesländer ab dem Schuljahr 2005/2006 von der Kultusministerkonferenz beschlossen. Sie definieren Standards in Form prozess- und inhaltsbezogener Kompetenzen, die Schülerinnen und Schüler zum Ende der 4. Jahrgangsstufe erreicht haben sollen. Die Bildungsstandards sind somit Bestandteil eines umfassenden Systems zur Qualitätssicherung des Unterrichts (KMK, 2004). Bislang wurde in Untersuchungen überwiegend auf die inhaltsbezogenen Kompetenzen fokussiert, so dass es bislang wenige Studien zu den prozessbezogenen Kompetenzen gibt. Diese zeigen jedoch, dass einige der prozessbezogenen Kompetenzen (Argumentieren, Modellieren und Darstellen) kaum im deutschen Mathematikunterricht vorkommen (Neubrand et al., 2011; Neubrand, 2002). Es ist vorstellbar, dass die Aufgabeninstruktionen von Lehrkräften bei der Umsetzung der prozessbezogenen Kompetenzen eine entscheidende Rolle spielen. Da die prozessbezogenen Kompetenzen (Modellieren, Argumentieren, Kommunizieren, Problemlösen, Darstellen und Technische Grundfertigkeiten) als wichtige Voraussetzungen für erfolgreiches Lernen im Mathematikunterricht gelten (Walther, Selter & Neubrand, 2008), ist es Ziel dieser Videostudie zu untersuchen, ob und inwieweit die Instruktionen von Aufgaben durch Lehrkräfte auf die Entwicklung und Förderung prozessbezogener Kompetenzen bei den Schülerinnen und Schülern abzielen.

Die Videostudie ist im Rahmen von SINUS an Grundschulen (SGS) durchgeführt worden. SGS ist ein bundesweites Unterrichtsentwicklungsprogramm für den mathematischen und naturwissenschaftlichen Grundschulunterricht, das von 2009 bis 2013 in 10 Bundesländern durchgeführt wurde. Insgesamt arbeiteten 5440 teilnehmende Lehrkräfte an 870 Schulen mit Hilfe von 10 Modulen und 5 thematischen Schwerpunkten an der Weiterentwicklung ihres Unterrichts. Einer der fünf thematischen Schwerpunkte des Programms war die Umsetzung der Bildungsstandards Mathematik für den Primarbereich. Veränderungen der Unterrichtskultur sollten gemäß der Wirkungskette von Professionalisierungsmaßnahmen (Desimone, 2009) im unterrichtlichen Handeln feststellbar sein. Aus diesem Grund wird mittels Videobeobachtung auf die prozessbezogenen Kompetenzen in den Instruktionsphasen der Lehrpersonen fokussiert. Durch die Videoanalysen sollen Antworten auf folgende Forschungsfragen gefunden werden:

Fragestellung: (1) In welchem Umfang kann eine Orientierung an den Bildungsstandards im Grundschulunterricht beobachtet werden? Welche Bereiche der prozessbezogenen Kompetenzen werden im Unterricht gefördert und unterstützt und wie gestalten sich Dauer und Häufigkeit dieser Ereignisse? (2) Unterscheiden sich die Ergebnisse der SINUS-Klassen im Vergleich zu einer Kontrollgruppe in Bezug auf die Umsetzung der Bildungsstandards bzw. der prozessbezogenen Kompetenzen im Unterricht?

Methode: Die Datengrundlage bilden Videoaufzeichnungen des Mathematikunterrichts von 10 SINUS-erfahrenen Lehrpersonen (21 Unterrichtsstunden) zu ein bis drei Messzeitpunkten in den Jahren 2010-2013. Weiterhin wurde der Unterricht von 9 nicht am Programm teilnehmenden Lehrpersonen aufgezeichnet (9 Unterrichtsstunden). Für die Studie wurden die Aufnahmen der Lehrkräfte aus der Kontrollgruppe mit Aufnahmen von Lehrkräften aus der SINUS-Gruppe parallelisiert.

Die Videos wurden in 30-Sekunden-Intervallen mit einem Kategoriensystem zur Erfassung der prozessbezogenen Kompetenzbereiche in den Instruktionen der Lehrpersonen analysiert. Das niedrig-inferente Kategoriensystem umfasst die disjunkten Kategorien „Problemlösen“, „Technische Grundfertigkeiten“, „Darstellen“, „Modellieren“, „Kommunizieren“ und „Argumentieren“. Die Erprobung des Kategoriensystems und Berechnungen der Beobachterübereinstimmung erfolgten an einer Pilotstichprobe. Diese Vorarbeiten stellten sicher, dass das Instrument zur Analyse der Hauptstichprobe geeignet war (Cohens Kappa > .75).

Ergebnisse: Erste deskriptive Ergebnisse aus der Studie werden vorgestellt. Während Problemlösen und Technische Grundfertigkeiten verhältnismäßig oft vorkommen, werden Darstellen, Kommunizieren, Argumentieren und Modellieren seltener angesprochen. Die Befunde zeigen, dass die Lehrkräfte in beiden Stichproben prozessbezogene Kompetenzen in ihrem Unterricht ansprechen, jedoch in unterschiedlichem Grad. Diskutiert wird, inwieweit die Unterschiede auf die Wirksamkeit des SGS-Programms zurückgeführt werden können. Die Analysen liefern interessante Informationen über die erfolgreiche Umsetzung der Bildungsstandards im Unterricht und können zudem als wirkungsvolle Beispiele in der Lehrerbildung genutzt werden.

Freitag
13. März

11:15 Uhr

HGB 30

Pauli, Christine; Reusser, Kurt

Lerntheoretische Überzeugungen von Mathematiklehrkräften und ihre Bedeutung für die Qualität von Unterrichtsgesprächen

Theoretischer Hintergrund: Die Bedeutung des Unterrichtsgesprächs für das Lernen der Schüler ist heute unbestritten und in zahlreichen Publikationen theoretisch begründet worden, sowohl in Bezug auf das konzeptuelle, fachliche Lernen, als auch in Bezug auf die Entwicklung überfachlicher Kompetenzen (Mercer & Howe, 2012; Resnick, Michaels & O'Connor, 2010; Schwarz, Dreyfus & Hershkowitz, 2009). Allerdings sind diese positiven Effekte nur zu erwarten, wenn das Gespräch gewissen Qualitätsanforderungen genügt. Auf der Basis (sozial-)konstruktivistischer Theorieansätze zeichnen sich lernwirksame Unterrichtsgespräche u.a. durch bedeutsame Problemstellungen, kognitiv aktivierende Fragen und eine aktive und verantwortliche Partizipation der Lernenden im Sinne einer ko-konstruktiven Generierung von Wissen aus (vgl. u.a. Alexander, 2008; Lampert & Cobb, 2003; Myhill, Jones & Hopper, 2006; Resnick et al., 2010). Im alltäglichen Unterricht scheinen jedoch viele Klassengespräche solchen Qualitätskriterien nur unzureichend zu entsprechen (Begehr, 2006; Franke, Kazemi & Battey, 2007). Dies deutet darauf hin, dass die Gestaltung produktiver Unterrichtsgespräche besondere Anforderungen an die Kompetenz der Lehrkräfte stellt. Dabei ist zu vermuten, dass neben dem Fach- und fachdidaktischen Wissen auch die lernpsychologischen (konstruktivistischen vs. transmissiven) Überzeugungen (Reusser, Pauli & Elmer, 2011; Voss, Kleickmann, Kunter & Hachfeld, 2011) der Lehrkräfte eine Rolle spielen dürften, was insbesondere auch für die Aus- und Weiterbildung von Lehrpersonen bedeutsam ist. Dabei ist zu erwarten, dass Lehrkräfte mit einem konstruktivistischen Lernkonzept sich eher um diskursive Unterrichtsgespräche bemühen als Lehrpersonen mit einem eher transmissiven Lernverständnis.

Fragestellung und Methode: Dies wurde im Rahmen einer Videostudie zur Qualität der Unterrichtskommunikation im Mathematikunterricht untersucht: Lässt sich ein Zusammenhang zwischen konstruktivistisch geprägten lerntheoretischen Überzeugungen der Lehrkräfte und der Qualität ihrer Klassengespräche nachweisen?

Die datenbezogene Grundlage der Auswertungen stellt die schweizerisch-deutsche Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“ dar (Klieme, Pauli & Reusser, 2009), deren (nicht repräsentative) Stichprobe in Deutschland und der Schweiz je 20 Schulklassen (9. bzw. 8. Schuljahr) und deren Mathematiklehrkräfte umfasste.

Die Messung des konstruktivistischen Lernverständnisses der Lehrpersonen erfolgte mittels Lehrpersonenfragebogen aufgrund einer sechs Items umfassenden Skala (Klieme, Pauli & Reusser, 2005). Für die Erfassung der Qualität der Klassengespräche wurden sämtliche Klassengespräche während einer je drei Unterrichtsstunden umfassenden, inhaltlich standardisierten Unterrichtseinheit (Einführung des Satzes von Pythagoras) analysiert. Es wurde ein auf früheren Instrumenten (Hiebert & Wearne, 1993; Kobarg & Seidel, 2007) aufbauendes Kategoriensystem verwendet, das eine Kodierung sämtlicher Lehrer- und Schüleräußerungen beinhaltete, wobei jede Äußerung in Bezug auf mehrere Dimensionen kodiert wurde (Pauli, 2006; Pauli, Ineichen & Suhner, 2009): Alle Lehreräußerungen wurden hinsichtlich der Art der Äußerung und der Art der Feedbackkomponente kodiert, die Fragen/Impulse zusätzlich hinsichtlich ihres kognitiven Anforderungsgehalts. Alle Schüleräußerungen wurden hinsichtlich der Art und Funktion der Äußerung sowie des Vorkommens einer Begründung kodiert. Die Kodierung wurde durch zwei Kodiererinnen hinreichend reliabel (> 88% Übereinstimmung bei jeder Kategorie) durchgeführt. Um der zwischen den Klassen variierenden Gesprächsdauer gerecht zu werden, wurden prozentuale Anteile der einzelnen Kategorien ermittelt.

Ergebnisse: Auf der Basis einer explorativen Faktorenanalyse wurden drei Skalen gebildet, welche drei Qualitätsdimensionen abbilden: (1) „Ko-konstruktives Gespräch“ (2) „Interaktionsunterstützung“ und (3) „Schülerinitiativen“ (im Detail siehe Pauli & Reusser, in press; vgl. auch Reusser & Pauli, 2013). Auf der Basis dieser Skalen konnte der Zusammenhang zwischen den lernpsychologischen Überzeugungen und der Gesprächsqualität untersucht werden. Der erwartete positive Zusammenhang zwischen der Gesprächsqualität und konstruktivistischen Überzeugungen ließ sich für die Skala „ko-konstruktives Gespräch“ nachweisen (.476; $p=.002$), jedoch nicht für die beiden andern Skalen. Insgesamt deuten die vorliegenden Ergebnisse unter Einbezug weiterer Daten auf eine charakteristische Ausprägung konstruktivistischer Überzeugungen von Lehrkräften hin, die im Hinblick auf Implikationen für die Unterrichtsentwicklung und die Aus- und Weiterbildung von Lehrpersonen diskutiert werden.

Freitag
13. März

11:15 Uhr

HGB 30

Krumbacher, Christina; Fischer, Hans

Modelle zur Tiefenstrukturierung und ihre Eignung für den physikalischen Sachunterricht – eine quantitativ-empirische Videostudie

Theoretischer Hintergrund: Die kognitiven Strukturen und mentalen Verarbeitungsprozesse, die im Unterricht ermöglicht werden, gelten als wichtiger Aspekt von Unterrichtsqualität (z.B. Fischer et al., 2002). Sie werden auch Tiefenstrukturen genannt und sind als wichtiges Planungselement von (Sach)Unterricht in entsprechender didaktischer Literatur zu finden (z.B. Gläser & Sothmann, 2013; Mischke & Kiper, 2009). Anders als in didaktischen Dreischritt-Modellen, die meist pauschal aus einer Art Einleitung, Erarbeitung und Reflexion bestehen, werden die Schritte in Tiefenstruktur-Modellen spezifisch auf den jeweiligen Lehr- oder Lehrprozess abgestimmt.

Eine Möglichkeit zur lernprozessorientierten Tiefenstrukturierung bieten Oser und Baeriswyl (2001) an. Sie haben mehrere Basismodelle entwickelt, die jeweils aus vier bis fünf Handlungskettenschritten (HKS) bestehen. Für den Physikunterricht werden drei dieser Modelle als relevant erachtet: Lernen durch Eigenerfahrung, Konzeptbilden und Problemlösen (z.B. Reyer, 2004; Wackermann, 2008). Sie setzen sich aus folgenden Schritten zusammen:

- (1) Lernen durch Eigenerfahrung (Planung der Handlungen; Durchführung der Handlungen; Konstruktion von Bedeutung; Generalisierung der Erfahrung; Reflexion von ähnlichen Erfahrungen)
- (2) Konzeptbildung (Bewusstmachung des Vorwissens; Durcharbeiten eines Prototyps; Beschreiben der wichtigen Merkmale des neuen Konzepts; Aktiver Umgang mit dem neuen Konzept; Anwendung des neuen Konzepts in anderen Kontexten)
- (3) Problemlösen (Problemgenerierung und Problempräzisierung; Entwicklung von Lösungswegen; Testen von Lösungswegen; Evaluation und Anwendung der Lösungen)

Es wird vermutet, dass sich die o.g. Basismodelle auch für den physikalischen Sachunterricht eignen, was aber bisher nicht empirisch überprüft wurde. Allerdings fand Ohle (2010), dass die Tiefenstrukturierung im physikalischen Sachunterricht eher einseitig ausfällt: Über 50% der tatsächlichen Unterrichtszeit wird darauf verwandt, anhand von Hands-on-Aktivitäten Erfahrungen zu sammeln, weniger als 10% der Gesamtunterrichtszeit darauf, gemachte Erfahrungen in irgendeiner Weise zu reflektieren (ebd.). Dabei werden gerade die reflektierenden Phasen als wichtiges Element des Lernprozesses angesehen (vgl. z.B. Hofstein & Lunetta, 2004).

Um auch reflektierende Phasen in den Fokus der Planung von Sachunterricht zu bringen, bietet es sich an, entsprechende Modelle zur Orientierung zur Verfügung zu stellen. Allerdings muss dafür erst empirisch überprüft werden, ob sich die o.g. Basismodelle überhaupt für den Sachunterricht eignen.

Fragestellung und Methode: Dafür wurde eine Videostudie im Prä-Post-Design durchgeführt, in der die Eignung durch drei Ebenen definiert wird: (1) Lassen sich die drei o.g. Basismodelle in die physikalische Sachunterrichts-Gestaltung implementieren? (Angebots-Seite) Bei der Beantwortung dieser Frage werden als Kriterien die Fokussierung auf ein Basismodell angenommen, die Vollständigkeit der Schritte und ihre Reihenfolge sowie die Unterrichtszeit, die sich auf den tatsächlichen Lerngegenstand bezieht. (2) Können die Schülerinnen und Schüler den angebotenen Handlungskettenschritten der drei o.g. Basismodelle folgen? (Nutzungs-Seite) Die Kriterien für die Beantwortung dieser Frage sind äquivalent zu den vorherigen, nur dass hier die Nutzung der Schülerinnen und Schüler im Fokus steht. (3) Führt eine Unterrichtseinheit, die sich aus den drei o.g. Basismodellen zusammensetzt, zu einem Lernerfolg? (Wirkung/Ertrag) Als Kriterium hierfür wird ein Lernerfolg im deklarativen und konzeptionellen Bereich angesehen, der mindestens im Bereich einer mittleren Effektstärke liegt.

Zur Bearbeitung der Fragen wurde ein Unterrichtsangebot im Sinne der o.g. Basismodelle zum Thema „Verdunstung und Kondensation“ für die 4. Klassenstufe geplant. Die Unterrichtseinheit wurde über einen Zeitraum von jeweils fünf Doppelstunden in fünf verschiedenen Klassen durchgeführt (N = 107 SuS). Der gesamte Unterricht wurde videografiert, wobei eine Kamera die Lehrperson gefilmt hat und sechs Kameras die Gruppentische der SuS. Das gesamte Unterrichtsangebot wurde anhand eines Manuals intervallbasiert kodiert und analysiert. Die Nutzung des Unterrichtsangebots wurde ebenfalls intervallbasiert anhand von Einzelanalysen von 23 SuS untersucht.

Ergebnisse: Beide Kodierungen zeigten gute Interrater-Reliabilitäten ($.41 < \kappa < 1$, $MW = .89$). Der Lernzuwachs wurde mit bereits bestehenden Instrumenten (Kauertz et al., 2011) gemessen. Es zeigte sich ein hoher Lernzuwachs ($t(106) = 7.95$, $p = .00$, $d = .81$).

Die Kodierung und Analyse der Tiefenstruktur sowie die Ergebnisse sollen im Vortrag vorgestellt und diskutiert werden.

Freitag
13. März

11:15 Uhr

HGB 30

Chair(s)

Ohle, Annika
McElvany, Nele

"Lehrkräfte, Unterrichtsqualität & Lernerfolge im Kontext von Videostudien
– Perspektiven aus Didaktik, Erziehungswissenschaften und Psychologie"

Ohle, Annika; McElvany, Nele; Oerke, Britta; Horz, Holger; Ullrich, Mark

Unterrichtsqualität beim Lehren und Lernen mit Bild-Text-Material: Zusammenhänge zu Lehrkraft- und Schülerkompetenz

Texte mit instruktionalen Bildern werden häufig in Lehr- Lernmaterialien aller Schulstufen eingesetzt. Es gibt jedoch zahlreiche empirische Hinweise, dass diese Art der Informationsvermittlung nicht nur Vorteile sondern auch Herausforderungen für die Lernenden birgt (Schnotz & Bannert, 2003). Wie muss Unterricht gestaltet sein und welche Kompetenzen brauchen Lehrkräfte, damit erfolgreiche Lernprozesse mit Bild-Text-Material initiiert werden können? Dieser Vortrag leistet einen Beitrag zur Beantwortung dieser Fragen u.a. durch Videoanalyse.

Theoretischer Hintergrund: Die integrative Informationsverarbeitung von Bildern und Texten ist selten explizites Thema von Unterricht. Dennoch wird von Schülerinnen und Schülern (SuS) erwartet, mit dieser Art Material zu lernen. Um Lehr-Lernprozesse mit Bild-Text-Material abzubilden, können gängige Unterrichtsmodelle herangezogen werden, die neben Interaktionen im Klassenraum auch individuelle Voraussetzungen der Lehrenden und Lernenden als Einflussfaktoren auf den Kompetenzerwerb von Lernenden in den Blick nehmen (u.a. Lipowsky, 2006).

In der Unterrichtsforschung haben sich drei zentrale, lernwirksame Merkmalsbereiche von Unterrichtsqualität herauskristallisiert: a) kognitive Aktivierung, b) Klassenführung, Regelklarheit und Struktur und c) unterstützendes, schülerorientiertes Unterrichtsklima (Klieme & Rakoczy, 2008). Um qualitativ hochwertige Lernangebote bereitstellen zu können sind die professionellen Kompetenzen von Lehrkräften essentiell. Nach Baumert und Kunter umfassen diese a) professionelles Wissen, b) Einstellungen, Werthaltungen und Überzeugungen, c) motivationale Orientierungen und d) selbstregulative Fähigkeiten (Baumert & Kunter, 2006). Es ist anzunehmen, dass diese allgemeinen Erkenntnisse auch auf den Bereich des Unterrichts mit Bild-Text-Lernmaterialien übertragbar sind. Die Kompetenz der SuS zur Verarbeitung von Texten und Bildern (BiTe-Kompetenz) kann als Zielkriterium von Unterricht mit Bild-Text-Material interpretiert werden.

Fragestellung: Bisher ist weitgehend unbekannt, welche Zusammenhänge zwischen Lehrkraftkompetenzen, Unterrichtsgestaltung und BiTe-Kompetenz von SuS bestehen. Zwei Fragestellungen sind daher zentral: (1) Wie lässt sich Unterrichtsqualität bezüglich Bild-Text-Integration theoretisch und empirisch beschreiben? (2) Welche Zusammenhänge zwischen Unterrichtsqualität und (a) Voraussetzungen von Lehrenden und (b) der BiTe-Kompetenz der Lernenden sind empirisch identifizierbar?

Es wird angenommen, dass die oben genannten Merkmale von Unterrichtsqualität auch auf Lehr-Lernsituationen mit Bild-Text-Material übertragbar sind und dass sie in positivem Zusammenhang zur BiTe-Kompetenz der SuS stehen. Ebenso werden Zusammenhänge zwischen Lehrkraftkompetenz und der Qualität von Unterricht erwartet.

Methode: Die Daten zur Beantwortung der Forschungsfragen stammen aus dem DFG-geförderten Projekt „BiTe“, in dem N = 34 vierte Klassen jeweils in drei Unterrichtsstunden gefilmt wurden. Außerdem wurden die unterrichtenden Lehrkräfte zu ihren professionellen Kompetenzen befragt und die etwa 1.100 SuS hinsichtlich ihrer BiTe-Kompetenz getestet.

Die Unterrichtsvideos wurden bezüglich der Merkmale metakognitive Strategieunterstützung, Varianz/ Exploration von Lösungswegen (Merkmalsbereich „kognitive Aktivierung“), Strukturierung und Motivationsqualität mit guter bis sehr guter Interraterreliabilität für alle Merkmale ($g > .80$). untersucht. Darüber hinaus wurden die Lehrkräfte dieser Klassen zu ihrem Wissen, Einstellungen, Werthaltungen und Überzeugungen, motivationalen Orientierungen sowie selbstregulativen Fähigkeiten im Bereich Bild-Text-Integration befragt. Die BiTe-Kompetenzentwicklung der SuS wurde durch einen Multiple-Choice Test im ersten Halbjahr und am Ende des Schuljahres längsschnittlich erfasst.

Ergebnisse: Die Unterrichtsqualität bezüglich Bild-Text-Integration kann über die o.g. Merkmale beschrieben werden (Forschungsfrage 1); Ergebnisse einer konfirmatorischen Faktorenanalyse unterstützen ein 4-Faktorenmodell ($\chi^2 = 50.12$; $df = 47$; $p = .35$; CFI = 0.98; RMSEA = .03). Zusammenhänge der Unterrichtsqualitätsmerkmale mit Voraussetzungen der Lehrenden wurden durch manifeste Strukturgleichungsmodelle untersucht (Forschungsfrage 2a). Für die metakognitive Strategieunterstützung ergeben sich positive Zusammenhänge zur Einstellung der Lehrkräfte zum Üben und ihrer Vermeidung von Bild-Text-Material, während das Wissen der Lehrkräfte im Bereich Bild-Text-Integration in positivem Zusammenhang zur Varianz/ Exploration von Lösungswegen steht. Bei dem Merkmal Strukturierung ergeben sich ein positiver Zusammenhang zu Selbstwirksamkeitserwartungen und ein negativer Zusammenhang zum Engagement der Lehrkräfte im Bereich Bild-Text-Integration. Die Motivierungsqualität weist (tendenziell) negative Zusammenhänge zu den Zweifeln und den selbstregulativen Fähigkeiten der Lehrkräfte beim Lehren mit Bild-Text-Material auf.

Zusammenhänge von Lehrkraft- und Unterrichtsmerkmalen zur BiTe-Kompetenzentwicklung der SuS werden auf der Tagung präsentiert.

Freitag
13. März

11:15 Uhr

HGB 30

Kiemer, Katharina

PISA Plus 2012/2013: Kompetenzveränderungen von Schülerinnen und Schülern von der neunten zur zehnten Klassenstufe

Im Rahmen der nationalen Begleitforschung zu PISA wurden im Anschluss an die fünfte Erhebungsrunde 2012 die teilnehmenden neunten Klassen (N = 222 Klassen) im darauf folgenden Schuljahr erneut getestet (PISA Plus). Damit sind über die alle drei Jahre erhobenen Trenddaten hinaus auch längsschnittliche Betrachtungen des Kompetenzerwerbs der Schülerinnen und Schüler in Deutschland möglich.

Die Beiträge des Symposiums gehen zwei Hauptfragestellungen nach: Zum einen wie entwickeln sich Kompetenzen im Verlauf eines Schuljahres am Ende der Sekundarstufe I und zum anderen welche Merkmale auf Ebene des Elternhauses, der Schule, des Unterrichts und der Jugendlichen beeinflussen die Kompetenzentwicklung. Dabei werden Perspektiven aus der Mathematikdidaktik, der Didaktik der Naturwissenschaften, der Psychologie und der Erziehungswissenschaft miteinander verknüpft.

Der erste Beitrag berichtet Ergebnisse des Kompetenzzuwachses im Fach Mathematik im Verlauf eines Schuljahres. In bisherigen PISA-Erhebungen zeigte sich wiederholt, eine signifikant höhere mathematische Kompetenz von Jungen im Vergleich zu Mädchen. Dieser Fragestellung wird auch in der längsschnittlichen Analyse nachgegangen. Ein vertiefender Blick auf die Veränderung der Kompetenz der Gesamtskala mathematischer Kompetenz sowie der vier Subskalen ermöglicht die Identifikation spezifischer Stärken und Schwächen der deutschen PISA-Stichprobe im Bereich mathematischer Kompetenz.

Der zweite Beitrag vertieft die Betrachtung der Kompetenzentwicklung in Mathematik, indem er schulartspezifisch den Zusammenhang des individuellen Zugehörigkeitsgefühls der Schülerinnen und Schüler zur eigenen Schule mit der mathematischen Kompetenz untersucht. Aus bildungspolitischer Sicht sind diese Befunde von zentraler Bedeutung für die weitere Gestaltung und wichtige Veränderungsimpulse der deutschen Bildungslandschaft sowie einer gezielten Förderung von Schülerinnen und Schülern.

Die Bedeutung der naturwissenschaftlichen Kompetenz für das Treffen evidenzbasierter Entscheidungen und die aktive Teilhabe an der Gesellschaft wurde immer wieder betont (OECD, 2007). Vor diesem Hintergrund beschäftigt sich der dritte Beitrag mit der Veränderung der naturwissenschaftlichen Kompetenz und berücksichtigt dabei schulart- und geschlechtsspezifische Disparitäten. Darüber hinaus wird der Einfluss des naturwissenschaftlichen Unterrichts für die Kompetenzentwicklung untersucht.

Der vierte Beitrag betrachtet gezielt die Entwicklung der Lesekompetenz am Ende der Sekundarstufe I als eine wichtige Grundvoraussetzung für allgemeine Bildungserfolge und lebenslanges Lernen. Da sich bei PISA 2012 große Leistungsunterschiede zwischen Jungen und Mädchen zeigten, werden diese Entwicklungsverläufe hier besonders betrachtet. Darüber hinaus wird der Einfluss der Erstsprache der Schülerinnen und Schüler, die Zuhause gesprochene Sprache sowie basale Prozesse des Leseverstehens auf Wort- und Satzebene und Updating als exekutive Facette des Arbeitsgedächtnisses auf die Lesekompetenzentwicklung untersucht.

Abschließend werden die Beiträge mit Blick auf mögliche Implikationen der beschriebenen Entwicklungsverläufe für weiterführende Forschung sowie Ansatzpunkte für Schule und Unterricht diskutiert.

Freitag
13. März

11:15 Uhr

HGB 40

Chair(s)

Schiepe-Tiska, Anja
Kiemer, Katharina**PISA Plus 2012/2013: Kompetenzveränderungen von Schülerinnen und Schülern von der neunten zur zehnten Klassenstufe**

Reiss, Kristina; Heinze, Aiso; Sälzer, Christine; Nagy, Gabriel; Pant, Hans Anand

Veränderung der mathematischen Kompetenz von der neunten zur zehnten Klassenstufe

Theoretischer Hintergrund: Mathematische Kompetenz ist eine von drei Domänen, die im Rahmen von PISA alle drei Jahre bei Schülerinnen und Schülern gegen Ende der Schulpflicht überprüft wird. Bei PISA 2012 war die Mathematik nach 2003 zum zweiten Mal die Schwerpunktdomäne. In der PISA-Längsschnittstudie 2003/2004 zeigte sich, dass die durchschnittliche mathematische Kompetenz der Schülerinnen und Schüler auf Individualebene im Verlauf eines Schuljahres um 25 Punkte anstieg (Ehmke, Blum, Neubrand, Jordan & Ulfig, 2005). Zugleich traten in allen bisherigen PISA-Erhebungszyklen signifikante Kompetenzunterschiede zwischen Mädchen und Jungen hervor, die sich sowohl auf die Gesamtskala Mathematik als auch auf die untersuchten Subskalen bezogen. Ähnlich wie in PISA 2003 handelte es sich dabei um die Subskalen Quantität, Raum und Form, Veränderung und Beziehungen sowie Unsicherheit und Daten.

Fragestellungen: In Anknüpfung an die PISA-Längsschnittstudie 2003/2004 steht das Interesse im Vordergrund, damalige Befunde anhand der erneuten Messwiederholung einer PISA-Kohorte zu ergänzen. Folgende Fragestellungen leiten diesen Beitrag: (1) Was lernen Schülerinnen und Schüler im Schuljahr nach dem PISA-Test in Mathematik dazu? (2) Wie verändert sich die Kompetenz auf individueller und auf Klassenebene? (3) Verändert sich die Kompetenz von Mädchen und Jungen nach dem PISA-Test unterschiedlich? Neben der Veränderung der Gesamtskala mathematischer Kompetenz wird dabei jeweils auch nachgezeichnet, inwieweit die Schülerinnen und Schüler im Bereich der vier Subskalen einen Kompetenzzuwachs erreicht haben.

Methode: Die drei Fragestellungen werden mit Daten aus der Längsschnitterhebung PISA Plus untersucht, die im Frühjahr 2013 und damit ein Jahr nach der PISA-Haupterhebung 2012 stattfand. Im Rahmen von PISA Plus wurden Schülerinnen und Schüler der 9. Klasse, die an PISA 2012 teilgenommen hatten, in der 10. Klasse erneut befragt und getestet (etwa 200 Klassen).

Die erste Fragestellung wird anhand eines längsschnittlichen Veränderungsmodells bearbeitet. Darüber hinaus wird unter Bezug auf die Subskalen mathematischer Kompetenz, wie sie in PISA 2012 verwendet wurden, deskriptiv berichtet, ob sich relative Stärken und Schwächen in der Kompetenz auf den Subskalen aus PISA 2012 auch in der Messwiederholung im Frühjahr 2013 finden lassen. In PISA 2012 zeigten Schülerinnen und Schüler in Deutschland in den beiden Subskalen Veränderung und Beziehungen sowie Quantität bessere Leistungen als in den beiden Subskalen Raum und Form sowie Unsicherheit und Daten.

Die zweite Fragestellung unterscheidet die individuelle Schülerebene von der Klassenebene und vergleicht individuelle Kompetenzentwicklungen mit solchen im Klassenverband. Für die Analyse der Entwicklung auf Klassenebene werden die Schülerdaten entsprechend aggregiert. Von besonderem Interesse wird hierbei sein, inwieweit sich innerhalb von Schulklassen Heterogenität bei den Entwicklungsverläufen finden lässt.

Zur Beantwortung der dritten Fragestellungen werden die Entwicklungsverläufe von Mädchen und Jungen getrennt voneinander betrachtet und miteinander verglichen. Die Mädchen starten hier mit einer signifikant niedrigeren mittleren Kompetenz als die Jungen und die Analysen werden zeigen, ob und inwieweit sich diese Geschlechterunterschiede ein Jahr nach der PISA-Erhebung noch immer zeigen oder ob sie sich verändert haben.

Chair(s)

Schiepe-Tiska, Anja
Kierner, Katharina

PISA Plus 2012/2013: Kompetenzveränderungen von Schülerinnen und Schülern von der neunten zur zehnten Klassenstufe

Sälzer, Christine; Seidel, Tina; Steinert, Brigitte; Bayer, Sonja

Der Zusammenhang zwischen Kompetenzentwicklung und dem Gefühl der Zugehörigkeit zur Schule in unterschiedlichen Sekundarschulformen

Theoretischer Hintergrund: In Bezug auf schulische Rahmenbedingungen für die Entwicklung von Kompetenzen lieferte PISA 2012 eine Reihe relevanter Indikatoren. Auf schulischer Ebene sind für die meisten untersuchten Merkmale kaum Veränderungen zu erwarten, die sich bereits ein Schuljahr nach der PISA-Erhebung nachweisen ließen. Deshalb greifen wir für diesen Beitrag ein Merkmal der Schülerinnen und Schüler heraus, das die Schule aus deren Perspektive beschreibt: das Gefühl der Zugehörigkeit zur eigenen Schule (engl. school belonging). Dieses Gefühl der Zugehörigkeit wurde zu beiden Messzeitpunkten erfasst. In bisherigen PISA-Zyklen hat sich wiederholt gezeigt, dass ein hohes Gefühl der Zugehörigkeit auch mit guten Kompetenzwerten einherging (OECD, 2013). Im Detail kann dieser Zusammenhang jedoch allein mit einer einmaligen Untersuchung nicht festgestellt werden.

Das Gefühl der Zugehörigkeit zur eigenen Schule ist ein plausibler Indikator dafür, wie stark sich Schülerinnen und Schüler mit ihrer Schule identifizieren und wie gerne sie dort hingehen. Mit der Messwiederholung ein Jahr nach PISA 2012 ist es möglich, den Zusammenhang zwischen dem Gefühl der Zugehörigkeit zur Schule und der mathematischen Kompetenz in seinem zeitlichen Verlauf zu untersuchen und zu interpretieren. Dies ist beispielsweise für Überlegungen relevant, die nach den Verbesserungen in der mathematischen Kompetenz zwischen PISA 2003 und PISA 2012 bezüglich möglicher Gründe für diesen Fortschritt angestellt wurden. In diesem Kontext werden häufig die gezielte Vorbereitung auf bestimmte Testformate, oft als Teaching to the Test bezeichnet (Brinkmann, 2009; Schrader & Helmke, 2008), oder enormer Leistungsdruck in den Schulen genannt, die mit einem hohen Gefühl der Zugehörigkeit schwer zu vereinbaren sind.

Nach der Darstellung einiger deskriptiver Befunde in den fünf untersuchten Schulformen (Gymnasium, Realschule, Schule mit mehreren Bildungsgängen, Integrierte Gesamtschule, Hauptschule) wirft unser Beitrag einen Blick auf den Zusammenhang zwischen dem subjektiven Gefühl der Zugehörigkeit zur Schule und der mathematischen Kompetenz der PISA-Schülerinnen und –Schüler in den verschiedenen Schulformen.

Fragestellung: Folgende Fragestellungen leiten unseren Beitrag: (1) Hängt das Gefühl der Zugehörigkeit zur eigenen Schule mit der Entwicklung mathematischer Kompetenz im Verlauf eines Schuljahres zusammen? (2) Ist dieser Zusammenhang in verschiedenen Schulformen unterschiedlich ausgeprägt?

Methode: Die Daten werden in unserem Beitrag in drei Schritten analysiert.

Zunächst wird die Varianz der mathematischen Kompetenz in die Anteile innerhalb und zwischen Schulen zerlegt. Wir betrachten diese Varianzaufteilung sowohl insgesamt als auch für die einzelnen Schulformen.

Im zweiten Schritt betrachten wir die Veränderung mathematischer Kompetenz in den verschiedenen Schulformen. Die Vielzahl unterschiedlicher Sekundarschulformen sowie der Anteil der Varianz bei den Schülerleistungen zwischen den Schulen, der sowohl Unterschiede zwischen Schularten als auch Unterschiede zwischen Einzelschulen kombiniert, machen für Deutschland die Frage besonders interessant, inwieweit sich Kompetenzen an verschiedenen Sekundarschularten unterschiedlich entwickeln. In Mathematik haben die Jungen in allen bisherigen PISA-Runden signifikant besser abgeschnitten als die Mädchen. Deshalb untersuchen wir die Kompetenzentwicklung zwischen 2012 und 2013 auch getrennt nach Mädchen und Jungen in den verschiedenen Schularten.

Im dritten Schritt untersuchen wir den Zusammenhang zwischen dem individuellen Gefühl der Zugehörigkeit der Schüler zu ihrer Schule und ihrer mathematischen Kompetenz. Zunächst betrachten wir die Entwicklung des Zugehörigkeitsgefühls zwischen PISA 2012 und der Messwiederholung in den verschiedenen Schulformen. Anschließend spezifizieren wir ein Kreuzpfadstrukturgleichungsmodell zur Analyse zeitverschobener Kreuzkorrelationen. Auch dieses Modell wird für jede Schulform separat betrachtet.

Erwartete Ergebnisse: Die für diesen Beitrag verwendeten Daten liegen zum Zeitpunkt der Einreichung noch nicht in skaliert Form vor, daher erfolgen die Analysen erst nach der Einreichung. Wir erwarten schulformspezifische Unterschiede zwischen Mädchen und Jungen beim Gefühl der Zugehörigkeit zu ihrer Schule sowie differenzielle Zuwächse der mathematischen Kompetenz, die durch das Gefühl der Zugehörigkeit mit erklärt werden.

Freitag
13. März

11:15 Uhr

HGB 40

Chair(s)

Schiepe-Tiska, Anja
Kiemer, Katharina

PISA Plus 2012/2013: Kompetenzveränderungen von Schülerinnen und Schülern von der neunten zur zehnten Klassenstufe

Schiepe-Tiska, Anja; Schöps, Katrin; Rönnebeck, Silke; Seidel, Tina; Lüdtke, Oliver; Prenzel, Manfred

Die Veränderung der naturwissenschaftlichen Kompetenz im Verlauf eines Schuljahres und der Einfluss der Qualität des naturwissenschaftlichen Unterrichts

Naturwissenschaftliches Wissen ist notwendig, um Entscheidungen im Alltag oder Beruf evidenzbasiert treffen zu können und schafft somit eine wichtige Voraussetzung für die Teilhabe an einer Gesellschaft, in der Wissenschaft und Technik eine bedeutende Rolle spielen (OECD, 2013; Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2013). Dabei ist nicht nur das in der Schule vermittelte Wissen von Bedeutung, sondern vor allem dass Schülerinnen und Schüler in der Lage sind, ihr erworbenen naturwissenschaftlichen Fähigkeiten und Fertigkeiten in alltagsnahen persönlichen, sozialen und globalen Situationen anzuwenden (Bybee & McCrae, 2011).

Nur wenige Studien haben bisher die Entwicklung der naturwissenschaftlichen Kompetenz im Laufe eines Schuljahres in Deutschland in einem größeren Umfang untersucht (Beaton et al., 1996; vgl. Köller & Baumert, 2002). Zu diesen gehört die im Rahmen von PISA 2003 durchgeführte Messwiederholungsstudie, die einen Zuwachs der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler von der 9. zur 10. Klasse belegte (Walter, Senkbeil, Rost, Carstensen & Prenzel, 2006). Seither wurden jedoch in Deutschland großangelegte Programme, länderübergreifende Maßnahmen und Bildungsreformen (SINUS Programm, Kontextprojekte, Einführung der Bildungsstandards) durchgeführt, mit dem Ziel, vor allem den naturwissenschaftlichen Unterricht weiter zu entwickeln und somit zu einer Verbesserung der naturwissenschaftlichen Kompetenz beizutragen. Bisher liegen aber keine Erkenntnisse darüber vor, wie sich nach diesen zum Teil massiven Veränderungen die naturwissenschaftliche Kompetenz im Laufe eines Schuljahres am Ende der Sekundarstufe I entwickelt. Diese Perspektive soll im vorliegenden Beitrag aufgegriffen werden.

Die Entwicklung der naturwissenschaftlichen Kompetenz wird entscheidend durch die Qualität des naturwissenschaftlichen Unterrichts beeinflusst, da dort systematisch Lerngelegenheiten zur Verfügung gestellt werden. Für den Lernprozess der Schülerinnen und Schüler ist entscheidend, wie diese Lerngelegenheiten wahrgenommen und genutzt werden. Erst die aktive und konstruktive Auseinandersetzung mit den angebotenen Lerngelegenheiten ermöglicht den Aufbau kognitiver Kompetenzen (Angebots-Nutzungs-Modell; Helmke & Weinert, 1997). Im vorliegenden Beitrag soll deshalb des Weiteren der Einfluss von Unterrichtsmerkmalen auf die Entwicklung der naturwissenschaftlichen Kompetenz untersucht werden.

Es werden demnach zwei Fragestellungen beantwortet: (1) Wie verändert sich die naturwissenschaftliche Kompetenz im Verlauf eines Schuljahres von der 9. zur 10. Klasse und zeigen sich dabei schulart- oder geschlechtsspezifische Unterschiede? (2) Welche Merkmale des naturwissenschaftlichen Unterrichts beeinflussen die Kompetenzentwicklung der Schülerinnen und Schüler? Dazu werden zunächst die Struktur des naturwissenschaftlichen Unterrichts und die wahrgenommenen Lehr-Lernbedingungen in der 10. beschrieben und abschließend mit der Entwicklung der naturwissenschaftlichen Kompetenz in Zusammenhang gebracht.

Die Fragestellungen werden mit Hilfe der Längsschnitterhebung PISA Plus untersucht. Dabei wurden Schülerinnen und Schüler der 9. Klasse, die an PISA 2012 teilgenommen hatten, in der 10. Klasse erneut befragt und getestet (circa 200 Klassen). Die Jugendlichen bearbeiteten einen Naturwissenschaftstest und beantworteten darüber hinaus im Rahmen des Schülerfragebogens Fragen zu den besuchten naturwissenschaftlichen Fächern, zur Häufigkeit von Unterrichtsaktivitäten (interaktives Lehren und Lernen, Demonstrationsexperimente, selbständiges Experimenten, Forschen) in den Fächern Biologie, Physik und integrierter Naturwissenschaftsunterricht sowie zu unterstützenden Lehr-Lernbedingungen (wahrgenommene Instruktionsqualität, Interesse der Lehrperson). Durch die klassenbasierte Stichprobe bietet PISA Plus hervorragende Möglichkeiten, um die Effekte von Unterrichtsmerkmalen auf die Kompetenzentwicklung zu untersuchen.

Wir erwarten schulartspezifische jedoch keine geschlechtsspezifischen Unterschiede in der Entwicklung der naturwissenschaftlichen Kompetenz, die durch Merkmale des Unterrichts beeinflusst werden.

Freitag
13. März

11:15 Uhr

HGB 40

Chair(s)

Schiepe-Tiska, Anja
Kierner, Katharina

PISA Plus 2012/2013: Kompetenzveränderungen von Schülerinnen und Schülern von der neunten zur zehnten Klassenstufe

Hohn, Katharina; Schiepe-Tiska, Anja; Goldhammer, Frank; Hahnel, Carolin; Kröhne, Ulf; Nagy, Gabriel; Lüdtke, Oliver

Die Entwicklung der Lesekompetenz innerhalb eines Schuljahres

Das Verstehen geschriebener Texte ist gerade in einer Wissensgesellschaft von großer Bedeutung. Dabei reicht es nicht, Texte zu „entziffern“. Vielmehr müssen die wesentlichen Aussagen eines Textes verstanden und kritisch interpretiert bzw. reflektiert werden können. Dies befähigt den Leser eigene Ziele zu erreichen, eigenes Wissen und eigene Potenziale zu entwickeln sowie aktiv an der Gesellschaft teilzunehmen (OECD, 2013).

Die Entwicklung der Lesekompetenz ist nicht auf die ersten Schuljahre beschränkt, auch wenn sich ein Großteil von Forschungsarbeiten auf die grundlegende Entwicklung der Lesekompetenz im Rahmen von Vor- und Grundschulbildung konzentriert. Vielmehr gilt, dass im Verlauf der Bildungsbiographie Leseanforderungen über Lernbereiche hinweg ansteigen und immer wieder neue und auch anspruchsvollere Texte erschlossen werden müssen. Die in der Sekundarstufe erworbene Lesekompetenz stellt dabei nicht nur eine wichtige Grundvoraussetzung für allgemeine Bildungserfolge und lebenslanges Lernen dar, sondern sie ist auch von grundlegender Bedeutung für eine erfolgreiche berufliche Entwicklung (z.B. Bussière, Hébert, & Knighton, 2010). Daraus folgt, dass die Lesekompetenz von Schülerinnen und Schülern und ihre Entwicklung auch im Bereich der Sekundarstufe empirisch zu untersuchen und begünstigende Entwicklungsfaktoren zu identifizieren sind (Artelt, Stanat, Schneider, & Schiefele, 2001).

Die PISA-Erhebungen erfassen seit 2000 in einem kontinuierlichen Abstand von drei Jahren u.a. die Lesekompetenz von 15-jährigen Schülerinnen und Schülern in Deutschland. In der letzten Erhebungsrunde 2012 erzielten die Jugendlichen in Deutschland erstmals Ergebnisse im Lesetest, die signifikant über dem OECD-Durchschnitt lagen (Hohn, Schiepe-Tiska, Sälzer, & Artelt, 2013). Im Rahmen der regulären PISA-Erhebungen können jedoch keine Fragen nach der individuellen Entwicklung von Lesekompetenz beantwortet werden, da es sich bei dem Studiendesign lediglich um ein querschnittliches Design handelt. Um dennoch die Entwicklung der Lesekompetenz von Schülerinnen und Schülern und mögliche Einflussfaktoren auf die Leseentwicklung untersuchen zu können, wurden Schülerinnen und Schüler, die an PSIA 2012 teilgenommen haben, ein Jahr später (2013) im Rahmen von PISA Plus erneut getestet. In der vorliegenden Studie werden neben der grundlegenden Beschreibung der Entwicklung der allgemeinen Lesekompetenz auch schulartspezifische Veränderungen untersucht. Da bereits bei den Analysen des ersten Messzeitpunktes (2012) große Leistungsunterschiede zwischen Jungen und Mädchen deutlich wurden, werden deren differenzielle Leistungsveränderungen genauer betrachtet. Insbesondere sind die Veränderungen der Lesekompetenz der Schülerinnen und Schüler auf den untersten Kompetenzstufen („Risikoschüler“) sowie auf den obersten Kompetenzstufen („High performer“) von Interesse. Neben der Erstsprache der Schülerinnen und Schüler, sowie der Zuhause gesprochenen Sprache werden des Weiteren basale Prozesse des Leseverstehens auf Wort- und Satzebene sowie Updating als exekutive Facette des Arbeitsgedächtnisses für die Erklärung der Veränderungen in der allgemeinen Lesekompetenz herangezogen.

Implikationen für die weitere schulische und außerschulische Förderung der Lesekompetenz von Schülerinnen und Schülern in der Sekundarstufe werden diskutiert.

Freitag
13. März

11:15 Uhr

HGB 40

Chair(s)

Schiepe-Tiska, Anja
Kiemer, Katharina

PISA Plus 2012/2013: Kompetenzveränderungen von Schülerinnen und Schülern von der neunten zur zehnten Klassenstufe

Knapp, Mariella

Heterogenität im Kontext von sozialen Beziehungen und Lernmotivation

Empirische Studien der Bildungsforschung und Pädagogischen Psychologie verweisen nicht zuletzt auf die Bedeutung der Lernmotivation für den schulischen Erfolg und die Kontinuität schulischen Lernens, weshalb Reformen immer wieder auf die Förderung dieser Komponente abzielen. In diesem Zusammenhang stellt sich die Frage, inwieweit Lernmotivation eine individuelle Dimension darstellt und welche Relevanz auch dem sozialen Umfeld zukommt. Forschungsergebnisse verweisen auf Kompositionseffekte und dies betrifft nicht nur das akademische Selbstkonzept sondern auch die Lernmotivation (Ryan, 2010). Hierbei wird angenommen, dass Klassenkameraden und LehrerInnen in der Schule, aber auch gleichaltrige Freunde außerhalb der Schule und Verwandte, durch soziale Vergleichsprozesse und den damit verbundenen Emotionen, Entwicklungen in der Lernmotivation mitbestimmen. Das geplante Symposium nimmt sich den sozialen Beziehungen an und möchte herausfinden, welche Bedeutung verschiedensten inner- und außerschulischen Faktoren zukommt, sowie inwiefern sich dies in verschiedenen nationalen Kontexten unterschiedlich zeigt.

Der erste Beitrag zielt auf die Einführung der Neuen Mittelschule (NMS) in Niederösterreich ab und untersucht wie sich die Lernmotivation 10-14jähriger SchülerInnen (N=800) im Laufe der NMS verändert hat. Mittels Längsschnittanalyse wird dabei versucht herauszufinden, welche außer- und innerschulischen Gegebenheiten diese Entwicklungen beeinflussen.

Die NMS stellt auch Ausgangspunkt des folgenden Beitrages dar, wobei hier vor allem Kompositionseffekte im Vordergrund stehen. Auch anhand der Mehrebenenanalysen, die auf SchülerInnendaten der 1. und 4. Klasse (N=893) fußen, wird die Bedeutung des Klassenklimas und der Einfluss durch KlassenkollegInnen deutlich. Kompositionseffekte sind dabei nicht alleinig auf die Klassenzusammensetzung rückführbar, sondern auch auf die subjektive Sichtweise des aktuellen Leistungsstandes.

Ausgehend vom Konzept der „sozio-motivationalen Abhängigkeit“ untersucht der dritte Beitrag inter- und intraindividuelle Unterschiede von Motivationstypen. Die motivierende Funktion von Peers und LehrerInnen wurde hier im Rahmen des SELF-Projektes längsschnittlich bei SchülerInnen der 8.-10. Schulstufe (N=1088) erfasst. Latente Transaktionsanalysen verweisen darauf, dass vor allem Peers im Laufe der Adoleszenz für die Lernmotivation an Bedeutung gewinnen.

Daran anschließend werden im letzten Beitrag die Erweiterung der Datengrundlage anhand einer SchülerInnenbefragung in Kanada (N=1477) diskutiert und nationale Unterschiede herausgearbeitet. Die konfirmatorischen latenten Klassenanalysen (Multigruppen) lassen erkennen, dass die postulierte Motivationstypologie auch außerhalb des deutschsprachigen Raumes vorzufinden ist, kanadische SchülerInnen jedoch die motivierende Funktion von Peers und LehrerInnen anders wahrnehmen als deutsche.

Unter Einsatz vielfältiger statistischer Auswertungsmethoden berichten die Beiträge der Bildungswissenschaftlerinnen und pädagogischen Psychologinnen über aktuelle Ergebnisse aus verschiedenen Ländern (konkret aus Deutschland, Kanada und Österreich) und ermöglichen damit eine internationale Perspektive. Die Ergebnisse aller Beiträge bestärken die heterogene Bedeutung des inner-, aber auch außerschulischen sozialen Umfeldes für die Entwicklung der Lernmotivation von SchülerInnen.

Chair(s)

Beer, Rudolf

Kilian, Michaela; Katschnig, Tamara; Geppert, Corinna; Knapp, Mariella

„Ja, ich lerne noch weiter! Oder doch nicht?“. Lernmotivation von 10 – 14jährigen SchülerInnen unter Berücksichtigung des innerschulischen wie außerschulischen Umfeldes.

Ziel dieses Beitrages ist es, anhand von Längsschnittdaten von 10 – 14jährigen SchülerInnen aus Neuen Mittelschulen (NMS) in Niederösterreich herauszufinden, ob und welchen Einfluss Lernerfahrungen innerhalb der Klasse auf ihre Lernmotivation haben und welchen Einfluss darauf ihr außerschulisches Umfeld (Peers, Großeltern, HorterzieherInnen u.a. Personen im Umfeld der SchülerInnen) nimmt. MitschülerInnen prägen die Lernerfahrungen der SchülerInnen in den Klassen, nehmen Einfluss auf die individuelle Lernmotivation das Selbstkonzept und bieten Hilfestellung im gemeinsamen Lernen. Nach Marsh et al. (1998) und Köller et al. (2006) kommen innerschulischen Lernerfahrungen, Kompositionseffekten der Klasse wie auch dem durchschnittlichen Leistungsbild von Klassen Bedeutung zu. MitschülerInnen können als Kapital angesehen werden, das durch den verschiedenen Klassenverband einen speziellen Erfahrungszusammenhang darstellt. Im außerschulischen Bereich wird das Peerkapital durch Art und Dauer von gemeinsamen Freizeitaktivitäten mit FreundInnen aktualisiert. Peers können eine Bereicherung darstellen, als inhaltliche und emotionale Unterstützer aber auch als Risikofaktor das schulische Lernen beeinflussen (Köhler, 2012; Krüger & Deppe, 2010; Rohlf, 2010). Elterliche Unterstützung, familiäre Ressourcen, wie auch die "Kapital"-Ausstattung der Herkunftsfamilie verweisen auf die Bedeutung des außerschulischen Umfeldes (Bourdieu, 1983, 1992; Dollmann, 2010; Ditton, 2010).

Die NMS, welche in Österreich seit dem Jahr 2011 eingeführt wird und die Hauptschule ersetzen soll zielt auch darauf ab, die Bildungsmotivation aller SchülerInnen stärker zu fördern. In der Ausformung des niederösterreichischen Modells werden Schulautonomie und Standortkonzepte als Schulentwicklungsaspekte betont. Die Schulen selbst streben eine Veränderung des Unterrichts gegenüber der tradierten Hauptschule an, entwickeln standortspezifischen Umsetzungsstrategien und treffen Maßnahmen, damit die angestrebten Ziele des Schulmodells unter Berücksichtigung des jeweiligen Schulumfeldes erreicht werden können (Landesschulrat für Niederösterreich, 2012). Nicht nur die Berücksichtigung des außerschulischen Kontextes ist wichtig, sondern auch vor allem das, was innerhalb der Schule passiert und wie die SchülerInnen dieses für ihr Lernen aktualisieren. Daher beschäftigen wir uns in diesem Beitrag mit innerschulischen wie auch außerschulischen Faktoren welche die Lernmotivation von SchülerInnen an NMS beeinflussen.

Fragestellung: Der Beitrag geht der Frage nach, wie sich die Lernmotivation von SchülerInnen an NMS im Verlauf der Sekundarstufe I verändert. Darüber hinaus interessiert welche innerschulischen wie auch außerschulischen Einflussfaktoren die Entwicklung der Lernmotivation von 10 – 14jährigen SchülerInnen an NMS mitbestimmen und welche Veränderungen von inner- und außerschulischen Komponenten im Längsschnitt zu Veränderungen der Lernmotivation führen.

Methode: Als Analysegrundlage für diesen Beitrag dienen Daten der NOESIS-Längsschnitterhebung (Niederösterreichische Mittelschule in der Schulentwicklung), in der seit dem Schuljahr 2010/11 SchülerInnen ab der 4. Schulstufe in 3 Kohorten jährlich befragt werden. Für die Bearbeitung der Fragestellung wurden SchülerInnendaten N = 1900 der 5. – 8. Schulstufe (2010 – 2014) herangezogen. Für vertiefende längsschnittliche wie auch mehrbenenanalytische Auswertungen liegen Daten von N = 800 SchülerInnen der ersten Kohorte vor. Durch Einbeziehung von Skalen der innerschulischen Wahrnehmungen und Skalen des außerschulischen Kontextes lässt sich das Zusammenpiel mehrerer Faktoren in der längsschnittlichen Entwicklung von Lernmotivation beobachten und die Bedeutung der jeweiligen Klasse kann dadurch untersucht werden.

Ergebnisse: Mithilfe von Längsschnittdaten aus der NOESIS-Studie lassen sich im Zeitraum von 2010 bis 14 die Wahrnehmung von SchülerInnen beschreiben. Dabei zeigt sich, dass die SchülerInnen kurz nach dem Eintritt in die NMS eine hohe Lernmotivation aufweisen. Über 4 Jahre hinweg lässt die Motivation von Jahr zu Jahr nach. Die Beziehungen zu den MitschülerInnen, ein gutes Einvernehmen in der Klasse und eine positive Beziehung zwischen SchülerInnen und Lehrkräften wirken förderlich auf die Lernmotivation. Nehmen außerschulische Peerkontakte über 4 Jahre hinweg stetig zu, so wirken diese nicht gerade förderlich auf die Lernmotivation. Allerdings können die SchülerInnen auf familiäre Unterstützung zählen. Viele erhalten Unterstützung durch Geschwister bzw. Eltern oder Nachhilfe. Förderlich für eine höhere Ausprägung der Lernmotivation ist auch das kulturelle Kapital der Familie.

Die Tatsache, dass nicht nur innerschulische Gegebenheiten, sondern auch Einflüsse von außerschulischer Seite bedeutsam für die Lernmotivation sind, kann mit dieser Studie bestärkt werden.

Freitag
13. März

11:15 Uhr

HGB 50

Chair(s)

Beer, Rudolf

Knapp, Mariella; Katschnig, Tamara; Geppert, Corinna; Kilian, Michaela

Die Klasse als soziale Ressource - Kompositionseffekte der Schulklasse auf die akademische Lernmotivation in der Sekundarstufe I

Der Erhalt der Lernmotivation kann als ein wichtiger Bestandteil der pädagogischen Arbeit in Bildungseinrichtungen gesehen werden. Studien verweisen nicht zuletzt darauf, welche Bedeutung der Lernmotivation im Zusammenhang mit schulischem Erfolg, mit Bildungsentscheidungen (Wigfield & Eccles, 2000) und der späteren beruflichen Leistungs- und Weiterbildungsbereitschaft zukommt (Stuhlmann, 2005). Mit dem postulierten Wandel von einer Industrie- zu einer Wissensgesellschaft (Europäischer Rat, 2000), bei gleichzeitiger augenscheinlicher geringer Studienbeginn- und Studienabschlussrate (OECD, 2013), kann auch vonseiten der Bildungspolitik die Förderung und der langfristige Erhalt von Bildungsmotivation als wünschenswert angesehen werden.

Die – in Österreich, als Reaktion auf die ersten PISA Ergebnisse im Schuljahr eingeführte – Schulform „Neue Mittelschule“ (neue Schulform der Sekundarstufe I – Schule der 10-14jährigen) setzt sich zum Ziel eine Schule für alle zu sein und Marginalisierungsprozesse zu minimalisieren. Dabei sollen vor allem auch die schulische Aspiration und Bildungsmotivation langfristig aufgebaut und erhalten bleiben. Vielfach zeigt sich aus der Forschung, dass, noch stärker als die individuelle Leistung, das akademische Selbstkonzept, also die persönliche Wahrnehmung der eigenen Fähigkeiten im fachspezifischen Kontext (Byrne & Shavelson, 1986), einen sehr einflussreichen Prädiktor darstellt, um die Lernmotivation vorherzusagen (Guay et al., 2010; Marsh & Yeung, 1997). Dennoch muss davon ausgegangen werden, dass sich die Lernmotivation nicht nur aus individuellen Merkmalen der einzelnen SchülerInnen erklären lässt, sondern auch durch Faktoren der schulischen Zusammensetzung und des schulischen Kontextes. Im Rahmen des Big-fish-little-pond-Effekts (Marsh, 1987), lässt sich die negative Wirkung eines allgemein hohen Selbstkonzepts in der gleichaltrigen Bezugsgruppe auf die Lernmotivation einzelner SchülerInnen durch soziale Vergleichsprozesse erklären. Obwohl der Einfluss von Kompositionseffekten entsprechend häufig auf die akademische Leistung und das Schulbesuchsverhalten untersucht worden zu sein scheint, legt die bisherige Forschung nur wenig Aufmerksamkeit auf kontextuelle Schul- und Peereffekte (Nagengast & Marsh, 2011). Hier gestaltet sich auch interessant, wie sich beispielsweise Peer-Beziehungen auf die Lernmotivation auswirken (Ryan, 2010).

Der Beitrag geht der Frage nach, welche inner- und außerschulischen Faktoren die Lernmotivation von niederösterreichischen SchülerInnen im inklusiven Setting Neue Mittelschule beeinflussen. Dabei interessieren nicht nur intraindividuelle Einflussfaktoren, sondern auch welche Bedeutung Kompositionseffekten durch die Bezugsgruppe von KlassenkollegenInnen zukommt und ob sich hier zu Beginn und am Ende der Neuen Mittelschule Unterschiede zeigen. Als Analysegrundlage für diesen Beitrag dienen Daten der NOESIS-Längsschnitterhebung (Niederösterreichische Mittelschule in der Schulentwicklung, www.noesis-projekt.at), in der seit dem Schuljahr 2010/11 drei SchülerInnenkohorten von der 4. Klasse Volksschule bis zum Ende ihrer Pflichtschulzeit begleitet werden. In die Analyse fließen die ersten Erhebungszeitpunkte der ersten Kohorte ein (Schulstufe 5 und 8, N=893).

Die Ergebnisse zeigen, dass sich innerschulisch vor allem das Klassenklima, die LehrerInnen-SchülerInnen-Beziehung sowie das akademische Selbstkonzept im Unterrichtsgegenstand Deutsch, aber auch die Ausprägung in der Lernzielorientierung (mastery goal orientation, vgl. Ames & Archer, 1988) und außerschulisch gemeinsame familiäre Tätigkeiten als wirkmächtige Prädiktoren zur Vorhersage der Lernmotivation erweisen. Im Sinne eines Kompositionseffektes scheinen SchülerInnen, die eine Klasse besuchen, in welcher die Kooperation als äußerst gut beschrieben wird und wo regelmäßige außerschulische Aktivitäten mit Familien stattfinden, auch eine erhöhte Lernmotivation aufzuweisen. Interessant zeigt sich im Rahmen der Mehrebenenanalyse zudem, dass diese Kompositionseffekte nicht allein auf die Klassenzusammensetzung der SchülerInnen rückführbar sind, sondern auch die subjektive Sichtweise des aktuellen Standes der eigenen schulischen Fähigkeiten eine signifikante Rolle einnimmt. Die Ergebnisse verweisen damit nicht zuletzt darauf, dass sowohl die Zusammensetzung der SchülerInnen, als auch der Vergleich mit der Peergruppe im Setting des Klassenzimmers einen nicht zu vernachlässigenden Stellenwert bei der Entwicklung der Lernmotivation einnimmt. Im Rahmen der Neuen Mittelschule kann damit in Verbindung vor allem die Qualität von inner-, aber auch außerschulischen sozialen Beziehungen als weiterer Förderansatz verfolgt und die Klasse als soziale Ressource zum Aufbau und Erhalt von akademischer Lernmotivation gesehen werden.

Freitag
13. März

11:15 Uhr

HGB 50

Chair(s)

Beer, Rudolf

3

Rauenfelder, Diana; Jagenow, Danilo

Die Entwicklung sozi-motivationaler (Un-)Abhängigkeit im Jugendalter

Theoretischer Bezugsrahmen: Basierend auf den empirischen Befunden einer ethnographischen Feldstudie, in der sich interindividuelle Unterschiede in der Rolle von Peers und Lehrern/-innen als Motivationsquelle abgezeichnet haben, haben Raufelder (2007) und Kollegen ein Konzept sozio-motivationaler Abhängigkeit postuliert (Raufelder, Jagenow, Drury & Hoferichter, 2013; Hoferichter & Raufelder, 2014). Demzufolge gelten Individuen als sozio-motivational abhängig, wenn ihre eigene Motivation durch die Motivation und das Lernverhalten anderer, sowie durch den Grad, in dem sie von diesen unterstützt werden, beeinflusst wird. Im Gegensatz dazu liegt sozio-motivationale Unabhängigkeit vor, wenn die eigene Motivation unabhängig von der Motivation, dem Lernverhalten und der sozialen Unterstützung anderer ist (Raufelder, Regner, Drury, & Eid, submitted for review). Im schulischen Kontext sind insbesondere Mitschüler/-innen (Peers) und Lehrer/-innen essentielle Motivationsquellen (Wentzel, 1998). Dieses Konstrukt sozio-motivationaler Abhängigkeit konnte mittels einer konfirmatorischen Latenten-Klassen-Analyse (CLCA) an Hand der Daten von 1088 Schüler/-innen der 7ten und 8ten bestätigt werden, insofern vier unterschiedliche Motivationstypen (MT) identifiziert werden konnten (Raufelder, Jagenow, Drury, & Hoferichter, 2013): (1) ein Lehrer-abhängiger MT (9,5 %; 50 Mädchen, 57 Jungen), (2) ein Peer-abhängiger MT (36,5 %; 233 Mädchen, 161 Jungen), (3) ein Peer-und-Lehrer-abhängiger MT (27,8 %; 166 Mädchen, 126 Jungen) und (4) ein Peer-und-Lehrer-unabhängiger MT (26,3 %; 138 Mädchen, 157 Jungen). Um nicht nur den inter-individuellen Unterschieden gerecht zu werden, sondern auch mögliche intra-individuelle Unterschiede im Verlaufe der Adoleszenz zu berücksichtigen, galt es in der vorliegenden Studie mögliche Veränderungen in Hinblick auf das Konstrukt sozio-motivationaler (Un-)Abhängigkeit zu untersuchen.

Fragestellung: Folgende Hypothesen galt es im Detail zu überprüfen:

(1) Auch im weiteren Verlauf der Adoleszenz können mindestens 4 Motivationstypen (MT) unterschieden werden: (1) Lehrer-abhängiger MT, (2) Peer-abhängiger MT, (3) Lehrer-und-Peer-abhängiger MT, (4) Lehrer-und-Peer-unabhängiger MT.

(2) Da die Peers insbesondere im Zuge der Adoleszenz zunehmend an Bedeutung gewinnen und die Qualität der Beziehung zu Erwachsenen im Alter zwischen 12 und 18 Jahren abnimmt (Cook, Deng & Morgano, 2007; Fend, 1998; Brown, 1990; Savin-Williams & Berndt, 1990), ist zu erwarten, dass die Gruppe des Peer-abhängigen MT über die Jugendzeit zunimmt, während die Gruppe des Lehrer-abhängigen MT im Verlauf der Adoleszenz abnimmt.

Methode: Die empirische Grundlage der Studie bildet eine längsschnittliche Fragebogenstudie an Sekundarschulen in Brandenburg mit 1088 Schülern/-innen. Die Schüler/-innen wurden 2011 (Messzeitpunkt 1, 7te und 8te Klasse; Mage = 13.70; SD = 0.53) und 2013 (Messzeitpunkt 2, 9te und 10te Klasse; Mage = 14.86; SE = .57) zur Rolle ihrer Peers und Lehrer/-innen als positive Motivatoren befragt. Um die Zugehörigkeit zu den vier Typen über die Zeit bzw. Veränderungen in Hinblick auf das Konzept sozio-motivationaler (Un-)Abhängigkeit zu untersuchen, wurden latente Transaktionsanalysen mit Mplus 7.0 durchgeführt.

Ergebnisse: Die Ergebnisse der latenten Transaktionsanalysen haben gezeigt, dass die 4-Klassen-Lösung zu beiden Messzeitpunkten die statistisch beste Lösung darstellt. D.h. die Vermutung liegt nahe, dass die Existenz der vier Motivationstypen im Verlauf von früher zur mittleren Adoleszenz eine feste Konstante ist, was auch durch unseren nationalen Vergleich mit kanadischen Schülern/-innen bestätigt wurde (Hoferichter, Raufelder, Eid, & Bukowski, in press). Darüber hinaus zeigen unsere Ergebnisse – entgegen der gängigen Forschungsergebnisse, dass Peers mit zunehmender Adoleszenz wichtiger werden, wohingegen Erwachsenen an Bedeutung verlieren – dass die Gruppe des Lehrer-abhängigen MT im Verlaufe des Jugendalters zunimmt, wohingegen die Gruppe des Lehrer-und-Peer-abhängigen MT abnimmt. Des Weiteren haben wir einen starken Wechsel in den drei Gruppen sozio-motivational-abhängiger Schüler/-innen aufzeigen können, wohingegen die Gruppe sozio-motivational-unabhängiger Schüler/-innen stabil bleibt (nahezu keine Wechsel), was das Konzept sozio-motivationaler (Un-)Abhängigkeit bestätigt. Anhand der hier skizzierten Typologie lassen sich wichtige Hinweise auf Unterschiede in der Motivation adoleszenter Schüler/-innen finden, die in Methoden des individualisierten Unterrichts und in der Förderung sozialer Kompetenz innerhalb des schulischen Kontextes ihre Entsprechung finden können. Schulpraktische Implikationen im Sinne heterogener Motivationsmuster werden diskutiert.

Freitag
13. März

11:15 Uhr

HGB 50

Chair(s)

Beer, Rudolf

Hoferichter, Frances; Rauenfelder, Diana

The impact of peer and teacher relationships on motivation of Canadian and German secondary school students

Theoretical Background: Classroom settings function not only as educational arenas but also as a powerful social environment dominated by interactions between students and teachers (Anderson, Christenson, Sinclair & Lehr, 2004; Roeser, Eccles & Sameroff, 2000). Social interactions within the school context affect different academic factors including students' motivation (Wentzel, Battle, Russel & Looney 2010). Based on these premises, Raufelder, Jagenow, Drury and Hoferichter (2013) have described the importance of inter-individual differences concerning the role of peers and teachers for students' motivation. Through an analysis of student's scholastic motivation and social relationships with peers and teachers conducted in the state of Brandenburg, Germany, the researchers identified four different motivation types (MTs) with the following distribution in percentage: teacher-dependent MT (10%), peer-dependent MT (34%), teacher-and-peer-dependent MT (27%) and teacher-and-peer-independent MT (29%) (Raufelder et al., 2013).

Research Question: Following previous research (Raufelder et al., 2013), we were interested in potential differences in adolescent students' perception of peers and teachers as source of motivation between Canadian and German students – two nations labeled as individualistic societies (Triandis, 1995) but with significantly different educational systems, including different social settings (cf. Beckmann, 2000; Kopp & Schmitt, 2004). For example, German students tend to focus on peers as motivational support as well as on their individuality, whereas the relationships with teachers are described as informal and distant (Hesse, 2004). In contrast, Canadian secondary school students tend to perceive both peers and teachers as supportive. In addition, the Canadian educational system places more emphasis on school as an important social environment by focusing and maintaining the development of social skills (Wilson & Lam, 2004).

We tested the following hypotheses: (1) Four different motivation types (MT) can be identified among Canadian students. (2) The distribution of the students within the typology varies among the countries. In the Canadian sample, the teacher-and-peer-dependent MT is expected to be the largest group, followed by the teacher-and-peer-independent MT.

Method: A total of 1477 seventh and eighth grades students from Germany and Canada participated in this study during the school years 2011/2012. The study was conducted in the state of Brandenburg. A total of 1088 German secondary students (54% girls) (Mage = 13.7, SD = 0.53, age span 12 –15 years) answered questions about their socio-motivational relationships with peers and teachers.

A total of 389 Canadian secondary students (55.9% girls) (Mage = 13.4, SD = 0.80, age span 12-16 years) participated in the study. The Canadian students from randomly selected English schools in the Montréal region answered the same questionnaire as the German students but did so in English.

All schools were randomly selected with permission from the governmental Department of Education in Brandenburg and Québec and parental permissions.

Results: By using Multigroup Confirmatory Latent Class Analysis (MCLCA), we were able to confirm four different motivation types (MT) among Canadian. These results are in line with previous research (Raufelder, 2007; Raufelder et al., 2013; Hoferichter & Raufelder, 2014), in which inter-individual differences among students were found concerning the impact of teachers and peers on motivational patterns. The current study extends this research by addressing cross-national aspects within the motivational typology: Although the pattern of the typology is the same in both countries, German and Canadian students differ in the distribution of the typology. In Canada, the distribution was as follows: teacher-dependent MT (9%), peer-dependent MT (14%), teacher-and-peer-dependent MT (57%) and teacher-and-peer-independent MT (20%). Besides the existence and distribution of the motivational typology across Germany and Canada, the study clearly showed that the environment, particularly a certain orientation and frame of the educational system, may be linked to specific roles assigned to peers and teachers in the motivational context across countries.

George, Ann Cathrice

Kompetenzmodellierung mit kognitiv-diagnostischen Modellen: Validierung und empirische Prüfung

Dieses Symposium beleuchtet aus einer interdisziplinären Perspektive der Disziplinen Wirtschaftspädagogik, Linguistik, Psychometrie und Mathematikdidaktik Aspekte der Kompetenzmodellierung unter besonderer Berücksichtigung von kognitiv-diagnostischen Modellen. Kognitiv-diagnostische Modelle (CDMs; Rupp, Templin & Henson, 2010) sind mehrdimensionale Klassifikationsmodelle, die SchülerInnen in Profile beherrschter und nicht beherrschter Fähigkeiten klassifizieren. Im Unterschied zu mehrdimensionalen IRT-Modellen werden Fähigkeiten in CDMs meistens als dichotom (beherrscht vs. nicht beherrscht) angenommen.

Das Symposium streift Aspekte der Testkonstruktion und der Validierung vor der Durchführung statistischer Analysen (erster Beitrag: Bley), wie die Verwendung der Ergebnisse der Kompetenzmodellierung im Allgemeinen (letzter Beitrag: Leuders). Die empirische Prüfung von Hypothesen mit Hilfe von CDMs wird im zweiten Beitrag (Müller und George) sowie im dritten Beitrag (George) vorgenommen.

Das Symposium ist wie folgt strukturiert:

In einem ersten Beitrag (Bley) aus dem Gebiet der Wirtschaftspädagogik wird ein Verfahren zur Validierung einer Q-Matrix auf Basis verschiedener Informationsquellen (verschiedener Expertengruppen) zur Erfassung der „Intrapreneurship“-Kompetenzen vorgeschlagen. Die vorgeschlagene Methode erhöht die Validität der abgeleiteten Q-Matrix. Als Ergebnis resultiert eine finale Q-Matrix mit einer reduzierten Anzahl von Kompetenzfacetten.

Im zweiten Beitrag (Müller und George) werden die in IGLU/PIRLS spezifizierten Leseverstehensprozesse näher betrachtet. Es wird untersucht, ob die (nahezu) hierarchische Struktur der Leseverstehensprozesse durch verschiedene Anforderungen der Items auf Sprachebenen zurückführbar ist, die sich hierarchisch anordnen lassen (Wort->Satz->Text). Diese Hypothese wird im Rahmen von CDMs beantwortet.

Der dritte Beitrag (George) greift die Modellierung von Gruppenvergleichen mit Hilfe kognitiver-diagnostischer Modelle für mathematische Kompetenzen auf. Geschlechter- und Schulformvergleiche werden im Rahmen von CDMs für die verschiedenen inhaltlichen mathematischen Kompetenzen und Handlungskompetenzen durchgeführt. Dabei wird insbesondere auf die Herausforderungen der Anwendung von CDMs im large-scale assessment eingegangen. Ein alternativer Ansatz für Effektstärken des differentiellen Itemfunktionierens wird vorgeschlagen.

Im vierten Beitrag (Leuders) wird die Frage der Validität von Kompetenzmodellierungen aus fachdidaktischer Perspektive beleuchtet, wobei einige prototypischer Ansätze, deren Probleme und Lösungen diskutiert werden. Der Beitrag schließt mit einer vergleichenden Bewertung unterschiedlicher Typen von Kompetenzmodellierungen aus der Sicht ihrer Nutzung für den Mathematikunterricht.

Freitag
13. März

11:15 Uhr

HGC 20

Chair(s)

Bley, geb. Trost, Sandra
George, Ann Cathrice**Kompetenzmodellierung mit kognitiv-diagnostischen Modellen: Validierung und empirische Prüfung**

Bley, Sandra

Validierung eines kognitiven Diagnose-Assessments zur Messung der kaufmännischen Kompetenz „Intrapreneurship“

Theoretischer Hintergrund: In der aktuellen Diskussion zur Kompetenzmessung ist ein verstärktes Interesse an detaillierteren Kompetenzinformationen über Lerner (i.S. eines diagnostisch formativen Assessments) zu verzeichnen. Mit Hilfe von kognitiven Diagnosemodellen werden individuelle multiple Kompetenz-Profile erzeugt, die auf Stärken und Schwächen des Lerners schließen lassen (u.a. Rupp, Templin & Henson, 2010).

Die Anwendung kognitiver Diagnosemodelle erbrachte in der Vergangenheit häufig unerwünschte Schätzprobleme, die zumeist mit fehlender kognitiver Validität (Fehlspezifikation der Q-Matrix) oder fehlender Inhaltsvalidität (Vernachlässigung der Fachdidaktik bzw. einer systematischen Domänenanalyse) begründet wurden (u.a. Leuders, 2014).

Das zugrundeliegende Kompetenzverständnis dieses diagnostischen Ansatzes sieht vor, dass eine domänenrelevante Situation (hier im Bereich des Intrapreneurship: = Unternehmertum im Unternehmen) nur dann zu bewältigen ist, wenn der Lerner die dafür notwendigen domänenspezifischen Kompetenzen besitzt und diese anwendet. Für die Modellierung diagnostischer Tests ist zunächst die Identifikation der diagnostizierbaren Kompetenzfacetten notwendig. Weber et al., (2014) identifizierten 14 Intrapreneurship-Kompetenzfacetten, auf deren Basis sie mehr als 50 Aufgaben konstruierten. Jede der Aufgabe erfordert einen anderen kognitiven Lösungsprozess resp. ein anderes Bündel von Intrapreneurship-Kompetenzfacetten. Diese Zusammenhänge (Kompetenzfacetten-Aufgaben-Interaktionen) in Form der sogenannten Q-Matrix müssen für die statistische Analyse a priori definiert werden.

Fragestellung: Ziel des Beitrages ist die elaborierte Konstruktion der Q-Matrix zur Prüfung der kognitiven Validität. Dabei soll insbesondere gefragt werden, inwiefern (1) die identifizierten Intrapreneurship-Kompetenzfacetten sich über die konstruierten Items abbilden lassen und (2) ob die Items wiederum unterschiedliche Bündel an Kompetenzfacetten zur Lösung benötigen?

Methode: Zur Validitätsprüfung der Q-Matrix erfolgt eine Triangulation verschiedener Informationsquellen (1. Fachdidaktiker, 2. Experten aus Schule und Betrieb (N=9), 3. Zielgruppe: Auszubildende (N=26)), Erhebungsmethoden (1. Experteninterviews, 2. Studie lauten Denkens) und Auswertungsmethoden (1. konzeptionell-entwickelnd auf Basis fachdidaktischer Theorien sowie Ergebnissen einer vorangegangenen Domänenanalyse (Weber et al, 2014), 2. Inhaltsanalyse, Quantifizierung qualitativer Daten und Cut-Off-Bewertung, 3. Prüfung auf signifikante Zusammenhänge von Intrapreneurship-Kompetenzfacetten (genannt/nicht genannt) und Aufgabenlösung (richtig/falsch gelöst) für jede Aufgaben-Kompetenz-Interaktion auf Basis der quantifizierten Transkripte des Lauten-Denkens). Mit Hilfe dieses Vorgehens ergeben sich vier eigenständige Q-Matrizen: 1. Q-Fachdidaktiker, 2. Q-Experten, 3. Q-Auszubildende-Cut-off und 4. Q-Auszubildende-Signifikanzen, die im Anschluss übereinander gelegt die finale Q-Matrix ergeben. Als relevante Kompetenzfacetten-Aufgaben-Interaktionen sollen diejenige angesehen werden, die in mindestens 3 der oben genannten 4 eigenständigen Q-Matrizen zu finden sind.

Ergebnisse: Ergebnis der Studie ist eine elaboriert entwickelte Q-Matrix mit 10 Kompetenzfacetten und 51 Aufgaben für die Messung von Intrapreneurship-Kompetenz. Von den 14 ursprünglich identifizierten Kompetenzfacetten zum Intrapreneurship konnten 10 Facetten validiert werden. Als besonders wichtig – da sie laut der finalen Q-Matrix sehr häufig zum Lösen der Aufgaben herangezogen werden – sind u. a. die Facetten: „Fachbegriffe anwenden“ und „Informationen selektieren, bewerten, usw.“ Die vier Facetten: „Risikoanalyse“, „Umgang mit Störungen“, „Reflexion“ und „Anwenden von Kreativitätstechniken“ konnten nicht bestätigt werden. Dieser Effekt könnte sich entweder auf die Konstruktion der Aufgaben zurückführen lassen, die noch in unzureichendem Maße eine kognitive Aktivierung der o.g. Facetten ermöglichen oder auf den Kompetenzfortschritt der Zielgruppe. Zielgruppe sind Auszubildende zum/r Industriekaufmann/-frau am Ende ihrer beruflichen Erstausbildung. Die konstruierten Items benötigen zur richtigen Lösung mindestens einen der 10 final identifizierten Kompetenzfacetten. Über alle Items hinweg ergeben sich vielfältige Variationen an notwendigen Kompetenzfacetten-Bündeln. Weitere fachdidaktische Analysen zur Charakterisierung der Domäne (z. B: die Hierarchie oder die Kompensierbarkeit zwischen den Kompetenzfacetten) zur Auswahl eines geeigneten psychometrischen Verfahrens aus der Familie der kognitiven Diagnosemodelle laufen derzeit. Die Messung auf Basis einer Stichprobe von ca. 1200 Auszubildenden stellt einen nächsten Schritt dar.

Freitag
13. März

11:15 Uhr

HGC 20

Müller, Claudia; George, Ann Cathrice

Empirische Modellierung und Prüfung von Lesekompetenzmodellen mit Kognitiv Diagnostischen Modellen

Lesekompetenzen stellen in modernen gesellschaftlichen Zusammenhängen Grundfähigkeiten dar, die für Bildungserfolg und gesellschaftliche Teilhabechancen ausschlaggebend sind. Studien wie PISA oder IGLU haben ergeben, dass ein beträchtlicher Anteil der SchülerInnen in Deutschland über elementare Lesefähigkeiten nicht hinaus kommt (z.B. 13,2% in IGLU 2006). Auch in der Pilotierung der Bildungsstandards Deutsch in Österreich (D4P) zeigt sich diese Tendenz. Was allerdings unter Lesefähigkeit zu verstehen ist, wird von verschiedenen Fachdisziplinen unterschiedlich beantwortet (vgl. Venn-Brinkmann, 2012). So wird etwa in der Studie D4P zur Lesefähigkeit von Viertklässlern ein Lesekonzept verwendet, das drei Kompetenzfelder fokussiert (vgl. BIFIE, 2009): „Explizite Informationen ermitteln“, „allgemeines Textverständnis“, „textbezogenes Interpretieren und Reflektieren“. Diese Kompetenzfelder ähneln den vier Leseverstehensprozessen, wie sie in IGLU eingesetzt werden (Bos, Valtin, Voss, Hornberg, & Lankes, 2007, S. 85); sowohl von ihrer Kurzbeschreibung her gesehen als auch in ihren Deskriptoren. In der Forschung herrscht Uneinigkeit darüber, ob diese vier Leseverstehensprozesse als parallel oder hierarchisch aufeinander aufbauend zu sehen sind (vgl. auch Bremerich-Vos, 1996). Während die vier Leseverstehensprozesse in der Rahmenkonzeption zu IGLU ursprünglich als parallel angesehen wurden (Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001), wird in aktuellen Arbeiten aufgrund der festgestellten hohen Korrelationen zwischen den vier Leseverstehensprozessen (Voss, 2005) davon ausgegangen, dass die vier Bereiche empirisch nicht unterscheidbar seien. Unter anderem mit diesem Argument werden Analysen der Daten (sowohl in IGLU als auch in der D4P) auf Grundlage eines eindimensionalen Gesamtscores durchgeführt (Bos et al., 2007).

Um nach der quantitativen Modellierung beschreiben zu können, über welche Lesekompetenzen getestete SchülerInnen verfügen, werden in einer Standardsetting-Prozedur qualitative Kompetenzstufen definiert, die mit bestimmten Bereichen des Gesamtscores übereinstimmen. Dieses Verfahren impliziert eine Hierarchie zwischen den Kompetenzstufen: SchülerInnen, die beispielsweise Lesefähigkeiten auf der Kompetenzstufe III haben, zeigen folglich auch Fähigkeiten auf den Stufen I und II. Aus linguistischer Sicht liegt die Vermutung nahe, dass die hohen Korrelationen zwischen den Kompetenzbereichen (und damit die postulierte Hierarchie) dadurch bedingt sein kann, dass in dem angenommenen Lesemodell die Items des ersten Kompetenzbereiches die Wortebene, Items des zweiten Kompetenzbereichs die Satzebene und Items des dritten und vierten Kompetenzbereichs vor allem die Textebene fokussieren. Die Ebenen können insofern als hierarchisch miteinander vernetzt interpretiert werden. Somit wäre die Sprachebene als Mediatoreffekt auf die Struktur zwischen den Kompetenzbereichen zu sehen.

Die hier vorgestellte Studie greift diesen Aspekt auf und zielt darauf ab, eine erste empirische Basis zu liefern, mit deren Hilfe eine Aussage über die Struktur zwischen den Kompetenzfeldern (bzw. Leseverstehensprozessen) unter Beachtung der Sprachebenen getroffen werden kann. Dazu werden mittels Kognitiv Diagnostischer Modelle (vgl. Rupp, Templin, & Henson, 2007) die beiden verschiedenen theoretischen Annahmen (parallel vs. hierarchisch) statistisch modelliert und mit unterschiedlichen Modellgütekriterien empirisch verglichen. Datengrundlage bilden die vom BIFIE erhobenen Daten zur D4P 2013 (N=2430) sowie die darin genutzten Items (vgl. BIFIE, 2010). Zur Konstruktion der Q-Matrix werden die Items gesichtet und nach linguistischen Kriterien den oben genannten drei Kompetenzfeldern, den vier Leseverstehensprozessen sowie den drei Sprachebenen zugeordnet. In einer Vorstudie anhand eines sehr kleinen Datensatzes ergab sich, dass Modelle mit der Annahme von 4 parallelen oder 4 hierarchisch aufeinander aufbauenden Leseverstehensprozessen eine signifikant bessere Anpassung an die Daten zeigen als ein Modell mit nur einem Leseskill. Außerdem zeigte sich eine sehr hohe Korrelation zwischen den Leseverstehensprozessen 2 und 3, so dass ein Modell mit nur 3 parallelen Leseverstehensprozessen die Daten signifikant besser beschrieb als ein Modell mit 4 parallelen Leseverstehensprozessen.

Die Ergebnisse der Studie ermöglichen es nicht nur, Lesekonzepte, wie sie in IGLU und D4P verwendet werden, auszudifferenzieren, sondern auch langfristig zu einer kompetenzorientierten Weiterentwicklung von Studien mit individual-diagnostischem Schwerpunkt beizutragen. Darüber hinaus lassen sich auf der Basis der Ergebnisse Förderkonsequenzen für den Unterricht formulieren, die an die einzelnen Sprachebenen anschließen (vgl. Venn-Brinkmann, 2012).

Freitag
13. März

11:15 Uhr

HGC 20

Chair(s)

Bley, geb. Trost, Sandra
George, Ann Cathrice**Kompetenzmodellierung mit kognitiv-diagnostischen Modellen: Validierung und empirische Prüfung**

George, Ann Cathrice

Herausforderungen bei Gruppenvergleichen mit kognitiv diagnostischen Modellen im Large-Scale Assessment

In jüngerer Zeit wird der Einsatz kognitiv diagnostischer Modelle (Cognitive Diagnosis Models, CDMs; z.B. Rupp, Templin, & Henson, 2010) in der Modellierung mehrdimensionaler Kompetenzen vorgeschlagen. Auf der Annahme basierend, dass die Testdomäne (z.B. Mathematik) in mehrere Grundkompetenzen (z.B. Zahlen, Variablen, Geometrie und Statistik) aufgegliedert ist, klassifizieren CDMs Schüler hinsichtlich ihrer Fähigkeiten in diesen Grundkompetenzen in kategoriale mehrdimensionale Skillklassen. Da mit diesem Vorgehen detailliertere Information über die Schülerleistungen gewonnen werden kann als mit einem eindimensionalen Testwert, werden die Skillklassen in Anwendungen als solide empirische Basis zur Entwicklung gezielter Fördersysteme gesehen.

Im Zusammenhang mit der einhergehenden Aufmerksamkeit auf die Modellklasse der CDMs wurden einige in der IRT-Methodik bereits verbreitete Methoden wie Mehrgruppenansätze (z.B. Xu & von Davier, 2008) und die Analyse differentiellen Itemfunktionierens (DIF; Hou, de la Torre, & Nandakumar, 2014) adaptiert. Damit können nicht nur die Grundkompetenzen einer Testpopulation analysiert werden, sondern auch Gruppenvergleiche detaillierter durchgeführt werden. Zum Beispiel kann die Aussage „Jungen besitzen höhere Mathematik-Kompetenz als Mädchen“ differenziert, d.h. unter Einbeziehung der Frage ob Jungen in jeder Grundkompetenz besser abschneiden als Mädchen, beantwortet werden.

Neben den genannten Stärken der Modellklasse stellt der Einsatz von CDMs im Large-Scale Assessment einige methodische Herausforderungen: Insbesondere steigt mit wachsender Dimensionalität der Kompetenzen die Anzahl der Modellparameter, woraus Probleme bei der Schätzung der Modelle entstehen. Weiterhin herrscht ein Defizit an etablierten Effektstärken (z.B. bei DIF-Analysen).

Im Vortrag werden die Möglichkeiten von Gruppenvergleichen mit CDMs anhand von Daten aus der flächendeckenden österreichischen Testung der Bildungsstandards 2012 in Mathematik der achten Schulstufe illustriert. In der Anwendung wird auf die Herausforderungen der Anpassung von Mehrgruppen-CDM-Modellen an den Datensatz mit 71464 Schülern und 72 Items eingegangen und Lösungsansätze präsentiert. Substanziell zeigt sich, dass Geschlechterunterschiede wenig spezifisch auf den vier zugrundeliegenden mathematischen Inhalts- und vier mathematischen Handlungsbereichen ausgeprägt sind. Dagegen fallen Schulformunterschiede in den Inhaltsbereichen verschieden stark aus.

Freitag
13. März

11:15 Uhr

HGC 20

Chair(s)

Bley, geb. Trost, Sandra
George, Ann Cathrice**Kompetenzmodellierung mit kognitiv-diagnostischen Modellen: Validierung und empirische Prüfung**

Leuders, Timo

Modellierungen mathematischer Kompetenzen – Kriterien für eine Validitätsprüfung aus fachdidaktischer Sicht

Die theoretischen und empirischen Bemühungen um die Modellierung, Erfassung und Rückmeldung von Kompetenzen in pädagogischen Kontexten beschäftigen die Bildungssystemforschung, die pädagogische Psychologie und die Fachdidaktiken bereits seit mehreren Jahrzehnten (z.B. Wilson, 2005; Hartig, Klieme & Leutner, 2008; Stecker, Fuchs & Fuchs, 2005; Mislevy, 2006; Bernholt, Parchmann & Commons, 2009; Niss & Højgaard, 2011). Als Kompetenzmodellierung werden dabei Ansätze verstanden, die den mehrschrittigen Prozesscharakter der theoretischen Beschreibung, der psychometrischen Modellierung, der Messung und schließlich der Nutzung in Diagnostik und Assessment explizit machen (Klieme & Leutner 2006, Pellegrino, Chudowsky & Glaser, 2001).

Ziel dieses Beitrags ist es, Ansätze der Kompetenzmodellierung unter zwei einschränkenden Perspektiven darzustellen: Die Diskussion konzentriert sich einerseits auf die Modellierung mathematischer Kompetenzen und andererseits auf die Frage der Validität solcher Kompetenzmodellierungen. Dies soll anhand einiger prototypischer Ansätze geschehen, die jeweils für einen bestimmten Strukturtyp von Kompetenzmodellen stehen und die jeweils typische Probleme und Lösungen hinsichtlich ihrer Validität aufzeigen. Die Frage der Validität wird dann vor allem aus fachdidaktischer Perspektive beleuchtet. Die dafür nötige theoretische Präzisierung des Validitätsbegriffes findet man in dem umfassenden Validitätskonzept von Messick (1995), der Validität als die Gesamtbewertung der theoretischen Argumente und empirischen Evidenzen für die Angemessenheit einer Leistungsmessung (Assessment) und zwar sowohl ihrer Interpretation als auch der Konsequenzen ihrer Anwendung, versteht.

Gegenstand der abschließenden Diskussion ist die vergleichende Bewertung unterschiedlicher Typen von Kompetenzmodellierungen aus der Sicht ihrer Nutzung für den Mathematikunterricht. Hervorgehoben wird dabei die Rolle der interdisziplinären Zusammenarbeit und der fachdidaktischen Fundierung kognitiver Prozesse, unter anderem anhand einer aktuell laufenden eigenen Studie zur Modellierung von arithmetischen Kompetenzen am Beginn der Klasse 5 (vgl. Schulz & Leuders auf dieser Tagung).

Freitag
13. März

11:15 Uhr

HGC 20

Chair(s)

Bley, geb. Trost, Sandra
George, Ann Cathrice**Kompetenzmodellierung mit kognitiv-diagnostischen Modellen: Validierung und empirische Prüfung**

Klapproth, Florian

Effekte von Klassenwiederholung auf die leistungsbezogene und psychoemotionale Entwicklung von Schülerinnen und Schülern: Heterogenität in empirischen Zugängen und Befunden?

Der Anteil 15-Jähriger, die bis zum Ende ihrer Pflichtschulzeit mindestens einmal eine Klasse wiederholt hatten, lag 2009 bundesweit bei 14,2 Prozent (Eurydice, 2011). Begründet werden Klassenwiederholungen als Maßnahme zur Reduktion von Leistungsheterogenität innerhalb von Lerngruppen (Tillmann, 2008) damit, den Leistungsschwächsten eine Chance zu bieten, nach einer Klassenwiederholung wieder an das Leistungsniveau ihrer neuen Lerngruppe aufzuschließen. Vor dem Hintergrund der durch Klassenwiederholungen verursachten Kosten für das deutsche Bildungssystem, die rund 1 Milliarde Euro jährlich betragen, wird das Sitzenbleiben hingegen als „teuer und unwirksam“ (Klemm, 2008) angesehen.

Theoretisch und empirisch sind die Auswirkungen von Klassenwiederholungen jedoch stark umstritten: So zeigen einige Studien in Bezug auf die erzielten Schulleistungen weder für die Repetenten noch für die verbleibenden Klassenverbände mittel- und langfristig Vorteile des zweimaligen Durchlaufens derselben Klassenstufe (z. B. Bless, Schüpbach & Bonvin, 2004; Einsiedler & Glumpler, 1989; Roßbach & Tietze, 2010). Hinsichtlich der psychoemotionalen Entwicklung zeigt sich, dass eine Nicht-Versetzung die betroffenen Schülerinnen und Schüler demotivieren und zu Schulunlust und gesteigerten Selbstzweifeln führen kann (Martin, 2011).

Das Symposium geht der Frage nach, welche Auswirkungen Klassenwiederholungen auf schulische Leistungen und die psychoemotionale Entwicklung der Schülerinnen und Schüler haben. Dabei wird zum einen die Heterogenität der hierzu vorliegenden theoretischen und empirischen Arbeiten aufgegriffen. Zum anderen will das Symposium auf Grundlage der in den Beiträgen angewandten Analysemethoden auf die Problematik des empirischen Nachweises kausaler Effekte von Klassenwiederholungen hinweisen. Ziel des Symposiums ist es daher, interdisziplinär aus erziehungswissenschaftlicher, psychologischer, soziologischer und bildungsökonomischer Perspektive nationale und internationale Studien und Befunde zur Effektivität von Klassenwiederholungen zu berichten und zu kontrastieren.

Der Beitrag von Schüpbach gibt einen Überblick über aktuelle empirische Studien zu Effekten von Klassenwiederholungen auf die Leistungsentwicklung und den weiteren Schul-, Ausbildungs- und Erwerbsverlauf der Repetenten.

Fabian, Scharenberg und Bos prüfen anhand der NEPS-Daten unterschiedliche methodische Ansätze zur Analyse der Frage, inwiefern Klassenwiederholungen die Leistungsentwicklung beeinflussen und über welche Schüler- und Kontextmerkmale deren Auswirkungen mediiert werden.

Klapproth et al. präsentieren Ergebnisse einer luxemburgischen Studie, in der anhand von Propensity-Score-Matching Effekte von Klassenwiederholungen auf unterschiedliche Leistungsindikatoren (Schulnoten, Testergebnisse) und die psychoemotionale Entwicklung (Selbstkonzept, Schulangst) der Schülerinnen und Schüler geschätzt werden.

Schwerdt und West stellen eine Studie aus Florida (USA) vor, in der ein quasi-experimenteller bildungsökonomischer Ansatz gewählt wurde, um Auswirkungen von Klassenwiederholungen in der Grundschule auf die Leistungsentwicklung bis zum Ende der Sekundarstufe I aufzudecken.

Neuenschwander diskutiert die Beiträge hinsichtlich der methodischen Herausforderungen bei der empirischen Analyse von Effekten der Klassenwiederholung und ihrer Implikationen für die pädagogische Praxis.

Freitag
13. März

11:15 Uhr

HGC 40

Chair(s)

Klapproth, Florian
Scharenberg, Katja

Effekte von Klassenwiederholung auf die leistungsbezogene und psychoemotionale Entwicklung von Schülerinnen und Schülern: Heterogenität in empirischen Zugängen und Befunden?

Schüpbach, Marianne

Kurz-, mittel- und längerfristige Wirkungen der Klassenwiederholung – Theoretische und empirische Zugänge

Klassenwiederholungen werden in der Regel dann angeordnet, wenn bei einem Kind ungenügende Schulleistungen festgestellt werden. Ein zusätzliches Schuljahr – die Wiederholung der gleichen Klassenstufe im darauffolgenden Schuljahr entweder im gleichen oder in einem anderen Schultypus – soll ihm ermöglichen, die Ziele des Lehrplanes zu erreichen. Durch eine Klassenwiederholung wird dem Kind mehr Lernzeit gewährt. Gleichzeitig manifestiert eine Klassenwiederholung aus soziologischer Perspektive den selektiven und allokativen Charakter des Bildungssystems. In diesem Zusammenhang wird auch von „formalem Schulversagen“ gesprochen, wobei es sich um eine „nicht (mehr) tolerierbare Diskrepanz zwischen schulisch gesetzten Entwicklungsnormen und den Lernvoraussetzungen bzw. dem Leistungsstand von Schülern“ handelt (Hildeschmidt, 1998, S. 990). Bei der Klassenwiederholung handelt es sich also um einen von verschiedenen Schullaufbahnübergängen während der gesamten Schulzeit. In diesem Zusammenhang wird auch von Transition gesprochen. Traditionellerweise werden in der Soziologie schulische Ausbildung und Bildungsabschlüsse unter einer Lebenslaufperspektive erklärt, die life transitions, also Transitionen, Lebensübergänge, als Wendepunkte betrachtet (Entwisle, Alexander & Olson, 2003). Individuen entwickeln im Laufe ihrer Lebenszeit soziale Rollen in der Arbeit und der Familie, die sich über die Zeit auch wieder verändern. Ausschlagend für deren Entwicklung sind vor allem markante Übergangsphasen wie eine Klassenwiederholung. Gerade solche Transitionen in der Schule erweisen sich als entscheidend für die künftige Entwicklung der Kinder, weil sie sowohl die künftige Schul- als auch die Ausbildungs- und Berufslaufbahn prägen (Elder & Shanahan, 2003). Wenn ein Kind in den ersten Jahren repetiert, beeinträchtigt dies somit nicht nur das entsprechende Schuljahr, sondern auch den weiteren Lebensverlauf des Kindes.

Ältere, weitgehend US-amerikanische Untersuchungen zeigen insgesamt negative mittelfristige Effekte der Klassenwiederholung auf die Schulleistung (Smith & Shepard, 1987). Insgesamt geht hervor, dass Repetierende in den darauffolgenden Schuljahren leistungsmäßig schlechter abschneiden als die regulär promovierten Peers, auch wenn die relevanten Leistungsvariablen und familialen Variablen kontrolliert werden. In neuere Untersuchungen (seit 1990) sind weitgehend Längsschnittstudien mit unterschiedlichen Vergleichen, same-age- und same-grade-Vergleiche, zu finden. Die Ergebnisse der überwiegend US-amerikanischen Längsschnittstudien dokumentieren größtenteils ein negatives Bild der Klassenrepetition (Hess, 2010). Besorgniserregend scheint die Langzeitwirkung dieser Maßnahme zu sein. Die Klassenwiederholung wird zu einer wichtigen Prädiktorvariable für einen späteren negativen Verlauf der Schulkarriere. Nach Alexander et al. (2004) kann die Klassenwiederholung Probleme in jedem Alter verursachen, aber besonders in der Adoleszenz, zu Beginn der Pubertät und beim Übergang in die Middle School.

Im Rahmen dieses Beitrags soll der Frage nachgegangen werden, (1) wie sich Repetierende in neueren Studien kurz- und mittelfristig hinsichtlich ihrer Schulleistung entwickeln und (2) welche Langzeitwirkungen hinsichtlich des weiteren Schul-, Ausbildungs- und Erwerbsverlaufs eine Klassenwiederholung zeigt. Dazu soll einerseits ein Forschungsüberblick über die neusten Studien zu diesen Fragestellungen gegeben und andererseits das methodische Vorgehen und die Ergebnisse der eigenen Studie Intsep-R, einer quasi-experimentellen Längsschnittuntersuchung bei 4.248 Grundschülerinnen und -schülern in der Deutschschweiz und in der französischsprachigen Schweiz, vorgestellt werden. Im Rahmen dieser Studie (Bless, Schüpbach & Bonvin, 2004) wurden zwei Versuchsgruppen – Repetierende des zweiten Schuljahres der deutschsprachigen und der französischsprachigen Schweiz – gebildet. Dies ist bis heute eine der wenigen Studien im deutschsprachigen Raum geblieben. Pro Versuchsgruppe wurde zudem durch Matching eine Kontrollgruppe von vergleichbaren schulleistungsschwachen, aber in das dritte Schuljahr promovierten Schülerinnen und Schülern, zusammengestellt. Abschließend sollen Forschungsdefizite herausgearbeitet werden.

Freitag
13. März

11:15 Uhr

HGC 40

Chair(s)

Klapproth, Florian
Scharenberg, Katja

Effekte von Klassenwiederholung auf die leistungsbezogene und psychoemotionale Entwicklung von Schülerinnen und Schülern: Heterogenität in empirischen Zugängen und Befunden?

Fabian, Paul; Scharenberg, Katja; Bos, Wilfried

Die Leistungsentwicklung von Schülerinnen und Schülern nach einer Klassenwiederholung – Übersicht über längsschnittliche Analysemöglichkeiten mit den Daten der NEPS-Studie

Klassenwiederholungen sind eine national wie international höchst umstrittene Homogenisierungsmaßnahme. In Deutschland befinden sich einige Bundesländer (z. B. Hamburg oder Berlin) im Prozess der Abschaffung von Klassenwiederholungen. In anderen Bundesländern (z. B. Bayern oder Bremen) wiederholen jedoch weiterhin überdurchschnittlich viele Schülerinnen und Schüler eine Klasse.

Auch die empirische Forschungslage ist diffus: Im englischsprachigen Raum wurden Effekte von Klassenwiederholungen auf die Leistungsentwicklung umfangreich erforscht (z. B. Jimerson, 2001a, 2001b; Penfield, 2010). In Deutschland finden sich allerdings nur wenige belastbare Studien, die konkrete Aussagen zur Effektivität von Klassenwiederholungen zulassen (z. B. Ehmke, Drechsel & Carstensen, 2008). In den letzten zehn Jahren prägte allerdings eine umfassende Methodendiskussion und -kritik die Forschungsarbeiten über Klassenwiederholungen.

Es gibt verschiedene theoretische Ansätze, die Effekte von Klassenwiederholungen zu überprüfen. Unser Beitrag geht deshalb der Frage nach, ob eine Klassenwiederholung zu einer Leistungskonsolidierung oder einer Leistungssteigerung führen kann (Fabian, 2012). Ein weiterer Analyseansatz ist gerechtigkeits-theoretisch geprägt: Wer wiederholt eine Klasse und warum? Einige Studien (z. B. Bless, Schüpbach & Bonvin, 2004) konnten Hinweise darauf liefern, dass leistungsfremde Merkmale wie die Einstellung von Lehrkräften zur Klassenwiederholung einen signifikanten Anteil an der Repetitionsentscheidung haben.

Die hier vorgestellte Studie soll den Aspekt der Leistungsentwicklung von Repetentinnen und Repetenten aufgreifen und dabei der o. g. Methodendiskussion Rechnung tragen. Deshalb ist eine Sekundäranalyse geplant, die folgende Kritikpunkte aufgreift:

(1) Vergleichsgruppenbildung:

Um Verzerrungen in den Analysen interpretierbar zu machen, müssen Repetentinnen und Repetenten sowohl mit Schülerinnen und Schülern des abgebenden Jahrgangs (same-age Vergleich) sowie mit denen der aufnehmenden Klasse (same-grade Vergleich) verglichen werden (Karweit, 1999).

Um die Leistungsentwicklung zu kontextualisieren, müssen die Repetentinnen und Repetenten mit vergleichbar leistungsschwachen Schülerinnen und Schülern ohne Klassenwiederholung verglichen werden, da sich die Ausgangsleistung als stärkster Prädiktor für spätere Schulleistungen erwiesen hat (z. B. Gröhlich, Guill, Scharenberg & Bos 2010a, 2010b).

(2) Längsschnittliche Analysen: Zur Analyse von Leistungsentwicklungen sind Längsschnittdaten daher unerlässlich.

(3) Analysemethodik: Studien, die auf Regressionsanalysen und anderen OLS-Verfahren basieren, finden maßgeblich neutrale bis negative Effekte der Klassenwiederholung (z. B. Allen, Chen, Willson & Hughes, 2009; Fabian, 2012). Studien, die auf Verfahren zur Schätzung von Kausalitäten (z. B. Instrumentvariablen, Regression Discontinuity, Differences-in-Differences-Analysen) basieren, belegen kurz- und mittelfristig positive und mittel- bis langfristig neutrale Effekte (Schwerdt & West, 2013). Dieser Diskrepanz soll in den Sekundäranalysen Rechnung getragen werden.

(4) Fokus auf Wirkmechanismen von Klassenwiederholungen: Die meisten Studien untersuchen lediglich, ob eine Klassenwiederholung positive, neutrale oder negative Effekte bezüglich der Schulleistung hat. Seltener jedoch wird nach der konkreten Wirkungsweise von Klassenwiederholungen gefragt.

Die im Vorfeld genannten Überlegungen sollen in die Sekundäranalysen des NEPS-Datensatzes einfließen. Dabei soll folgende Forschungsfrage beleuchtet werden: Wie entwickelt sich die Schulleistung der Repetentinnen und Repetenten in Lesen und Mathematik?

Zu beiden Messzeitpunkten (in Jahrgangsstufe 5 und 6) befinden sich 3.000 Schülerinnen und Schüler in der Analytestichprobe. Von diesen wiederholten 174 Schülerinnen und Schüler (5,8 %) mindestens einmal eine Klasse – der größte Teil von ihnen in den Jahrgangsstufen 1 bis 4.

In späteren Analysen soll für die Gruppe der Repetentinnen und Repetenten mittels Propensity-Score-Matching die Leistungsentwicklung in Lesen und Mathematik mit vergleichbaren Schülerinnen und Schülern, die keine Klasse wiederholt haben, verglichen werden. So soll herausgefunden werden, ob die Klassenwiederholung einen signifikanten Effekt auf die Entwicklung der Lese- und Mathematikleistung ausübt.

Ausgehend von den theoretischen Überlegungen und bislang vorliegenden empirischen Forschungsbefunden wird vermutet, dass es (1) kurzfristig schwache, jedoch negative Effekte und (2) mittel- bis langfristig keine Unterschiede zu vergleichbar leistungsschwachen Nicht-Wiederholern gibt.

Freitag
13. März

11:15 Uhr

HGC 40

Chair(s)

Klapproth, Florian
Scharenberg, Katja

Effekte von Klassenwiederholung auf die leistungsbezogene und psychoemotionale Entwicklung von Schülerinnen und Schülern: Heterogenität in empirischen Zugängen und Befunden?

Klapproth, Florian; Schaltz, Paule; Brunner, Martin; Keller, Ulrich; Fischbach, Antoine; Ugen, Sonja; Martin, Romain

Kurz- und mittelfristige Effekte von Klassenwiederholungen in der luxemburgischen Sekundarstufe I auf die leistungsbezogene und psychoemotionale Entwicklung der Schülerinnen und Schüler

In den Meta-Analysen von Holmes (1989), Jimerson (2001) und Hattie (2009) finden sich nur wenige Hinweise darauf, dass Klassenwiederholungen positive Auswirkungen auf die Leistung der Schülerinnen und Schüler und ihre psychoemotionale Entwicklung haben. Häufiger als positive Effekte konnten dagegen negative Effekte in Bezug auf die Leistung und das Ausmaß des Wohlbefindens der Schülerinnen und Schüler aufgezeigt werden. Mögliche Gründe für das Auftreten negativer Effekte sehen einige Autoren in der Stigmatisierung der nicht-versetzten Kinder und in der Verhinderung der Erfahrung neuartiger Lernmöglichkeiten (Pagani, Tremblay, Vitaro, Boulerice & McDuff, 2001). Allerdings konnte verschiedentlich gezeigt werden, dass Klassenwiederholung zumindest kurzfristig zu einer Leistungssteigerung führt (Holmes, 1989).

In Luxemburg sind traditionell viele Schülerinnen und Schüler von Klassenwiederholungen betroffen, gut kontrollierte Studien zur Wirksamkeit von Klassenwiederholungen in Luxemburg fehlen jedoch bislang. Daher war das Ziel dieser Studie die Untersuchung von kurz- und mittelfristigen Effekten von Klassenwiederholungen auf die leistungsbezogene und psychoemotionale Entwicklung der Schülerinnen und Schüler in der luxemburgischen Sekundarstufe I.

Wir verwendeten das Propensity-Score-Matching zur Bildung einer Vergleichsgruppe von versetzten Schülerinnen und Schülern, die sich in einer Reihe von Merkmalen den nicht-versetzten Schülerinnen und Schülern ähnelten. Darüber hinaus untersuchten wir zwei Gruppen von Klassenwiederholern, nämlich diejenigen Schülerinnen und Schüler, die die 7. Klasse wiederholt haben, und diejenigen, die die 8. Klasse wiederholt haben. Dem Vergleich lag ein same-grade, same-age cohort, different times of measurement Versuchsplan zugrunde. Die Wirkung der Klassenwiederholungen wurde anhand unterschiedlicher Leistungsvariablen (Schulnoten und Testergebnisse) und psychoemotionaler Variablen (z.B. Selbstkonzept, Schulangst) gemessen. Da die Schulnoten von Klasse 7 bis Klasse 9 für verschiedene Fächer bekannt waren, konnten kurz- und mittelfristige Effekte der Klassenwiederholung auf die Schulnoten erfasst werden.

Was die Schulnoten betrifft, konnte gezeigt werden, dass Klassenwiederholer im Vergleich zu versetzten Schülerinnen und Schülern im ersten Jahr nach der Wiederholung einen Leistungsvorsprung aufwiesen, dieser jedoch bereits nach zwei Jahren nicht mehr nachweisbar war. In Bezug auf die Testwerte konnte zwischen nicht-versetzten und versetzten Schülerinnen und Schülern kein Unterschied festgestellt werden. Im Hinblick auf die psychoemotionalen Variablen konnte lediglich beim akademischen Selbstkonzept ein Unterschied zugunsten der versetzten Schülerinnen und Schüler beobachtet werden.

Freitag
13. März

11:15 Uhr

HGC 40

Chair(s)

Klapproth, Florian
Scharenberg, Katja

Effekte von Klassenwiederholung auf die leistungsbezogene und psychoemotionale Entwicklung von Schülerinnen und Schülern: Heterogenität in empirischen Zugängen und Befunden?

Schwerdt, Guido; West, Martin

Die Effekte von Klassenwiederholung auf zukünftige Schülerleistungen – Quasi-experimentelle Evidenz aus Florida

Sollten Schülerinnen und Schüler, die keine ausreichenden Leistungen zeigen, die Klasse wiederholen müssen? In Deutschland wird derzeit intensiv über diese Frage debattiert. Vielfach wird eine Abschaffung des „Sitzenbleibens“ gefordert. Im Gegensatz dazu führen immer mehr Bundesstaaten in den USA die Pflicht zur Klassenwiederholung überhaupt erst ein. Welche bildungspolitische Strategie letztlich zielführender ist, hängt entscheidend davon ab, ob betroffene Schülerinnen und Schüler von der Klassenwiederholung im Hinblick auf ihren langfristigen Schulerfolg profitieren.

Die empirische Evidenz zu dieser Frage deutet überwiegend auf negative Konsequenzen einer Klassenwiederholung für die betroffenen Schüler hin (Allen et al., 2009; Jimerson, 1999; Jimerson et al., 2000, 2002; McCoy & Reynolds, 1999). Allerdings basiert diese Evidenz fast ausschließlich auf Untersuchungen, die versuchen, für beobachtbare Determinanten des Schulerfolgs zu kontrollieren. Diese Ergebnisse sind jedoch möglicherweise verzerrt, sofern nicht-beobachtbare Determinanten des derzeitigen Schulerfolgs mit der Wahrscheinlichkeit einer früheren Klassenwiederholung korreliert sind. Neuere Untersuchungen, die auf quasi-experimentellen Schätzverfahren basieren, weisen durchaus gemischte Ergebnisse auf (Jacob & Lefgren, 2004, 2009; Manacorda, 2012; Winters & Greene, 2012).

Wir schätzen die Effekte der Wiederholung der dritten Klasse und damit verbundener Fördermaßnahmen im US-Bundesstaat Florida im Hinblick auf zukünftige Schülerleistungen. Seit dem Schuljahr 2002/2003 gilt an öffentlichen Schulen in Florida eine testbasierte Versetzungspolitik, die eine verpflichtende Klassenwiederholung vorsieht, falls Schülerinnen und Schüler am Ende der dritten Schulklasse kein ausreichendes Leseverständnis auf dem landesweiten Lesetest demonstrieren können.

Unser Untersuchungsdesign nutzt eine Diskontinuität im Zusammenhang zwischen dem Ergebnis des Lesetests in der dritten Klasse und der Wahrscheinlichkeit der Klassenwiederholung aus, um die kausalen Effekte des Sitzenbleibens auf zukünftige Schülerleistungen zu schätzen. Unsere Schätzungen basieren auf administrativen Daten für alle Schülerinnen und Schüler an öffentlichen Schulen im Bundesstaat Florida in den Schulklassen drei bis zehn zwischen den Jahren 2001 und 2009. Die Daten beinhalten jährliche Testergebnisse in den Bereichen Mathematik und Leseverständnis für jeden einzelnen Schüler. Während herkömmliche Schätzverfahren auf negative Effekte hindeuten, offenbart der Regressionsdiskontinuitätenansatz positive Effekte des Sitzenbleibens auf zukünftige Schülerleistungen und negative Effekte auf die Wahrscheinlichkeit einer Klassenwiederholung in höheren Schulklassen. Die positiven Effekte auf Schülerleistungen verringern sich jedoch mit der Zeit und werden nach 6 Jahren statistisch insignifikant.

Freitag
13. März

11:15 Uhr

HGC 40

Chair(s)

Klapproth, Florian
Scharenberg, Katja

Effekte von Klassenwiederholung auf die leistungsbezogene und psychoemotionale Entwicklung von Schülerinnen und Schülern: Heterogenität in empirischen Zugängen und Befunden?

Kuger, Susanne

Zum Umgang mit kultureller Heterogenität in Fragebogenstudien: Methodenaspekte in internationalen (computerbasierten) Erhebungen

International vergleichende Studien erfreuen sich in der Bildungsforschung immer größerer Beliebtheit. Zum einen können mit ihrer Hilfe ganz neue Forschungsfragestellungen in Sekundäranalysen bearbeitet werden, zum anderen sind sie in einer immer stärker international ausgerichteten Welt zentrale Mittel, um systematisch Ähnlichkeiten und Unterschiede zwischen Ländern, Bildungssystemen und Kulturkreisen zu beschreiben. Aus Gründen der Erhebungsökonomie und der Standardisierung werden in diesen Studien immer häufiger computerbasierte Erhebungen eingesetzt. Man erhofft sich standardisiertere Erhebungsbedingungen, weniger kulturell bedingte Verzerrungen und damit einhergehend eine bessere Vergleichbarkeit der resultierenden Daten verschiedener Stichproben unterschiedlicher kultureller Herkunft.

In den Methodendiskussionen zu diesen Studien kursieren verschiedene Themen (Van de Vijver & Leung, 1997), von denen das Symposium vier aufgreift, die sich alle darauf beziehen, ob jenseits psychometrischer Fragebogenmerkmale Unterschiede zwischen dem Antwortverhalten von Teilnehmern verschiedener Stichproben in verschiedenen Ländern bestehen, welcher Art und Stärke diese Unterschiede sind und wie man ihnen schon bei der Erhebungsplanung oder nach der Datensammlung begegnen kann. Zwei Beiträge sprechen dabei dezidiert Fragestellungen der computerbasierten Erhebung an, die beiden anderen befassen sich mit dem Ausmaß und einer möglichen Lösung zur Verringerung internationaler Antwortverzerrungen.

Der erste Beitrag untersucht, inwiefern Primingeffekte innerhalb einer Frage und zwischen Fragen international vergleichbar ausfallen und von welchen Personenmerkmalen sie in den verschiedenen Ländern besonders abhängen. Der zweite Beitrag vergleicht daraufhin inwiefern der Einsatz von Ankervignetten (eine der Möglichkeiten, kulturspezifischen Antwortbias zu reduzieren) für eine Stichprobe von Schülern und eine andere Stichprobe von Lehrern vergleichbare Ergebnisse zeigt. Im dritten Beitrag untersuchen die Autoren, ob bestimmte computerbasierte Fragebogenformate reliablere und validere Daten produzieren als herkömmliche Formate und damit Effekte unterschiedlicher Antwortstile reduzieren können. Der vierte und letzte Beitrag analysiert, wie anhand von log-Daten computerbasierter Erhebungen eventuell vorhandenes kultur- und subgruppenspezifisches Antwortverhalten gefunden und in seiner Größe abgeschätzt werden können.

Alle vier Beiträge beziehen sich auf Daten des im Sommer 2014 durchgeführten Feldtests von PISA 2014, der in 61 Ländern computerbasiert durchgeführt wurde. Zum Zeitpunkt der Einreichung konnten knapp 2/3 der Daten – aus 40 Staaten – in erste Auswertungen einbezogen werden; bis zur Tagung werden alle Analysen auf den vollständigen Datensatz ausgeweitet.

Die Studien tragen einen wichtigen Baustein zum Verständnis interkultureller und subgruppenspezifischer Unterschiede bei der Beantwortung von Fragebogenstudien in der Bildungsforschung bei. Sie liefern damit notwendige Voraussetzung um genauer zu verstehen, wie sehr heterogene Teilnehmergruppen in ihrem Verhalten divergieren und wie etwaig dadurch entstehende Verzerrungen in Sekundäranalysen berücksichtigt werden können, um zu reliablen und validen Ergebnissen zu kommen (Vandenberg & Lance, 2000; Poortinga & Klieme, in Druck).

Freitag
13. März

11:15 Uhr

HGC 50

Chair(s)

Kuger, Susanne

Zum Umgang mit kultureller Heterogenität in Fragebogenstudien:
Methodenaspekte in internationalen (computerbasierten) Erhebungen

Roczen, Nina; Klingebiel, Franz

Priming zwischen und innerhalb von Fragen in internationalen Large Scale Assessment Studien

Hintergrund und Untersuchungsgegenstand: Bei der Erfassung von psychologischen Konstrukten wie zum Beispiel Einstellungen in verschiedenen Ländern spielen Antwortbiases eine entscheidende Rolle. Fehlende Messinvarianz beeinträchtigt den interkulturellen Vergleich dieser Konstrukte oft erheblich, verringert das Potenzial für Sekundäranalysen oder führt zu Über- und Fehlinterpretationen der Daten.

Ein typischer, international gefundener Antwortbias ist das Priming: Die Reihenfolge einer gestellten Frage in einem Fragebogen oder die Reihenfolge von Items innerhalb einer Frage kann die Beantwortung folgender Fragen und Items beeinflussen. Primingeffekte fallen für verschiedene Subgruppen unterschiedlich aus und sind somit an Merkmale der beantwortenden Person gekoppelt. Es wird vermutet, dass vor allem die Verarbeitungskapazität einer Person, d.h. ihre kognitive Leistungsfähigkeit mit dem Ausmaß des Primingeffekts zusammenhängt. Mögliche Mechanismen könnten eine bessere oder weniger gute Aktivierung und Unterdrückung relevanter und irrelevanter Informationen sein.

Ziel dieses Beitrags ist es, das Ausmaß von Primingeffekten zwischen Fragen und innerhalb von Fragen systematisch zu untersuchen und dabei insbesondere subgruppenspezifische (z.B. geschlechtsassoziierte) und kulturabhängige Muster zu finden, die die internationale Vergleichbarkeit der Daten beeinträchtigen. Der Beitrag untersucht dabei auch ob Primingeffekte von der kognitiven Leistungsfähigkeit der Teilnehmer abhängt. Um Hinweise auf mögliche Bedingungen von unterschiedlich starken Priming-Effekten zu erhalten, analysieren wir darüberhinaus verschiedene Merkmale von Ländern, bei welchen dieser Bias ähnlich stark ausgeprägt ist.

Methode und Vorgehen: Als Datengrundlage für die Analysen stehen uns Stichproben von mehr als N = 90.000 Schülern und mehr als N = 6.000 Lehrern zur Verfügung, die im Rahmen des Feldtests für PISA 2015 gesammelt wurden. In Anlehnung an Morgan & Poppe (2012) wurde im Lehrerfragebogen ein Primingexperiment über Fragen hinweg eingebracht. Die Hälfte der Befragten Lehrer erhielt zunächst eine Frage zur allgemeinen Einschätzung der Entwicklung des Bildungssystems im eigenen Land (im Vergleich zu den anderen sehr erfolgreichen Ländern in internationalen Schulvergleichsstudien) und anschließend eine Frage zu seiner Einschätzung des Management der eigenen Schule vorgelegt. Bei der anderen Hälfte der Befragten war die Reihenfolge umgekehrt. Nach Morgan & Poppe nahmen wir leichte Primingeffekte für die erste Gruppe an, d.h. eine negativere Einschätzung der eigenen Situation. Im Schülerfragebogen wurden in einer Fragenbatterie zur situativen Beurteilung von Herangehensweisen an eine wissenschaftliche Fragestellung die einzuschätzenden Einzelitems rotiert. Je einem Viertel der Befragten wurde eine andere Reihenfolge der Items präsentiert. Die Daten liegen bisher zu zwei Dritteln vor und werden momentan explorativ, deskriptiv und varianzanalytisch untersucht.

Erste Ergebnisse: Erste Ergebnisse auf internationaler Ebene deuten an, dass Primingeffekte in diesen Untersuchungsdesigns über die Gesamtheit an Schülern und Lehrern nur ein geringes Ausmaß haben und somit kaum oder gar nicht zu Antwortverzerrungen beitragen. Detailliertere Analysen zu einzelnen Subgruppen und im internationalen Vergleich stehen noch aus. Die Ergebnisse sollen im Hinblick auf mögliche Ursachen für die Abweichung von den Ergebnissen der Vorbildstudie diskutiert werden.

Freitag
13. März

11:15 Uhr

HGC 50

Chair(s)

Kuger, Susanne

Zum Umgang mit kultureller Heterogenität in Fragebogenstudien:
Methodenaspekte in internationalen (computerbasierten) Erhebungen

Marksteiner, Tamara; Klingebiel, Franz

Funktionsweise von Ankervignetten bei Schülern und Lehrern im internationalen Setting

Theoretischer Hintergrund: Internationale Vergleichsstudien sind vor das Problem kulturell bedingter Antworttendenzen gestellt, die die Vergleichbarkeit der Ergebnisse über Kulturräume hinweg erschwert. In der Vergangenheit ist die Methode der Ankervignetten in computerbasierten Befragungen bei Schülern erfolgreich verwendet worden, um diese Art der Verzerrung zu reduzieren. Das Prinzip der Ankervignetten beruht darauf, dass in einem ersten Schritt die kulturbedingten Antworttendenzen erfasst werden (Ankervignetten). Vignetten beinhalten Beschreibungen von Personen mit einer Ausprägung (niedrige, mittlere, hohe) einer bestimmten Eigenschaft (z.B. Leistungsmotivation), auf der auch die Studienteilnehmer ihre Selbstberichte vornehmen bzw. sich selbst einschätzen. Mit dieser Information werden in einem zweiten Schritt die interessierenden Faktoren (Selbstberichte) um diese Verzerrung korrigiert. Allerdings gibt es Hinweise darauf, dass verschiedene Zielgruppen unterschiedliches Antwortverhalten (z.B. in Bezug auf die Bearbeitungszeit) bei computerbasierten Befragungen zeigen.

Fragestellung: Daher gehen wir davon aus, dass die Funktionsweise der Ankervignetten bei Schülern und Lehrern unterschiedlich geprägt ist. Die vorliegenden Analysen haben zum Ziel, die Unterschiede zwischen Lehrern und Schülern hinsichtlich der Reduktion der Verzerrungen zu untersuchen. Dabei wird die empirische Güte hinsichtlich Deskriptiva, Reliabilität und Validität der geankerten Selbstberichte betrachtet.

Methode: Zunächst einmal wurden – basierend auf der Einschätzung der Personen, die in den Vignetten beschrieben werden – die Selbsteinschätzungen geankert, d.h. die Selbsteinschätzungen werden in Relation zu den Einschätzungen der Personen in den Vignetten gesetzt. Anschließend wurden Skaleneigenschaften zwischen geankerten und nicht-geankerten Selbstberichten verglichen und geprüft, ob sich die Befundmuster zwischen den beiden Zielgruppen unterscheiden. Dabei wurden Deskriptiva (Mittelwerte), Reliabilität (Cronbach's Alpha, Item-Total-Korrelation, Muster der Faktorladungen) und Validität (Korrelationen mit theoretisch zusammenhängenden Outputs) der Selbstberichtsskalen untersucht.

Ergebnisse: Erste Analysen zeigen, dass sich die Funktionalität von Ankervignetten zwischen den beiden Zielgruppen tatsächlich unterscheidet. Geankerte im Vergleich zu nicht-geankerten Selbsteinschätzungen von Schülern zeigen eine höhere empirische Güte und reduzieren die Antwortverzerrungen, was sich u.a. im Zusammenhang dieser mit Leistung zeigt. Auch bei der Zielgruppe der Lehrer zeigen sich Verbesserungen der Skaleneigenschaft nachdem die Einzelitems durch die Ankerinformationen korrigiert wurden. Allerdings fallen die Verbesserungen nicht sehr stark aus. Als Grund hierfür können die bereits sehr guten (ungeankerten) Skaleneigenschaften gelten. Innerhalb verschiedener Kulturräume finden sich differenzielle Befunde. Tiefergehende Analysen werden zur Diskussion gestellt.

Freitag
13. März

11:15 Uhr

HGC 50

Chair(s)

Kuger, Susanne

Zum Umgang mit kultureller Heterogenität in Fragebogenstudien:
Methodenaspekte in internationalen (computerbasierten) Erhebungen

Kuger, Susanne; Bayer, Sonja

Computerbasierte interaktive Fragenformate im Test

Hintergrund und Untersuchungsgegenstand: Die Einführung computerbasierter Erhebungen erlaubt die Nutzung neuer Fragenformate, die herkömmliche Formate ablösen oder ergänzen können. Ein Ziel des Einsatzes dieser neuen Formate ist die Optimierung der Datenqualität bzw. die Minimierung von Zeiten für die nachträglich Datenaufbereitung, eine Verbesserung der Präzision und Validität sowie der Untersuchungsökonomie.

Ein in diesem Zusammenhang viel diskutiertes und fehleranfälliges Antwortformat herkömmlicher Papierfragebogen sind offene Antworten. Diese sind besonders nachbearbeitungsintensiv in der Datenaufbereitung und bedeuten zudem beim computerbasierten Testen, dass Teilnehmer das Eingabemedium wechseln müssen, von der Maus zur Tastatur. Verschiedene Lösungsansätze sollen diesen Problemen vorbeugen, darunter auch alternative Fragenformate, die durch ihre Interaktivität den Studienteilnehmer besonders zum Ausfüllen motivieren sollen. Der Beitrag untersucht drei Alternativformate im direkten Vergleich mit einer Texteingabe: Drop-down Menü, Schieberegler und ein interaktives Kreisdiagramm.

Vorgehen und Methode: Ausgewählt wurden dabei Fragen, deren Beantwortung den unmittelbaren Abruf verfügbarer oder nachschlagbarer numerischer Fakten (bzw. numerische Prozentangaben) bedarf. Somit sollen Effekte unterschiedlicher intrapsychischer Abrufprozesse ausgeschlossen werden. Die drei untersuchten Formate variieren dabei die Spanne der realistisch möglichen Antworten (17, 30 und 100 numerische Kategorien). Angesichts der Übersichtlichkeit der Eingabe wurde für die erste Frage mit 17 Antwortkategorien ein drop-down Menü genutzt, für die Frage mit 30 möglichen Antwortkategorien wurde ein Schieberegler und für die dritte Frage eine Visualisierung mittels Kreisdiagramm eingeführt.

Genutzt werden Daten des Feldtests der PISA 2015 Studie, die im Frühjahr und Sommer 2014 in 61 computerisiert erhebenden Ländern weltweit durchgeführt wurde. Die Fragen wurden so angelegt, dass jeweils eine Hälfte der Teilnehmer die Frage in einem offenen Antworttextfeld vornehmen sollte und die andere Hälfte der Teilnehmer die identische Frage mit dem neuen Format ausfüllte. Dabei nahm jeder Teilnehmer nur an einer der drei Vergleiche teil. Die Zuteilung zu den Fragenalternativen erfolgte computergesteuert aufgrund einer semi-randomisierten Zuweisung.

Die Analysen betrachten sowohl die Auswertung der letztendlich gespeicherten Antworten der Teilnehmer, als auch detaillierte Analysen zu Item-Non-Response, der Bearbeitungsdauer und Verhaltensdaten aus den Log-Daten der Computerplattform. Die separate Betrachtung der Daten einzelner Länder erlaubt Rückschlüsse auf kulturelle Unterschiede bei der Bearbeitung von Computer basierten Fragebögen und einzelnen Formaten. Neben kulturellen Unterschieden wird der Einfluss kulturübergreifende Personenmerkmale auf die Beantwortung der Fragen untersucht (Geschlecht, finanzieller Hintergrund (oder besser ICT-Ressourcen, Leistung).

Ergebnisse und Diskussion: Es zeigt sich, dass die Anzahl der fehlenden Werte unter Anwendung des drop-down Menus geringer sind als bei der offenen Antworteingabe. Mittelwerte und Streuung sind nahezu gleich. Anders verhält es sich beim Schieberegler. Hier sind die fehlenden Werte höher als in der Antwortalternative mit dem Texteingabefeld. Zugleich zeigt sich die Datenpräzision uneinheitlich. Während sie in Bezug auf Ausreißer und unplausible Werte mit dem Schieberegler etwas besser ist, sind zugleich die ICCs der Schülerangaben mit dem Schieberegler kleiner als erwartet und auch kleiner als mit dem herkömmlichen Antwortformat. Die Log-Daten zeigen dagegen ein einheitliches Muster längerer Bearbeitungszeiten mit den neuen Formaten.

Besonders auffällig ist, dass diese Befundmuster im internationalen Vergleich variieren. Während es erwartbare Ergebnisse von geringeren Unterschieden zwischen den Fragenformaten in einigen (vorwiegend östlichen) Industrieländern gibt und größere Unterschiede z.B. in südamerikanischen Ländern, finden sich auch größere Unterschiede zwischen den Formaten in einigen anderen Industriestaaten z. B. in Europa. Bei den kulturübergreifenden Personenmerkmalen haben das Geschlecht sowie der Besitz von ICT-Ressourcen einen Effekt auf die Befundmuster.

Die Diskussion konzentriert sich zum einen auf die Konsequenzen dieser Ergebnisse für die Weiterentwicklung moderner, ökonomischer Erhebungsverfahren und zum anderen auf Aspekte des Umgangs mit dieser Vielfalt unterschiedlicher Effektmuster im Bestreben, international äquivalente Datenerhebungen und –qualität zu erreichen.

Klieme, Eckhard; Kröhne, Ulf

Lese- und Antwortverhalten beim Bearbeiten computerbasierter Fragebogen: Inter- und intranationale Unterschiede

Hintergrund: Computerbasierte Fragebogenerhebungen bieten den Vorteil, während der Erhebung auch als zusätzliche Datenquelle unmittelbare Verhaltensdaten der Teilnehmer wie z.B. Zeiten, Klicks und die Navigation innerhalb des Fragebogens zu erfassen.

Hieraus lassen sich Schlüsse sowohl auf Personenmerkmale wie Motivation und Selbstregulation beim Lesen und Bearbeiten elektronischer Texte ziehen als auch auf mögliche Beeinträchtigungen der Datenqualität wie Gründlichkeit vs. Oberflächlichkeit der Bearbeitung, schematisches Antwortverhalten, Übe- und Ermüdungs-Effekte sowie Verzerrungen des Bearbeitungsprozesses und der Antwortmuster durch konstrukt fremde Faktoren wie Lesefähigkeiten, ICT-Fähig- und Fertigkeiten, und Aspekte des persönlichen Hintergrunds (Geschlecht, Herkunftssprache usw.). Auch eine mögliche Beeinträchtigung durch technische Probleme wie übermäßig lange Ladezeiten kann anhand von automatisch aufgezeichneten Prozessdaten erkannt werden. Die Nutzung solcher Verhaltensdaten kann daher eine wertvolle Ergänzung bisheriger üblicher psychometrischer Kriterien bei der Überprüfung der Datenqualität von Fragebogen (Blasius & Thiessen, 2012) sein.

Für die Konstruktion internationaler Fragebogen ergeben sich aus den Log-Daten noch zusätzlich neue Aspekte, da auch die Verhaltensmerkmale der Personen aus sehr unterschiedlich stark an den Umgang mit ICT Geräten gewöhnten Ländern vergleichbar werden. Daraus ergeben sich neue Forschungsfragen, die für die Zukunft der international vergleichenden Bildungsforschung zentral sind.

Forschungsfragen: Der Beitrag betrachtet zunächst die technische Qualität der Fragebogendaten, die in der internationalen PISA-Studie gewonnen werden konnten. Hierzu werden Lade- und Bearbeitungszeiten dokumentiert und Outlier untersucht.

Nach Ausschluss von Outliern wird sodann das Bearbeitungsverhalten rekonstruiert. Die Gesamtbearbeitungszeit lässt sich in eine Lese- und eine Antwortphase unterteilen. Es soll untersucht werden, von welchen Materialfaktoren (Inhalt, Länge und Format der Fragen sowie Position im Fragebogen) sowie von welchen Personenmerkmalen (ICT-Fähigkeiten und -Erfahrungen, allgemeine kognitive Fähigkeiten, Geschlecht usw.) die Dauer der Lese- und Antwortzeit abhängen. Diese Modelle werden getrennt für einzelne Länder aufgestellt und interkulturell vergleichend untersucht.

Eine spezielle Fragestellung gilt der Vergleichbarkeit des Antwortverhaltens bei Personengruppen verschiedenen Alters. Während für die Generation der „digital natives“, der heutigen Schülerinnen und Schüler, angenommen werden kann, dass die meisten schon sehr früh in Kontakt mit digitalen Medien kamen und nur geringe Unterschiede zwischen Kinder verschiedenen Geschlechts oder unterschiedlicher kultureller und sozioökonomischer Hintergründe bestehen, gelten diese Annahmen nicht zugleich für Erwachsene wie z.B. ihre Lehrer und Schulleiter. Hier sollten viel stärkere Effekte verschiedener Personenmerkmale (Alter, unterrichtetes Schulfach) und der allgemeinen Geläufigkeit des Umgangs mit ICT eine Rolle spielen, so dass auch kulturelle Unterschiede noch stärker in den Vordergrund rückendürften.

Vorgehen: Als Datengrundlage für die Analysen werden Daten der Befragungen von Schülern, Lehrern und Schulleitern des Feldtests für PISA 2015 genutzt, der im Frühjahr und Sommer 2014 in insgesamt 61 Ländern computerbasiert stattgefunden hat. Die Analyse bezieht dabei Daten von ca. 90.000 Schülern, mehr als 6.000 Lehrern und 1.500 Schulleitern ein, die an mindestens 30-minütigen Onlinebefragungen teilgenommen haben. Analysiert werden Beantwortungszeiten und Häufigkeiten einzelner Fragen sowie das Navigationsverhalten im Fragebogen insgesamt. Die Analysen betrachten dabei sowohl länderspezifische Aspekte als auch Details der Fragebogengestaltung wie Fragenlänge und Fragenformat.

In multivariaten Mehr-Ebenen-Regressionsanalysen (mit den Ebenen Land und Person) werden Lese- und Antwortzeiten in Abhängigkeit von Materialfaktoren, Personenmerkmalen und kulturellem Kontext modelliert.

Ergebnisse und Diskussion: Erste Ergebnisse der Analysen an etwa 2/3 der Daten weisen auf stark unterschiedliche Verhaltensmuster zwischen Schülern und Lehrern hin. Während Schüler ihre Antworten eher schneller abgeben, achten die Erwachsenen Teilnehmer besonders auf die Genauigkeit ihrer Angaben, korrigieren diese häufiger und navigieren auch häufiger zu verschiedenen Fragen zurück. Zudem zeigen sich sowohl hardwareabhängige als auch personenabhängige internationale Unterschiede in den Verhaltensdaten. Inwieweit diese Unterschiede eine Beeinträchtigung der Gesamtbearbeitung von computerbasierten Fragebogen darstellt, soll anschließend diskutiert werden.

Freitag
13. März

11:15 Uhr

HGC 50

Chair(s)

Kuger, Susanne

Zum Umgang mit kultureller Heterogenität in Fragebogenstudien:
Methodenaspekte in internationalen (computerbasierten) Erhebungen

Göbel, Kerstin; Vieluf, Svenja

Unterrichtsqualität und Klassenzusammensetzung im Hinblick auf die interkulturellen Lernergebnisse im Englischunterricht in der 9. Klasse

Die Förderung von interkultureller Kompetenz ist ein wichtiges Ziel von schulischen Bildungsprozessen. Dem Fremdsprachenunterricht kommt im Hinblick auf die Zielerreichung dabei eine besondere Rolle zu. Trotz seiner bildungspolitischen Bedeutung gibt es bislang wenig empirische Forschung, die untersucht, wie dieses Lernziel erreicht werden kann und welche Art von Unterricht hierbei hilfreich sein könnte. Erste empirische Analysen im Kontext der DESI-Studie weisen darauf hin, dass sowohl allgemeine didaktische Kompetenzen, wie Classroom-Management und eine positive Fehlerkultur als auch spezifische Aspekte des interkulturellen Englischunterrichts, wie die Erfahrungs- und Diskursorientierung im Hinblick auf interkulturelle Themen bedeutsam zu sein scheinen (Göbel & Hesse, 2008). Darüber hinaus ist anzunehmen, dass die Zusammensetzung der Klassen im Hinblick auf verschiedene Dimensionen wie Lernvoraussetzungen, Geschlecht und Migrationsstatus sowohl das Unterrichtsangebot als auch die Lernergebnisse der Klassen beeinflusst (Helmke & Klieme, 2008).

Der vorliegende Beitrag versucht, die Zusammenhänge zwischen Unterrichtsqualität und interkulturellen Lernergebnissen sowie die Interaktion zwischen Unterrichtsqualität und der Klassenzusammensetzung im Hinblick auf die Lernergebnisse zu beleuchten und nutzt hierzu Daten sowie Reanalysen von Daten der DESI-Studie (Deutsch-Englisch-Schülerleistungen International). Die Zieldimension interkulturelle Kompetenz wurde im Sinne des Modells von Bennett, Bennett, und Allen (2003) als subjektive Konstruktion kultureller Differenzen operationalisiert. Unter Nutzung der Critical Incidents Methode wurden die subjektiven Deutungen interkultureller Situationen durch Schülerinnen und Schüler abgebildet. Dieses Testverfahren wurde im Rahmen der DESI-Studie von insgesamt ca. 11.000 Schülerinnen und Schülern der 9. Klasse bearbeitet. Im Anschluss an die Erhebung wurden die Probanden im Rahmen latenter Klassenanalysen in drei Gruppen interkultureller Kompetenz eingeordnet (acceptance/adaptation, minimization, denial/defense; Hesse, Göbel & Jude, 2008). Im Rahmen des DFG-Projektes zu "Lehr-/Lernprozessen im interkulturellen Englischunterricht der 9. Klasse" im Jahr 2011-2013 wurden insgesamt 93 DESI-Unterrichtsvideos aus den Jahren 2004-2005 hochinferent von zwei Beurteilern eingeschätzt. Für den vorliegenden Beitrag wurden die genannten Video-Analysen sowie die Fragebogendaten von insgesamt n=2.071 Schülerinnen und Schülern in den videografierten Klassen herangezogen.

Die Ergebnisse der Mehrebenenanalysen weisen folgende Qualitätsmerkmale als bedeutsam für den interkulturellen Englischunterricht aus: Regelklarheit, positive Fehlerkultur, Classroommanagement, das Einnehmen kritischer Perspektiven, die Behandlung von Vorurteilen sowie die Haltung der Lehrperson als interkultureller Mediator. In Klassen mit hohem Anteil von Schülerinnen und Schülern unterschiedlicher Erstsprachen ist die interkulturelle Qualität des Englischunterrichts höher als in Klassen mit geringerem Migrantenanteil. Vor allem fungieren Lehrkräfte in diesen Klassen häufiger als interkulturelle Mediatoren und sie thematisieren häufiger kritische Perspektiven. Eine höhere Qualität des interkulturellen Unterrichts geht wiederum einher mit einer höheren interkulturellen Kompetenz der SuS. Die Klassenkomposition hat folglich einen indirekten (mediierten) Effekt auf die interkulturelle Kompetenz der SuS.

Neben den genannten Qualitätsindikatoren scheint eine gemischtsprachliche Zusammensetzung der Klasse also eine Ressource für die Vermittlung interkultureller Kompetenz im Unterricht darzustellen und es Lehrkräften zu erleichtern eine hohe interkulturelle Qualität zu realisieren.

Strohn, Meike

Binnendifferenzierung im Englischunterricht – die Lehrerperspektive

Jüngste wissenschaftliche Erkenntnisse, wie die Ergebnisse des EVENING-Projekts (Evaluation Englisch in der Grundschule; 2005-2007) und der DESI-Studie (Deutsch Englisch Schülerleistungen International; 2001-2006), machen deutlich, dass die Schüler eines Jahrgangs, sowohl in der Primar- als auch in der Sekundarstufe, über sehr heterogene Englischleistungen verfügen. Des Weiteren belegen eine repräsentative Befragung von Englischlehrkräften innerhalb des Schulentwicklungsprojekts Ganz In (Mit Ganzttag mehr Zukunft. Das neue Ganzttagsgymnasium NRW; 2009-2015), sowie Rückmeldungen zum fachdidaktischen Beratungsbedarf, dass der Themenkomplex Umgang mit Heterogenität im Englischunterricht aus Sicht der Englischlehrkräfte hoch relevant ist und hier zudem vergleichsweise großer Unterstützungsbedarf besteht. Zugleich ist jedoch speziell für den Bereich der Englischdidaktik ein Forschungsdesiderat auf dem Gebiet zu konstatieren. Das vorliegende Forschungsprojekt setzt sich daher zum Ziel, diese bestehende Forschungslücke zumindest in Teilen zu schließen, indem die Perspektive von Englischlehrerinnen und -lehrern in Nordrhein-Westfalen in Bezug auf Binnendifferenzierung im Englischunterricht empirisch erforscht wird. Die Forschungsergebnisse sollen wiederum der Innovation fremdsprachenunterrichtlicher Praxis dienen und konkrete, fremdsprachenspezifische Forschungsdesiderate aufzeigen.

Kernfrage 1: Welche subjektiven didaktischen Theorien haben Englischlehrer zum Konzept Binnendifferenzierung im Englischunterricht?

Unterfragen: (1) Wie wird der Begriff „Binnendifferenzierung“ allgemein und in Bezug auf den Englischunterricht von den befragten Englischlehrkräften definiert? (2) Wie wirkt sich das jeweilige pädagogisch-didaktische Gesamtverständnis der Englischlehrer auf ihre Sichtweise von Binnendifferenzierung im Englischunterricht aus? (3) Welche externen Rahmenbedingungen (auf Meta-, Meso- und Mikroebene) werden von den Forschungspartnern explizit oder implizit als förderlich oder hinderlich genannt? (4) Welche internen Determinanten werden, bei sich selbst oder auch bei Kollegen, von den Forschungspartnern explizit als förderlich oder hinderlich genannt? (5) Welche nicht explizit genannten, internen Determinanten beeinflussen das Konzept, das die jeweilige Englischlehrkraft von Binnendifferenzierung hat?

Kernfrage 2: Welche Konsequenzen lassen sich aus den rekonstruierten subjektiven Theorien für die universitäre Lehrerbildung, die Lehrerfortbildung und die fremdsprachenspezifische Unterrichtsforschung ableiten?

Das Erkenntnisinteresse legt einen qualitativ-rekonstruktiven Ansatz nahe, der sich am Forschungsprogramm Subjektive Theorien (Groeben et al. 1988; 2010) orientiert, in Teilen jedoch auch davon abweicht. Entsprechend wurden von mir zur Datenerhebung qualitative, leitfadengestützte Interviews mit zwölf Englischlehrkräften verschiedener Schulformen geführt, sowie Hintergrundfragebögen und Post-Scripts eingesetzt. Die Interviews wurden von mir transkribiert und mit Hilfe der integrativen texthermeneutischen Analyseverfahren (Kruse et al. 2007, 2010, 2014) einzeln ausgewertet. Im Anschluss erfolgten jeweils eine kollegiale Validierung im Forscherteam sowie eine kommunikative Validierung mit Hilfe einer ausformulierten Kurzinterpretation. Ein zweiter Schritt der Auswertung bestand in der Querauswertung, indem die wichtigsten Thematisierungsregeln und zentralen Motive, die sich abzeichnenden Muster, sowie die internen und externen Hinderungsgründe und Gelingensbedingungen für Binnendifferenzierung im Englischunterricht herausgearbeitet wurden.

Die Ergebnisse meiner Forschung sind einerseits eine umfassende Definition von Binnendifferenzierung im Englischunterricht aus Perspektive der Englischlehrkräfte. Andererseits wurden sich abzeichnende Muster, sowie Implikationen für die fremdsprachenspezifische Unterrichtsforschung und für die Lehrer(fort)bildung herausgearbeitet. Exemplarisch seien hier genannt:

- Zusammenhänge zwischen dem von der Lehrkraft verfolgten lehr-lerntheoretischen Ansatz (eher instruktivistisches oder konstruktivistisches Bild vom Englischunterricht), dem Bild vom Fremdsprachenlernen und der Sichtweise von Binnendifferenzierung als primär einer Differenzierung von oben durch die Lehrkraft oder einer Individualisierung von unten durch die Schüler (zur Terminologie vgl. Brügelmann 2002)
- genereller Trend von einer Differenzierung von oben hin zu einer Individualisierung von unten
- Englischlehrkräfte benötigen bessere Diagnose- und Differenzierungskompetenz sowie pädagogische Haltung
- potentiell hinderlich: das in vielen subjektiven Theorien zentrale Motiv Effizienz sowie das zentrale Motiv teaching to the test
- förderlich: Schülerorientierung mit Blick auf den Einzelschüler statt auf die Lerngruppe
- förderlich: Implementation von Mindeststandards statt Regelstandards innerhalb der curricularen Vorgaben (Transparenz Fundamentum/ Additum)
- in Lehreraus- und fortbildung notwendig: zunehmende Reflektion der eigenen subjektiven Theorien in Bezug auf die Sichtweise von und den Umgang mit Heterogenität im Englischunterricht

Freitag
13. März

11:15 Uhr

GC 03/142

Chair(s)

Göbel, Kerstin

Michalak, Magdalena; Kölzer, Carolin; Lemke, Valerie

Umgang mit diskontinuierlichen Darstellungsformen: Kompetenzen von Lernenden mit Deutsch als Zweitsprache (DaZ)

Ausschlaggebend für schulisches Lernen ist neben der allgemeinen Lesekompetenz auch die Fähigkeit, mit diskontinuierlichen Darstellungsformen angemessen umzugehen. In PISA-Studien wird der Umgang mit ihnen als Bestandteil naturwissenschaftlicher Kompetenz deklariert (Baumert et al. 2001). Tabellen, Karten, Diagramme etc. werden jedoch in allen Unterrichtsfächern als Lernmedium eingesetzt. So spielen sie auch im gesellschafts- und sozialwissenschaftlichen Unterricht, der für soziale, politische und ökonomische Lebenssituationen kompetent machen will, eine große Rolle.

Während die Kernaussagen eines linearen Textes mithilfe verbaler Mittel explizit hervorgehoben werden, sind diese bei nichtlinearen Formaten aus der Darstellungsform abzuleiten (Weidenmann 1994). Im Rezeptionsprozess solcher logischen Bilder im Sinne von Schnotz (2002) kann der ‚Leser‘ nicht auf kognitive Schemata zurückgreifen, die der Alltagswahrnehmung entsprechen. Je größer das Vorwissen und je höher die Anzahl an bereits kennengelernten Diagrammen, desto eher sind die Leser in der Lage, komplexe Diagramme zu interpretieren (Schnotz; Dutke 2004). Zudem bedarf das echte Verstehen von Diagrammen semantischer Verarbeitungsprozesse (Ulrich et al. 2012): Die visuell-räumlichen Zusammenhänge werden auf semantische Relationen übertragen. Die Transformation der erschlossenen Informationen erfolgt in einen kontinuierlichen Text, was eine hohe Sprachkompetenz voraussetzt und insbesondere für DaZ-Lernende eine Herausforderung darstellen kann (Oleschko 2012).

Systematische empirische Untersuchungen dazu, wie sich unterschiedliches sprachliches Vorwissen auf die Auswertung diskontinuierlicher Darstellungsformen auswirkt und wie insbesondere Lernende mit DaZ mit diesen umgehen, stehen noch aus. Vor diesem Hintergrund ist das Ziel unseres interdisziplinären Projektes (Fachdidaktiken Deutsch, DaZ und Sozialwissenschaften), dieses Forschungsdesiderat anzusprechen. Am Beispiel eines sozialwissenschaftlichen Diagramms wird untersucht, wie DaZ-Lernende nichtlineare Darstellungsformen fachlich und sprachlich erschließen und welche Rezeptionswege sie bevorzugen. Von Interesse ist die Frage, wie sie eine über das Beschreiben hinausgehende Aufgabenstellung fachlich und (schrift-)sprachlich bewältigen.

Diese Fragen wurden explorativ-qualitativ an einem kleinen Sample (n=10, je fünf Mädchen und Jungen mit DaZ im Alter von 12 bis 15 Jahren) im Rahmen des Kooperationsprojekts Sprachliche Bildung der Universität zu Köln untersucht. Die qualitative Stichprobe ist heterogen bezogen u.a. auf das Sprachniveau, die Lesekompetenz, den Bildungs- und Migrationshintergrund. Diese Merkmale wurden über einen Sprachstandtest in Deutsch (C-Test und Zuordnung zum GER), einen Lesetest (LGVT 6-12) und einen halbstandardisierten Fragebogen zu soziodemografischen Daten erhoben.

Die Untersuchung wurde mit Adaptionen der Methoden Fokussiertes Interview und Lautes Denken sowie einer Schreibprobe mit Diagrammauswertung durchgeführt. Zunächst wurde der visuelle Reiz, eine Statistik, kriterienorientiert ausgewählt und einer Bedarfsanalyse unterzogen. Den Lernenden wurde dieses Diagramm einzeln mit der Instruktion vorgelegt, rezipierte Teile in der Reihenfolge der Wahrnehmung zu nummerieren. Die hierbei gemachten Beobachtungen zu Orientierung und Vorgehensweise der Lernenden wurden währenddessen vom Forscherteam notiert. Nach dem ersten Rezeptionsdurchgang markierten die Lernenden die für sie besonders wichtigen Stellen im Diagramm. Sowohl die vorgenommenen Nummerierungen als auch Markierungen dienten als Einstieg in die retrospektive Introspektion, die die Rezeptionsweise und das erste Verstehen der Grafik fokussierte. Anschließend werteten die Lernenden das Balkendiagramm mündlich aus. Alle verbalen Äußerungen wurden aufgenommen und später transkribiert. In einem letzten Schritt erfolgte die Auswertung der Grafik schriftlich. Dies dient der Überprüfung, inwiefern sich die mündlichen und schriftlichen Leistungen der Schüler unterscheiden.

Die vielfältigen empirischen Daten – Interviewtranskripte, Schreibproben, Ergebnisse aus Sprach- und Lesetests u.a. – wurden trianguliert und offen kodiert. Erste Ergebnisse geben Hinweise darauf, dass der kompetente Umgang mit Grafiken nicht mit den sprachlichen Kenntnissen der Lernenden korreliert, sondern mit der Intensität der bisherigen Auseinandersetzung und der systematischen Arbeit an Diagrammen im Unterricht zusammenhängt. Eine entscheidende Rolle spielt Vorwissen bezogen auf Fachinhalte der Grafik sowie auf methodische Vorgehensweise mit Diagrammen aus anderen Fächern. Bei der Rezeption des Diagramms zeigte sich, dass die unerfahrenen Lernenden assoziativ vorgehen und sich an vertrauten Begriffen orientieren. Diese und weitere Ergebnisse werden in dem Beitrag präsentiert.

Busse, Vera; Krause, Ulrike-Marie

Förderung sozialen und interkulturellen Lernens: Effekte zweier unterrichtsmethodischer sowie zweier unterrichtssprachlicher Ansätze

Gesellschaftliche Pluralisierungsprozesse und kulturelle Heterogenität erfordern eine konstruktive Auseinandersetzung mit unterschiedlichen Sichtweisen und Werten. Bereits in der Schule, so betont der Europarat, müssen die Grundlagen für einen positiven Umgang mit kultureller Diversität gelegt werden, wobei die Förderung sogenannter interkultureller Kompetenz als besonders wichtig gilt (Council of Europe, 2008). In den Empfehlungen zur „Interkulturellen Bildung und Erziehung in der Schule“ der KMK (2013) wird ferner die Bedeutung einer fächerübergreifenden Einbindung interkultureller Lerninhalte sowie einer „interkulturell sensiblen Dialog- und Konfliktkultur“ (S. 3) betont. Die Ergebnisse der DESI-Studie (Deutsch Englisch Schülerleistungen International) deuten jedoch darauf hin, dass viele Schüler/innen Probleme haben, kulturelle Unterschiede in die Interpretation interkultureller Begegnungen miteinzubeziehen und dementsprechend sensibel auf kommunikative Unterschiede einzugehen (Hesse, Göbel & Jude, 2008). Ziel dieses Projektes ist es, soziale und interkulturelle Kompetenzen im Bereich des Umgangs mit kultureller Vielfalt im Schulunterricht fächerübergreifend zu fördern. Ausgehend von Arbeiten zu interkulturellen Kompetenzen (z.B. Deardorff, 2006) und von Studien zur Wirksamkeit verschiedener Lehr- und Lernformen, insbesondere zum problembasierten kooperativen Lernen (z.B. Krause, 2007), wurde eine Interventionsstudie mit Prä- und Posttest konzipiert, in der sich Schüler/innen in verschiedenen Unterrichtsfächern mit einer problembasierten interkulturellen Lerneinheit auseinandersetzen. Der Studie liegt ein 2x2-faktorielles quasiexperimentelles Design zugrunde mit den Faktoren Unterrichtsansatz (eher affektiv-kreativ/eher kognitiv-analytisch) und Unterrichtssprache (Deutsch/Englisch). Die Lerneinheit umfasste 6 x 45 Minuten und wurde im Rahmen des regulären Schulunterrichts in der Sekundarstufe I und II in neun Schulklassen (n = 143) durchgeführt. Als Kontrollgruppe dienten fünf Schulklassen (n = 66), die nicht an der Intervention teilnahmen. Anhand von Fallbeispielen in Form von sog. Critical Incidents wurden Missverständnisse oder Konflikte präsentiert, die als Folge von kulturellen Unterschieden entstehen können. In der individuellen und kooperativen Auseinandersetzung mit den kritischen Ereignissen wurden die Schüler/innen zur Perspektivenübernahme und zur Reflexion ihrer eigenen kognitiven und affektiven Prozesse (wie Kategorisierung, Beurteilung und Einfühlung) angeregt. Im Rahmen der Unterrichtseinheit entwickelten die Schüler/innen außerdem Strategien zum Umgang mit den dargestellten Situationen. Mit Hilfe von Tests, Fragebögen, Kurzinterviews und Tonbandmitschnitten wurden kognitive, motivationale und einstellungsbezogene Effekte der Intervention untersucht. Der Vortrag stellt relevante Ergebnisse vor und vergleicht diese mit Ergebnissen aus der ersten Erhebungsphase (Busse & Krause, in Druck). Hervorzuheben ist insbesondere, dass Lernende aus allen Experimentalgruppen im Mittel signifikant bessere Ergebnisse erzielten als die Lernenden aus den Kontrollgruppen. Zudem war der Lernerfolg in allen Experimentalgruppen deutlich größer als in der ersten Erhebungsphase.

Harren, Inga

Interaktive Verfahren im Unterrichtsgespräch – Eine qualitative Studie zur mündlichen Unterstützung bildungssprachlicher Praktiken

In sprachlich heterogenen Lerngruppen stellt sich – über alle Unterrichtsfächer hinweg – immer auch die Frage nach der sprachlichen Förderung und Unterstützung aller Schüler/innen und der Möglichkeiten zur spezifischen Anknüpfung an ihre jeweiligen Fertigkeiten. Unter anderem Wellington und Osborne betonen: "science teachers are (among other things) language teachers. This requires a range of strategies and skills, some of which are at a high level" (Wellington/Osborne 2001: 5).

Die Studie, deren Ergebnisse in diesem Vortrag vorgestellt werden, befasst sich zum einen mit verschiedenen bildungssprachlichen Praktiken (Morek/Heller 2012; Feilke 2012; Gogolin/Lange 2011), die Schüler/innen bei der Versprachlichung fachlicher Inhalte zu bewältigen haben. Zum anderen werden die interaktiven Verfahren untersucht, mit denen Biologielehrer/innen die mündlichen Darstellungen fachlicher Inhalte unterstützen, sowohl in Bezug auf die Strukturierung der Schülerbeiträge als auch in Bezug auf die Wahl adäquater Formulierungen.

Die qualitative Studie wurde mit konversationsanalytischer Methodik (Hutchby/Wooffitt 2005; Sidnell/Stivers 2013; Richards/Seedhouse 2005) durchgeführt und beruht auf 18 Stunden Videoaufnahmen gymnasialen Biologieunterrichts (meist aus zwei Videoperspektiven) von 9 verschiedenen Lehrkräften in 10 verschiedenen Klassen. Die Analysen umfassen das verbale Lehrer- und Schülerverhalten, ihr interaktionales Zusammenspiel sowie zum Teil multimodale Ebenen der Interaktion, beispielsweise bei der Arbeit mit grafischen Repräsentationen fachlicher Zusammenhänge (Infografiken sowie von Schüler/innen erstellte Poster).

Im Vortrag werden Ergebnisse dieser Studie präsentiert: Hierzu werden exemplarisch einige sprachliche Hürden sowie Unterstützungsverfahren vorgestellt, die speziell bei der unterrichtlichen Besprechung von (biologischen) Vorgängen zum Tragen kommen.

- Um detaillierte und präzise Darstellungen zu erreichen, müssen Schüler/innen jeweils entscheiden, welche Informationen als eigene Darstellungsschritte formuliert werden und an welchen Stellen spezifizierende Zusätze (Relativsätze, Adverbiale, zusätzliche satzförmige Einheiten...) integriert werden.
- Bei biologischen Abläufen ändert sich im Darstellungsverlauf z. T. die Benennung des jeweiligen Referenten. Während in der Alltagsinteraktion häufig ein Verweisen mittels Pro-Formen und Zeigegesten auf das gemeinsame Sichtfeld ausreicht, ist es in der Darstellung fachlicher Inhalte meist notwendig, Referenten explizit zu benennen und auch im Laufe des Gesprächs wiederholt explizit zu benennen.
- In die Versprachlichung des durch fachliche Abbildung Sichtbaren müssen Hintergrundinformationen eingeflochten werden. Auf diese Weise können Schüler/innen über eine bloße Beschreibung des Bildes hinausgehen und die biologischen Zusammenhänge erklären bzw. erläutern.

Wollen Lehrkräfte diese Aspekte im Unterrichtsgespräch unterstützen, verwenden einige unter ihnen sowohl reformulierende und/oder ergänzende Darstellungen, die eng an die Bezugsäußerungen der Schüler/innen angebunden sind, als auch Verfahren, mit denen Schüler/innen selbst zu Reformulierungen und Ergänzungen aufgefordert und angeleitet werden. Anhand von good-practice-Fällen wurden diese interaktiven Verfahren rekonstruiert und in Bezug auf ihre sequenzielle Entfaltung innerhalb des Unterrichtsgesprächs beschrieben. Sie weisen zum Teil Ähnlichkeiten mit dem Gesprächs-Verfahren auf, das bereits aus der Eltern-Kind-Interaktion (Bruner 1978) und aus dem Fremdsprachenunterricht als Scaffolding (Seedhouse 2004) bekannt ist.

Die Ergebnisse dieser Studie können sowohl innerhalb der biologiedidaktischen Ausbildung von Lehrkräften als auch unter sprach-/ grammatikdidaktischen Fragestellungen Anwendung finden. Sie könnten in der empirischen Bildungsforschung außerdem eine Basis bilden, um Kodierleitfäden für die quantifizierende Beurteilung sprachsensibler Interaktionen und sprachlicher Unterstützungshandlungen im Unterrichtsgespräch zu entwerfen.

Spengler, Marion; Brunner, Martin; Martin, Romain; Lüdtkke, Oliver

Der Einfluss von Persönlichkeit auf die Veränderung von schulischem Erfolg über vier Jahre

Erfolgreiche Bildungswege stellen eine der wichtigsten Stellschrauben für eine erfolgreiche berufliche Karriere dar. Der Grundstein für einen erfolgreichen Lebensweg wird bereits in den frühen Phasen einer schulischen Laufbahn gelegt (Almlund, Duckworth, Heckman & Kautz, 2011; Blossfeld, Roßbach & von Maurice, 2011). Studien im schulischen Kontext belegen hierbei die wichtige Rolle kognitiver und nicht-kognitiver Prädiktoren - wie beispielsweise IQ, Motivation und auch Persönlichkeitseigenschaften - für die Vorhersage von akademischem Erfolg (Gottfredson, 2002; Gustafsson & Undheim, 1996; Hailikari, Nevgi & Komulainen, 2007; Kuncel, Hezlett & Ones, 2004; Spengler, Lüdtkke, Martin & Brunner, 2013; Steinmayr & Spinath, 2009). Vor allem die Rolle von Persönlichkeitseigenschaften ist in den letzten Jahren in den Mittelpunkt vieler Untersuchungen gerückt. Es zeigte sich, dass gewissenhafte Schüler erfolgreicher in der Schule waren (Spengler et al., 2013; Porpoat, 2009).

Jedoch gibt es bisher wenige längsschnittliche Studien zum Zusammenhang der Big Five Persönlichkeitseigenschaften (O, C, E, A, N) und akademischem Erfolg. Darüber hinaus haben nur wenige Längsschnittstudien die inkrementelle Validität von Persönlichkeit über Vorwissen, Intelligenz und akademisches Selbstkonzept untersucht. Weiterhin gibt es keine empirische Studie, die den Zusammenhang von kognitiven und nicht-kognitiven Prädiktoren und Veränderung von akademischem Erfolg untersucht, nachdem für die Stabilität des akademischen Erfolgs kontrolliert wurde.

Wir haben daher untersucht, ob

- (1) Persönlichkeitsfaktoren Schulerfolg in drei Kernfächern (Mathematik, Deutsch und Französisch) über einen Zeitraum von vier Jahren vorhersagt
- (2) Persönlichkeitsfaktoren einen inkrementellen Beitrag über Intelligenz und fachspezifisches akademisches Selbstkonzept hinaus leistet
- (3) Persönlichkeitsfaktoren valide Prädiktoren darstellen, nachdem für die Stabilität der Noten kontrolliert wurde (bzw. ob Persönlichkeitsfaktoren Veränderung in den Noten vorhersagen).

Um diese Fragestellungen zu untersuchen, haben wir Daten aus der luxemburgischen Erweiterung der 2009er Welle von PISA (Programme for International Student Assessment) genutzt. Die Stichprobe stellte eine repräsentative Substichprobe des PISA Zyklus aus dem Jahr 2009 dar (53,8% der vollen PISA Stichprobe). Wir haben den Analysen hierbei Daten von zwei Schülerkohorten zugrunde gelegt, die über einen Zeitraum von 4 Jahren verfolgt wurden. Die erste Kohorte war zum Zeitpunkt der PISA-Erhebung in Klasse 9 (N = 240, M = 15,63 Jahre; SD = 0,25), die zweite Kohorte war in Klasse 10 (N = 276; M = 15,97 Jahre; SD = 0,22).

Wir haben drei Sets von Strukturgleichungsmodellen mit steigender Anzahl an Prädiktoren analysiert. In jedem Set wurde die fachspezifische Note (Deutsch, Mathematik und Französisch) für jeweils beide Kohorten vorhergesagt. In Modelset 1 wurde die Stabilität der Noten über vier Jahre modelliert. In Modelset 2 wurden zusätzlich die Persönlichkeitseigenschaften hinzugefügt, die eine signifikante bivariate Korrelation mit mindestens zwei der vier Noten aufwiesen. In Modelset 3 wurde zusätzlich das fachspezifische Selbstkonzept und Intelligenz als Prädiktoren aufgenommen. Alle finalen Modelle wiesen einen guten Fit auf. Die CFI Werte variierten zwischen .95 und .99. Der RSMEA variierte zwischen .02 und .05.

Die korrelativen Zusammenhänge haben gezeigt, dass vor allem Gewissenhaftigkeit und Offenheit jeweils positiv mit den Noten über vier Jahre hinweg zusammenhängen. Dieser Befund konnte über beide Kohorten hinweg repliziert werden. Weiterhin wiesen beide Persönlichkeitsfaktoren einen inkrementellen Beitrag zur Vorhersage der Noten über fachspezifisches akademisches Selbstkonzept und Intelligenz auf. Wenn für die Stabilität der Noten kontrolliert wurde, gab es lediglich kleine Effekte aller Prädiktoren auf die späteren Noten. Zusammenfassend lässt sich daraus ableiten, dass die Persönlichkeit eines Schülers (aber auch Intelligenz und akademisches Selbstkonzept) eher den stabilen Anteil an Noten vorhersagen als ihre Veränderung.

Jurkowski, Susanne; Hänze, Martin

Soziale Fertigkeiten und Schulleistung unter der Lupe: der Zusammenhang sozialer Fertigkeiten mit schriftlichen und mündlichen Noten

Soziale Fertigkeiten stellen Verhaltensweisen dar, die es der handelnden Person ermöglichen, sozialen Anforderungen in ihrem Alltag gerecht zu werden (Owens & Johnston-Rodriguez, 2010). Für Kinder und Jugendliche identifizierten Caldarella und Merrell (1997) fünf Dimensionen sozialer Fertigkeiten, von denen sich zwei auf das Verhalten gegenüber Peers beziehen: Beziehungsfähigkeit und Durchsetzungsfähigkeit. Die Beziehungsfähigkeit kennzeichnet eine Schülerin bzw. einen Schüler mit guten Beziehungen zu Peers und schließt Verhaltensweisen wie beispielsweise helfendes Verhalten gegenüber Anderen und das Verstehen der Sichtweise von Anderen ein. Die Durchsetzungsfähigkeit charakterisiert eine aufgeschlossene und selbstbewusste Person, die z.B. Gespräche mit Mitschülerinnen bzw. Mitschülern beginnt und sich Gruppenaktivitäten anschließt.

Eine Reihe von Studien konnte bereits zeigen, dass soziale Fertigkeiten positiv mit der Schulleistung zusammenhängen, wobei diese entweder über standardisierte Schulleistungstests oder den Notendurchschnitt im Endjahreszeugnis erhoben wurden (Elias & Haynes, 2008; Green, Forehand, Beck & Vosk, 1980; Oberle, Schonert-Reichl, Hertzman & Zumbo, 2014; Welsh, Parke, Widaman & O'Neil, 2001; Wentzel, 1991). O' Connor und Paunonen (2007) gehen jedoch davon aus, dass Zeugnisnoten ein relativ breites Kriterium für Schulleistung sind und sich aus spezifischeren Maßen wie den schriftlichen und mündlichen Leistungen zusammensetzen; je nach spezifischem Maß können sich andere Zusammenhänge zwischen den Charakteristika von Lernenden und ihrer Schulleistung ergeben. Hinweise für diese Annahme stammen aus Studien, die einen differentiellen Zusammenhang zwischen breiten Persönlichkeitsmerkmalen und der akademischen Leistung je nach spezifischem Leistungsmaß fanden (Kanfer, Wolf, Kantowitz & Ackerman, 2010; Kappe & van der Flier, 2012). Studien über den Zusammenhang von sozialen Fertigkeiten und spezifischen Maßen der Schulleistung stehen allerdings noch aus.

In der vorliegenden Studie wurde der Zusammenhang von sozialen Fertigkeiten mit der Schulleistung genauer untersucht, indem die schriftlichen und mündlichen Noten der Schülerinnen und Schüler als spezifische Leistungsmaße herangezogen wurden. In einem prospektiven Design wurden zwischen dem ersten und dem zweiten Schulhalbjahr die Beziehungsfähigkeit und die Durchsetzungsfähigkeit von 94 Schülerinnen und Schülern aus vier Klassen der neunten Klassenstufe anhand der Einschätzung ihrer Klassenlehrerinnen bzw. Klassenlehrer erhoben. Der Fragebogen war angelehnt an den Fragebogen zur Erfassung der sozialen Kompetenzen von Kindern und Jugendlichen von Jurkowski und Hänze (2014). Außerdem wurden zu diesem Messzeitpunkt als Kontrollvariablen ein standardisierter Schulleistungstest in Deutsch, und zwar die Lesetestbatterie für die Klassenstufen 8-9 (Bäuerlein, Lenhard & Schneider, 2012), und ein standardisierter Schulleistungstest in Mathematik, der Deutsche Mathematiktest für neunte Klassen (Schmidt, Ennemoser & Krajewski, 2013) erhoben. Im Verlauf des zweiten Schulhalbjahres wurden die schriftlichen Noten in Klassenarbeiten und die mündlichen Noten für die Unterrichtsbeteiligung, die von den jeweiligen Fachlehrern in Deutsch und Mathematik gegeben wurden, gesammelt. Aus den einzelnen schriftlichen Noten und den einzelnen mündlichen Noten wurde jeweils der Durchschnitt gebildet, einmal für das Fach Deutsch und einmal für das Fach Mathematik.

Eine pfadanalytische Auswertung unter Berücksichtigung der standardisierten Schulleistungstests und der dummykodierte Klassenzugehörigkeit als weitere Prädiktoren ergab, dass die Beziehungsfähigkeit sowohl die schriftlichen als auch die mündlichen Leistungen in Deutsch und Mathematik vorhersagte. Für die Durchsetzungsfähigkeit zeigte sich dagegen ein signifikanter Effekt auf die mündlichen Leistungen in Deutsch und Mathematik, jedoch war der Pfad von der Durchsetzungsfähigkeit auf die schriftlichen Leistungen nicht signifikant.

Die vorliegenden Resultate werden dahingehend interpretiert, dass gute Beziehungen zu Peers den Wissenserwerb fördern, z.B. indem Peers Ressourcen bereitstellen (Wentzel & Watkins, 2011), was sich positiv sowohl auf die schriftlichen als auch die mündlichen Noten auswirkt. Die Durchsetzungsfähigkeit dagegen unterstützt wahrscheinlich das Einbringen bereits vorhandenen Wissens, so dass aufgeschlossene und selbstbewusste Schülerinnen und Schüler bessere mündliche Noten zeigen als weniger durchsetzungsfähige Schülerinnen und Schüler. Für die Förderung schulischer Leistungen verweisen die vorliegenden Ergebnisse darauf, auch Ansätze zur Stärkung sozialer Fertigkeiten zu berücksichtigen.

Freitag
13. März

11:15 Uhr

GC 03/149

Chair(s)

Spengler, Marion

3

Bergold, Sebastian; Wirthwein, Linda; Rost, Detlef H.; Steinmayr, Ricarda

Schulbesuchsdauer und Intelligenz: Ein Vergleich zwischen G8- und G9-Schülern

Studien zur Intelligenzentwicklung zeigen, dass die Schulbesuchsdauer auf die Intelligenzentwicklung einwirkt (z.B. Cahan & Cohen, 1989; Cliffordson & Gustafsson, 2008; Rost & Wild, 1995). Hierbei konnte gezeigt werden, dass die Schulbesuchsdauer nicht nur spezifische, curriculumsgebundene Basiskompetenzen fördert, sondern auch allgemeine, curriculumsunabhängige Fähigkeiten wie das logische Denkvermögen (Artman, Cahan & Avni-Babad, 2006; Cahan, Greenbaum, Artman, Deluya & Gappel-Gilon, 2008). Es wird daher angenommen, dass die Steigerung der Leistungen in Intelligenztests auch und vor allem durch eine tatsächliche Steigerung intelligenznaher Fähigkeiten wie dem logischen Schlussfolgern bewirkt wird.

Somit stellt sich die Frage, ob sich vor dem Hintergrund der um ein Jahr verkürzten Unter- und Mittelstufe im Zuge der G8-Reform Intelligenzunterschiede zwischen G8- und G9-Gymnasiasten im ersten Jahr des Oberstufenbesuchs ergeben. Diese Fragestellung ist umso interessanter, als G8- und G9-Gymnasiasten bis dahin das gleiche Curriculum absolvieren, so dass positive Effekte der längeren Schulzeit von G9-Gymnasiasten auf eine bessere Förderung curriculumsunabhängiger Fähigkeiten zurückzuführen wären. Hierzu wurden zwei Studien (A und B) durchgeführt.

In Studie A wurden $n = 81$ G8-Schüler (55.6% Mädchen; Alter: $M = 15.16$ Jahre, $SD = 0.37$) und $n = 80$ G9-Schüler (55.0% Mädchen; $M = 16.39$ Jahre, $SD = 0.49$) der ersten Jahrgangsstufe der Sekundarstufe II eines Gymnasiums in NRW hinsichtlich ihrer Leistungen im Berliner Intelligenzstruktur-Test (Jäger, Süß & Beauducel, 1997) verglichen. Bei den Schülern handelte es sich um die erste G8- und die letzte G9-Kohorte. Sie wurden in der Oberstufe gemeinsam unterrichtet. Die Testung fand kurz nach der Zusammenführung der beiden Stufen statt. Das Alter und das elterliche Bildungsniveau wurden in den durchgeführten MANCOVAs und nachfolgenden ANCOVAs als Kovariaten berücksichtigt.

Die G9-Schüler schnitten bei statistischer Kontrolle der Kovariaten in allen Denkinhalten bedeutsam besser ab als die G8-Schüler (verbal: $F(1, 157) = 6.20$, $p = .014$, $d = 0.40$; numerisch: $F(1, 157) = 10.60$, $p = .001$, $d = 0.52$; figural-bildhaft: $F(1, 157) = 4.63$, $p = .033$, $d = 0.35$). In den Denkopoperationen „Bearbeitungsgeschwindigkeit“ ($F(1, 157) = 10.41$, $p = .002$, $d = 0.51$) und „Verarbeitungskapazität“ ($F(1, 157) = 5.74$, $p = .018$, $d = 0.38$) wiesen sie ebenfalls höhere Werte auf. In der Merkfähigkeit ($d = 0.23$) und im Einfallsreichtum ($d = 0.26$) erreichten die Unterschiede zugunsten der G9-Schüler keine Signifikanz. Im BIS-Gesamtwert zeigte sich wiederum ein statistisch und praktisch bedeutsamer Unterschied zugunsten der G9-Schüler ($F(1, 157) = 9.57$, $p = .002$, $d = 0.49$). Um die Ergebnisse zu replizieren, wurde in einer zweiten Studie (B) eine weitere Stichprobe von G8- und G9-Schülern unter Einsatz eines anderen Instruments untersucht.

In Studie B wurden die Leistungen von $n = 244$ Schülern zweier im Herbst 2012 bzw. 2013 erhobenen G8-Stufen (61.5% Mädchen; Alter: $M = 15.23$ Jahre, $SD = 0.42$) und $n = 204$ Schüler zweier im Herbst 2007 bzw. 2008 erhobenen G9-Stufen (55.4% Mädchen; $M = 16.33$ Jahre, $SD = 0.47$) der ersten Jahrgangsstufe der Sekundarstufe II eines Gymnasiums in NRW im Intelligenz-Struktur-Test 2000 R (Liepmann, Beauducel, Brocke & Amthauer, 2007) analog zum methodischen Zugang in Studie A miteinander verglichen.

Die G9-Schüler schnitten unter Konstanzhaltung des Alters und des elterlichen Bildungsniveaus in allen Skalen des IST 2000 R statistisch und praktisch bedeutsam besser ab als die G8-Schüler (verbale Intelligenz: $F(1, 444) = 37.59$, $p < .001$, $d = 0.58$; numerische Intelligenz: $F(1, 444) = 36.54$, $p < .001$, $d = 0.57$; figural-räumliche Intelligenz: $F(1, 444) = 22.62$, $p < .001$, $d = 0.45$). Im Gesamtwert zeigte sich ein mittlerer bis großer Effekt ($F(1, 444) = 57.13$, $p < .001$, $d = 0.72$).

Lotz, Christin; Sparfeldt, Jörn; Greiff, Samuel

Komplexes Problemlösen und die Vorhersage von Schulleistung. Prädiktiv trotz „gutem g“ ?

Komplexes Problemlösen (KPL), verstanden als die Fähigkeit aktiv und erfolgreich mit einem unbekanntem dynamischen System zu interagieren, hat in der empirischen Bildungsforschung einen hohen Stellenwert, wie dessen Erfassung in internationalen Schulleistungsuntersuchen (z.B. im PISA 2012 Zyklus) zeigt. Auch wenn KPL und Intelligenz hoch miteinander zusammenhängen (z.B. Kröner et al., 2005; Greiff et al., 2013), konnten jüngere empirische Studien zeigen, dass KPL über Intelligenz hinaus bedeutsame Varianzanteile in Schulnoten aufklärte (z.B. Greiff et al., 2013). Allerdings wurde Intelligenz in bisherigen Studien entweder durch einige wenige Subtests und/oder durch vergleichsweise schmale Intelligenzgruppenfaktoren (z.B. figurales Reasoning) erfasst. Eine breitere Operationalisierung der Intelligenz im Sinne eines „guten g“ (Jensen & Weng, 1994) böte dabei die Möglichkeit, spezifische Intelligenzgruppenfaktoren (figural, numerisch und verbal) im Zusammenspiel mit KPL bei der Vorhersage von Schulleistungen in den Blick zu nehmen.

In dieser Arbeit soll daher die statistische Vorhersage von Schulleistungen (operationalisiert durch Schulnoten und standardisierte Schulleistungstests) aus KPL und Intelligenz gemeinsam unter Berücksichtigung einer breiteren Intelligenzoperationalisierung als in bisherigen Untersuchungen im Vordergrund stehen. Durch die Intelligenzmodellierung im nested-factor model kann über die allgemeine Intelligenz g hinaus die Vorhersagekraft spezifischer g-unabhängiger und mit den Schulfächern Deutsch bzw. Mathematik affinen materialartspezifischen Gruppenfaktoren (verbal bzw. numerisch) inspiziert werden.

Dafür bearbeiteten N = 496 10.-Klässler (Gymnasium) oder 11.-Klässler (Gesamtschule) zehn Subtests des Berliner Intelligenzstruktur-Tests (Jäger et al., 1997), verschiedene computergestützte Szenarien zum komplexen Problemlösen (MicroDYN; Greiff et al., 2012) und einen standardisierten Schulleistungstest in entweder Deutsch (Leseverständnistest; n = 251) oder Mathematik (n = 245). Zusätzlich wurden die Zensuren in Deutsch und Mathematik von den Schülerinnen und Schülern erfragt. Es wurden für Deutsch und Mathematik getrennte Strukturgleichungsmodelle zur simultanen Vorhersage des Schulleistungstests und der Schulnote aus KPL und Intelligenz (im nested-factor model) unter Berücksichtigung der hierarchischen Datenstruktur berechnet.

Die Ergebnisse für beide Modelle (Deutsch bzw. Mathematik) zeigen, dass KPL und Intelligenz hoch miteinander korrelieren. Im Modell für Mathematik wird die Mathematiktestleistung insbesondere durch die allgemeine Intelligenz g und in geringerem Maße durch KPL prädiziert. Darüber hinaus zeigten die g-unabhängigen, spezifischen numerischen und figuralen Intelligenzfaktoren inkrementelle Varianzaufklärung der Mathematiktestleistung. Auf die Mathematiknote zeigte sich lediglich von der spezifisch-numerischen Intelligenzfacette ein kleiner direkter bedeutsamer Pfad. Im Modell für Deutsch wird die Lesetestleistung durch den spezifisch-verbalen Intelligenzfaktor vorhergesagt, die Deutschnote ebenfalls durch den spezifisch-verbalen Intelligenzfaktor und den spezifisch-numerischen Intelligenzfaktor. Abschließend werden die Ergebnisse in Bezug auf die Vorhersagekraft des KPL über die allgemeine und spezifische Intelligenz hinaus diskutiert.

Freitag
13. März

11:15 Uhr

GC 03/149

Chair(s)
Spengler, Marion

Hülsmann, Carolin; Walpuski, Maik

Gymnasiale Kurswahlen unter dem Blickwinkel des Erwartungs-mal-Wert Modells

Theoretischer Rahmen: Wie in vielen anderen Ländern, in denen naturwissenschaftliche Fächer in der Oberstufe nicht verpflichtend unterrichtet werden, wählt auch in Deutschland ein Großteil der Schülerschaft das Fach Chemie schon zu Beginn der Sekundarstufe II ab (MSW NRW, 2013). In Anbetracht des sinkenden Interesses der Schülerinnen und Schüler (SuS) an naturwissenschaftlichen Karrieren und der hohen Studienabbruchquoten dieser Fächer (Heublein et al., 2012) beschäftigen sich seit Längerem einige Arbeiten mit der Ermittlung von Einflussfaktoren auf den Entscheidungsprozess.

Zur Vorhersage des Kurswahlverhaltens in der Oberstufe wird der Prozess der Entscheidungsfindung in viele Studien aus dem Blickwinkel des expectancy-value model of achievement-related choices von Eccles und Wigfield (2002) betrachtet. Dabei wird leistungsbezogenes Wahlverhalten direkt auf eine Erwartungs- und eine Wertkomponente der anstehenden Entscheidung zurückgeführt. Unter der Erwartungskomponente werden die individuell an eine Handlung geknüpften Erfolgserwartungen verstanden. Die Wertkomponente umfasst hingegen alle subjektiven mit einer Leistung oder Entscheidung verbundenen Werte, die sich darauf beziehen wie Freude berekend (interest-enjoyment value), persönlich wichtig (attainment value), nützlich (utility value) oder kostspielig (relative cost) ein Ergebnis empfunden wird.

Fragestellung & Zielsetzung: Um der Entwicklung sinkender Kursteilnehmerzahlen im Fach Chemie in der Oberstufe entgegen zu wirken, ist es das zentrale Ziel dieser Studie zu ermitteln, welche Motive für die Wahl bzw. Abwahl des Faches Chemie von Bedeutung sind und welche Faktoren sich als Prädiktoren für den Erfolg im Fach identifizieren lassen. Mit Hilfe der gewonnenen Erkenntnisse sollen Möglichkeiten erschlossen werden, mit denen die Situation verbessert werden kann.

Methode: Basierend auf der Erwartungs- und Wertkomponente werden Fachwissen, Chemienoten, Fähigkeitsselbsteinschätzungen und Selbstwirksamkeitserwartungen sowie das chemische Fachinteresse und die Berufswünsche der SuS erfasst. Für die Überprüfung des Fachwissens wurde ein Multiple-Choice Test zum Inhaltsbereich chemische Reaktionen entwickelt, der Items zu verschiedenen Themenfeldern der 9. - 12. Klasse umfasst und im Multi-Matrix Design eingesetzt wurde. Die anderen Einflussgrößen wurden über standardisierte Skalen von Jerusalem und Satow (1999) zur Selbstwirksamkeitserwartung und Dickhäuser (2002) zum Fähigkeitsselbstkonzept sowie Schulz (2010) zum Interesse ermittelt. Als Kontrollvariable wurden nonverbale Skalen des Kognitiven Fähigkeitstest von Heller und Perleth (2000) verwendet.

Eingesetzt wurden die Testinstrumente in einem Quasi-Längsschnitt an drei Jahrgangsstufen (9. - 12. Klasse), indem die SuS zu zwei Messzeitpunkten im Abstand von einem Jahr befragt wurden. Alle SuS wurden am Ende der Sekundarstufe I getestet, wobei diejenigen SuS, die Chemie in der Oberstufe weiterhin belegten, zum Ende der Einführungsphase der Oberstufe erneut befragt wurden (Teilstichprobe I). Nach getroffener Kurswahlentscheidung für die Qualifikationsphase nahmen anschließend diejenigen SuS, die das Fach Chemie fortführten, jeweils zum Schuljahresende an einer Befragung teil (Teilstichprobe II). Die beiden Teilstichproben wurden in einem echten Längsschnitt erhoben und bilden über Mittelwertvergleiche einen Quasi-Längsschnitt über die gesamte Oberstufe ab. An der Studie nahmen insgesamt 2.137 SuS aus 14 nordrhein-westfälische Gymnasien und Gesamtschulen in den Schuljahren 2012/13 und 2013/14 teil.

Ergebnisse: Ersten Ergebnissen zufolge weisen die Skalen zum Interesse, Fähigkeitsselbstkonzept, den Selbstwirksamkeitserwartungen und kognitiven Fähigkeiten sehr gute Reliabilitäten mit Cronbach's α Werten von .89 bis .97 auf. Die Ergebnisse der Rasch-Analyse für den Fachwissenstest zeigen eine EAP/PV-Reliabilität von .61, wenn die SuS zu gleichen Anteilen aus allen vier Jahrgängen in die Analyse einbezogen werden. Betrachtet man ausschließlich Lernende der letzten beiden Schuljahre dann werden Werte von .63 bis .71 erreicht.

In Bezug auf die Motive der Wahlentscheidung zeigen Korrelationsanalysen, dass Interessen ($r=.611$, $p\leq.001$) und Kompetenzüberzeugungen der SuS ($r=.522$, $p\leq.001$) sowie deren Berufswünsche ($r=.510$, $p\leq.001$) eine bedeutende Rolle bei der Kurswahlentscheidung spielen, während Chemienoten ($r=.386$, $p\leq.001$) und kognitive Fähigkeiten ($r=.213$, $p\leq.001$) der Lernenden weniger wichtig zu sein scheinen. Bedauerlicherweise wählt ein Großteil der SuS, der diese Voraussetzungen nicht erfüllen, das Fach bereits zu einem sehr frühen Zeitpunkt ab.

Koch, Stefan

DFG-Förderung in der Empirischen Bildungsforschung: Aktuelle Informationen zu Förderprogrammen, Antragsstellung und Begutachtungs- und Entscheidungsverfahren

Der workshop informiert über die verschiedenen Förderprogramme der DFG, die für die Empirische Bildungsforschung relevant sind. Es wird das gesamte Verfahren - von der Antragstellung über die Begutachtung bis zu endgültigen Entscheidung - vorgestellt. Im Anschluss wird es viel Raum für Fragen der Teilnehmerinnen und Teilnehmer geben. Die Veranstaltung richtet sich vorrangig, aber nicht ausschließlich an Erstantragsteller/-innen.

W03 Workshop

Freitag
13. März

11:15 Uhr

HGC 30

3

Maag Merki, Katharina
Heterogenität und Schulentwicklung

Aktuelle Angebots-Nutzungsmodelle oder dynamische Modelle der Schuleffektivität gehen davon aus, dass Lernangebote in Abhängigkeit der kognitiven, motivationalen, sozialen oder familiären Voraussetzungen der Schülerinnen und Schüler differieren bzw. differieren müssen, sollen sie ertragreich für das Lernen der Schüler/innen sein. Zudem zeigen Befunde aus der Professionalisierungs- und Unterrichtsentwicklungsforschung, dass Lernangebote ihr Potenzial für effektives Lernen der Schüler/innen nur dann entfalten, wenn es gelingt, sie in einem systematischen und zielorientierten Entwicklungsprozess neuen Kontextbedingungen oder Konstellationen in der Schule / im Unterricht anzupassen und dabei die entsprechenden Rekontextualisierungsprozesse sowie Voraussetzungen der beteiligten Akteure zu berücksichtigen.

Dies impliziert eine doppelte Heterogenität:

Heterogenität auf der Nutzungsseite erfordert ein differenzielles Angebot auf Klassen- und Schulebene, um den unterschiedlichen Voraussetzungen der Schüler/innen gerecht zu werden. Angebote in Klassen oder Schulen mit einem hohen Anteil an Kindern mit Migrationshintergrund werden sich beispielsweise von jenen unterscheiden mit einem geringen Anteil an Kindern mit Migrationshintergrund. Zudem werden binnendifferenzierende Angebote zwischen den Schüler/innen innerhalb der Schule oder Klasse sichtbar werden. Unter dieser Perspektive kann als Analysegegenstand die Heterogenität der Schüler/innen in Klassen und Schulen und ihre Relation zum Lernangebot betrachtet werden, wobei die Varianz zwischen Schulen / Klassen aufgrund unterschiedlicher Nutzungskonstellationen von besonderem Interesse ist.

Heterogenität auf der Angebotsseite wiederum bedeutet, dass trotz vergleichbarer Kontextbedingungen bzw. Nutzungsvoraussetzungen unterschiedliche Strategien auf Schul- oder Klassenebene eingesetzt werden. Dabei wird berücksichtigt, dass Lernangebote Ergebnisse eines mehr oder weniger expliziten Rekontextualisierungs- und Transformationsprozesses in der Schule, im Kollegium oder von einzelnen Lehrpersonen sind, welcher beeinflusst wird von den Kompetenzen, Einstellungen, Orientierungen oder Motivationen der verschiedenen Akteure. Unter dieser Perspektive kann als Analysegegenstand die Heterogenität zwischen den Schulen, Kollegien oder Fachteams bezüglich der Gestaltung des Lernangebots in Relation zu den Nutzungsmöglichkeiten der Schüler/innen betrachtet werden, wobei vor allem die Varianz zwischen den Schulen / Klassen aufgrund ähnlicher Nutzungskonstellationen interessiert.

Diese zweite Perspektive wird im Fokus des Referates stehen. Neben der theoretischen und empirischen Erörterung des Problemzusammenhangs werden das Forschungsdesign sowie erste Ergebnisse eines aktuellen, kontrastiv angelegten Forschungsprojektes zur Diskussion gestellt. In diesem werden u.a. schulinterne Orientierungen von Schulleitungen in Primarschulen der Stadt Zürich mit einem vergleichbaren sozialen Kontext (hoher Anteil an Schüler/innen mit Migrationshintergrund), aber grosser Heterogenität bezüglich der eingesetzten Schulentwicklungsstrategien hinsichtlich adaptiv-kompensatorischer Förderstrategien untersucht.

Gorges, Johannes-Franz; Schneider, Michael; Mayer, Anne-Kathrin; Krampen, Günter
“Konzeptuelles Verständnis der Funktionsweise des menschlichen Gedächtnisses” (KVG): Konstruktion und erste Validierung eines Tests

Theoretischer Hintergrund und Fragestellung: Das menschliche Gedächtnis ist eines der zentralen Konstrukte der Lebens- und Sozialwissenschaften; seine Funktionsprinzipien zu verstehen ist in einer Vielzahl von Kontexten relevant, so z.B. im Kontext von Kompetenzerwerbsprozessen im Bildungssektor (Neale, Smith & Johnson, 1990) oder der Beurteilung und Erklärung von Hirnfunktionen in der Neurologie und Psychiatrie und Zeugenaussagen in der Forensik (Loftus, Doyle & Dysert, 2008). Bisher existiert jedoch kein Instrument, das ein umfassendes konzeptuelles Verständnis der Funktionsweise des menschlichen Gedächtnisses erfasst. Das Ziel der Studie besteht daher in der Entwicklung und Validierung eines solchen Instruments, das neben der Untersuchung inhaltlicher Fragen zur Rolle des Gedächtnisverständnisses auch die quantitative Analyse von Konzeptwandelprozessen und ihren Determinanten ermöglichen sollte (vgl. Schneider & Hardy, 2013). Bei der Testentwicklung wurden neueste Erkenntnisse der Konzeptwandelersforschung berücksichtigt. So wird Konzeptwissen multidimensional operationalisiert (Schneider & Stern, 2010) und stets in Bezug auf konkrete Situationen und Beispiele erfasst (diSessa, Gillespie & Esterly, 2004).

Der Test wurde psychometrisch überprüft. Die konvergente und diskriminante Validität wurde durch Bestimmung der Zusammenhänge zu Referenzvariablen ermittelt.

Methode: Der Test „Konzeptuelles Verständnis des menschlichen Gedächtnisses“ (KVG) wurde theoriegeleitet entwickelt. Er bildet das Verständnis eines basalen Gedächtnisprinzips ab, wonach das Gedächtnis Informationen nicht unverändert speichert und wiedergibt, sondern sie aktiv verarbeitet und dadurch modifiziert (Lynn & McConkey, 1998). Dieses Prinzip wird im Kontext von drei grundlegenden Gedächtnismechanismen thematisiert: Chunking, Interferenz und Quellenkontrolle. Zu jedem der drei Gedächtnismechanismen wurden jeweils drei klassische Experimente aus der Gedächtnisforschung allgemeinverständlich beschrieben ohne Nennung der Ergebnisse. Zu jedem der Experimente wurden den Probanden im Test sechs mögliche Ausgänge mit unterschiedlichen Begründungen vorgegeben. Eine Aussage beschreibt das tatsächliche Ergebnis der Studie, die anderen fünf beschreiben populäre Misskonzepte, die aus qualitativen Interviews und der Literatur zu Misskonzepten entnommen wurden.

Stichprobe 1 umfasste 137 bzw. 126 Psychologiestudierende (Alter: $M = 20.4$ Jahre zu Messzeitpunkt 1), die vor (t_1) und nach (t_2) dem Besuch einer Einführungsvorlesung zur Gedächtnispsychologie untersucht wurden. Die Ersterhebung erfolgte zu Beginn des ersten Studiensemesters. Stichprobe 2 umfasste 202 Studierende der Bildungswissenschaften aus allen Semestern (Alter: $M = 23.31$ Jahre).

Ergebnisse: Die Lösungsrate des Gesamttests lag in beiden Stichproben und an beiden Messzeitpunkten zwischen 66 Prozent ($SD = 8$; $min = 45\%$; $max = 90\%$) und 77 Prozent ($SD = 9$; $min = 55\%$; $max = 95\%$). Die Werte von Cronbachs Alpha lagen in den drei Stichproben zwischen .80 und .85 und sind somit höher als erwartet. Wie erwartet führte der Besuch einer Vorlesung zum Thema menschliches Gedächtnis zu einem hochsignifikanten Anstieg der Gesamtwerte um $d = .71$ und Psychologiestudierende haben ein besseres Verständnis des Gedächtnisses als Studierende der Bildungswissenschaften ($d = .63$).

Erwartungsgemäß korreliert das KVG mit impliziten Theorien über Gedächtnisleistung (Dweck, Chiu & Hong, 1995; Spinath, 1998). Schwache Korrelationen mit Intelligenz (APM; Raven, Raven & Court, 1993) und der Deutschnote weisen auf eine gute diskriminante Validität hin. Wie erwartet ist das KVG unkorreliert mit epistemologischen Überzeugungen (CAEB; Stahl & Bromme, 2007) und Persönlichkeitsfaktoren (NEO-FFI-30; Körner et al., 2007). Mit dem akademischen Selbstkonzept (Dickhäuser, Schöne, Spinath & Stiensmeier-Pelster, 2002) korreliert das KVG nur nach Besuch der Vorlesung Gedächtnis. Zu weiteren Schulnoten wurden wie erwartet keine hohen Zusammenhänge gefunden.

Der KVG kann zum Beispiel bei Lehrern, Psychotherapeuten, Polizisten und Richtern eingesetzt werden, um typische falsche Annahmen zu diagnostizieren und ggf. zu korrigieren, sodass diese professionellem beruflichem Handeln nicht mehr entgegenstehen. Nachfolgende Studien können den hier vorgestellten Test nutzen, um Konzeptwandelprozesse und den Zusammenhang von konzeptuellem Wissen zu Referenzvariablen quer- und längsschnittlich zu untersuchen.

Freitag
13. März

15:30 Uhr

GA 03/49

Chair(s)

Gorges, Johannes-Franz

Zehner, Fabian; Christine, Sälzer; Frank, Goldhammer

Wie automatisches Kodieren offener Kurztextantworten zur Verbesserung von Assessments beitragen kann

Beim automatischen Kodieren kategorisiert oder bewertet ein Computer Kurztextantworten. Solche offenen Antwortformate können entgegen geschlossenen die Konstruktvalidität eines Erhebungsinstruments verbessern (z. B. für Mathematik: vgl. Bridgeman, 1991; für Lesen: vgl. Millis, Magliano, Wiemer-Hastings, Todaro & McNamara, 2011). Trotzdem sind geschlossene Formate heutzutage geläufiger, weil sie die Datenverarbeitung vereinfachen und keiner menschlicher Kodierer bedürfen, die wiederum Nachteile mit sich bringen (vgl. Bejar, 2012). Die Präsentation berichtet erste Ergebnisse eines Projekts, in dem eine Software aus bestehenden Bibliotheken zur Verarbeitung natürlicher Sprache implementiert wurde, um das automatische Kodieren durch den Computer zu ermöglichen. Es hat zum Ziel, Wissenschaftlern die Nutzung von offenen Antwortformaten durch automatische Auswertung zu erleichtern, indem eine niederschwellig benutzbare Software bereitgestellt wird. Auf dieser Grundlage ließen sich dann auch moderne Testkonzepte wie computerisiertes adaptives Testen mittels offenen Antwortformats realisieren. Die Software wird voraussichtlich 2015 offen und frei veröffentlicht. Die präsentierte Studie zeigt anhand von Daten aus der PISA-Studie 2012, dass die Methoden zu reliablen und validen Ergebnissen führen können.

Damit der Computer zu seiner Kodierung kommt, wird zuerst jede Kurzantwort vorverarbeitet und anhand eines zuvor generierten semantischen Raums in eine quantifizierte Repräsentation ihrer Semantik transformiert. Zweitens wird ein Clustermodell gebildet, indem diese numerischen Repräsentationen genutzt werden, um ähnliche Antworten zu gruppieren. Drittens wird jeder dieser Gruppen durch maschinelles Lernen ein interpretierbarer Code zugewiesen. Die Präsentation beleuchtet zwei Forschungsfragen: Einerseits wird die allgemeine Leistungsfähigkeit der Methoden sowie deren Abhängigkeit von verschiedenen Items geprüft und andererseits wird die Leistungsfähigkeit bei unterschiedlichen Stichprobengrößen simuliert. Insgesamt wurden $N = 41.990$ Antworten von zehn dichotomen PISA-Items in den Analysen verwendet (auf deren Charakterisierung hier nicht eingegangen werden kann). Zur Evaluierung der Modelle wurde die wiederholte zehnfache Kreuzvalidierung angewendet (vgl. Witten, Frank & Hall, 2011).

Die Ergebnisse zeigen, dass die automatische Kodierung zu guter bis zu exzellenter Übereinstimmung zwischen Mensch und Computer führen kann, wobei die Leistung zwischen Items variiert. Für alle Items wird eine hohe prozentuale Übereinstimmung von 76 bis 98 Prozent wie auch eine exzellente interne Konsistenz ($,750 \leq \kappa_{c:c} \leq ,983$) erreicht. Lediglich der Kappa-Wert des Items 3•SCHWIERIG zur Übereinstimmung zwischen Mensch und Computer ist niedrig ($,047$), sieben weitere weisen eine gute Übereinstimmung von $\kappa_{m:c} = ,503$ bis $,728$ auf. Die zwei Items 10•NATURWISSENSCHAFTEN und 4•AUFLISTUNG erreichen exzellente Übereinstimmung von $\kappa_{m:c} = ,761$ und $\kappa_{m:c} = ,955$. Betrachtet man zusätzlich den relativen Zugewinn an Übereinstimmung durch die Clusteranalyse, fallen zusätzlich noch die Items 1•EINFACH und 5•STILELEMENT ERKLÄREN auf. Dieses eine schlecht und die zwei nicht optimal automatisch kodierten Items haben gemein, dass die Richtigkeit ihrer Antworten teilweise durch synsemantische Wörter (z. B. Funktionswörter wie wofür) beeinflusst wird. Zur Richtigkeit der Antworten in den gut funktionierenden Items tragen vor allem autosemantische Wörter bei (z. B. semantisch bedeutsame Verben wie entschuldigen); als Extremfall kann das nahezu perfekt kodierte Item 4•AUFLISTUNG genannt werden, bei dem die Schüler lediglich vier Wörter aus dem Stimulustext nennen müssen. Dass dieser Antworttyp nicht nur für das Lese-Konstrukt relevant ist, konnte mit den Items 9•MATHE und 10•NATURWISSENSCHAFTEN gezeigt werden, die gut bzw. exzellent automatisch kodiert werden. Die Datenmengensimulation zeigt, dass die Methoden im Bereich von $n = 1.661$ bis 4.152 gleichwertig und mit etwas niedrigerer jedoch akzeptablen Leistung bis hinunter zu $n = 249$ genutzt werden können und somit nicht nur Large-Scale-Studien, sondern auch sozialwissenschaftlichen Studien üblicher Größenordnung zur Verfügung stehen.

In der Präsentation werden exemplarisch acht Möglichkeiten aufgezeigt, mit denen das automatische Kodieren zur Verbesserung von Assessments beitragen kann. So etwa neben dem Ermöglichen offenen Antwortformats für adaptives Testen auch das Reduzieren des Kodieraufwands für menschliche Kodierer durch Gruppierung der Antworten.

Engelmann, Katharina; Fischer, Frank

Eine metaanalytische Betrachtung der Effekte von Interventionen auf wissenschaftliches Denken und Argumentieren

Theoretischer Hintergrund: Die Thematik des wissenschaftlichen Denkens und Argumentierens (WDA) im Bildungsbereich hat in den letzten Jahren erheblich an Bedeutung gewonnen. Gleichzeitig wurden Defizite im wissenschaftlichen Denken von Lernenden aller Altersgruppen festgestellt (z.B. Zimmerman, 2000). Die Förderung von WDA erscheint daher als wichtige Aufgabe.

Dem Konstrukt von WDA liegen eine Vielzahl von Wissenskomponenten und Fertigkeiten zugrunde, die sich heuristisch drei Komponenten von WDA zuordnen lassen. Als ersten Komponente werden Aktivitäten der Erkenntnisgewinnung (z.B. Klahr and Dunbar, 1988) betrachtet, wie der Generierung von Forschungsfragen und der Evaluation von Evidenzen (Fischer et al., 2014). In der zweiten Gruppe liegt der Fokus auf wissenschaftlicher Argumentation (z.B. Kuhn, 2010; Osborne, 2010). Als dritte Komponente werden die Annahmen zum Wesen der Wissenschaft untersucht (z.B. Lederman & Lederman, 2014; Lederman & Zeidler, 1987).

Interventionen zur Förderung von WDA binden Lernende typischerweise in Aktivitäten ein. Diese Aktivitäten lassen sich in passive, aktive, konstruktive und interaktive Lernaktivitäten einteilen (Chi, 2009). Der Hypothese von Chi (2009) nach führen interaktive Lernaktivitäten zu den besten Lernergebnissen, gefolgt von konstruktiven Lernaktivitäten. Förderansätze zu WDA wurden bisher allerdings weder daraufhin untersucht, welche Komponente von WDA und welche Lernaktivitäten sie einbinden, noch wie sich diese Unterschiede auf den Lernerfolg auswirken.

Fragestellungen: (1) Inwieweit kann WDA gefördert werden? (2) Inwieweit hängt der Erfolg von Interventionen von den geförderten Komponenten von WDA ab? (3) Inwieweit hängt der Erfolg von Interventionen von den geförderten Lernaktivitäten ab?

Methode: Zur Beantwortung der Fragenstellungen wurde eine systematische Literatursuche in den Datenbanken PsycINFO und ERIC durchgeführt. Eine Suche der Begriffe scientific reasoning, scientific thinking, etc. in Publikationstiteln erbrachte 664 Treffer. 15 Studien entsprachen den folgenden Kriterien und wurden in die Metaanalyse inkludiert: In der Studie wurde eine Intervention berichtet, deren Wirksamkeit in einem Between Group Design mit Posttest quantifiziert und zwischen 1988 – 2012 publiziert wurde.

Anhand der angebenen deskriptiven und inferenzstatistischen Ergebnisse der Studien wurde jeweils eine Effektgröße berechnet. Die Interventionen wurden bezüglich der geförderten Komponenten von WDA und Lernaktivitäten gruppiert. Die Metaanalyse wurde auf Grundlage des Random-Effekt-Modells (Borenstein, Hedges, Higgins, & Rothstein, 2010) berechnet.

Ergebnisse und Diskussion: Zur Beantwortung der ersten Fragestellung wurde die mittlere Effektgröße und Heterogenität zwischen den Studien berechnet. Es wurde ein großer mittlerer Effekt ($g = 0.85$, $se = 0.13$, $p < 0.01$) der Interventionen auf WDA gefunden, allerdings auch eine hohe Heterogenität zwischen den Studien ($Q(14) = 96.34$, $v = 0.22$, $p < 0.01$). WDA lässt sich also erfolgreich fördern, es gibt allerdings große Unterschiede zwischen den Studien.

Zur Beantwortung der zweiten Fragestellung wurde eine Moderationsanalyse mit den Komponenten von WDA durchgeführt. Deskriptiv zeigten sich Unterschiede zwischen Interventionen, die Aktivitäten der Erkenntnisgewinnung förderten ($g = 1.40$, $se = 0.34$, $p < 0.01$), Interventionen, die Annahmen zum Wesen der Wissenschaft förderten ($g = 0.86$, $se = 0.32$, $p < 0.01$) und Interventionen, die wissenschaftliches Argumentieren förderten ($g = 0.68$, $se = 0.23$, $p < 0.01$). Der Unterschied zwischen den Interventionen erwies sich allerdings als nicht signifikant ($Q(2) = 3.07$, $p > 0.05$).

Zur Beantwortung der dritten Fragestellung wurde eine Moderationsanalyse mit der Lernaktivität durchgeführt, deren Ergebnis einen signifikanten Unterschied ($Q(1) = 4.31$, $p < 0.05$) zeigte. Interventionen mit konstruktiven Lernaktivitäten ($g = 1.25$, $se = 0.22$, $p < 0.01$) hatten einen größeren Effekt auf WDA als Interventionen mit interaktiven Lernaktivitäten ($g = 0.68$, $se = 0.16$, $p < 0.01$).

Aus den Ergebnissen kann geschlussfolgert werden, dass WDA durch Interventionen gefördert werden kann. Die Unterstützung konstruktive Lernaktivitäten ist, entgegen der ICAP-Hypothese (Chi, 2009), erfolgreicher als die Unterstützung interaktiver Lernaktivitäten. Es stellt sich die Frage, ob die ICAP-Hypothese im Bereich von WDA keine Gültigkeit besitzt oder ob die interaktiven Lernaktivitäten eine gezieltere Unterstützung brauchen um höhere Effekte zu erzielen.

Gess, Christopher

Messung sozialwissenschaftlicher Forschungskompetenz von Studierenden

Theoretischer Hintergrund: Studierende für eigene Forschungsaufgaben zu qualifizieren, ist neben der Vermittlung von Fachwissen ein vorrangiges Ziel des Universitätsstudiums (Wissenschaftsrat, 2006). Um zu beantworten, ob die Universität dieser Aufgabe gerecht wird und wie das Studium aufgebaut sein muss, um diese Fähigkeiten zu vermitteln, bedarf es eines Instrumentariums zur Messung von Forschungskompetenz. Theoretische Modelle von Forschungskompetenz und Tests zur Messung von Forschungskompetenz liegen bislang bis auf eine Ausnahme lediglich für Teilbereiche des Forschungsprozesses (information literacy, z. B. Katz, 2007; statistical thinking, z. B. Pfannkuch und Wild, 2004) oder für naturwissenschaftliche Domänen (scientific reasoning, z. B. Schunn & Anderson, 1999; scientific discovery, z. B. Hammann et al., 2007) vor. Eine Ausnahme bildet der im LeScEd-Projekt entwickelte Test zur bildungswissenschaftlichen Forschungskompetenz, bei dem die Konstrukte der information literacy, statistical literacy und critical thinking in einem Instrument kombiniert werden (Groß-Ophoff et al., im Druck). Die Autoren konzentrieren sich dabei jedoch auf die Fähigkeit, Forschungswissen zu verstehen und anzuwenden. Die Fähigkeit, Forschungsergebnisse methodisch gestützt zu generieren, wird hier ebenso ausgeklammert wie die qualitative Forschungstradition. Die vorliegende Studie will diese Lücke schließen.

Fragestellung: Komplementär zum LeScEd-Projekt soll in der vorliegenden Studie ein Instrument zur Messung von Forschungskompetenz entwickelt werden, das auf den Prozess der Wissensproduktion mittels sozialwissenschaftlicher Forschungsmethoden fokussiert und Items sowohl zu qualitativen und quantitativen Methoden beinhaltet. Die zentrale Fragestellung lautet dabei: Wie lässt sich die Fähigkeit, Forschungsergebnisse methodisch gestützt zu generieren in einem Kompetenztest erfassen? Zudem wird der Frage nachgegangen, ob und wie sich Aufgaben zu qualitativen und quantitativen Forschungsmethoden in einem gemeinsamen Testinstrument integrieren lassen.

Methode: Der Kompetenztest wurde in mehreren Stufen auf Basis von Experteninterviews (n=4, Grounded Theory nach Charmaz, 2006), einer Expertenbefragung (n=8) und eines Item-Panels (Wilson, 2005) entwickelt. Um die Inhaltsvalidität des späteren Tests sicher zu stellen, wurden zwei Expertenratings durchgeführt. Das erste Rating wurde durch Experten aus Forschung und Lehre vorgenommen (n=24). Den Experten wurden je eine Bewertungsfrage zur Repräsentativität des Items für den getesteten Wissensbereich und zur Generalisierbarkeit der Antworten über das konkret verwendete Item hinaus gestellt (Jenßen, Dunekacke & Blömeke, im Druck). Das zweite Rating wurde durch Experten aus der beruflichen Praxis der Wissenskonstruktion vorgenommen (n=54). Sie sollten auf bewerten, ob das getestete Wissen in ihrer beruflichen Praxis benötigt wird. Die Bewerbungsfragen wurden für jedes Item separat auf einer Skala von 1 bis 4 bewertet. Parallel zum Expertenrating wurde ein Pretest (n=202) mit zwei verankerten Testheften durchgeführt und im 1PL- und 2PL-Modell ausgewertet. Die Ergebnisse aus Pretest und Ratings wurden zusammengeführt und Items entsprechend der Mittelwerte aus den Expertenratings (Jenßen et al., im Druck), der Diskriminationsparameter (Wilson, 2005), der Selektionskennwerte (Lienert & Raatz, 1998) und von Wald-Tests zur Subgruppeninvarianz (Strobl, 2012) in Einzelfallentscheidungen selektiert.

Ergebnisse: Die Items wurden von den Experten als weitgehend repräsentativ (MW=3,07, SD=0,41) und praxisrelevant (MW=3,02, SD=0,41) bewertet. Die Generalisierbarkeit der Items wurde als mittelmäßig bewertet (MW=2,87, SD=0,39). Insgesamt 9 von 54 Items mussten aufgrund niedriger Bewertungen ausgeschlossen werden. Bei 17 Items konnte kein signifikant positiver Diskriminationsparameter nachgewiesen werden und 10 Items verletzen die Annahme der Subgruppeninvarianz. Die Gründe für diese Befunde werden im Vortrag anhand von Beispielitems diskutiert. Für den finalen Test lässt sich eine akzeptable Reliabilität erwarten (Cronbach's $\alpha=0,76$, Mc Donald's $\omega=0,80$, jeweils nach Spearman-Brown-Korrektur). Mehr als die Hälfte der befragten Praxisexperten (32 von 54) gaben an, in der beruflichen Praxis sowohl qualitative als auch quantitative Forschungsmethoden anzuwenden, was zeigt, dass der Test beide Methodentraditionen berücksichtigen sollte. Gleichwohl wiesen auffallend viele Items zu qualitativem Methodenwissen schlechte Kennwerte auf. Dies könnte darauf zurückzuführen sein, dass sich kaum Standards für die qualitative Forschung identifizieren lassen, die über alle qualitativen Methodologien hinweg gelten. Dies wird im Vortrag anhand konkreter Items diskutiert.

Freitag
13. März

15:30 Uhr

GA 03/49

Chair(s)

Gorges, Johannes-Franz

Schladitz, Sandra; Groß Ophoff, Jana; Wirtz, Markus

Konstruktvalidierung eines Tests zur Messung bildungswissenschaftlicher Forschungskompetenz

Ein angemessener Umgang mit wissenschaftlicher Evidenz sowie die Fähigkeit, daraus basierend Entscheidungen zu treffen, sind zentrale Ziele einer Hochschulausbildung. Das Projekt LeScEd (Learning the Science of Education) integriert Ansätze aus Bereichen wie Informationswissenschaften, Mathematikdidaktik und Psychologie in ein gemeinsames Strukturmodell bildungswissenschaftlicher Forschungskompetenz (BFK). Für eine verlässliche Messung spezifischer Kompetenzen ist eine differenzierte Konstruktvalidierung insbesondere hinsichtlich konvergenter und diskriminanter Validierung anhand inhaltlich relevanter Referenzkonstrukte erforderlich. In diesem Sinne werden Zusammenhänge der BFK zu selbsteingeschätzter Kompetenz und fluider Intelligenz untersucht. Die Daten wurden im Rahmen einer Large-Scale-Erhebung gesammelt, wobei jeweils kleineren Teilstichproben eine Subskala zur verbalen Intelligenz (n = 461) und sprachfreien Intelligenz (n = 434) vorgelegt wurde. In Strukturgleichungsmodellen zeigen sich schwache positive Zusammenhänge zu den Intelligenzskalen, aber keine Zusammenhänge zur Selbsteinschätzung. Es liegen keine differentiellen Effekte für die Intelligenzarten (verbal vs. sprachfrei) oder die Kompetenzfacetten (Informationskompetenz, statistische Kompetenz, kritisches Denken) vor. Diese Ergebnisse zeigen, dass BFK mit Intelligenz verwandt, aber abgrenzbar ist und dass die Selbsteinschätzung keinen geeigneten Indikator für die tatsächliche Kompetenz darstellt. Entsprechend scheint sich eine Erfassung spezifischer Kompetenzen auch über Intelligenz hinaus zu lohnen; Kompetenzmessungen per Selbsteinschätzungen sollten möglichst durch eine objektive Testung ergänzt werden.

Freitag
13. März

15:30 Uhr

GA 03/49

Chair(s)

Gorges, Johannes-Franz

Neuenschwander, Markus P.; Niederbacher, Edith

Lehrpersonenerwartungen und Leistungsentwicklung in Deutsch und Mathematik von deutschsprachigen und fremdsprachigen Kindern

Studien berichteten über tiefere Leistungen und eine ungünstigere Leistungsentwicklung von fremdsprachigen Kindern in Deutsch und Mathematik im Vergleich zu Kindern mit der Schulsprache als Erstsprache (z.B. Helmke & Weinert, 1997; Müller, 2000). Dieser Effekt wurde auch nach Kontrolle der Schichtzugehörigkeit der Kinder gefunden. Weil der Effekt fachübergreifend auftritt, kann er nicht nur mit Sprachdefiziten des Kindes erklärt werden. Vielmehr dürften Erwartungen von Lehrpersonen für die Leistungsentwicklung von Kindern eine wichtige Rolle spielen (vgl. Pygmalion-Effekt, Rosenthal & Jacobson, 1974). Diese Erwartungen sind schicht- und migrationsspezifisch geprägt (z.B. Jussim & Eccles, 1992; McKown & Weinstein, 2002). Aufgrund gesellschaftlicher Diskurse über Migrantinnen und Migranten entwickeln Lehrpersonen Vorstellungen und Stereotype, die sich im schulischen Alltag zeigen. Kinder mit Migrationshintergrund werden bei gleicher Leistung signifikant schlechter eingeschätzt als ihre einheimischen Mitschülerinnen und Mitschüler (z.B. Jungbluth, 1994; Kronig, Haeblerlin & Eckhart, 2000). Entsprechend richten sie vermutlich tiefere Erwartungen an fremdsprachige Kinder als an Kinder mit der Schulsprache als Familiensprache (=deutschsprachige Kinder). Daraus folgt, dass fremdsprachige Kinder aufgrund tieferer Leistungserwartungen ihrer Lehrpersonen tiefere Leistungen in Deutsch und Mathematik zeigen und dass ihre Leistungen langsamer zunehmen, als diejenigen von deutschsprachigen Kindern.

Im vorliegenden Beitrag wird die Leistungsentwicklung von Schülerinnen und Schülern mit Migrationshintergrund (operationalisiert über die Familiensprache) vom 5. zum 6. Schuljahr untersucht. Ziel ist, zur Aufklärung von Disparitäten in der Leistungsentwicklung von zugewanderten Kindern mit nicht-deutscher Familiensprache beizutragen. Folgenden Forschungsfragen wird nachgegangen: Wie verläuft die Leistungsentwicklung in Deutsch und Mathematik von Schülerinnen und Schülern mit nicht deutscher Familiensprache (=fremdsprachige Kinder) im Vergleich zu Kindern mit deutscher Familiensprache (=deutschsprachige Kinder)? Medieren leistungsbezogene Lehrpersonenerwartungen den Effekt der Fremdsprachigkeit?

Die Schweizer Längsschnittstudie Wirkungen der Selektion WiSel untersucht die Leistungsentwicklung vom 5. zum 7. Schuljahr. Die Fragen wurden mit einer Teilstichprobe von 925 Schülerinnen und Schülern aus 102 zufällig ausgewählten Klassen aus fünf Deutschschweizer Kantonen, die im 5. und 6. Schuljahr an der Studie teilnahmen, analysiert. Zum Einsatz kamen zu beiden Messzeitpunkten standardisierte, reliable Leistungstests in den Promotionsfächern Deutsch und Mathematik, die nach der Item-Response Theorie ausgewertet wurden. In einem standardisierten Fragebogen für Schülerinnen und Schüler wurden die Umgangssprache in der Familie (deutschsprachig vs. nicht-deutschsprachig) und die Berufe der Eltern, die nach dem System von ISCO-08 kodiert wurden, erfasst. Der ISCO-Code wurde in den ISEI-Code transformiert. Jedem Kind wurde der höhere ISEI-Code der beiden Elternteile zugeordnet. Zusätzlich schätzten die Klassenlehrpersonen für alle Kinder ihrer Klasse im 5. Schuljahr einzeln die Leistung ein.

Die Ergebnisse einer Mehrebenenanalyse mit der Klasse als Level 2-Variable zeigten, dass Lehrpersonen bei gleichen Leistungen in Deutsch und Mathematik höhere Leistungserwartungen an deutschsprachige Kinder richteten als an fremdsprachige. Die Leistungserwartungen unterschieden sich zwischen den Klassen signifikant. Weitere stufenweise Mehrebenenanalysen zeigten, dass fremdsprachige Schülerinnen und Schüler tiefere Leistungen in den Fächern Deutsch und Mathematik erbrachten als deutschsprachige Kinder, auch wenn ihre Schichtzugehörigkeit (ISEI) kontrolliert wurde. Die Leistungen unterschieden sich signifikant zwischen den Klassen. Dieser querschnittliche Effekt verschwand nach Einbezug der leistungsbezogenen Lehrpersonenerwartungen (Mediation, getestet mit Bootstrappingverfahren zu multiplen Mediationen nach Preacher und Hayes 2008). Im Längsschnitt erklärten nach Kontrolle der Schicht die leistungsbezogenen Lehrpersonenerwartungen die Leistungsentwicklung in Mathematik vom 5. zum 6. Schuljahr und sie mediieren den Effekt der Fremdsprache auf die Leistungsentwicklung. In Deutsch erklärten die Lehrpersonenerwartungen die Leistungsentwicklung nach Kontrolle der Schicht signifikant, mediieren den Effekt der Fremdsprache jedoch nicht.

Die Bedeutung von Lehrpersonenerwartungen für die differenzielle Leistungsentwicklung in der Schulsprache und in Mathematik von deutschsprachigen und fremdsprachigen Schülerinnen und Schülern wird diskutiert.

Freitag
13. März

15:30 Uhr

GA 03/142

Chair(s)

Neuenschwander, Markus

Walter, Oliver

Oliver Walter: Herkunftsassoziierte Disparitäten im Lesen – revisited: Intergenerationale Kompetenzveränderungen zwischen PISA 2000 und PISA 2009

Theoretischer Hintergrund: Die ersten PISA-Studien haben im Mittel geringere Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund aufgezeigt, insbesondere von Jugendlichen aus den ehem. Anwerbestaaten (Türkei, ehemaliges Jugoslawien und Italien). Zeitweise bestand sogar die Befürchtung, daß entgegen der klassisch-assimilationstheoretischen Annahme eine Kompetenzabnahme von erster zu zweiter Einwanderergeneration zumindest für die türkischstämmige Gruppe vorliegen könnte (Baumert & Schümer, 2001; Ramm et al., 2004). Als ein bildungspolitisches Ziel wurde daher die Verringerung der Disparitäten zu Jugendlichen ohne Migrationshintergrund, insbesondere der beiden Einwanderergenerationen, formuliert. Bereits nach PISA 2006 ließen sich für die großen Herkunftsgruppen drei Trends aufzeigen: (1) eine Zunahme der relativen Anteile der höheren Einwanderergenerationen gegenüber den jüngeren Einwanderergeneration für alle Herkunftsgruppen („Alterungseffekt“); (2) eine Abnahme der Disparitäten zwischen Jugendlichen der ersten Generation und Jugendlichen ohne Migrationshintergrund in allen Kompetenzbereichen, insbesondere in der Lesekompetenz sowie (3) eine Abnahme der Disparitäten zwischen Jugendlichen, deren Eltern aus der ehemaligen Sowjetunion stammen, und Jugendlichen ohne Migrationshintergrund in allen Kompetenzbereichen, insbesondere in der Lesekompetenz (Walter, 2008).

In neueren Untersuchungen wird die tendenzielle Verringerung der Kompetenzunterschiede zwischen Jugendlichen mit und solchen ohne Migrationshintergrund behauptet, die in PISA 2009 die erste und in PISA 2012 auch die zweite Generation betreffen soll (Stanat, Rauch & Segeritz, 2010; Gebhart et al., 2013). Eine empirisch validierte Erklärung für diese Kompetenzveränderungen ist bislang noch nicht hinreichend dargelegt worden (Ehmke, Klieme & Stanat, 2013 gehen z.B. nur der allgemeinen Frage nach, worauf die Gesamtveränderungen in der Lesekompetenz bis PISA 2009 zurückzuführen sein könnten). Es werden jedoch die Wirksamkeit bildungspolitisch initiierten Interventionen z.B. von Sprachtrainings (Stanat et al., 2010) gegenüber einer veränderten Zusammensetzung der Migrantengruppe diskutiert (Walter, 2008, 2014).

Fragestellungen: Daher wird folgenden Fragestellungen nachgegangen: (1) Wie hat sich in Deutschland die Zusammensetzung von bedeutenden Migrantengruppen nach Einwanderergeneration seit PISA 2000 entwickelt? (2) Wie haben sich die Disparitäten in der Lesekompetenz zwischen den Herkunftsgruppen differenziert nach Einwanderergeneration und Jugendlichen ohne Migrationshintergrund seit PISA 2000 verändert? (3) Lassen sich Disparitätsveränderungen durch Veränderungen in der Zusammensetzung der Migrantengruppen erklären?

Methode: Die empirische Untersuchung basiert auf den internationalen Datensätzen von PISA 2000, PISA 2003, PISA 2006 und PISA 2009. Die Daten von PISA 2012 wurden aufgrund der hohen Anzahl von fehlenden Daten (17,6%) ausgeschlossen. Als abhängige Variable wurde die Lesekompetenz verwendet, da die Sprachkompetenz in der Verkehrssprache des Einwanderungslandes eine entscheidende Bedeutung für die Integration von Migranten hat. Als unabhängige Variable wurde der Migrationshintergrund differenziert nach Herkunftsland der Eltern und Einwanderergeneration herangezogen. Um hinreichend große Stichprobenumfänge für die Migrantengruppen zu erhalten, wurden die Herkunftsländer in zwei Großkategorien zusammengefaßt, die sich auch vor dem Hintergrund der deutschen Einwanderungsgeschichte rechtfertigen lassen (ehem. UdSSR/Polen: Approximation der Gruppe „Aussiedler, Spätaussiedler“; ehem. Anwerbestaaten: Türkei, ehem. Jugoslawien, Italien, Griechenland). Die Erfassung der Disparitäten fand über z-standardisierte PISA-Punkte statt, da sich über die PISA-Erhebungen hinweg sowohl Mittelwerte als auch Standardabweichungen der Skalen verändert haben und der Untersuchungsansatz eine normorientierte Betrachtung nahelegt. Die statistischen Analysen schließen lineare Regressionsanalysen mit Interaktionstermen sowie Mittelwertvergleiche ein.

Ergebnisse: Die Befunde bestätigen den Trend abnehmender Anteile der ersten Generation innerhalb der Migrantengruppen, es zeigen sich aber überwiegend keine signifikant verringerten Disparitäten zwischen Jugendlichen verschiedener Herkunftsgruppen und solchen ohne Migrationshintergrund in der Lesekompetenz. Die einzige Ausnahme bildet die Disparität zwischen Jugendlichen, deren Eltern aus der ehem. UdSSR oder Polen stammen, und solchen ohne Migrationshintergrund, die sich um etwa 0,3 Standardabweichungen verringert hat. Innerhalb dieser Gruppe ist nur für die erste Generation eine Abnahme der Disparität festzustellen. Daher läßt sich die Veränderung für die gesamte Migrantengruppe aus der ehem. Sowjetunion und Polen zu 71,4% durch die starke Verringerung des Anteils der kompetenzschwächeren ersten Generation, also die Veränderung der Zusammensetzung der Migrantengruppe, erklären. Damit bedarf es zur Erklärung dieser Disparitätsentwicklungen prima facie nicht der Hypothese der Wirkung bildungspolitischer Interventionen.

Seiz, Johanna; Decristan, Jasmin; Kunter, Mareike

Differenzielle Effekte von Klassenführung – wer profitiert besonders?

Klassenführung stellt eine zentrale Dimension der Unterrichtsqualität dar (Kunter & Voss, 2011; Pianta & Hamre, 2009) und ist für sowohl den Lernerfolg von Schülerinnen und Schülern (z.B. Hattie, 2009) als auch für ihre motivationale Entwicklung (Kunter, Baumert, & Köller, 2007) bedeutsam. Außerdem weisen einige Studien darauf hin, dass gute Klassenführung zu weniger antisozialem Verhalten und Konflikten führt (Freiberg & Lapointe, 2006). In den bisherigen Studien wird angenommen, dass Klassenführung für alle Schülerinnen und Schülern in gleichem Maße wirksam ist. Einige Studien zeigen jedoch für den Bereich Grundschule, dass bestimmte Risikoschüler in besonderem Maße von gutem Unterricht profitieren (z.B. Hamre & Pianta, 2005). Wir widmen uns der Fragestellung, ob Klassenführung eine spezifische Wirksamkeit für Risikoschüler und Risikoschülerinnen in der Sekundarschule aufweist und somit einen Beitrag zur Adaptivität von Unterricht leistet. Dabei analysieren wir die differenzielle Wirksamkeit von Klassenführung für die Risikogruppe der Schülerinnen und Schülern mit Migrationshintergrund. Die geringeren Bildungserfolge dieser Risikogruppe speziell in Deutschland wurden mittlerweile in vielen Studien nachgewiesen (Stanat & Edele, 2011). So zeigen beispielsweise die PISA-Studien der Jahre 2003 und 2012, dass auch unter Kontrolle des sozioökonomischen Status und des Bildungsniveaus der Eltern diese Schülergruppe schlechtere Leistungen in der mathematischen Kompetenz aufweisen als Schülerinnen und Schülern ohne Migrationshintergrund (Gebhardt, Rauch, Mang, Sälzer, & Stanat, 2013).

Die Frage der differenziellen Wirksamkeit von Klassenführung für diese Risikogruppe wird im Hinblick auf die abhängigen Variablen Mathematikleistung, Interesse an Mathematik sowie dem physischen Gewaltverhalten analysiert, um neben der akademischen Leistung noch weitere Bildungsziele in Betracht zu ziehen.

Methode: Die Grundlage unserer Untersuchung bildet ein längsschnittlicher Datensatz mit Daten von 3357 Neuntklässlern und ihren 155 Mathematiklehrkräften aus der COACTIV-Studie (Kunter et al., 2011). Eine Folgerhebung erfolgte am Ende der 10. Klasse. Der Migrationsstatus wurde über die Herkunft definiert und galt, sobald mindestens ein Elternteil nicht in Deutschland geboren wurde. Der Anteil der Schülerinnen und Schülern mit einem Migrationshintergrund betrug demnach 15,5 %. Klassenführung wurde als Klassenvariable über die Lehrer- sowie die Schülereinschätzung der Aspekte Unterrichtsstörungen und Zeitverschwendung operationalisiert (Baumert, Gruehn, Heyn, Köller, & Schnabel, 1997). Zu beiden Messzeitpunkten wurde die Mathematikleistung über einen Leistungstest erfasst. Weiterhin wurde Interesse an Mathematik und das physische Gewaltverhalten per Selbstbericht zweifach erfasst (Ramm et al., 2006).

In latenten Cross-Level-Interaktionen wurde analysiert, ob Klassenführung den Zusammenhang zwischen Migrationsstatus und den abhängigen Variablen moderiert. Dabei wurden die abhängigen Variablen für den ersten Messzeitpunkt kontrolliert, weiterhin erfolgte eine Kontrolle der Schulform und des sozioökonomischen Status. Die Güte der drei Messmodelle war zufriedenstellend (CFI > .95; TLI > .90; RMSEA < .05).

Ergebnisse: Klassenführung sagte als signifikanter Prädiktor sowohl einen Leistungszuwachs, als auch eine Abnahme an aggressivem Verhalten über den Verlauf eines Schuljahrs vorher. Zudem moderierte Klassenführung den Zusammenhang zwischen Migrationsstatus und den abhängigen Variablen. In Klassen mit guter Klassenführung weisen Schülerinnen und Schülern mit Migrationshintergrund am Ende des Schuljahres eine bessere Mathematikleistung auf, als in Klassen mit weniger effektiver Klassenführung. Weiterhin zeigten Schülerinnen und Schülern mit Migrationshintergrund in Klassen mit besserer Klassenführung mehr Interesse an Mathematik und weniger gewalttätiges Verhalten.

Bis zur Tagung im März werden die Analysen mit imputierten Datensätzen wiederholt. Außerdem werden zusätzliche Mediationsanalysen angestrebt, um zu überprüfen, ob Aspekte des Lernverhaltens (z.B. Selbstregulation) als vermittelnde Prozesse dieser Ergebnisse agieren.

Ausblick: Unsere Ergebnisse weisen darauf hin, dass speziell Schülergruppen mit bildungsrelevanten Risikofaktoren von einer effektiven Klassenführung profitieren können. Diese Ergebnisse verdeutlichen, dass alltägliche Unterrichtsprozesse dazu geeignet sind, Bildungsdisparitäten zu entgegnen und als protektive Faktoren zu wirken. Diese Befunde betonen umso mehr die Relevanz von effektiver Klassenführung sowie einer systematischen Integration dieses Unterrichtsmerkmals in Lehreraus- und Fortbildung.

Haag, Nicole; Roppelt, Alexander; Heppt, Birgit

Effekte von Bildungssprache auf Mathematikleistungen: Vertiefende Analysen zur Bedeutung der Klassenstufe und des Sprachhintergrunds

In nationalen und internationalen Schulleistungsstudien erreichen Schülerinnen und Schüler nicht-deutscher Herkunftssprache in der Regel deutlich schwächere Leistungen in Mathematik als Gleichaltrige deutscher Herkunftssprache. Frühere Studien konnten differenzielle Zusammenhänge zwischen nicht-mathematischen sprachlichen Anforderungen von Testaufgaben und Mathematikleistungen von Erst- bzw. Zweitsprachlernenden feststellen (z.B. Haag, Heppt, Stanat, Kuhl, & Pant, 2013; Martiniello, 2009). Unklar bleibt jedoch, inwiefern sich die Effekte der sprachlichen Komplexität über Klassenstufen hinweg unterscheiden. Verwendet man dieselben Testaufgaben in unterschiedlichen Klassenstufen, sollten die Effekte sprachlicher Komplexität in den höheren Klassenstufen wegen der besseren sprachlichen Fähigkeiten der Schülerinnen und Schüler geringer ausfallen (vgl. Shaffel, Belton-Kocher, Glasnapp, & Poggio, 2006). Mit dem vorliegenden Beitrag soll überprüft werden, inwieweit sich die Effekte bildungssprachlicher Merkmale von Mathematikaufgaben auf die Mathematikleistungen von Grundschülerinnen und -schülern zwischen der dritten und der vierten Klasse verringern. Hierbei werden insbesondere die Effekte für Kinder mit nicht-deutscher Herkunftssprache untersucht.

Die Studie basiert auf Daten zu 126 Mathematikitems, die im Rahmen zweier Large-Scale Assessment Studien von insgesamt 28 812 Dritt- und Viertklässlern bearbeitet wurden. Die sprachlichen Merkmale der Items wurden nach dem Konzept der Bildungssprache (z.B. Bailey, 2007) von acht geschulten Kodierern eingeschätzt (eine ausführlichere Beschreibung der Prozedur findet sich in Heppt, Haag, Böhme, & Stanat, 2014). Die Itemantworten der Schülerinnen und Schüler wurden mit Hilfe explanatorischer Itemresponsemodelle gleichzeitig durch Merkmale der Items und der Kinder vorhergesagt. In die Analysen flossen (a) die bildungssprachliche Komplexität der Items, (b) die Familiensprache der Kinder, (c) die Klassenstufe der Kinder sowie (d) die Interaktion der Variablen ein. Zusätzlich wurde in den Analysen für den sozioökonomischen Status der Schüler kontrolliert.

Die Analysen zeigen einen etwas größeren mittleren Leistungsunterschied zwischen Schülern deutscher Herkunftssprache und Schülern nicht-deutscher Herkunftssprache in der vierten Klasse als in der dritten Klasse ($\beta = 0.023$, $p < .05$, OR = 1.023). Über alle Personen und Items hinweg ist die bildungssprachliche Komplexität der Items kein signifikanter Prädiktor der Itemschwierigkeit; allerdings fanden sich signifikante Interaktionen zwischen der bildungssprachlichen Komplexität der Aufgaben und den Variablen Herkunftssprache ($\beta = -0.008$, $p < .05$, OR = 0.992) und Klassenstufe ($\beta = -0.014$, $p < .01$, OR = 0.986). Diese zeigen, dass Kinder in der dritten Klassenstufe – über den mittleren Leistungsunterschied hinaus – etwas schwächere Leistungen bei bildungssprachlich komplexen Items erzielen als Kinder der vierten Klassenstufe. Tatsächlich verschwindet der Zusammenhang zwischen bildungssprachlicher Komplexität und Mathematikleistung für die Viertklässler sogar vollständig. Ebenso erzielen Kinder nicht-deutscher Herkunftssprache etwas schwächere Leistungen als Kinder deutscher Herkunftssprache. Die Dreifachinteraktion zwischen bildungssprachlicher Komplexität, Klassenstufe und Herkunftssprache verfehlt jedoch die statistische Signifikanz, sodass sich keine klassenstufenspezifischen Effekte bildungssprachlicher Komplexität in Abhängigkeit von der Herkunftssprache der Schülerinnen und Schüler identifizieren lassen.

Zusammenfassend deuten die Ergebnisse auf einen geringen Einfluss bildungssprachlicher Merkmale hin, der in der vierten Klasse sowohl für Schüler deutscher Herkunftssprache als auch für Schüler nicht-deutscher Herkunftssprache deutlich geringer ausgeprägt ist als in der dritten Klasse. In beiden Klassenstufen sind Schüler nicht-deutscher Herkunftssprache etwas stärker durch bildungssprachlich komplexe Aufgaben benachteiligt als Schüler deutscher Herkunftssprache. Die Ergebnisse legen den Schluss nahe, dass die Benachteiligung von Schülerinnen und Schülern nicht-deutscher Herkunftssprache nicht primär auf die sprachlichen Anforderungen der Testaufgaben zurückzuführen ist, sondern dass eine Benachteiligung möglicherweise bereits in der Erwerbssituation entsteht. In diesem Zusammenhang könnte der höhere Leistungsrückstand von Schülern nicht-deutscher Herkunftssprache in der vierten Klasse darauf hindeuten, dass Schüler nicht-deutscher Herkunftssprache aufgrund ihrer geringeren Sprachkenntnisse nicht in gleichem Maße vom Unterricht profitieren wie Schüler deutscher Herkunftssprache.

Freitag
13. März

15:30 Uhr

GA 03/142

Chair(s)

Neuenschwander, Markus

Burrichter, Rita; Herzig, Bardo; Schaper, Niclas; von Stosch, Klaus; Woppowa, Jan; Isik, Tuba; Kamcili-Yildiz, Naciye; Nelles, Wolfgang

Entwicklung eines Kompetenzmodells und Testinstruments für die Domäne "Islamischer Religionsunterricht"

Motivation: Bislang fehlen der jungen Fachdisziplin Islamische Religionspädagogik wissenschaftliche Erkenntnisse über die Ausbildung von Lehrenden für den Islamischen Religionsunterricht. Die Qualifikation der Lehrenden ist entscheidend, um zu vermeiden, dass unzureichend ausgebildete Lehrerinnen und Lehrer, deren theologische und religionspädagogische Ausbildung nicht die erforderliche Qualität ausweist, dieses Schulfach unterrichten. Daher ist es besonders wichtig, den Prozess der Einführung und Etablierung des Schulfachs Islamischer Religionsunterricht wissenschaftlich zu begleiten, um den Ausbildungsprozess der Lehrenden in überschaubarer Zeit optimieren zu können.

Theoretischer Hintergrund: Dem Projekt IRU liegt das Kompetenzverständnis von Weinert (Weinert, 2001) zugrunde. So werden neben den kognitiven Leistungsvoraussetzungen ebenso motivationale, volitionale und soziale Fähigkeiten, die zur Bewältigung komplexer Anforderungssituationen benötigt werden, bestimmt. Das Kompetenzmodell für Lehrerinnen und Lehrer des Schulfachs Islamischer Religionsunterricht hat sowohl die Funktion, das Gefüge von relevanten Kompetenzen zu beschreiben als auch zusätzlich eine fundierte Vorstellung darüber zu geben, welche Abstufungen die besagten Kompetenzen annehmen können und welche Niveaustufen sich bei den Probanden feststellen lassen (Anderson, L.H. & Krathwohl, 2001).

Fragestellung: Das Projekt IRU verfolgt vor allem zwei Ziele: (1) Die Entwicklung eines wissenschaftlich fundierten Kompetenzmodells für Lehramtsstudierende und Lehrende eines bekenntnisgebundenen, islamischen Religionsunterrichts. (2) Die Entwicklung und Überprüfung eines Testinstruments zur Messung der Lehrkompetenz hinsichtlich der Zieldomäne (IRU). Vor dem Hintergrund der Einführung und Etablierung des Schulfachs Islamischer Religionsunterricht liefern das Kompetenzmodell und das Testinstrument einen wichtigen Beitrag zu den Standards in der Lehrerbildung. Die Entwicklung des Kompetenzmodells fungiert zur Bestimmung und Strukturierung der benötigten Leistungsanforderungen innerhalb der Zieldomäne; desweiteren liefert es eine Grundlage zur Itementwicklung eines Instruments zur Messung dieser Kompetenzen.

Methode: Expertinnen und Experten aus unterschiedlichen Bezugsdisziplinen entwickelten in gemeinsamer Arbeit das Kompetenzmodell. Hierbei wurden neben der Religionspädagogik und der islamischen Theologie auch die Wissenschaftsdomänen der Pädagogischen Psychologie, der Katholischen und Evangelischen Theologie berücksichtigt. Ferner flossen Erfahrungen im Hinblick auf das methodische Vorgehen bei der Kompetenzmodellierung von Domänen anderer Schulfächer (Evangelischer Religionsunterricht, Informatik) ein. Das resultierende Modell wurde ausformuliert: Das heißt, es wurden Kompetenzprofile formuliert, die mit Hilfe geeigneter Operatoren beschreiben, welche Kompetenzen adressiert werden. Auf Grundlage dieser präzisen Kompetenzprofile wurde in einem darauffolgenden Schritt ein Testinstrument zur Messung der Kompetenzen der Lehramtsstudierenden und Lehrenden des Islamischen Religionsunterrichts zu entwickelt. Neben einer umfangreichen Befragung zur Biografie, inklusive der persönlichen religiösen Entwicklung, misst dieses Instrument motivationale Bereitschaften, soziale Fähigkeiten, Beliefs und Selbstwirksamkeitserwartungen innerhalb der Domäne des Unterrichtens Islamischer Religion an der Schule. Der Hauptteil des Instruments fokussiert die Messung der kognitiven Dimensionen: fachwissenschaftliche und fachdidaktische Kompetenzen. Es werden unterschiedliche Aufgabenformate verwendet. Vor der eigentlichen, derzeit laufenden Vorerprobung des Messinstruments fand bereits eine erste Pilotuntersuchung an 10 Probanden statt, um die generelle Geeignetheit des Instruments zu prüfen. Neben den eigentlichen, erhobenen Daten wurden vor allem die Rückmeldungen der Probanden intensiv genutzt, um die Aufgabenitems zu verbessern (z.B. Aufgabeninstruktionen).

Ergebnisse: Das Modell untergliedert die relevanten Kompetenzen in zwei kognitive Dimensionen (fachwissenschaftliche und fachdidaktische Kompetenzen) und eine nicht-kognitive Dimension (Motivation, Volition, soziale Fähigkeiten, Beliefs). Die fachwissenschaftliche Dimension beinhaltet u.a. Kompetenzkomponenten zu den Bereichen (1) Genese und Exegese des Koran, (2) Hadithwissenschaften, (3) Islamisches Recht usw. Die verschiedenen Kompetenzniveaus der Probanden werden über Anforderungskategorien in vier lern-taxonomischen Stufen erfasst (Anderson & Krathwohl, 2001). Die fachdidaktische Dimension beinhaltet u.a. folgende Kompetenzkomponenten: (1) Kennzeichen, Ansprüche, Prozesse islamischer Bildungsprozesse, (2) wirkungsvolle und aktivierende Unterrichtsmethoden und -ansätze des IRUs, (3) Berücksichtigung kulturspezifischer Aspekte im Islamischen Religionsunterricht. Auch diese Dimension wird zusätzlich in den oben erwähnten lern-taxonomischen Stufen erfasst. Derzeit wird das Messinstrument an ca. 30 Probanden erprobt, und die ersten zurückgelaufenen Fragebögen werden statistisch erfasst. Die Datenanalyse zielt darauf ab, die Aufgabenitems weiterzuentwickeln und der Frage nachzugehen, inwieweit sich zukünftig Empfehlungen für die Curricula der Aus- und Weiterbildungskonzepte ableiten lassen.

Paul, Johanna; Limpinsel, Carlotta; Bromme, Rainer; Stadler, Marc

“Ich könnte es, aber keiner sagt mir, dass ich es soll“

Schülerinnen und Schüler stehen heute oft vor der Aufgabe, sich über wissenschaftliche Themen im Internet zu informieren, beispielsweise um Präsentationen oder Aufsätze vorzubereiten. Dabei ist es nicht ungewöhnlich, dass Schülerinnen und Schüler auf widersprüchliche Wissensbehauptungen stoßen. Ein wichtiges Verstehensziel besteht dann unter anderem darin, die konkurrierende Wissensbehauptungen im Hinblick auf ihre Gültigkeit zu bewerten.

Bromme, Kienhues und Porsch (2010) schlagen hierfür zwei Strategien vor: Bei Plausibilitätsurteilen bewerten Schülerinnen und Schüler die Gültigkeit der Wissensbehauptung basierend auf ihrem eigenen Themenverständnis, während bei Vertrauensurteilen über die Bewertung von Quellenmerkmalen wie Expertise und Benevolenz auf die Informationsgültigkeit rückgeschlossen wird. Insbesondere wenn Schülerinnen und Schüler nur wenig Vorwissen zum Recherchethema verfügen, mögen Vertrauensurteile ein wichtiger Zugang zur Validitätsbeurteilung sein, da das für Plausibilitätsurteile notwendige Wissen nicht oder nur unzureichend vorliegt.

Dennoch zeigen empirische Studien, dass Lernerinnen und Lerner verschiedener Altersstufen selten spontane Quellenbewertungen vornehmen (Gerjets, Kammerer & Werner, 2011; Stadler, 2006) und selbst dann nicht auf Quellen achten, wenn sie dazu aufgefordert werden (Britt & Aglinskas, 2002). Überraschenderweise zeigen Schülerinnen und Schüler der neunten Klasse jedoch Kompetenzen zur Quellenidentifikation, -bewertung und -nutzung in einfachen Leseaufgaben (Stadler, Babel, Rouet & Bromme, 2014). Zum Beispiel gelingt es ihnen in kurzen Texten, Quellenparameter zu identifizieren und die Expertise und Intention von Autoren angemessen einzuschätzen. Um diese vermeintliche Diskrepanz aufzulösen, führten wir eine Interviewstudie durch, in der wir die Perspektiven der betroffenen Schülerinnen und Schüler auf ihre eigene Quellennutzung beim Lesen von Internetdokumenten näher beleuchteten. Dabei fokussierten wir auf schülerseitig generierte Gründe für die geringe Beachtung von Quellenmerkmalen sowie den subjektiven Nutzen von Quellevaluation für den Lernprozess.

Teilnehmer der Studie waren 20 Schülerinnen und Schüler der neunten Klasse eines Gymnasiums (65% weiblich, M = 14.85 Jahre). Zunächst führten die Teilnehmerinnen und Teilnehmer eine kurze Internetrechercheaufgabe zu einem kontrovers diskutierten medizinischen Thema durch (Aspartam). Im Folgenden wurden die Teilnehmerinnen und Teilnehmer hinsichtlich des potentiellen Nutzens von Quelleninformationen befragt. Außerdem wurden ihnen Fragen zu ihrem Umgang mit Quellen im persönlichen sowie im schulischen Alltag gestellt und sie wurden gebeten, Gründe anzuführen, warum sie Quellen oft wenig Beachtung schenken.

Die inhaltsanalytische Auswertung der Interviews erfolgte über zwei unabhängige Beurteiler, die eine gute Übereinstimmung in der Kategorisierung der Schülerantworten aufwiesen (Cohen's Kappa = .77 - 1). Die Ergebnisse zeigen, dass Neuntklässlerinnen und Neuntklässler differenzierte Konzepte über die Nützlichkeit von Quellenmerkmalen aufweisen. Sie geben an, dass Quelleninformationen nützlich sein können, um die Schritte einer Internetrecherche zu planen (55% der Schüler gaben mindestens eine Antwort, die dieser Kategorie zugeordnet wurde), und die Angabe von Quellen dazu dient, bestimmte (wissenschaftliche) Regeln einzuhalten (55%). Besonders reichhaltig waren die Antworten, die sich auf die Gültigkeitsbewertung des Inhaltes bezogen. So nannten viele Schüler beispielsweise die Möglichkeit, Merkmale wie Benevolenz (90%) und Expertise (95%) aus Quelleinformationen abzuleiten.

Antworten auf die Frage, warum Quellen häufig nicht beachtet werden, konnten drei Kategorien zugeordnet werden: individuellen, kontextuellen und materialbezogenen Begründungen. Auf der individuellen Ebene bezogen sich die Antworten auf die Bevorzugung des Inhaltes gegenüber den Quellen (50%) und persönlichen Motivationsmangel (45%). Kontextuelle Begründungen bezogen sich beispielsweise auf den Mangel an externer Aufforderung (90%), die Art der Aufgabe (80%) und mangelnde Verstärkung (40%). Materialbezogene Antworten drehten sich vorrangig um die Sichtbarkeit von Quellenangaben (30%) und die Widersprüchlichkeit der Inhalte (25%).

Die Antworten der Schüler zeigen differenzierte Konzepte über die Nützlichkeit von Quelleninformationen und legen nahe, dass Schüler in der Lage sind, diese zu Gültigkeitsbewertungen heranzuziehen. Jedoch scheinen die unterrichtlichen Rahmenbedingungen das Beachten von Quellen nicht zu fördern. Unsere Ergebnisse machen den Bedarf an Interventionen sichtbar, die Schüler darin unterstützen, ihre bereits vorhandenen Bewertungskompetenzen in realen Lernaufgaben einzusetzen, um für das Lernen im Internet gewappnet zu sein.

Stadtler, Marc; Paul, Johanna; Globoschütz, Silke; Bromme, Rainer

Evaluation eines Trainings zur Förderung der Quellenbewertungskompetenzen in der gymnasialen Sekundarstufe

Für viele Lernerinnen und Lerner ist das Internet zu einer wichtigen Quelle für wissenschaftsbasierte Informationen geworden. Diese nutzen sie vielfach, um Hausaufgaben zu erledigen oder Präsentationen für die Schule vorzubereiten (Medienpädagogischer Forschungsverbund Südwest, 2012). Das Lernen im Internet ist jedoch keine triviale Aufgabe. Besonders wenn Dokumente sich inhaltlich widersprechen, stehen Lernerinnen und Lerner vor der Herausforderung, Konflikte zu entdecken, ihren Verständnisprozess zu regulieren und Konflikte aufzulösen, indem sie die Gültigkeit konkurrierender Wissensbehauptungen bewerten (Stadtler & Bromme, 2014).

Bromme, Kienhues und Porsch unterscheiden zwischen zwei Strategien der Konfliktauflösung, die sie als Plausibilitätsurteile und Vertrauensurteile bezeichnen. Für Plausibilitätsurteile beurteilen Lernerinnen und Lerner die Gültigkeit einer Wissensbehauptung unmittelbar auf Basis ihres eigenen Vorwissens. Bei Vertrauensurteilen beurteilen Lernerinnen und Lerner anstelle dessen die Glaubwürdigkeit der Quelle und leiten hieraus Annahmen über die Gültigkeit der Wissensbehauptung ab. Insbesondere wenn Lernerinnen und Lerner nur über wenig Hintergrundwissen zum Recherchethema verfügen, erscheinen Vertrauensurteile ein geeigneter Zugang zur Gültigkeitsbeurteilung.

Empirische Studien legen jedoch nahe, dass Lernerinnen und Lerner verschiedener Altersstufen selten spontan Quelleninformationen beachten und zur Gültigkeitsbeurteilung heranziehen (Gerjets, Kammerer & Werner, 2011; Stadtler, 2006). Dies gilt selbst dann, wenn sie direkt dazu aufgefordert werden (Britt & Aglinskas, 2002; Stadtler & Bromme, 2007). Diese Beobachtungen legen nahe, dass es umfassenderer Interventionen bedarf, um die Aufmerksamkeit auf und Nutzung von Quelleninformationen im Zuge der Konfliktauflösung nachhaltig zu fördern. Gegenstand dieser Studie ist deshalb die empirische Evaluation einer Intervention zur Förderung der Quellenbewertungskompetenzen im Schulunterricht der gymnasialen Sekundarstufe.

An der Studie nahmen 113 Neuntklässlerinnen und Neuntklässler (82 weiblich, $M = 14.74$ Jahre) aus drei Gymnasien teil. Die Schüler wurden zufällig der Interventions- oder Kontrollgruppe zugeteilt. Beide Gruppen führten eine Rechercheaufgabe mit sechs Internetdokumenten zu einem kontrovers diskutierten medizinischen Thema durch. Alle Schülerinnen und Schüler lasen die Internetdokumente um zu entscheiden, welche der widersprechenden Wissensbehauptungen sie für „wahr“ halten (epistemisches Leseziel). Vor der Lese-phase erhielt die Interventionsgruppe zusätzlich einen kurzen Instruktionstext mit dem Ziel, ihre epistemische Vigilanz zu fördern (Sperber et al., 2010). In dem Instruktionstext erhielten die Schülerinnen und Schüler eine kurze Einführung in das Phänomen der ungleichen Verteilung von Wissen in der Gesellschaft („kognitive Arbeitsteilung“) (Keil, Stein, Webb, Billings & Rozenblit, 2008). Es wurde erklärt, wie die kognitive Arbeitsteilung uns von der Expertise und der Intention der Informanten abhängig macht und wie wichtig vor diesem Hintergrund eine sorgfältige Evaluation dieser Quellenmerkmale in offenen Medien wie dem Internet ist. Während des Lesens wurde das Navigationsverhalten der Probanden aufgezeichnet. Im Anschluss an die Lese-phase erfassten wir mit einem Multiple-Choice Test die Erinnerungsleistung für Quelleninformationen. Zudem verfassten die Probanden eine begründete Stellungnahme zur Kontroverse, die wir inhaltsanalytisch in Bezug auf die Verwendung von Quellenreferenzierungen auswerteten.

Die Ergebnisse zeigen, dass die Schülerinnen und Schüler in der Interventionsbedingung häufiger Impressumsseiten besuchten ($p < .001$, $\eta^2 = .23$) und die Inhalte nach dem Lesen häufiger korrekt ihren Quellen zuordneten ($p = .042$, $\eta^2 = .04$). Zudem verwiesen die Probanden der Interventionsgruppe in ihren Stellungnahmen häufiger auf Quelleinformationen als die Probanden der Kontrollgruppe ($p < .001$, $\eta^2 = 0.12$). Schließlich erwähnten die Schülerinnen und Schüler der Interventionsgruppe häufiger die Argumente einer besonders reliablen Quelle als die Schülerinnen und Schüler der Kontrollgruppe ($\chi^2 = 5.09$, $p = .02$).

Unsere Ergebnisse weisen darauf hin, dass eine kurze schriftliche Einführung in das Prinzip der kognitiven Arbeitsteilung und seiner Relevanz bei der Internetrecherche Schülerinnen und Schülern darin unterstützt, Quellen zur Gültigkeitsbewertung von Wissensbehauptungen heranzuziehen. Effekte der Intervention wurden auf der Ebene des Navigationsverhaltens, der kognitiven Repräsentation von Quelleninformationen und ihrer Nutzung im Zuge der Konfliktauflösung beobachtet. Zukünftige Studien sollten sich der Frage widmen, wie nachhaltig die gefundenen Effekte im Sinne des Transfers der gezeigten Verhaltensweisen auf neue Lernsituationen sind.

Balzer, Lars

Evaluation des Jugendprojektes LIFT

Hintergrund: In der Schweiz besteht die politische Zielvorgabe, dass 95% aller Jugendlichen einen qualifizierenden Abschluss auf Sekundarstufe II erwerben sollen (EDK, 2006). Da rund 2/3 aller Jugendlichen eines Jahrgangs eine Berufsbildung in Angriff nehmen (SBFI, 2014), aber die Zahl derjenigen, die trotz allem keine Lehrstelle finden mit 9000 Jugendlichen (6% einer Kohorte; BBT, 2009) zu hoch ist, sind Anstrengungen an dieser Stelle angebracht, um die Zielvorgabe zu erreichen.

Es wurden unterschiedliche Projekte lanciert, um dieser Herausforderung zu begegnen. Der vorliegende Beitrag beschäftigt sich mit der Evaluation eines dieser Projekte: Der Evaluation des Jugendprojektes LIFT ("Leistungsfähig durch individuelle Förderung und praktische Tätigkeit").

Evaluationsgegenstand: Kernelemente des Jugendprojektes LIFT sind: (1) Frühzeitig individuell erfassen: LIFT konzentriert sich auf den Einbezug von Jugendlichen, welche aufgrund ihrer schulischen und sozialen Situation überdurchschnittlich Mühe haben könnten, nach der obligatorischen Schulzeit eine valable Anschlusslösung in der Berufswelt zu finden. Hierzu werden die Jugendlichen bereits ab der 7. Schulklasse erfasst, um sie daraufhin im Projekt für die Arbeitswelt zu sensibilisieren, zu motivieren und zu qualifizieren. (2) Wöchentlich praktische Einsätze an Wochenarbeitsplätzen (WAP): In kleinen und middle-ren Unternehmen der Region können Jugendliche erste Erfahrungen in der Arbeitswelt sammeln, sich regelmässig bewähren und dazu noch ein Taschengeld verdienen. In einem WAP arbeiten die Jugendlichen 3 bis 5 Stunden pro Woche (max. 3 Stunden/Tag). (3) Gezielt fördern und fordern: In Modulkursen werden soziale, methodische und personale Kompetenzen der jungen Leute gestärkt. (4) Professionelle Vorbereitung und Begleitung: In enger Kooperation mit Schule und Betrieben werden die Jugendlichen in die Arbeitswelt eingeführt. Ergänzende Trainingsmodule fördern ihre Kompetenzen. Im Gruppen- und Einzelcoaching werden laufend anstehende Herausforderungen angegangen.

Evaluationszweck: Das Projekt LIFT wurde mit unterschiedlichen Zugängen evaluiert, um einerseits Erfahrungen und Verbesserungsvorschläge zum Projekt einzuholen (formativer Aspekt) und um andererseits die Wirksamkeit des Projektes zu überprüfen (summativer Aspekt) (vgl. Balzer, 2010, 2012, 2013a, 2013b).

Evaluationsvorgehen:

1. Befragung von Teilnehmenden und Stakeholdern zur Zufriedenheit mit dem Projekt, zu Verbesserungsvorschlägen und zu ihrer Wahrnehmung von Projektwirkungen, mittels 32 halbstrukturierten Interviews sowie diversen Fragebogen im Verlauf einer dreijährigen Projektdauer pro Kohorte
2. Definition und Bewertung von zentralen Projekterfolgskriterien durch 111 verantwortliche LIFT-Organisatoren in 44 Standorten in allen drei Sprachregionen der Schweiz
3. Überprüfung und Bewertung der Anschlusslösungen, die LIFT-Jugendliche nach Abschluss des Projektes in der Arbeitswelt gefunden haben
4. Vergleich der Anschlusslösungen mit vor Projektbeginn definierten Projektzielen
5. Überprüfung der Nachhaltigkeit der Anschlusslösungen durch telefonische Follow up-Interviews wenige Monate und 3,5 Jahre nach Projektabschluss
6. Vergleich der mit LIFT erzielten Anschlusslösungen mit offiziellen kantonalen Statistiken zu Anschlusslösungen gesamter Schülerkohorten
7. Quasiexperimentelles Längsschnitt-Vergleichsgruppendesign. In diesem Design wurden 78 LIFT-Jugendliche der ersten LIFT-Kohorte und 38 vergleichbare Jugendliche ausserhalb von LIFT vor Projektstart, nach einem Jahr, nach 2 Jahren und nach Projektabschluss mit Fragebogen zur u.a. Soziodemographie, persönlichen Situation, Schulleistung, Sprachfähigkeiten, überfachlichen Kompetenzen und Motivation befragt

Zusammenfassung ausgewählter Ergebnisse: Die Daten zeigen, dass LIFT insgesamt betrachtet ein recht erfolgreiches Projekt ist.

- Projektteilnehmende, Verantwortliche und Vertreter der Arbeitswelt stellen dem Projekt ein sehr gutes Zeugnis aus und sind mit Projektverlauf und -ergebnissen zufrieden.
- Zentrale Erfolgskriterien von LIFT (z.B. Kooperation von Keyplayern innerhalb und ausserhalb von LIFT, gute Betreuung der LIFT-Jugendlichen, Steigerung des Selbstbewusstseins der Jugendlichen) werden positiv beurteilt.
- LIFT-Jugendliche zeigen eine positivere Entwicklung einzelner Dimensionen überfachlicher Kompetenzen als Jugendliche ausserhalb von LIFT
- Die vordefinierten Erfolgskriterien von erreichten Anschlusslösungen werden durch LIFT erreicht
- Anschlusslösungen von LIFT-Jugendlichen sind nie schlechter, aber für viele Standorte besser als diejenigen von Jugendlichen ausserhalb von LIFT. Diese Anschlusslösungen sind über die Zeit recht stabil.

Freitag
13. März

15:30 Uhr

GA 03/149

Chair(s)

Paul, Johanna

Müller, Christina

Schafft Homophilie zwischen Arbeitnehmern und Vorgesetzten Einkommensunterschiede?

Das Homophilie-Prinzip bezeichnet die Neigung von Personen sich mit gleichen Anderen zu umgeben (vgl. McPherson et al. 2001). In der Forschung konnte durch vielerlei Studien nachgewiesen werden, dass ähnliche Personen zum Beispiel Freunde werden, heiraten oder sich zu Gruppen zusammenschließen (vgl. McPherson et al. 2001; Ibarra 1992). Homophilie auf dem Arbeitsmarkt ist jedoch weitgehend unerforscht. Vor allem gibt es kaum Ergebnisse über den Einfluss von Homophilie bei Vorgesetzten und Arbeitnehmern auf das Einkommen. Dabei kann eine homophile Beziehung von Arbeitnehmern und Vorgesetzten vorteilhaft sein. Die Gleichheit von Vorgesetzten und Arbeitnehmern verringert deren soziale Distanz, wodurch der Informationsaustausch zwischen ihnen verbessert wird und die Vertrauensbeziehungen gefördert werden (vgl. Ibarra 1992). Auf dem Arbeitsmarkt können sich die Vorteile im sozialen Umgang mit einem gleichen Anderen durch eine gemeinsame Kommunikationsbasis in Produktivität umwandeln. Arbeitsaufträge werden besser verstanden und Lösungswege sind für den Vorgesetzten leichter nachvollziehbar (vgl. Rivera 2012). Dadurch stellt sich die Frage, ob Homophilie zwischen Vorgesetzten und Arbeitnehmern zu einer höheren Einkommensentwicklung führt als für nicht-homophile Arbeitnehmer. Als Merkmal für Homophilie wird das Geschlecht herangezogen.

Obwohl es in der Bevölkerung einen gleich hohen Anteil von Frauen und Männern gibt, ist der Arbeitsmarkt in frauendominierte, männerdominierte und gleichgeschlechtliche Segmente unterteilt (vgl. Leuze/Strauß 2009). Nach Kanter (1977) wirkt sich die ungleiche Geschlechterverteilung auf die Einkommensentwicklung durch das Homophilie-Prinzip unterschiedlich aus. Bezieht man Überlegungen aus der Social Identity Theorie und aus Stuserwartungstheorien mit ein, kann man klare Aussagen über die Wirkrichtung der Homophilie je nach Geschlecht und Gruppenstruktur machen. Minderheitsgruppen werden von Kanter als Token bezeichnet. Wenn Mitglieder einer niedrigeren Statusgruppe Token sind, dann versuchen diese sich von der eigenen Gruppe zu distanzieren und sich der Mehrheitsgruppe anzuschließen. Dadurch wird die Solidarität der Minderheitengruppe aufgeweicht. Aufgrund unterschiedlicher Statuszuweisungen von Männern und Frauen stellen sich die Auswirkungen der Segregation für die beiden Geschlechter unterschiedlich dar. Männern wird ein höherer Status zugeschrieben, weshalb sie keine negativen Folgen als Token befürchten müssen. Die Gruppensolidarität bei Männern in der Minderheit bleibt durch den höheren Status erhalten (vgl. Roth 2004).

Aufbauend auf diesen theoretischen Überlegungen wird angenommen, dass Frauen in der Minderheit mit einem weiblichen Vorgesetzten im Vergleich zu solchen mit einem männlichen Vorgesetzten keine höheren Einkommenszuwächse im Zeitverlauf haben. Sind Frauen jedoch in der Mehrheit, dann führt ein homophiler Vorgesetzter zu einem höheren Einkommen. Bei den Männern wird davon ausgegangen, dass es unabhängig vom Anteil männlicher Kollegen innerhalb ihrer Stelle grundsätzlich höhere Einkommen für Männer mit einem männlichen Vorgesetzten gibt, als für Männer mit einem weiblichen.

Die Analysen werden anhand des Bayerischen Absolventenpanels durchgeführt. Zunächst wird ein multivariates Regressionsmodell herangezogen, um die Ergebnisse mit anderen Querschnittstudien vergleichbar zu machen. Zur Berechnung des innerbetrieblichen Einkommensanstiegs wird ein Growth-Curve Modell verwendet. Hierbei kann man Gruppenunterschiede von homophilen und nicht-homophilen Vorgesetzten-Arbeitnehmer-Paaren im Zeitverlauf darstellen. Der Fokus liegt dabei auf der Einkommensentwicklung innerhalb der ersten Stelle und in den ersten Berufsjahren.

Erste vorläufige Ergebnisse zeigen, dass ohne Berücksichtigung der Geschlechterverteilung Männer mit einem männlichen Vorgesetzten beim Berufseinstieg keine signifikanten Gehaltsunterschiede im Vergleich zu Männern mit einer weiblichen Vorgesetzten haben. Frauen hingegen verdienen weniger, wenn sie einen weiblichen Vorgesetzten haben.

Ebenso zeigt sich ein negativer Effekt für Frauen, wenn der Frauenanteil mindestens so hoch ist wie der der Männer. Männer und Token-Frauen haben keine Einkommensunterschiede, wenn der Vorgesetzte homophil ist. Die Ergebnisse spiegeln jedoch die theoretischen Überlegungen nicht wider. Nur für Frauen zeigen sich signifikante Effekte, allerdings in der gegenteiligen Wirkungsrichtung. Dies lässt vermuten, dass weibliche Vorgesetzte bei größerer weiblicher Konkurrenz am Arbeitsplatz andere Frauen seltener unterstützen. Männer werden hingegen von weiblichen und männlichen Vorgesetzten gleich behandelt, genauso wie Frauen, wenn sie in der Minderheit sind.

Smidt, Wilfried; Roux, Susanna

Prädiktoren für den objektiven und subjektiven Berufserfolg von frühpädagogischen Fachkräften

Theoretischer Hintergrund: Objektiver und subjektiver Berufserfolg sind aus gesellschaftlicher und ökonomischer Sicht bedeutsam und gelten als Merkmal einer gelungenen Lebensgestaltung (Abele, 2002). Objektiver Berufserfolg umfasst direkt beobachtbare und gut operationalisierbare Merkmale (z.B. Einkommen, Einnahme von Leitungspositionen) (Abele et al., 2011; Fuchs-Rechlin, 2007). Subjektiver Berufserfolg bezieht sich auf wahrgenommene Erfolgskriterien, die von der betroffenen Person selbst definiert werden (z.B. Arbeitszufriedenheit, Belastungserleben) (Fuchs-Rechlin, 2007; Judge & Kammeyer-Mueller, 2007). Welche Merkmale den Berufserfolg beeinflussen, ist Gegenstand zahlreicher Untersuchungen, die sich auf heterogene Berufsgruppen beziehen (Gelissen & de Graaf, 2006; Seibert & Kraimer, 2001). In einer Metaanalyse (Ng et al., 2005) werden Prädiktoren für Berufserfolg zu den Merkmalsblöcken Soziodemographie, Humankapital, stabile individuelle Unterschiede und organisationale Unterstützung zusammengefasst. Berufserfolg wird auch im Kontext der Akademisierungsdiskussion in der Frühpädagogik diskutiert (Thole, 2010). Frühpädagogische Fachkräfte werden seit 2004 nicht mehr nur an Fachschulen, sondern zunehmend an Hochschulen ausgebildet (Pasternack & Schulze, 2010). Untersuchungen zu der Frage, welche Merkmale den Berufserfolg von frühpädagogischen Fachkräften beeinflussen, liegen bisher allerdings kaum vor. Der Beitrag greift dieses Forschungsdesiderat auf, indem die Prädiktion des Berufserfolgs fachschul- und hochschulausgebildeter frühpädagogischer Fachkräfte durch theoretisch und empirisch bedeutsame Einflussmerkmale untersucht wird.

Fragestellung: Bisherige Untersuchungen belegen, dass individuelle und strukturelle Merkmale prädiktiv für Berufserfolg sind. In Bezug auf frühpädagogische Fachkräfte liegen bisher kaum Ergebnisse vor. Daher werden folgende Forschungsfragen gestellt: Welche Merkmale beeinflussen den objektiven Berufserfolg von frühpädagogischen Fachkräften? Welche Merkmale beeinflussen den subjektiven Berufserfolg von frühpädagogischen Fachkräften?

Methode: Die Analysen basieren auf zwei Befragungszeitpunkte der bundesweiten Längsschnittstudie „Übergang von fachschul- und hochschulausgebildeten Fachkräften in den Arbeitsmarkt“. Die Absolventenbefragung wurde am Ende der Ausbildung/des Studiums durchgeführt und fand in zwei Kohorten von Juni bis Oktober 2012 sowie Januar bis April 2013 statt. Die Berufsbefragung wurde im Herbst 2013 durchgeführt, 15 Monate (erste Kohorte) bzw. drei Monate (zweite Kohorte) nach Abschluss der Ausbildung/des Studiums.

Objektiver Berufserfolg wurde durch die Merkmale wöchentliche Arbeitszeit, unbefristete Stelle und Leitungsposition erhoben. Subjektiver Berufserfolg wurde anhand einer Skala zur psychischen Belastung (Siegerist et al., 2004) und vier Skalen zur Arbeitszufriedenheit (berufliche Gestaltungsmöglichkeiten, Arbeitsklima, berufliche Sicherheit, Beziehung zu Vorgesetzten; Krüger et al., 2003) erhoben. Als Prädiktorvariablen wurden berücksichtigt: Soziodemographische Merkmale (Geschlecht, Migrationshintergrund, Alter, Familienstand, Elternschaft), Humankapitalmerkmale (höchster Schulabschluss, Fachschul- oder Hochschulabschluss, bereits abgeschlossene pädagogische Ausbildung vor der Fachschul-/Hochschulausbildung, Beschäftigungsbeginn vor bzw. während der Fachschul-/Hochschulausbildung, Kohortenzugehörigkeit), stabile individuelle Unterschiede (Big Five Persönlichkeitsmerkmale, Gerlitz & Schupp, 2005; berufliche Selbstwirksamkeitserwartung, Abele et al., 2000) und Merkmale der organisationalen Unterstützung (Arbeitsfeld, Größe der Einrichtung).

Zur Prädiktion des Berufserfolgs wurden lineare Regressionsanalysen und binär-logistische Regressionsanalysen gerechnet. In die Analysen gingen Daten von Personen ein, von denen für alle Merkmale valide Werte vorlagen (n = 653 bis n = 659).

Ergebnisse: Die Hauptbefunde zeigen, dass ein höheres Alter der Befragten mit einer höheren Wahrscheinlichkeit einhergeht, eine unbefristete Stelle einzunehmen. Zudem haben akademisch ausgebildete Frühpädagoginnen im Vergleich zu Fachkräften mit Fachschulabschluss eine erhöhte Wahrscheinlichkeit, eine Leitungsposition einzunehmen. Weiter zeigt sich, dass frühpädagogische Fachkräfte, die sich als vergleichsweise verträglich einschätzen, mit einer geringeren Wahrscheinlichkeit Leitungspositionen innehaben. Hinsichtlich der organisationalen Unterstützung wird deutlich, dass in Kindertageseinrichtungen beschäftigte Fachkräfte im Vergleich zu Fachkräften, die in anderen Einrichtungen tätig sind, eine höhere Chance haben, unbefristet beschäftigt zu sein. Bezüglich des subjektiven Berufserfolgs hat keines der soziodemographischen Merkmale einen Einfluss. Hinsichtlich des Humankapitals zeigt sich, dass frühpädagogische Fachkräfte, die bereits vor oder während ihrer Ausbildung oder ihres Studiums die aktuelle Stelle angetreten haben, sich zufriedener über ihre beruflichen Gestaltungsmöglichkeiten äußern als Fachkräfte, die erst nach Abschluss ihrer Ausbildung ihre aktuelle Stelle angetreten haben. Zudem sind berufliche Selbstwirksamkeitserwartungen deutlich positiv mit Aspekten der Arbeitszufriedenheit und negativ mit dem Belastungserleben assoziiert.

Freitag
13. März

15:30 Uhr

GA 03/149

Chair(s)

Paul, Johanna

Linberg, Tobias; Wenz, Sebastian

Ausmaß und Verteilung sozialer Ungleichheiten in den Kompetenzen von Kindergartenkindern

Während das Ausmaß sozialer Ungleichheiten im Kompetenzstand für Schülerinnen und Schüler im Primar- und Sekundarbereich gut dokumentiert ist, lassen sich entsprechende Auswertungen für Kindergartenkinder vergleichsweise selten finden. Unstrittig ist, dass die Möglichkeiten von Familien in ihre Kinder zu investieren und anregungsreichere Lernumwelten bereitzustellen in Abhängigkeit ihrer sozialen, ökonomischen und kulturellen Ressourcen variieren (vgl. Boudon 1974; Bourdieu 2012; Betz 2008; Lareau 2011) und die Varianz im Entwicklungsstand von Kindergartenkindern maßgeblich durch familiäre Merkmale aufgeklärt wird (vgl. Roßbach 2011). Als theoretisch und empirisch besonders bedeutsam erweisen sich in diesem Zusammenhang interfamiliäre Unterschiede nach sozialer Klasse, formalem Bildungsniveau, Haushaltseinkommen und Migrationshintergrund eines oder beider Eltern (Baumert et al. 2006; Becker & Lauterbach 2010; Faust und Roßbach 2014).

Doch obwohl vorschulische Förderung darauf abzielt, soziale Ungleichheiten von Beginn der Bildungskarriere von Kindern entgegenwirken, liegt bislang keine systematische bundesweite Bestandsaufnahme zum Ausmaß sozialer Disparitäten im Kompetenzstand von Kindergartenkindern in Deutschland vor. Ziel dieses Beitrags ist es daher, einen explorativen, deskriptiven aber dennoch detaillierten Überblick zu Disparitäten in sprachlichen und naturwissenschaftlichen Kompetenzen von Kindergartenkindern entlang zentraler Ungleichheitsdimensionen zu geben, der auf den aktuellen und für Deutschland repräsentativen Daten des Nationalen Bildungspanels (NEPS) beruht. Um diese Disparitäten möglichst detailliert darstellen zu können, visualisieren wir die Kompetenzverteilungen mittels Kerndichteschätzer und berichten insbesondere Ergebnisse bivariater Quantilsregressionen. Diese erlauben es, Unterschiede zwischen den verschiedenen Gruppen an unterschiedlichen Stellen der Kompetenzverteilung zu schätzen, wurden aber in der deutschen Bildungsforschung bislang kaum eingesetzt.

Die Ergebnisse zeigen insbesondere beim Wortschatz deutlich größere Unterschiede im unteren Bereich der Leistungsverteilung als am Median oder im oberen Bereich. Der mittlere Vorsprung der sozioökonomisch und kulturell – also nach sozialer Klasse, Bildungsherkunft und elterlichem Einkommen – besser gestellten Kinder beruht hier weniger auf einer hohen Anzahl sehr kompetenter Kinder, sondern vielmehr auf einer vergleichsweise niedrigen Zahl dieser Kinder mit einem geringen oder sehr geringen Wortschatz. Auch der Vorsprung der Kinder ohne vor denen mit Migrationshintergrund zeigt dieses Muster. Einen signifikanten Geschlechterunterschied finden wir an keiner Stelle der Kompetenzverteilung. Bei den naturwissenschaftlichen Kompetenzen ist dagegen kein derart eindeutiges Muster zu erkennen. Insgesamt sind die Disparitäten bei den Naturwissenschaften etwas gleichmäßiger über das Spektrum der Kompetenzwerte verteilt als beim Wortschatz und auch nicht so stark ausgeprägt. Das fällt auch beim Blick auf die berichteten Effektstärken auf. Die Werte für Cohen's d liegen größtenteils unter den entsprechenden Werten der Gruppenunterschiede im Wortschatz. Während aber diese nominellen Unterschiede in der Regel klein sind, wird das Bild größerer sozialer Disparitäten im Wortschatz insbesondere dann deutlich, wenn jeweils die Relation zum zwölfmonatigen Alterseffekt berücksichtigt wird: Die größten Differenzen zeigen sich in beiden Domänen beim Vergleich von Kindern, deren Eltern beide im Ausland geboren sind mit Kindern ohne Migrationshintergrund. Während diese in naturwissenschaftlichen Kompetenzen den zwölfmonatigen Alterseffekt nur knapp übersteigt, wird der Alterseffekt im Wortschatz etwa um das Zweieinhalbfache überschritten.

Zusammenfassend kann festgehalten werden, dass es bereits im Kindergartenalter große Unterschiede zwischen Kindern unterschiedlicher Herkunft gibt. Die Ergebnisse bivariater Quantilsregressionen zeigen insbesondere beim Wortschatz deutlich größere Unterschiede im unteren Bereich der Leistungsverteilung, die durch eine herkömmliche lineare Regression nicht erkennbar sind.

Freitag
13. März

15:30 Uhr

GB 03/49

Biedinger, Nicole; Kolb, Jan-Philipp; Klein, Oliver

Ethnische Unterschiede bei der Wahl des Kindergartens: Einfluss der Entfernung

Theoretischer Hintergrund: Der Kindergarten ist die erste bildungsrelevante Institution, mit denen Kinder in Kontakt treten. Daher gibt es zahlreiche Forschung zum Einfluss des Kindergartenbesuchs auf verschiedene Entwicklungsbereiche. Die Forschung beschäftigt sich dabei mit der Frage inwieweit der Kindergarten anfängliche Ungleichheiten kompensieren kann. Relativ wenig Forschung beschäftigt sich jedoch mit der Frage welcher Kindergarten besucht wird. Theoretisch ist davon auszugehen, dass sowohl familiäre als auch kindspezifische Faktoren, Einstellungen und Kontextvariablen einen bedeutenden Einfluss ausüben. Bei den Familien- und Kindmerkmalen ist vor allem an das Alter, der Erwerbstatus der Eltern, Anzahl der Geschwister, der Bildungsstand und das Einkommen zu denken. Bei den Einstellungen können die Bedeutung der frühkindlichen Betreuung, Bildungsaspirationen oder auch das Wissen über das Bildungssystem eine Rolle spielen. Und schließlich spielt die Anzahl der erreichbaren Kindergärten, die Kosten des Kindergartens oder auch spezifische Kindergartenmerkmalen (z.B. Qualität, ethnische Zusammensetzung) als Kontextvariablen eine wichtige Rolle.

Fragestellung: Empirisch zeigt sich, dass fast jedes Kind einen Kindergarten besucht. Daher beschäftigen wir uns mit der Frage welcher Kindergarten besucht wird. Wir gehen davon aus, dass die Entfernung zum Kindergarten bedeutsam ist und vermuten, dass sich deutsche Eltern und Familien mit türkischem Migrationshintergrund in der Wahl des Kindergartens unterscheiden. Genauer gesagt interessiert uns, ob der nächstgelegene oder ein Kindergarten, der weiter entfernt ist, gewählt wird. Der vorliegende Beitrag fragt spezifisch wie sich die ethnischen Unterschiede bei der Wahl des Kindergartens anhand der Eltern-/Kindmerkmale, der Einstellungen und auf Basis des Kontextes erklären lassen. Die Hauptfragestellung lautet: Werden Familien mit Migrationshintergrund durch andere (vor allem kontextuelle) Merkmale beeinflusst als deutsche Familien?

Methoden: Mit den Daten aus dem Projekt „Erwerb von sprachlichen und kulturellen Kompetenzen von Migrantenkindern in der Vorschulzeit“ kann zunächst überprüft werden, ob ethnische Unterschiede in der Wahl des Kindergartens existieren. Deskriptiv können ethnische Unterschiede bei der Wahl des Kindergartens festgestellt werden.

Mit logistischen Regressionen wird danach untersucht, inwieweit ethnische Unterschiede bei der Wahl des nächstgelegenen Kindergartens bestehen. In weiterführenden linearen OLS-Regressionen wird betrachtet wodurch die ethnischen Unterschiede erklärt werden können. Im Rahmen dieser Analysen kann neben den theoretisch vorgestellten Erklärungsvariablen auch ein Vergleich zwischen eher städtischen und ländlichen Gebieten durchgeführt werden.

Ergebnisse: Bei der Wahl des nächstgelegenen Kindergartens können keine ethnischen Unterschiede gefunden werden, sobald nach den Eltern-/Kindmerkmalen kontrolliert wird. Die Wahl des nächstgelegenen Kindergartens hängt vor allem vom Erwerbstatus bzw. dem Bildungsstand ab. Hierin unterscheiden sich deutsche Familien und Familien mit türkischem Migrationshintergrund nicht.

Betrachtet man die zusätzliche Entfernung, die die Familien in Kauf nehmen, um einen anderen Kindergarten zu besuchen, so lassen sich ethnische Unterschiede feststellen. Familien mit Migrationshintergrund nehmen einen wesentlich geringeren Weg in Kauf als vergleichbare deutsche Familien. Diese Unterschiede können vor allem durch den Erwerbstatus und die Einstellung erklärt werden.

Berücksichtigt man nun stärker den Wohnkontext, so kann interessanter Weise kein ethnischer Unterschied bei städtisch-lebenden Familien festgestellt werden, wohingegen in eher ländlichen Regionen ein starker Unterschied gefunden wird. Dies bedeutet, dass in Städten auch von Familien mit Migrationshintergrund die Wahl des Kindergartens nach ähnlichen Kriterien abläuft wie bei deutschen Familien (oder die Entscheidung liegt beim Kindergarten). Dies könnte eventuell auch so zu interpretieren sein, dass in Städten keine wirkliche Wahl stattfindet, sondern die Plätze so begrenzt sind, dass einfach ein Kindergartenplatz genommen wird. In ländlicheren Gebieten zeigt sich dies jedoch nicht. Hier spielen die vorgestellten Modellvariablen eine zentrale Rolle und ethnische Unterschiede bleiben auch nach Kontrolle der Modellvariablen bestehen. Es zeigt sich, dass Familien mit türkischem Migrationshintergrund stärker durch die eigenen Einstellungen und durch die Angebotsstruktur beeinflusst werden als vergleichbare deutsche Eltern.

Arndt, Petra A.; Sturmhöfel, Nicole; Fäsche, Anika; Rehm, Anja

Soziale und emotionale Belastungen und Veränderungen sozio-emotionaler Kompetenzen bei Kindern am Übergang in die Grundschule – Ergebnisse einer Elternbefragung im Modellprojekt „Bildungshaus 3 – 10“

Die Bedeutung des Übertritts vom Kindergarten in die Grundschule als entscheidendes Ereignis für die Entwicklung von Kindern und ihren Bildungserfolg wird in den letzten Jahren kontrovers diskutiert. Vertreter des Transitionsansatzes betrachten den Übergang in die Institution Schule aufgrund der damit verbundenen Veränderungen als kritisches Lebensereignis, das die Kinder, aber auch ihr Umfeld vor veränderte Herausforderungen stellt (z.B. Griebel & Niesel 2004). Die mit dem Schuleintritt einhergehenden Anpassungsleistungen sowie Aufbau und Neugestaltung von Beziehungen werden als potentiell Risiko für die Kinder betrachtet. Nach Griebel und Niesel (2004) treten bei 30 bis 50 Prozent der Kinder in Deutschland Übergangsprobleme auf. Andere Untersuchungen dagegen weisen auf eine weitaus geringere Anzahl von Kindern mit Problemen hin: Beelmann (2006) berichtet von fünf Prozent (aus Sicht der Eltern) bzw. vier Prozent (aus Sicht der pädagogischen Fach- und Lehrkräfte) der Kinder, die sechs Monate vor und nach dem Schuleintritt Probleme zeigen. Im Rahmen der „BiKS“-Studie ergaben Befragungen von Eltern, pädagogischen Fach- und Lehrkräften keine Hinweise für eine Steigerung von Problembelastungen durch den Schuleintritt (z.B. Faust, 2008). Die aus den beschriebenen Ergebnissen abgeleiteten Konsequenzen für die pädagogische Praxis sind dementsprechend unterschiedlich.

Im Rahmen der wissenschaftlichen Begleitung des Modellprojekts „Bildungshaus 3 – 10“ (Land Baden-Württemberg), wird daher, gefördert mit Mitteln des BMBF und des ESF der EU, u.a. untersucht, welche möglichen Belastungen durch den Schuleintritt auftreten und ob die Arbeit im „Bildungshaus 3 – 10“ diese Belastungen ggf. verringert oder zu ihrer Bewältigung beiträgt. Seit 2007/2008 arbeiten an 32 Bildungshausstandorten Grundschulen mit einem oder mehreren Kindergärten intensiv zusammen. Es finden gemeinsame, institutions- und jahrgangsübergreifende Spiel- und Lernangebote statt, mit dem Ziel Kindern eine bruchlose Bildungsbiographie und einen fließender Übergang zu ermöglichen. Neben den Bildungshausstandorten nehmen 28 Vergleichsstandorte an der wissenschaftlichen Begleitung teil. Dadurch wird es möglich, sowohl eventuelle Belastungen als auch mögliche Effekte übergangsbegleitender Maßnahmen zu untersuchen.

Einer der Gründe für die unterschiedlichen Ergebnisse der o.g. Publikationen ist sicherlich methodischer Natur. So werden in den Studien unterschiedliche Kriterien etwa für einen gelungenen Übergang sowie unterschiedliche Erhebungsinstrumente verwendet. Teilweise entstammen die eingesetzten Erhebungsinstrumente dem klinischen Bereich. Der Fokus solcher Instrumente und ein möglicherweise dadurch hervorgerufener Antwortbias der Befragten könnten zu einer systematischen Unterschätzung der Anzahl betroffener Kinder führen. Daher wurde für die hier vorgestellte Untersuchung ein eigenes Erhebungsinstrument entwickelt, das vermehrt auf Stärken und Kompetenzen der Kinder fokussiert und weniger auf (klinisch relevante) Problematiken. Dieser Fragebogen zu sozialen und emotionalen Kompetenzen (FSEK), der u.a. auf Basis des SDQ und der CBCL 4-18 entwickelt wurde, umfasst die Bereiche soziale Einbindung, positive Emotionalität, emotionale Stabilität, Ängstlichkeit, Empathie und prosoziales Verhalten. Der Einsatz eines stärkenorientierten Fragebogens ist umso wichtiger, als mit dem Begriff „kritische Lebensereignisse“, der in der Transitionsforschung verwendet wird, ursprünglich solche Ereignisse gemeint sind, die sowohl Entwicklungschancen als auch Risiken beinhalten (Filipp 1981) und nicht nur rein risikobehaftete Ereignisse, wie es die aktuelle Diskussion nahelegt (vgl. Faust 2004). Ziel ist es, ggf. auch genutzte Entwicklungschancen zu erfassen.

Die Eltern wurden jeweils ein halbes Jahr vor und nach der Einschulung mit dem FSEK zu sozialen und emotionalen Kompetenzen und Verhaltensweisen ihrer Kinder befragt. Für vier Kohorten am Übergang vom Kindergarten in die Grundschule liegen Daten von insgesamt 3134 Kindern vor. Da sowohl Beelmann (2006) als auch Faust (2008) zu dem Befund, dass bestimmte Gruppen von Problemen beim Schuleintritt verstärkt betroffen sind, übereinstimmen, wurden in Mehrebenenanalysen entsprechende Faktoren insbesondere der familiären Situation und der Migrationshintergrund einbezogen. Insgesamt zeigen die Ergebnisse, dass für einzelne Kohorten und für unterschiedliche Bereiche der Übergang eine Herausforderung darstellt. Ebenso zeigen sich in einzelnen Bereichen Unterschiede bei der Bewältigung des Übergangs zwischen Kindern in Modell- und Vergleichseinrichtungen sowie Kompetenzgewinne besonders in den sozialen Fertigkeiten bei allen Kindern.

Sommer, Anja; Sechtig, Jutta

Bedeutung einer erweiterten Altersmischung in Kindergärten für die sozio-emotionale Interaktionsqualität

Die Bedeutung pädagogischer Interaktionen wird durch verschiedene Studienergebnisse unterstrichen, die auf positive Auswirkungen auf den kognitiven und sozio-emotionalen Entwicklungsbereich sowie Kompetenzen der Lebensführung hinweisen und sich zum Teil bis ins Jugend- und Erwachsenenalter nachweisen lassen (Peisner-Feinberg et al., 2001; Vandell, Belsky, Burchinal, Steinberg & Vandergrift, 2010). Als Schlüsselvariable qualitativ hochwertiger Kindertageseinrichtungen gelten vor allem Prozesse zwischen pädagogischen Fachkräften und Kind (Viernickel, 2008). In der vorschulischen Bildungsforschung werden üblicherweise drei Dimensionen der Prozessqualität unterschieden: kognitive Anregung, Strukturierung pädagogischer Prozesse und sozio-emotionale Unterstützung (Bäumer, Aßmann, von Maurice & Blossfeld, 2013; Pianta, La Paro & Hamre, 2008). Die vorliegenden Analysen beziehen sich dabei auf die Dimension der sozio-emotionalen Unterstützung, unter der ein Interaktionsverhalten pädagogischer Fachkräfte verstanden wird, das auf Bedürfnisse aller Kinder der Gruppe angemessen reagiert und von emotionaler Wärme und Strukturierung gekennzeichnet ist (Ahnert, Pinquart & Lamb, 2006; Remsperger, 2013).

Die sozio-emotionale Interaktionsqualität ist nicht losgelöst von strukturellen Rahmenbedingungen zu betrachten. So sind kleinere Gruppen (Gerber, Whitebook & Weinstein, 2007), erfahrenere pädagogische Fachkräfte (Suchodoletz, Fäsche, Gunzenhauser & Hamre, 2014) sowie günstige Aspekte der Gruppenzusammensetzung (Gerber et al., 2007; Tietze et al., 2013) mit dem Ausmaß sozio-emotionaler Interaktionsqualität assoziiert. Das Feld der vorschulischen institutionellen Betreuung ist jedoch seit geraumer Zeit von großen strukturellen Veränderungen geprägt. Während ein Überangebot von Betreuungsplätzen in regulären Kindergartengruppen für Kinder von drei bis sechs Jahren besteht (Nied, Niesel, Haug-Schnabel, Wertfein & Bense, 2011), zeigt sich eine große Differenz zwischen Betreuungsquote und Betreuungsbedarf bei der Altersgruppe der Ein- bis Dreijährigen (Bundesministerium für Familie, Senioren, Frauen und Jugend, 2013). Dieser Herausforderung wird zum Teil mit einer Erweiterung der Altersmischung in regulären Kindergartengruppen durch Aufnahme jüngerer Kinder begegnet.

Einzelne Befunde verweisen auf einen möglichen negativen Zusammenhang einer solchen erweiterten Altersmischung in Kindergartengruppen (Kuger & Kluczniok, 2008; Tietze et al., 2013). Jedoch gibt es für den deutschen Kontext keine differenzierten Ergebnisse dazu, welche Merkmale von Kindergartengruppen und pädagogischen Fachkräften zur Reduktion bzw. Aufrechterhaltung sozio-emotionaler Interaktionsqualität beitragen. Dementsprechend sind folgende Forschungsfragen für den vorliegende Beitrag leitend: (1) Wie gestaltet sich die Beziehung von Strukturmerkmalen und sozio-emotionaler Interaktionsqualität unter Berücksichtigung der erweiterten Altersmischung? (2) Lassen sich Strukturmerkmale identifizieren, die den Zusammenhang von erweiterter Altersmischung und sozio-emotionaler Interaktionsqualität verringern?

Zur Beantwortung dieser Fragen wurde auf Videoaufnahmen von Erzieherin-Kind-Interaktionen (Dauer: 54-147 Minuten) während der Freispielzeit und auf ergänzende Fragebogendaten zurückgegriffen. Die Einschätzungen der sozio-emotionalen Interaktionsqualität erfolgte durch zwei unabhängige Beobachterinnen mit Hilfe der Caregiver Interaction Scale (Arnett, 1989) (Beobachterübereinstimmung ICC = .84). 39 Erzieherinnen aus 22 Kindergärten konnten so hinsichtlich ihres Interaktionsverhaltens eingeschätzt werden. In linearen OLS-Regressionen wurde die Stärke des Zusammenhangs zwischen erweiterter Altersmischung und sozio-emotionaler Interaktionsqualität unter Kontrolle von Merkmalen der Strukturqualität (Fragestellung 1) berechnet, um im Anschluss durch die Berücksichtigung von Interaktionseffekten herauszufinden, ob die Stärke der Beziehung zwischen erweiterter Altersmischung und sozio-emotionaler Interaktionsqualität von Merkmalen der Strukturqualität abhängt (Fragestellung 2).

Insgesamt stellt die erweiterte Altersmischung den wichtigsten Prädiktor der sozio-emotionalen Interaktionsqualität dar. Zudem lassen sich Aspekte der Strukturqualität identifizieren, die den negativen Zusammenhang einer erweiterten Altersmischung und sozio-emotionaler Interaktionsqualität – wenn auch nicht vollständig – abmildern. So zeigten sich kleinere Gruppen sowie pädagogische Fachkräfte mit einer mittleren Berufserfahrung als günstige Rahmenbedingungen. Abschließend werden, vor dem Hintergrund der Limitationen dieser Studie, die zentralen Befunde und daraus resultierende Implikationen diskutiert.

Schneider, Rebecca; Sparfeldt, Jörn

Je älter, desto genauer? - Akkuratess selbstberichteter Zensuren bei Grundschulkindern

Theorie: In der empirischen Bildungsforschung greift man häufig auf die von Schülerinnen und Schülern selbstberichteten Zensuren als Indikatoren der erhaltenen Zensuren und damit der Schulleistungen zurück. Dieses ökonomische Vorgehen kann, insgesamt gesehen, für höhere Klassenstufen und für verschiedene Schulfächer als angemessen beurteilt werden (Dickhäuser & Plenter, 2005; Sparfeldt, Buch, Rost & Lehmann, 2008; zsf. Kuncel, Credé & Thomas, 2006). Ob dieses Befundmuster jedoch auch auf Grundschulkindern mit u.a. anderem kognitiven Entwicklungsstand generalisiert werden kann, ist weitgehend unklar. Ersten Ergebnissen zufolge erinnerten sich 92% der Viertklässler korrekt an ihre Note auf dem Abschlusszeugnis der 3. Klasse in Mathematik (Lese-, Sport- und Kunstnote: 85%), bei Drittklässlern waren die Selbstberichte weniger akkurat; entsprechend nahm der Anteil (leicht) überschätzter selbstberichteter Zensuren mit steigender Klassenstufe ab (Ostrop, Schmude & Valtin, 2002).

Fragestellungen: Im Vortrag betrachten wir systematisch alle vier Grundschulklassenstufen und gehen den Fragen nach, (1) wie hoch die von den Schülerinnen und Schülern selbstberichteten mit den von den Lehrkräften gegebenen Zensuren in den Fächern Deutsch, Mathematik und Sport in den einzelnen Klassenstufen korrelieren, (2) inwieweit die schülerberichteten von den erhaltenen Zensuren in den jeweiligen Klassenstufen abweichen (betrachtet werden Mittelwertdifferenzen und die Anteile an Über- bzw. Unterschätzungen der selbstberichteten Zensuren), und ob sich (3) differentielle kriteriumsbezogene Validitätskoeffizienten zwischen selbstberichteten und erhaltenen Zensuren mit der Intelligenz und schulischen Selbstkonzepten zeigen.

Methode: Die Stichprobe umfasst 48 Klassen aus 10 Schulen mit N = 592 Schülerinnen und Schülern der ersten bis vierten Klasse. Die Erhebung erfolgte vier Monate nach der Ausgabe der Halbjahreszeugnisse; erfasst wurden die Zeugniszensuren in Deutsch, Mathematik und Sport – zum einen im Selbstbericht der Schülerinnen und Schüler und zum anderen im Lehrerprotokoll; in den unteren Klassenstufen ohne Ziffernzeugnis (Klassenstufe 1 und teilweise 2) wurden Schülerinnen und Schüler sowie Lehrkräfte um die Angabe von Quasinoten gebeten. Zusätzlich wurden die (nonverbale) Intelligenz und schulische Selbstkonzepte (Lesen, Mathematik, Sport) erfasst.

Ergebnisse: (1) Die Korrelationen schülerberichteter mit lehrerberichteten (Quasi-)Zensuren steigen über die Klassenstufen hinweg an (Klassenstufe 1: $r \approx .30$; Klassenstufe 4: $r \approx .80$). (2) Die Mittelwerte der selbstberichteten Zensuren bzw. Quasinoten fallen stets besser aus als der erhaltenen; die Notendifferenzen zwischen den Datenquellen (Lehrer, Schüler) werden mit höherer Klassenstufe kleiner. Entsprechend nimmt der Prozentsatz korrekter Angaben in allen Fächern über die Klassenstufen hinweg zu (von rund 40% auf ca. 80%); auch die Über- und Unterschätzungsanteile nehmen über die Klassenstufen ab. (3) In allen vier Klassenstufen korrelieren die lehrerberichteten Zensuren höher (negativ) mit der Intelligenz als die schülerberichteten. Die Zusammenhänge mit den jeweils korrespondierenden schulischen Selbstkonzepten fallen – zumindest ab der zweiten Klasse mit Ziffernzensuren – für selbstberichtete Zensuren vergleichbar oder etwas enger als für lehrerberichtete aus.

Diskussion: Die Angemessenheit der Verwendung selbstberichteter Zensuren als Schulleistungsindikatoren bei Grundschulkindern wird diskutiert, insbesondere wird auf die differentielle Eignung selbstberichteter Zensuren in den verschiedenen Klassenstufen eingegangen.

Lohse-Bossenz, Hendrik

Empirische Zugänge zur Erfassung von Kompetenzen im Bereich früher naturwissenschaftlicher Bildung

Frühe naturwissenschaftliche Bildung gewinnt in Wissenschaft und Praxis zunehmend an Bedeutung. Es existiert eine Vielzahl an Ansätzen und Zugängen, wie frühpädagogische Fachkräfte Angebote für Kinder in diesem Bildungsbereich gestalten können. Parallel dazu existieren vielfältige Initiativen und Projekte, die Professionalisierungsangebote für pädagogische Fachkräfte konzipieren und anbieten. Ein bislang wenig thematisierter Bereich ist die empirische Untersuchung von pädagogischen Prozessen und Lernerfolgen innerhalb dieser Domäne. Erkenntnisse solcher Untersuchungen können u.a. das Verständnis kindlicher Lernprozesse bereichern, aber auch Ausgangspunkt zur Optimierung bestehender Fortbildungsinitiativen darstellen. Ziel des Symposiums ist es, verschiedene Ansätze und Methoden zur Erfassung professioneller Kompetenz und professionellen Verhaltens frühpädagogischer Fachkräfte (Fragebögen, Wissenstests, standardisierte Beobachtung sowie Videoanalysen) sowie empirische Zugänge zur Erfassung früher naturwissenschaftlicher Kompetenz von Kindern darzustellen und vergleichend zu betrachten.

Der erste Beitrag untersucht mittels Pfadanalysen den Zusammenhang von Selbstwirksamkeit (als kontextspezifischer Aspekt professioneller Kompetenz) mit zwei Maßen professionellen Verhaltens (pädagogische Handlungen und kognitive Aktivierung). Im zweiten Beitrag wird ein Beobachtungsverfahren vorgestellt, das die pädagogische Interaktionsqualität auf sechs Dimensionen erfasst (Alltagsbezug, Selbsttätigkeit, inhaltliche Orientierung, naturwissenschaftliche Aktivitäten, Kommunikationsgeschehen und kognitive Aktivierung) und im Rahmen einer experimentellen Feldstudie eingesetzt wird. Das Konstrukt der kognitiven Anregung steht auch im Zentrum des dritten Beitrags. In der Analyse von videographierten Angebotssituationen werden Zusammenhänge zwischen kognitiver Anregung, Konzeption, Merkmalen der Lernumgebung und dem Professionshintergrund der pädagogischen Fachkräfte untersucht. Der letzte Beitrag beschäftigt sich mit Möglichkeiten der empirischen Erfassung emotionaler und motivationaler Aspekte naturwissenschaftlicher Kompetenz bei Kindern im Vorschulalter. Die Ergebnisse der Einzelbeiträge werden übergreifend miteinander verglichen und diskutiert.

Dieses Symposium ist darauf ausgerichtet, die Vielfalt der empirischen Zugänge darzustellen. Der erste und letzte Beitrag orientiert sich an pädagogisch-psychologischen Merkmalen und setzt diese zueinander in Beziehung. Dabei werden sowohl pädagogische Fachkräfte als auch Vorschulkinder in den Blick genommen. Der zweite Beitrag rekurriert auf den in der Pädagogik verankerten Ansatz der Prozessqualität. Im dritten Beitrag werden Angebote aus naturwissenschaftsdidaktischer Sicht analysiert und mit Kontext- und Individualmerkmalen in Beziehung gesetzt. Die hier dargestellten Beiträge spiegeln somit die Heterogenität des Feldes wider, verbunden mit dem Ziel, konzeptuelle und methodische Schnittmengen zu identifizieren.

Freitag
13. März

15:30 Uhr

HGB 20

Chair(s)

Lohse-Bossenz, Hendrik

Empirische Zugänge zur Erfassung von Kompetenzen im Bereich früher naturwissenschaftlicher Bildung

Lohse-Bossenz, Hendrik; Zimmermann, Monika; Janke, Miriam; Müller, Stefan

Selbstwirksamkeit von fröhpädagogischen Fachkräften im Bereich früher naturwissenschaftlicher Bildung – Kontextspezifität und Beziehung zum professionellen Verhalten

Theoretischer Hintergrund: Selbstwirksamkeit als die Erwartung von Personen, Herausforderungen aufgrund eigener Fähigkeiten bewältigen zu können (Guo et al., 2010), wird im Bereich früher Bildung vielfach aufgegriffen. Kinder sollten in verschiedenen pädagogischen Angeboten die eigene Selbstwirksamkeit erleben und darin gestärkt werden (u.a. M. Zimmermann, 2011). Vor allem Studien aus der empirischen Schulforschung zeigen dabei den positiven Zusammenhang von Selbstwirksamkeit von Kindern auf deren Engagement in Lernprozessen und den damit verbundenen Lernerfolgen (B. J. Zimmermann, 2000). In gleichem Maße wurde das Konstrukt der Selbstwirksamkeit auch auf pädagogische Fachkräfte übertragen. Dabei zeigte sich, dass Selbstwirksamkeit positiv mit Interaktionsqualität und Lernerfolgen bei Kindern assoziiert ist (Guo, Justice, Sawyer, & Tompkins, 2011; Guo, Piasta, Justice, & Kaderavek, 2010). Trotz einer Vielzahl an positiven Befunden aus der Schulforschung sowie einiger empirischer Befunde in der internationalen Forschungsliteratur lassen sich in der deutschsprachigen Forschungslandschaft wenige empirische Zugänge zur bereichsspezifischen Erfassung von Selbstwirksamkeitserwartungen für frühe naturwissenschaftliche Bildung (fnB) identifizieren.

Ziel des vorliegenden Beitrags ist es daher, Selbstwirksamkeit im Bereich früher naturwissenschaftlicher Bildung von allgemeiner Selbstwirksamkeit und Selbstwirksamkeit für frühe Bildung empirisch abzugrenzen und deren Beitrag zur Erklärung professionellen Verhaltens von fröhpädagogischen Fachkräften bei Angeboten zur frühen naturwissenschaftlichen Bildung aufzuzeigen.

Methode.

Stichprobe: Zur Beantwortung der Fragestellung wurden Daten aus zwei unabhängigen Erhebungen vergleichend analysiert. Die erste Erhebung umfasst 88 und die zweite Erhebung 97 fröhpädagogische Fachkräfte aus verschiedenen Regionen Deutschlands.

Instrumente: Zur Erfassung der Selbstwirksamkeit wurden drei Fragebögen eingesetzt, die einerseits die allgemeine Selbstwirksamkeit (Schwarzer & Jerusalem, 1999; 10 Items, $\alpha = .86$) sowie die Selbstwirksamkeit im Bereich früher Bildung (Herrmann, Frei, Sticca & Perren, 2014; 10 Items, $\alpha = .84$) und andererseits die Selbstwirksamkeit für frühe naturwissenschaftliche Bildung (Eigenentwicklung; 10 Items, $\alpha = .79$) abbilden. Gleichzeitig wurden die pädagogischen Fachkräfte gebeten, ihr eigenes Verhalten bei Angeboten zu fnB einzuschätzen. Eine von zwei Skalen zielt auf unterstützende Verhaltensweisen der Fachkräfte beim Entdecken von Naturphänomenen ab (Adaption nach M. Zimmermann, 2011; 5 Items, $\alpha = .67$). Innerhalb der anderen werden Verhaltensweisen der pädagogischen Fachkraft zur kognitiven Aktivierung der Kinder erfasst (Eigenentwicklung; 7 Items, $\alpha = .84$). Alter und Berufserfahrung wurden als Kontrollvariablen in die Regressionsanalysen eingeschlossen.

Ergebnisse: Korrelationsanalysen für die Selbstwirksamkeitsskalen zeigen geringe bis mittelhohe Zusammenhänge ($.19 < r > .62$). Gleiches gilt für die Zusammenhänge zum professionellen Verhalten. Um die komplexen Wirkungszusammenhänge näher zu betrachten, wurden manifeste Strukturmodelle gerechnet, bei denen die Selbstwirksamkeitsmerkmale mit den Verhaltensmerkmalen in einem Pfadmodell geschätzt wurden. Als Kontrollvariablen wurden dabei das Alter und die Berufserfahrung aufgenommen. Für die erste Stichprobe lässt sich sagen, dass allgemeine Selbstwirksamkeit professionelles Verhalten in Angeboten zur frühen naturwissenschaftlichen Bildung nicht vorhersagt ($\beta = -.19 / .13$). Selbstwirksamkeit für frühe Bildung ist prädiktiv für die Unterstützung der Kinder beim Entdecken von Naturphänomenen ($\beta = .37^*$), nicht aber für kognitiv aktivierendes Verhalten ($\beta = .18$). Für die Selbstwirksamkeit im Bereich frühe naturwissenschaftliche Bildung zeigen die Befunde eine hohe prädiktive Kraft sowohl für unterstützende Verhaltensweisen ($\beta = .26^*$) als auch für Verhaltensweisen, die Kinder zur tiefer gehenden gedanklichen Auseinandersetzung anregen ($\beta = .34^*$). Dieses Befundmuster lässt sich in der zweiten Stichprobe replizieren.

Diskussion: Zusammenfassend belegen die Befunde die Kontextspezifität von Selbstwirksamkeitserwartungen dahingehend, dass allgemeine Selbstwirksamkeitserwartungen professionelles Verhalten im Bereich fnB nicht vorhersagen. Darüber hinaus zeigt sich, dass kognitiv aktivierende Verhaltensweisen von den Fachkräften häufiger gezeigt werden, wenn diese über höhere Selbstwirksamkeit in den Naturwissenschaften verfügen. Daraus ergibt sich die Notwendigkeit, einerseits Selbstwirksamkeit differenziert zu erfassen und andererseits im Rahmen von Fortbildungen Selbstwirksamkeitserwartungen für fnB zu stärken.

Die hier berichteten Befunde beruhen ausschließlich auf Selbstauskünften der Fachkräfte. Inwieweit die eigenen Auskünfte mit beobachtetem Verhalten zusammenhängen und welche prädiktive Kraft Selbstwirksamkeitserwartungen für beobachtbares Verhalten haben, ist u.a. Gegenstand eines von der Klaus Tschira Stiftung gGmbH geförderten längsschnittlichen Projekts.

Freitag
13. März

15:30 Uhr

HGB 20

Chair(s)

Lohse-Bossenz, Hendrik

Empirische Zugänge zur Erfassung von Kompetenzen im Bereich früher naturwissenschaftlicher Bildung

von Thaden, Dorina; Lohse-Bossenz, Hendrik; Zimmermann, Monika; Nattefort, Rebecca; Janke, Miriam; Kremer, Isabelle; Müller, Stefan; Tietze, Wolfgang

Interaktionsqualität in Angeboten zur frühen naturwissenschaftlichen Bildung – Erfassung und Verwendung zur Fortbildungsevaluation

Theoretischer Hintergrund: Die Erfassung der Qualität früher naturwissenschaftlicher Bildung (fnB) stellt sich aus mehreren Gründen als schwierig dar. Zum Ersten existieren aktuell keine Verfahren, die diesen Entwicklungsbereich hinreichend differenziert beschreiben. In der curricularen Erweiterung zur Kindergartenskala KES-E (Sylva, Siraj-Blatchford, & Taggart, 2009) wird der Bereich Natur- und Sachwissen aufgegriffen, an strukturellen Gegebenheiten orientiert allerdings nicht auf spezifische Angebotsstrukturen zugeschnitten. Zum Zweiten existiert aktuell kein normativer Konsens, der Ziele und Inhalte früher naturwissenschaftlicher Bildung beschreibt und als Richtwert zur Instrumentenentwicklung dienen kann. Verschiedene theoretische Auffassungen reichen dabei vom reinen Erleben und Entdecken naturwissenschaftlicher Phänomene und kindlicher Umwelten (Zimmermann, 2011) bis hin zur systematischen Ausbildung schulischer Vorläuferfähigkeiten (Möller & Steffensky, 2010; Steffensky, Lankes, Carstensen, & Nölke, 2012). Damit einhergehend existieren auch keine allgemein akzeptierten curricularen Beschreibungen darüber, welche Naturerfahrungen Kinder gemacht und welche Phänomene sie kennengelernt haben sollten (Ginsburg & Golbeck, 2004). Das Ziel der vorliegenden Studie besteht in der Entwicklung eines Beobachtungsverfahrens, das im Rahmen einer längsschnittlich experimentellen Evaluationsstudie zu einem spezifischen Fortbildungsprogramm „Mit Kindern die Welt entdecken“ (Zimmermann, 2011) eingesetzt wird. Damit verbunden sind die Forderungen nach der inhaltlichen Passung zum Fortbildungskonzept (im Sinne inhaltlicher/curricularer Validität) und nach der Möglichkeit, Veränderungen im Interaktionsverhalten abzubilden.

Methode.

Verfahrensentwicklung: Aufbauend auf einer Analyse des Fortbildungskonzepts sowie dem damit verknüpften Leitbild der „kompetenten“ Erzieherin wurden zehn Merkmale identifiziert. Durch eine inhaltliche Verknüpfung verschiedener Merkmale und die Sichtung weiterer fnB-Konzepte (z.B. Fthenakis, Wendell, Eitel, Daut, & Schmitt, 2012) wurden Merkmale entwickelt, die die Qualität eines strukturierten Angebots im Bereich früher naturwissenschaftlicher Bildung beschreiben: Alltagsbezug, Orientierung des Angebots an kindlichen Bedürfnissen, Fragen und Interessen, Selbsttätigkeit, Naturwissenschaftliche Aktivitäten, Kognitive Aktivierung, Kommunikationsgeschehen. In einem weiteren Schritt wurden zu jedem Merkmal Beobachtungsaspekte entwickelt, die dieses Merkmal charakterisieren und Inhalte verschiedener theoretischer Ansätze aufgreifen. Die Formulierung der Beobachtungsaspekte orientiert sich an der Konstruktion der Kindergartenskala (KES-R; Tietze, Schuster, Grenner, & Roßbach, 2007). Die Einschätzung der Qualität erfolgt für jede Beschreibungsdimension separat anhand einer 7-stufigen Skala, bei der den Beurteilern einzelne Beobachtungsaspekte auf den Stufen 1, 3, 5, und 7 zur Verfügung stehen. Die inhaltliche Struktur (Merkmale) sowie ein Großteil der Beobachtungsaspekte wurden verschiedenen Personen aus dem Feld zur Begutachtung vorgelegt: zwei Hochschulprofessor(inn)en, einer pädagogische Fachberatung sowie einer Erzieherin mit mehreren Jahren Erfahrung im Bereich früher naturwissenschaftlicher Bildung. Die Anmerkungen wurden diskutiert und entsprechend im Verfahren berücksichtigt.

Design und Stichprobe: Die Verfahrensentwicklung ist Bestandteil eines durch die Klaus Tschira Stiftung gGmbH geförderten Verbundprojekts. Insgesamt nehmen 64 frühkindliche Bildungseinrichtungen in einem längsschnittlichen Evaluationsprojekt teil. Bisher liegen bzgl. des Beobachtungsverfahrens Daten von 50 Einrichtungen vor. In diesen wurde jeweils eine pädagogische Fachkraft bei der Durchführung eines pädagogischen Angebots zum Thema „Seifenblasen“ beobachtet.

Ergebnisse.

Beobachterübereinstimmung: Im Vorfeld der Untersuchung fanden mehrtägige Beobachtertrainings statt, an deren Abschluss eine Reliabilitätsprüfung stand. Hierfür wurde eine Praxisbeobachtung durchgeführt. Die Interraterreliabilität lag bei Kendall's $W = .81$.

Deskriptive Daten und Interkorrelationen: Die Ergebnisse zeigen eine zufriedenstellende Verteilung der Beobachtungswerte. So wurde für 3 von 5 Dimensionen der komplette Wertebereich von 1-7 ausgenutzt. Bei den anderen wurden Werte von 2-7 vergeben. Die Mittelwerte lagen um den theoretischen Mittelwert (3.36 – 4.44) bei einer ausreichenden Streuung (1.32 – 2.48). Da die Interkorrelationen der Dimensionen zwischen .20 und .51 liegen, können diese als ausreichend diskriminant angesehen werden.

Diskussion und Ausblick: Zusammenfassend kann festgehalten werden, dass es gelungen ist, ein reliables Beobachtungsverfahren zu entwickeln, das keinerlei Boden- bzw. Deckeneffekte zeigt und somit geeignet erscheint, Fortbildungseffekte abzubilden. Weitere Analysen können darauf ausgerichtet werden, einerseits Unterschiede zwischen pädagogischen Fachkräften zu erklären (z.B. durch bisherige Fortbildungserfahrungen oder Aspekte professioneller Kompetenz) und andererseits die hier beschriebenen Dimensionen mit anderen Merkmalen pädagogischer Qualität in Beziehung zu setzen (z.B. sprachliche Interaktionen, Fried & Briedigkeit, 2008).

Freitag
13. März

15:30 Uhr

HGB 20

Chair(s)

Lohse-Bossenz, Hendrik

Empirische Zugänge zur Erfassung von Kompetenzen im Bereich früher naturwissenschaftlicher Bildung

Gierl, Katharina; Kauertz, Alexander

Professionalisierung pädagogischer Fachkräfte im Elementarbereich - Eine Videostudie zur Analyse naturwissenschaftlicher Bildungsangebote

Theoretischer Hintergrund: Die Qualität frühkindlicher Bildung wird an Strukturmerkmalen der Einrichtung, Orientierungsmerkmalen der pädagogischen Fachkraft und Prozessmerkmalen des Angebots festgemacht (Kuger & Kluczniok, 2009). Der Interaktionsprozess zwischen pädagogischer Fachkraft und Kind ist dabei maßgeblich für die Qualität der Bildung (Kontos & Wilcox-Herzog 2002). Kognitiv anregende Interaktionen eröffnen den Kindern Bildungschancen im naturwissenschaftlichen Bereich (Siraj-Blatchford, I. & Manni, L 2008). Naturwissenschaftliche Denk- und Arbeitsweisen regen das Kind kognitiv an, seine Wissensstrukturen zu verbalisieren, zu erweitern und auf andere Kontexte anzuwenden (Steffensky & Lankes 2011).

Forschungsfragen: Ein vielfältiges materielles und räumliches Umfeld der Kinder erhöht die Chance kognitiver Anregungen. Das Konzept der Einrichtung sollte Freiräume für das selbstständige Entdecken und Ausprobieren vorsehen. Die pädagogische Fachkraft muss Entdecken und Ausprobieren anregen und begleiten können. Daher wird korrelativ der Einfluss der Konzeption, der räumlichen und materiellen Ausstattung und personenspezifische Merkmale der Fachkraft auf naturwissenschaftliche kognitive Aktivierung untersucht.

Methode: Die Studie fand im Rahmen des BMBF-geförderten Verbundvorhabens PRIMEL statt. Aufgrund der Vielzahl erhobener Konstrukte, waren die domänenspezifischen Itemzahlen stark eingeschränkt. Es wurden N=30 Fachkräfte einer Gelegenheitsstichprobe bei der Durchführung nicht standardisierter, naturwissenschaftlicher Angebote (t=30 min) videografiert.

Der personale Kontext wurde durch selbstentwickelnde Kurzfragebögen erhoben: Wissen über Naturwissenschaften (4 Items; $\alpha = .55$), Fähigkeitsselbstkonzept (5 Items; $\alpha = .84$), eigener Bezug zur Domäne (9 Items; $\alpha = .85$) und Wissen über naturwissenschaftliche Lernprozesse (5 Items; $\alpha = .10$). Mit der geringen Itemzahl konnte das vielschichtige Konstrukt „Wissen über naturwissenschaftliche Lernprozesse“ nicht konsistent erfasst werden. Der Ausbildungshintergrund wurde als distales Merkmal des Professionswissens (Baumert & Kunter 2006) erhoben (fachschulisch, akademisch, schweizerische Kindergartenlehrpersonen).

Der situationale Kontext (materielle und räumliche Ausstattung, Konzeption) wurde durch Fragebögen mit dichotomen Antwortformat erfasst (ja/nein bzw. vorhanden/nicht vorhanden) und durch ein selten genutztes offenes Item ergänzt („sonstiges“).

Die kognitive Aktivierung wurde in den Videos im 10s-Timesampling kodiert. Das Kategoriensystem wurde iterativ (Seidel et al. 2006) aus CLASS und KES (Pianta et al. 2014, Tietze et al. 2005) abgeleitet und induktiv angepasst. Die Skala zur naturwissenschaftlichen kognitiven Aktivierung umfasst vier Items ($\alpha = .60$). Aufgrund der hochinferenten Items wurden alle Videos doppelt kodiert (Interkoderübereinstimmung 61).

Ergebnisse: Die Studie gibt Hinweise, welche Merkmale lohnenswert für die weitere Untersuchung von Bedingungen naturwissenschaftsspezifischer kognitiver Aktivierung sind. Der situationale Kontext zeigt teilweise einen Zusammenhang zur Häufigkeit kognitiver Aktivierungen. Die Schweizer Konzeption zeigt einen negativen Effekt gegenüber den offenen, teiloffenen und geschlossenen Konzeptionen ($F(1, 30) = 7,921$; $p = .009$). Die Anzahl einschlägiger räumlicher Möglichkeiten wirkt positiv ($r = .485$, $p = .007$). Beim Material zeigen zwei Items mittlere Effekte, die bei der Stichprobengröße nicht signifikant werden (strukturierte Ablage: $T(28) = 1.469$; $p = .153$, Verfügbarkeit zusätzlicher Informationen: $T(28) = -1.530$; $p = .137$). Bei personalen Merkmalen zeigt nur die Skala „Anregen von Lernprozessen“ einen mittleren Effekt ($b = .395$; $p = .041$; $R^2 = .156$). Aufgrund der geringen internen Konsistenz der Skala ist jedoch noch zu klären, welche Aspekte dieses Konstrukts tatsächlich die naturwissenschaftliche kognitive Aktivierung hervorrufen.

Freitag
13. März

15:30 Uhr

HGB 20

Chair(s)

Lohse-Bossenz, Hendrik

Heinig, Elisa; Anders, Yvonne; Maria, Kleinert

Wie lassen sich nicht-leistungsbezogene Aspekte frühkindlicher naturwissenschaftlicher Kompetenz erfassen?

Theoretischer Hintergrund: Naturwissenschaftliche Bildung beginnt vor der Einschulung. Aktuelle Forschungsergebnisse zeigen, dass Kinder bereits im vorschulischen Alter in der Lage sind, sich mit naturwissenschaftlichen und technischen Themen auseinanderzusetzen (Koerber, Sodian, Thoermer, & Nett, 2005; Krahn, 2005; Pauen, 2003; Sodian, Koerber, & Thoermer, 2004). Die naturwissenschaftliche Kompetenz jüngerer Kinder wird in der internationalen und nationalen Diskussion in Anlehnung an Konzeptionen der Scientific Literacy als mehrdimensionales Konstrukt aufgefasst, das neben Wissenskomponenten auch nicht leistungsbezogene Komponenten umfasst (Bybee, McCrae, & Laurie, 2009; Norris & Phillips, 2003). Hierzu zählen emotionale Merkmale, wie das Interesse und die Freude an der Beschäftigung mit Naturphänomenen sowie selbstbezogene Kognitionen, z.B. Selbstkonzept und Selbstwirksamkeitserwartung. Obgleich das Interesse und die Motivation in Bezug auf die Naturwissenschaften bei jüngeren Kindern allgemein als hoch eingeschätzt werden (Lück, 2010), fehlen hierzu umfangreiche empirische Untersuchungen. Dies ist u.a. der Tatsache geschuldet, dass sich nicht-leistungsbezogene Komponenten frühkindlicher naturwissenschaftlicher Kompetenz nur schwer erfassen lassen. Im Gegensatz zu den Wissensaspekten lassen sich emotionale Kompetenzaspekte sowie selbstbezogene Kognitionen nicht durch Testverhalten erschließen, sondern erfordern eine differenzierte Selbsteinschätzung der Kinder. Bisher ist jedoch kaum erforscht, ob Kinder im Vorschulalter in der Lage sind, differenziert über ihre emotionalen Einstellungen sowie Fähigkeitsselbstkonzept im Umgang mit Naturphänomen Auskunft zu geben. Zudem ist unklar, ob Kinder bereits in diesem jungen Alter domänenspezifische Interessen und selbstbezogenen Kognitionen besitzen. Vor diesem Hintergrund stellt sich die Frage nach den Möglichkeiten und Grenzen empirischer Erfassung und Beschreibung nicht-leistungsbezogener Aspekte frühkindlicher naturwissenschaftlicher Kompetenz.

Fragestellungen: (1) Lassen sich Interesse und selbstbezogenen Kognitionen in Bezug auf den Umgang mit Naturphänomenen bei Kindern im Vorschulalter reliabel und empirisch abgrenzbar erfassen?(2) Sind Kinder in diesem Alter in der Lage über ihr Interesse und ihre Freude bei der Beschäftigung mit naturwissenschaftlichen Inhalten sowie ihre eigenen Fähigkeiten im Umgang mit Naturphänomen differenziert Auskunft zu geben?(3) Lassen sich domänenspezifische Ausprägungen von Interesse und selbstbezogenen Kognitionen bei Kindern im Vorschulalter nachweisen?

Methode: Die Entwicklung von Fragebögen zur Erfassung der nicht-leistungsbezogenen Aspekte frühkindlicher naturwissenschaftlicher Kompetenz erfolgte im Rahmen der Studie EASI-Science (Steffensky et al., 2014). In einem mehrschrittigen Prozess wurden verschiedene Itemvarianten und Antwortformate erprobt. Insgesamt wurden 89 Kinder in 20-minütigen Einzelinterviews zu ihrer Freude bei der Beschäftigung mit Naturphänomenen sowie ihrer subjektiven Einschätzung der eigenen Fähigkeiten im Umgang mit Naturphänomenen befragt. Die Antworten wurden auf einer 4-stufigen Rating-Skala kodiert (1 = geringes Interesse/Selbstkonzept bzw. Selbstwirksamkeitserwartung; 4= hohes Interesse/Selbstkonzept bzw. Selbstwirksamkeitserwartung).

Ergebnisse: (1) Ausgehend von den Ergebnissen der Befragungen konnten zwei empirisch abgrenzbare Skalen zum kindlichen Interesse sowie zu den selbstbezogenen Kognitionen mit jeweils 3 inhaltspezifischen Unterskalen (Natur allgemein, belebte und unbelebte Natur) entwickelt werden. Die Reliabilitäten der Skalen weisen zufriedenstellende bis gute Werte auf ($\alpha \geq .77$). (2) Die hohe Übereinstimmung (92%) der Antworten verschiedener Antwortformate zeigt, dass die Fragen und die Antwortmöglichkeiten von den Kindern verstanden werden. Die Häufigkeitsverteilungen der Antworten lassen zudem erkennen, dass die Kinder in der Lage sind, zwischen vier verschiedenen Antwortmöglichkeiten zu unterscheiden. (3) Die Korrelationen der Unterskalen innerhalb der Kompetenzbereiche Interesse ($.58^{**} \leq r \leq .61^{**}$) und selbstbezogene Kognitionen ($.38^{**} \leq r \leq .61^{**}$) liegen größtenteils im moderaten Bereich. Zusammenhänge zwischen Unterskalen desselben Inhaltsbereiches (belebte bzw. unbelebte Natur) fallen teilweise höher aus als innerhalb der beiden Kompetenzbereiche. So hängt bspw. das kindliche Interesse an Phänomenen der unbelebten Natur stärker mit den Selbstwirksamkeitserwartung im selben Inhaltsbereich zusammen ($r=.78^{**}$), als mit dem Interesse im Bereich der belebten Natur ($r=.61^{**}$). Dies ist ein erster Hinweis dafür, dass Kinder in Ihrem Interesse und ihren selbstbezogenen Kognitionen bereits im Vorschulalter zwischen den Inhaltsbereichen der belebten und unbelebten Natur unterscheiden können.

Becker, Michael

Psychologische und strukturelle Faktoren des Karrierewegs von Promovierenden und Promovierten

Die Rekrutierung und Förderung des wissenschaftlichen Nachwuchses zählt zu den wichtigsten und dringlichsten Aufgaben in einer Gesellschaft, „deren wirtschaftliche Prosperität und Zukunftsaussichten immer stärker von wissenschaftlicher Qualifikation, Leistung und Innovation abhängen“ (Kerst & Wolter, 2010, S. 109). Ein besonderer Stellenwert liegt dabei auf der vertieften wissenschaftlichen Weiterqualifizierung im Anschluss an das Studium im Rahmen der Promotion und einer sich daran evtl. anschließenden Habilitation, nicht zuletzt vor dem Hintergrund des sich gegenwärtig vollziehenden Generationenwechsels beim wissenschaftlichen Personal an deutschen Hochschulen. Trotz des hohen Stellenwertes, den diese Bildungsetappe in Politik, öffentlichem Diskurs und nicht zuletzt in der Wissenschaft in Hinblick auf ihren eigenen Nachwuchs besitzt, wurde der Gegenstand international wie national, insbesondere im Vergleich zu anderen Etappen der Bildungsbiographie, vergleichsweise wenig intensiv beforscht bzw. sind erst seit den letzten Jahren eine größere Zahl von Arbeiten im Entstehen (Enders & Bornmann, 2001; Kerst & Wolter, 2010). Entsprechend gibt es bislang nur eingeschränkt vertiefende Informationen zu Promovierenden und Habilitierenden, etwa wie genau sich die Gruppe der promovierenden bzw. promovierten Personen hinsichtlich individueller und sozialer Faktoren charakterisieren lässt, insbesondere im Vergleich zur Gruppe der Hochschulabsolventen ohne Promotion, oder wie Karriereplanung und Werdegang von Promovierenden und Habilitierenden erlebt und gestaltet werden.

Hier setzt das Symposium an und stellt Arbeiten zu Karriereverläufen des wissenschaftlichen Nachwuchses vor. Die Beiträge beschäftigen sich mit unter psychologischer, soziologischer und erziehungswissenschaftlicher Perspektive mit individuellen und strukturellen Faktoren der Aufnahme und des Verlaufs einer Promotion sowie der Post-Doc-Zeit. Susanne Radmann und Koautoren stellen eine Arbeit zur Bedeutung psychologischer Einflussfaktoren für die Aufnahme einer Promotion vor. Anna Bachsleitner und Koautoren beschäftigen sich mit der Frage, inwieweit die Aufnahme einer Promotion mit vorherigen Bildungsetappen seit dem Erwerb der Hochschulreife in Zusammenhang steht. Mit dem Arbeitsverhalten während und nach der Promotion ist der Beitrag von Daniel Spurr, Anneke Dubbel und Simone Kauffeld befasst. Die Autoren analysieren hierbei individuelle und institutionelle Prädiktoren zwanghaften Arbeitsverhaltens und Workaholismus des wissenschaftlichen Nachwuchses. Ramona Schürmann und Kolja Briedis präsentieren eine Arbeit, in der individuelle und strukturelle Faktoren, die für die weitere Karriereplanung im Zeitraum nach der Promotion entscheidend sind, vertiefend untersucht werden. Die Beiträge werden von Harm Kuper diskutiert.

Freitag
13. März

15:30 Uhr

HGB 30

Chair(s)

Becker, Michael

Psychologische und strukturelle Faktoren des Karrierewegs von Promovierenden und Promovierten

Radmann, Susanne; Becker, Michael; Neumann, Marko; Maaz, Kai

Die Bedeutung psychosozialer Merkmale für die Promotionsentscheidung

Mit dem Studienabschluss stehen Hochschulabsolventen vor der Entscheidung, ins Erwerbsleben überzugehen oder eine Promotion anzuschließen. Zur Sicherung wirtschaftlichen Wachstums, Wettbewerbsfähigkeit und Zukunftsaussichten sind Wissensgesellschaften auf wissenschaftliche Qualifikation und Leistung sowie auf Innovationen angewiesen. Die Rekrutierung und Förderung des wissenschaftlichen Nachwuchses stellt somit eine wichtige Aufgabe moderner Gesellschaften dar (Kerst & Wolter, 2010).

Welche Faktoren für die Promotionsentscheidung relevant sind, ist bisher vergleichsweise wenig erforscht bzw. die vorliegenden Befunde beziehen sich zumeist auf die Promotions- oder Laufbahnabsicht und nur vereinzelt auf die tatsächliche Promotionsentscheidung. Ein Schwerpunkt lag dabei bislang vor allem auf der Untersuchung sozialer oder geschlechtsbezogener Disparitäten (vgl. z.B. Jaksztat, 2014), während psychosoziale und motivationale Faktoren in den meisten Untersuchungen unberücksichtigt blieben.

Eine prominente Rolle in der Untersuchung von Bildungsentscheidungen nehmen Wert-Erwartungstheorien (Eccles et al., 1983) ein. Dabei steht die Wert-Komponente für den persönlichen Wert bzw. Nutzen, der mit einer Bildungsentscheidung verbunden wird. Der Wert eines Bildungsganges wird häufig über Bildungsaspirationen, der Nutzen unter anderem über motivationale Orientierungen erfasst. Motivationale Orientierungen sind beispielweise die Möglichkeit, seinen Fähigkeiten und Interessen nachzugehen oder die Verbesserung beruflicher Chancen. Die Erwartungs-Komponente stellt die Einschätzung der Realisierungschancen dar, die sich über Selbstwirksamkeitserwartungen oder auch Fähigkeitsselbstkonzepte operationalisieren lässt. Befunde vorangehender Studien konnten bereits Hinweise auf die Relevanz derartiger psychologischer Merkmale für die Promotions- bzw. Laufbahnintention vorlegen (Abele & Krüsken, 2003; Briedis, Jaksztat, Preßler, Schürmann & Schwarzer, 2014).

Von diesen Befunden ausgehend geht die vorliegende Untersuchung der Frage nach, inwieweit sich Einflüsse psychosozialer Faktoren für die tatsächliche Entscheidung zur Aufnahme einer Promotion feststellen lassen. Es wird davon ausgegangen, dass sich die Promotionsentscheidung sowohl durch Wert- als auch die Erwartungskomponenten vorhersagen lässt. Die Erwartungs-Komponente wird über Selbstwirksamkeitserwartungen und das Begabungsselbstkonzept während des Studiums operationalisiert. Die Wert-Komponente wird über vier motivationale Orientierungen abgebildet, die für unterschiedliche Nutzenfacetten stehen: Der Möglichkeit, sich mit persönlich interessanten Inhalten auseinanderzusetzen zu können (intrinsische Motivation), gute Leistungen zeigen zu können (Leistungsorientierung), zu den Besten gehören zu wollen (Wettbewerbsorientierung) und die Berufschancen zu verbessern (Berufsorientierung).

Die Grundlage der Untersuchung stellen 2215 Hochschulabsolventen der Längsschnittuntersuchung Bildungsverläufe und psychosoziale Entwicklung im Jugend- und jungen Erwachsenenalter (BIJU) dar, von denen 363 Personen im Jahr 2010 eine Promotion begonnen hatten. Für die Hochschulabsolventen liegen Daten zu psychosozialen, familiären, leistungsbezogenen und strukturellen Merkmalen von der Sekundarstufe I bis zum Alter von rund 30 Jahren vor. Anhand logistischer Regressionen wurde der Einfluss motivationaler Orientierungen (Wert-Komponente) sowie Selbstwirksamkeitserwartungen und Fähigkeitsselbsteinschätzungen (Erwartungs-Komponente) auf die Promotionsentscheidung untersucht.

In bivariaten Analysen zeigte sich für die Selbstwirksamkeitserwartungen erwartungswidrig kein Zusammenhang mit der Promotionsentscheidung. Bei simultaner Betrachtung der motivationalen Orientierungen, Selbstwirksamkeitserwartungen und des Selbstkonzeptes, sagten ein hohe Ausprägung des Selbstkonzeptes, eine ausgeprägte intrinsische und wettbewerbsorientierte Motivation einhergehend mit einer geringen Berufsorientierung die Promotionsentscheidung vorher.

Die Untersuchung weist darauf hin, dass psychosoziale Merkmale im Entscheidungsprozess für oder gegen einen Promotionsbeginn eine bedeutende Rolle spielen. Sie erweitert die bisherigen Befunde zur Relevanz psychosozialer Merkmale für die Promotionsabsicht. In einem weiteren Schritt sollen leistungsbezogene Merkmale sowie individuelle Hintergrundmerkmale miteinbezogen und in ihrem Zusammenspiel mit den psychosozialen und motivationalen Merkmalen untersucht werden.

Bachsleitner, Anna; Becker, Michael; Neumann, Marko; Maaz, Kai

Bildungsverläufe von Promovierenden und Hochschulabsolventen ab dem Abitur im Vergleich

Forschung zu Bildungsverläufen und damit verbundenen Ungleichheiten hat in der Bildungssoziologie in den letzten Jahrzehnten verstärkt an Relevanz gewonnen. Während sich der Großteil der Arbeiten auf den schulischen Bereich und den Übergang ins Hochschulstudium bezieht, gibt es bisher vergleichsweise wenige, insbesondere längsschnittliche Untersuchungen über den späteren Bildungsweg bis zur Promotion (Enders/Bornmann, 2001), vor allem zur Frage, inwiefern spätere Promovierende nach dem Abitur einen anderen Bildungsweg aufweisen als andere Hochschulabsolventen. Unterschiede im nachschulischen Verlauf wären einerseits zu erwarten, da Personen mit frühzeitiger Promotionsintention in ihrem Bildungsweg gradliniger sein könnten und daher einen schnelleren Bildungsweg aufweisen. Durch einen bereits längeren Bildungsweg, z.B. in Folge einer absolvierten Berufsausbildung, könnte eine weitere mehrjährige Investition in Bildung in Zusammenhang mit einer Promotion stärker gescheut werden (Jaksztat, 2014). Andererseits ist denkbar, dass bestimmte Aspekte des Bildungsweges, wie die Studiendauer oder ein Auslandsaufenthalt, für Arbeitgeber bzw. betreuende Hochschullehrer neben der Studienabschlussnote als Signal dafür dienen könnten, dass eine Person leistungsfähig und daher für eine Promotion besonders geeignet wäre (Signaling-Theorie, Spence, 1973) und somit auch ein Stellenangebot zur Promotion begünstigen könnten.

Informationen über ausgewählte Abschnitte des Bildungsweges lassen sich vor allem auf Grundlage von Absolventenbefragungen entnehmen, die zeigen, dass Promovierende im Vergleich zu nichtpromovierenden Hochschulabsolventen bessere Abitur- und Studienabschlussnoten haben und zudem gradlinigere Bildungsbiographien aufweisen (u.a. Hauss et al., 2012). Promovierende haben außerdem seltener als andere Hochschulabsolventen vor dem Studium bereits eine Berufsausbildung absolviert (Lewin et al., 1996). Trotz einiger bereits vorliegender Arbeiten ist der bisherige Forschungsstand insgesamt betrachtet jedoch nach wie vor lückenhaft. Hier setzt der vorliegende Beitrag an und betrachtet die Bildungsverläufe ab dem Abitur für Hochschulabsolventen mit und ohne spätere Aufnahme einer Promotion. Dabei wird untersucht, ob sich die beiden Gruppen in ihrem nachschulischen Bildungsweg unterscheiden und welche Aspekte des Bildungsweges eine Promotion voraussagen können. Zudem wird geprüft, inwieweit etwaige Effekte auch bei Berücksichtigung der Studiennote bestehen bleiben, was als ein Hinweis für den Signalwert von Verlaufsetappen gesehen werden könnte.

Die Forschungsfrage wird auf Grundlage der BIJU-Daten bearbeitet, einer Längsschnittstudie, die die Bildungsverläufe und Entwicklung von Jugendlichen von der Sekundarstufe I bis zum Alter von knapp 30 Jahren untersucht hat. Die Analysen beziehen sich auf N=2.004 Personen, die ein Studium abgeschlossen haben, davon 1.726 Hochschulabsolventen ohne begonnene Promotion und 278 Personen mit begonnener und zum Teil bereits abgeschlossener Promotion, für die Angaben zum Verlauf der Studien- und Promotionszeit vorlagen.

Die Ergebnisse deskriptiver Analysen zeigten wie vermutet, dass sich Hochschulabsolventen ohne Promotion und Promovierende bzw. Promovierte signifikant in ihrer Studiendauer, einer dem Studium vorgelagerten Berufsausbildung sowie längeren Auslandsaufenthalten unterscheiden. Während die Unterschiede in der Studiendauer zwischen Hochschulabsolventen ohne Promotion und Promovierenden geringer ausfielen, war der Anteil an Personen mit Berufsausbildung unter den Hochschulabsolventen fast dreimal so hoch wie bei den Promovierenden. Von den Promovierenden hatte mehr als jede zweite Person mindestens drei Monate im Ausland verbracht, bei den nichtpromovierenden Hochschulabsolventen waren es nur etwa ein Drittel der Personen. Bei der multivariaten Vorhersage einer Promotionsaufnahme mittels logistischer Regression zeigte sich, dass eine Berufsausbildung und die Studiendauer in einem signifikant negativen Zusammenhang mit einer Promotionsaufnahme standen, während ein Auslandsaufenthalt positiv darauf wirkte. Studiendauer, Auslandsaufenthalt und Berufsausbildung hingen zudem mit der Studiennote zusammen; die Effekte der einzelnen Verlaufsetappen blieben jedoch auch nach Aufnahme der Studienleistung bedeutsam. Von einer schlechteren Studiennote ging wie erwartet ein negativer Effekt auf eine Promotionsaufnahme aus.

Spurk, Daniel; Dubbel, Anneke; Kauffeld, Simone

Zwanghafte und exzessive Arbeit trotz oder gerade wegen einer unsicheren Laufbahn? Prädiktoren von Workaholismus bei Nachwuchswissenschaftlern

Insbesondere der wissenschaftliche Nachwuchs in Deutschland ist mit unsicheren Laufbahnperspektiven konfrontiert. Befristete Verträge mit kurzen Laufzeiten, Kürzungen von Mitteln und die relativ geringe Chance auf eine unbefristete Professorenstelle machen eine erfolgreiche Laufbahn in der Wissenschaft immer ungewisser (BUWIN 2013; Wissenschaftsrat 2014). Wie reagieren Nachwuchswissenschaftler/innen auf diesen Umstand? Kann diese Unsicherheit zu exzessiven um zwanghaften Arbeiten bzw. einer „Arbeitssucht“ führen? Die vorliegende Studie untersucht den Zusammenhang zwischen psychologischen Laufbahnvariablen mit dem Phänomen Workaholismus bei Nachwuchswissenschaftlern. Ziel ist die Überprüfung eines theoretischen Modells, in welchem angenommen wird, dass persönliche (z.B. Laufbahnunsicherheit, extrinsische Laufbahnziele, Laufbahncommitment) und kontextuelle Variablen (z.B. Laufbahnhindernisse, wahrgenommene Unterstützung durch die Organisation) mit zwanghaftem und exzessivem Arbeiten in Verbindung stehen. Zudem wird die Ausprägung von Laufbahnunsicherheit und Laufbahnhindernissen in verschiedenen Gruppen von Nachwuchswissenschaftlern untersucht.

Die Daten wurden in einer groß angelegten Befragung gewonnen, welche im Rahmen eines Forschungsprojektes zur Laufbahnentwicklung des wissenschaftlichen Nachwuchses in Deutschland durchgeführt wurde. Die Stichprobe besteht ca. 550 Nachwuchswissenschaftler/innen, die an deutschen Universitäten und Forschungsinstituten tätig sind. Sie umfasst Doktorand/innen und Post-Docs sowohl aus unterschiedlichen MINT-Fachbereichen als auch den Sozial-, Geistes- und Wirtschaftswissenschaften. Die Variablen wurden einmalig mit etablierten Skalen erhoben und die internen Konsistenzen lagen alle im angemessenen Bereich. Die Modellannahmen wurden mittels einer multiplen hierarchischen Regressionsanalyse überprüft.

Nach statistischer Kontrolle mehrerer Variablen, die zudem Workaholismus beeinflussen könnten, zeigte sich, dass Laufbahnunsicherheit, extrinsische Laufbahnziele, Laufbahncommitment und Laufbahnhindernisse positiv mit Workaholismus assoziiert sind. Die wahrgenommene Unterstützung durch die Organisation hängt negativ mit Workaholismus zusammen. Mögliche Unterschiede zwischen Fachdisziplinen, Karrierestufen und Geschlecht sowohl bei Workaholismus als auch bei einzelnen Prädiktoren werden noch genauer untersucht und berichtet.

Die Studie bietet einen ersten Überblick darüber, wie persönliche und kontextuelle Laufbahnvariablen mit unterschiedlichen Komponenten von Workaholismus in Beziehung stehen. Da Workaholismus eher mit negativen Konsequenzen wie Burnout oder physischen Beschwerden verbunden ist (Clark, Michel, Zhdanova, Pui, & Baltes, 2014), ist es wichtig Ursachen für dieses Verhalten zu finden. Die Ergebnisse können somit wichtige Hinweise zur Reduktion von Workaholismus liefern und im Rahmen der persönlichen Laufbahnplanung von Nachwuchswissenschaftler/innen, innerhalb der Konzeption von wissenschaftlichen Laufbahnmodellen und für die Laufbahnberatung verwertet werden.

Freitag
13. März

15:30 Uhr

HGB 30

Chair(s)

Becker, Michael

Psychologische und strukturelle Faktoren des Karrierewegs von Promovierenden und Promovierten

Schürmann, Ramona; Briedis, Kolja

Laufbahnentscheidungen und Verbleib der Promovierten

Insgesamt ist die Promotion gesellschaftlich hoch anerkannt, wenngleich sie für die jeweilige Laufbahn von unterschiedlicher Bedeutung ist. Für eine wissenschaftliche Hochschullaufbahn ist die Promotion notwendige Voraussetzung, für eine Laufbahn außerhalb der akademischen Forschung und Lehre kann die Promotion in bestimmten Fächern Türen öffnen.

Warum entscheidet sich der wissenschaftliche Nachwuchs für die akademische Wissenschaft oder für alternative Karrierewege? Welche Karrierewege schlagen Promovierte ein, wenn sie sich gegen die akademische Wissenschaft entscheiden? Zu welchem Zeitpunkt im Lebensverlauf werden die Laufbahnentscheidungen getroffen? Zwar geben 60-80 Prozent der Promovierenden an, auch künftig eine wissenschaftliche Tätigkeit anzustreben (Briedis et al. 2014), doch konnten bisherige Studien zeigen, dass der Wunsch in der wissenschaftlichen Forschung und Lehre zu verbleiben, nur für einen Teil des wissenschaftlichen Nachwuchses realisierbar ist (Enders/Bornmann 2001, dazu auch Enders/Kottmann 2009). Insbesondere für die Frage des „akademischen Frauensterbens“ sind die Gründe für das Aussteigen aus der akademischen Wissenschaft von hohem Interesse (vgl. Heintz et al. 2004; Leemann 2002; von Stebut 2003). Die Geschlechterschere geht mit steigendem Qualifikationsniveau zunehmend auseinander und zeigt sich gerade an dem Übergang von der Promotion in die Postdoc-Phase (Kahlert 2011).

Auf der Grundlage des Absolventen- und des WiNbus-Panels, durchgeführt vom Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW), werden die Gründe vorgestellt, die die Laufbahnentscheidungen der Promovierten beeinflussen. Darüber hinaus wird ein umfassendes Bild der Promovierten zehn Jahre nach Abschluss des Studiums dargestellt (wo und in welchen Beschäftigungsverhältnissen sind die Promovierten überwiegend). Neben motivationalen Faktoren zeigen sich deutliche Effekte für die Promotionsfächer, die die Laufbahnentscheidungen von Promovierten maßgeblich beeinflusst haben.

Freitag
13. März

15:30 Uhr

HGB 30

Chair(s)

Becker, Michael

Psychologische und strukturelle Faktoren des Karrierewegs von Promovierenden und Promovierten

Fuchs, Gesine

Heterogenität standardisierter Testverfahren im Kontext der KMK-Gesamtstrategie zum Bildungsmonitoring: Validität. Wert. Schätzen.

Mit der von der Kultusministerkonferenz beschlossenen Gesamtstrategie zum Bildungsmonitoring (KMK, 2006) wird neben der systematischen Erfassung von Informationen über das Bildungssystem im nationalen und internationalen Vergleich ebenso die Verknüpfung dieser Informationen mit Maßnahmen der Qualitätsentwicklung an Einzelschulen angestrebt. Zur Erfassung dieser Informationen werden Ergebnisse von Schüler/-innen bei international standardisierten Schulleistungstests (IGLU/PIRLS, TIMSS, PISA) sowie bei nationalen bildungsstandardbasierten Tests (IQB-Ländervergleich, Vergleichsarbeiten) herangezogen. Da weitreichende Entscheidungen auf Grundlage dieser Informationen getroffen werden, stellt sich eine zentrale Frage: Wie ist die Validität dieser Testverfahren zu bewerten?

Dieses Symposium nähert sich der Beantwortung dieser Frage in Form von vier Einzelbeiträgen, die im interdisziplinären Schnittbereich von Psychologie, Pädagogik, Fachdidaktik und empirischer Bildungsforschung liegen. Im Einklang mit den theoretischen Überlegungen zum Validitäts- bzw. Validierungsbegriff nach Kane (2013) bzw. Messick (1989), sehen wir hierbei die intendierte Interpretation und Verwendung der Testwerte als zentrale Validitätsaspekte an. Die Beiträge fokussieren auf unterschiedliche Testverfahren der KMK-Gesamtstrategie und den damit einhergehenden heterogenen Testwertinterpretationen und -verwendungen, die sich aus abweichenden Ansprüchen unterschiedlicher Nutzergruppen ergeben.

Im ersten Beitrag vergleicht Katrin Böhme exemplarisch den nationalen IQB-Ländervergleich zur Überprüfung des Erreichens der Bildungsstandards in der Primarstufe 2011 mit den internationalen Schulleistungsstudien IGLU und TIMSS 2011. Hierbei werden Gemeinsamkeiten und Unterschiede der Studien und Testverfahren aufgezeigt. Ferner wird die empirische Ähnlichkeit der Tests geprüft und es werden Schlussfolgerungen bzgl. der Testwertinterpretationen für die mathematische Kompetenz und die Lesekompetenz dargestellt.

Im Rahmen des zweiten Beitrags wird Johanna Hildebrandt auf Grundlage eines Teildatensatzes der Studie ÜBERGANG Ergebnisse aus Kovarianzanalysen zur prognostischen Validität curricular gemessener Testleistungen am Ende der Grundschulzeit für Merkmale des Bildungserfolgs zur Sekundarstufe I (Noten Ende Klasse 6, elterliche Aspiration Beginn Klasse 7) berichten und dabei auf Leistungsdaten in Mathematik (TIMSS) und Deutsch (BiSta) zurückgreifen.

Im dritten Beitrag schließen sich vergleichbare Analysen von Tanja Graf und Gesine Fuchs an mit dem Fokus auf flächendeckend eingesetzte bildungsstandardbasierte Tests (Vergleichsarbeiten=VERA). Empirische Evidenz für die Vorhersagekraft von schulischen Leistungskriterien über die Grundschulzeit für Mathematik (KEGS-Kompetenztestergebnisse, Noten) und zum Ende der Sekundarstufe I für Mathematik, Deutsch und Englisch (Noten, MSA-Prüfungsnoten) werden präsentiert.

Im vierten Beitrag wird Peter Harych mit Analysen zur Passung der Vergleichsarbeiten und der Prüfungen zum Mittleren Schulabschluss fortfahren und dabei Konstrukte wie Ergebnisse (Kompetenzstufenzuordnungen und Prüfungsnoten) gegenüberstellen. Die Datenausgangsbasis liefern direkt vor der Prüfung eingesetzte VERA-Kompetenztests und MSA-Prüfungsergebnisse. Abschließend wird auf das Dilemma hingewiesen, welches sich aus der Nutzung von Daten zur Schul- und Unterrichtsentwicklung und der Verwendung der Prognose als Validierungsmerkmal ergibt.

Freitag
13. März

15:30 Uhr

HGB 40

Chair(s)

Fuchs, Gesine
Graf, Tanja**Heterogenität standardisierter Testverfahren im Kontext der KMK-Gesamtstrategie zum Bildungsmonitoring: Validität. Wert. Schätzen.**

Böhme, Katrin; Richter, Dirk; Weirich, Sebastian

Messen wir dasselbe? Eine Gegenüberstellung des IQB-Ländervergleichs 2011 und der internationalen Studien IGLU und TIMSS 2011

In regelmäßigen Abständen werden in Deutschland nationale und internationale Schulleistungsstudien durchgeführt, um die Kompetenzstände von Schülerinnen und Schülern in ausgewählten, teilweise gleichlautenden Kompetenzbereichen zu überprüfen. Inwiefern jedoch die zur Messung eines Kompetenzbereichs eingesetzten Instrumente zwischen den Studien vergleichbar sind, war bislang nur selten Gegenstand empirischer Forschung. In diesem Beitrag gehen wir deshalb der Frage nach, inwieweit der nationale IQB-Ländervergleich 2011 und die internationalen Studien IGLU und TIMSS 2011 miteinander vergleichbar sind und prüfen dies zunächst auf konzeptioneller und formaler Ebene anhand der folgenden Kriterien: (1) Testkonstrukte, (2) Zielpopulationen, (3) Anlage und Durchführung der Studien, (4) Testinstrumente, (5) Datenanalyse sowie (6) Berichtlegung und Schlussfolgerungen. Die Ergebnisse weisen darauf hin, dass die nationalen und internationalen Studien weitgehende Übereinstimmungen in der Zielpopulation und in den Durchführungsbedingungen aufweisen. (Kleinere) Unterschiede bestehen in der Rahmenkonzeption, den Testinstrumenten, den statistischen Analysemethoden und den aus den Ergebnissen ableitbaren Schlussfolgerungen.

Im Anschluss quantifizieren wir die empirische Ähnlichkeit der jeweils gemessenen Konstrukte mittels latenter Korrelationen und prüfen ferner ob die Daten der hier untersuchten Tests besser durch ein ein- oder ein zweidimensionales Modell repräsentiert werden können. Hinsichtlich der latenten Korrelationen ergab sich zwischen der im IQB-Ländervergleich 2011 und in TIMSS 2011 gemessenen mathematischen Kompetenz ein sehr hoher Zusammenhang von $r = .93$. Dieser fiel deutlich höher aus als die entsprechende Korrelation für die Lesekompetenz zwischen dem IQB-Ländervergleich 2011 und IGLU 2011, die bei $r = .85$ lag. Berücksichtigt man für den Bereich Lesen in der Modellierung, dass sich Items innerhalb einer Leseaufgabe auf denselben Stimulus beziehen und in diesem Sinne ein Testlet darstellen, ergeben sich auch hier höhere Korrelationen zwischen den Studien von $r \geq .89$. Sowohl für die Lesekompetenz als auch für die mathematische Kompetenz werden die Daten durch ein Modell, das für den IQB-Ländervergleich einerseits und die jeweilige internationale Studie (IGLU bzw. TIMSS) andererseits jeweils eine separate Konstruktdimension annimmt (zweidimensionale Skalierung), besser repräsentiert als in einem eindimensionalen Modell.

Zusammenfassend schlussfolgern wir, dass zwischen den betrachteten Studien zwar zahlreiche Ähnlichkeiten und Überschneidungen bestehen, dass es sich bei den im IQB-Ländervergleich 2011 sowie in IGLU bzw. TIMSS 2011 eingesetzten Tests allerdings nicht um identische, gegeneinander austauschbare Instrumente handelt.

Freitag
13. März

15:30 Uhr

HGB 40

Chair(s)

Fuchs, Gesine
Graf, Tanja

Heterogenität standardisierter Testverfahren im Kontext der KMK-Gesamtstrategie zum Bildungsmonitoring: Validität. Wert. Schätzen.

Hildebrandt, Johanna; Watermann, Rainer

Prognostische Validität von curricular gemessenen Testleistungen am Ende der Grundschulzeit

Schulleistungsuntersuchungen und Bildungsstandards stellen zentrale Maßnahmen des Bildungsmonitorings in Deutschland dar. Testleistungen aus Schulleistungsuntersuchungen werden genutzt, um die Entwicklung und Leistungsfähigkeit des Bildungssystems zu analysieren. Damit ist die Annahme verbunden, dass Testleistungen Ergebnisse schulischer Lernprozesse widerspiegeln, inhaltlich valide sind und darüber hinaus prognostische Validität besitzen. Die Validität von Testleistungen wurde in der Literatur breit diskutiert und kritisiert, wenngleich verhältnismäßig geringe empirische Evidenz vorliegt. Nach Fischbach et al. (2013) lassen sich drei Kritikpunkte ausmachen: (1) Es sei nicht möglich, Testleistungen zur Vorhersage von Merkmalen des späteren Bildungserfolgs zu nutzen, weil Testleistungen in hohem Maße für die Bearbeitung von Tests spezifische Fähigkeiten und nicht wie anvisiert domänenspezifische Fähigkeiten messen würden. Durch die fehlende inhaltliche Validität könne daher auch die prognostische Validität nicht gewährleistet werden. Für diese Argumentation findet sich jedoch kaum empirische Evidenz. (2) Testleistungen erfassen Fähigkeiten und Wissen nicht domänenspezifisch, stattdessen bildeten sie eine Leistungsdimension ab, die der Intelligenz sehr ähnlich bzw. mit ihr identisch sei, folglich seien Testleistungen nicht konstruktvalide. Während Lynn und Meisenberg (2010) eine Korrelation von 1.0 zwischen Schulleistungs- und Intelligenzdaten fanden, konnte Brunner (2008) aufzeigen, dass sich die Mathematikleistungen gemessen in der Sekundarstufe nicht durch einen einzigen g-Faktor abbilden lassen. Auch Winkelmann et al. (2012) zeigen für die Bildungsstandards der Primarstufe, dass nur etwa 50% der Varianzaufklärung auf die Intelligenz, jedoch substantielle Teile auf die Kompetenz zurückzuführen sind. (3) Testleistungen seien auch deshalb nicht zur Vorhersage des Bildungserfolgs geeignet, weil sie nur den sozioökonomischen Status abbildeten. In einer Meta-Analyse von Sirin (2005) wurden korrelative Zusammenhänge von $r = .32$ und $r = .35$ zwischen sozioökonomischem Status und verbaler bzw. mathematischer Kompetenz gemessen, was jedoch auch dahingehend interpretiert werden kann, dass Schüler aus Familien mit höherer sozioökonomischer Stellung über ein positiveres Unterstützungssystem bestehend aus Aspirationen, häuslicher und elterlicher Unterstützung verfügen.

Im Vortrag wird auf der Grundlage eines Teildatensatzes der Studie ÜBERGANG (N=1624) der prognostischen Validität von am Ende der vierten Klasse der Grundschule curricular gemessenen Testleistungen in Mathematik (vgl. TIMSS, Trends in International Mathematics and Science Study) und Deutsch (vgl. BiSta, Normierung von Testaufgaben für die Bildungsstandards Deutsch und Mathematik im Primarbereich) für domänenspezifische Leistungsmaße (Noten Ende Klasse 6) und für ein domänenunspezifisches nicht-leistungsbezogenes Maß (elterliche Bildungsaspirationen Beginn Klasse 7) nachgegangen. (1) Getrennt nach Bildungsgang durchgeführte Kovarianzanalysen (mit SES, Intelligenz, Geschlecht und der Note am Ende der Klasse 4 als Kovariaten) zeigten auf, dass Testleistungen in Mathematik und Deutsch die jeweils korrespondierenden Noten in Mathematik und Deutsch Ende der 6. Klasse der Sekundarstufe I vorhersagen. Die standardisierten Koeffizienten fielen gering bis mittel aus und lagen knapp unter den Koeffizienten der als Vorwissen gemessenen Noten in der vierten Klasse. Nicht korrespondierende Noten wurden nur im Bereich Mathematik und nur bei Gymnasiasten vorhergesagt. Der Koeffizient fiel geringer aus als im Modell der korrespondierenden Noten. Dass die Befunde auch unter Kontrolle von Intelligenz und SES stabil blieben, lässt sich als Hinweis auf die Konstruktvalidität der Testleistungen interpretieren. (2) Um der Frage nach prognostischer Validität hinsichtlich domänenunspezifischer und nicht-leistungsbezogener Merkmale nachzugehen, wurden elterliche idealistische Bildungsaspirationen zu Beginn der 7. Klassenstufe vorhergesagt. Es zeigte sich über die Bildungsgänge hinweg stabile Aspirationen, die nicht durch Testleistungen vorhergesagt werden konnten. Unter Kontrolle der in der Grundschule gemessenen Noten nahm die Stabilität der Aspirationen etwas ab, blieb jedoch auch unter Kontrolle von Intelligenz und SES annähernd stabil. Die Befunde können dahingehend interpretiert werden, dass Testleistungen hinsichtlich domänenübergreifender und nicht-leistungsbezogener Merkmale eine vergleichsweise geringere Rolle spielen. Aspirationen entwickeln sich zudem vermutlich eher in Abhängigkeit der Leistungsrealität und damit der Noten. Zusammengefasst untermauern die Befunde die Konstruktvalidität wie auch die prognostische Validität curricular valide gemessener Testleistungen für domänenspezifische Merkmale des Bildungserfolgs.

Freitag
13. März

15:30 Uhr

HGB 40

Chair(s)

Fuchs, Gesine
Graf, Tanja**Heterogenität standardisierter Testverfahren im Kontext der KMK-Gesamtstrategie zum Bildungsmonitoring: Validität. Wert. Schätzen.**

Fuchs, Gesine; Graf, Tanja; Brunner, Martin

Können Ergebnisse aus Vergleichsarbeiten wichtige Bildungserträge vorhersagen? Evidenz für die prognostische Validität bildungsstandardbasierter Testverfahren in der Primar- und Sekundarstufe I

Vergleichsarbeiten (VERA), als ein zentrales Instrument im Rahmen der Gesamtstrategie der KMK zum Bildungsmonitoring, wurden vorrangig mit dem Ziel eingeführt, ein Unterrichts- und Schulentwicklungsinstrument zu sein (KMK, 2012). Diese werden jährlich bundesweit in der 3. und 8. Klasse (VERA 3 und VERA 8) durchgeführt und orientieren sich inhaltlich an den nationalen Bildungsstandards für die 4. und 9. bzw. 10. Klasse. Bildungsstandards legen fest, welches Leistungsniveau die Schülerinnen und Schüler bis zum Ende der jeweiligen Jahrgangsstufe erreichen sollen. VERA-Ergebnisse können Hinweise darüber liefern, ob bzw. inwieweit Schülerinnen und Schüler in der 3. bzw. 8. Jahrgangsstufe diese definierten Leistungsniveaus bereits erreicht haben bzw. auf welchem Stand sich diese befinden. Anhand dieser Informationen sollen Lehrkräfte Entwicklungs- und Fördermaßnahmen für ihren Unterricht ableiten. VERA wird folglich vor dem Zeitpunkt durchgeführt, zu dem Schüler/-innen die Bildungsstandards erreicht haben sollen. Demnach stellt sich die grundlegende Frage, inwiefern auf Grundlage der VERA Ergebnisse eine Vorhersage darüber ermöglicht wird, ob Schülerinnen und Schüler in nachfolgenden Schuljahren, die in den Bildungsstandards formulierten Leistungserwartungen für die 4. Jahrgangsstufe (KMK, 2013) oder den Mittleren Schulabschluss (MSA) erreichen werden (KMK, 2011).

Die Frage der Prognosegüte bildungsstandardbasierter Tests wurde bislang nur wenig empirisch untersucht. Erste Untersuchungen in Schleswig Holstein (IQSH, 2010) weisen darauf hin, dass es einen Zusammenhang von $r = -.54$ zwischen dem Lösungsanteil VERA 8-Mathematik und der Prüfungsnote beim Hauptschulabschluss gibt. Weitere Fächer und andere Schulzertifikate wurden in diesem Zusammenhang nicht untersucht. In diesem Beitrag berichten wir daher die Ergebnisse aus zwei Einzelstudien zur prognostischen Validität bildungsstandardbasierter Tests.

In der ersten Studie wurden im Rahmen einer Längsschnitterhebung an Brandenburger Grundschulen zwischen den Schuljahren 2006/07 bis 2010/11 (KEGS; $N = 1.745$ Schüler/-innen, 75 öffentliche Grundschulen) jährlich, bildungsstandardbasierte Tests der mathematischen Kompetenz über die Grundschulzeit von der 2. bis zur 6. Jahrgangsstufe durchgeführt. Die Ergebnisse aus Regressionsanalysen zeigen, dass Kompetenztestergebnisse bspw. in der 3. Jahrgangsstufe Kompetenzerwartungen für die 4. Jahrgangsstufe in Bezug auf Kompetenztestergebnisse ($\beta = .59$) sowie Schulnoten ($\beta = -.53$) voraussagen. Unter der Betrachtung weiterer Erhebungszeitpunkte zur 4. und 6. Jahrgangsstufe zeigt sich ein prognostischer Mehrwert von Kompetenztestergebnissen über Schulnoten hinaus auf künftige Schülerleistungen (Kompetenztestergebnisse; Schulnoten).

Die zweite Studie basiert auf einer Längsschnittstichprobe von Berliner Gymnasien in der Sekundarstufe I über einen Zeitraum von drei Jahren (2011 bis 2013). Erster Messzeitpunkt der repräsentativen Stichprobe ($N = 1.294$ Schüler/-innen aus 45 Klassen/12 Schulen) ist der VERA 8 Durchgang 2011 in den Fächern Mathematik, Deutsch und Englisch. Die Erhebung der Ergebnisse zum mittleren Schulabschluss (MSA) der Schüler/-innen dient als letzter Messzeitpunkt der Studie. Ergebnisse aus korrelativen Analysen weisen für den Zusammenhang der Leistung in den VERA 8 Testdomänen Mathematik, Deutsch Lesen und Englisch mit den entsprechenden MSA-Prüfungsnoten mittlere bis starke Korrelationskoeffizienten auf. Für das Fach Mathematik sind die Zusammenhänge leicht höher ($r = -.56$) als für die Domänen Deutsch und Englisch ($r = -.43$ bzw. $r = -.55$).

Zusammenfassend zeigen beide Studien, dass bildungsstandardbasierte Testverfahren zukünftige Bildungserträge substantiell vorhersagen können. Dieser Befund unterstreicht damit, dass Ergebnisse solcher Testverfahren valide Interpretationen zur Vorhersage schulischer Entwicklungen ermöglichen. Die vorliegenden Ergebnisse sind somit auch für die pädagogische Praxis relevant, da sie die Auseinandersetzung von Lehrkräften mit den Rückmeldungen der bildungsstandardbasierten Tests wie VERA intensivieren können.

Freitag
13. März

15:30 Uhr

HGB 40

Chair(s)

Fuchs, Gesine
Graf, Tanja**Heterogenität standardisierter Testverfahren im Kontext der KMK-Gesamtstrategie zum Bildungsmonitoring: Validität. Wert. Schätzen.**

Harych, Peter; Brunner, Martin

Wie wahrscheinlich verfehlt ein Schüler die Prüfung zum Mittleren Schulabschluss, wenn er die Mindeststandards der Vergleichsarbeiten nicht erreicht?

Die Länder der Bundesrepublik Deutschland haben als Konsequenz aus PISA 2000 verbindliche Bildungsstandards beschlossen. Diese sind als domänenspezifische Kompetenzerwartungen für feste Zeitpunkte der Bildungsbiographien von Schülerinnen und Schüler formuliert, so zum Beispiel für das Fach Mathematik zum Ende der Sekundarstufe 1. Im Land Berlin findet eine landesspezifische Transformationen dieser Kompetenzerwartungen ihren Niederschlag sowohl in den Rahmenplänen wie in den zentral formulierten Prüfungen zum Mittleren Schulabschluss (MSA). Die bundesweiten Vergleichsarbeiten der 8. Klasse (VERA 8) sind mit dem Anspruch angetreten, das Erreichen von Kompetenzständen zwei Jahre vor ihrer Erwartung am Ende der zehnten Jahrgangsstufe festzustellen. Es wird unterstellt, dass diese Kompetenzfeststellung konkrete Anhaltspunkte für Fördernotwendigkeiten enthalten, damit das Ziel, den MSA zu erreichen, nicht verfehlt wird und eine Prognose für die Zielerreichung darstellt.

Diese Arbeit betrachtet drei Aspekte dieser Konstruktion genauer.

(1) Kann die vom Rahmenplan abgeleitete Prüfungsanforderung inhaltlich als bildungsstandardbasiert angesehen werden? Die fachlich-inhaltliche Struktur der Mathematik-Prüfung zum MSA für das Land Berlin (2013) wird mit jener der drei VERA-Tests (2011, 12, 13) deskriptiv verglichen. Parameter des Konstrukt-Vergleichs sind die Auswahl und Gewichtung der Aufgaben bezüglich der Anforderungsbereiche, der Kompetenzen, Leitideen und Aufgabenformate sowie die Kodierungsanforderungen. Da selbst unterschiedliche theoretische Konzepte bei der Messung mathematischer Kompetenz sehr ähnliche Ergebnisse erzielen, scheinen die Desiderata der Prüfungsarbeiten vernachlässigbar.

(2) Es wird untersucht, ob eine Rückmeldung auf Basis von Kompetenzstufen als Ergebnis der Vergleichsarbeiten ein valides Instrument für eine Prognose des Prüfungsergebnisses zum MSA darstellt bzw. hilfreiche Hinweise für eine angemessene Förderung in diesem Sinne bietet. Grundlage für Analysen ist die letzte Erhebungswelle des Projektes VERAMSA. Hier wurde über 1.000 Schüler/innen an Berliner Gymnasien kurz vor den MSA-Prüfungen (2013) je einer von drei unterschiedlichen VERA-Tests für das Fach Mathematik vorgelegt. Da zwischen VERA-Test und MSA-Prüfung nur wenige Wochen liegen, wird hier tatsächlich die Äquivalenz überprüft. Der Grad der Übereinstimmung beider Messungen stellt eine obere Schranke für die Qualität von Prognose wie Förderung dar, wie sie sich aus einem realen VERA-8-Durchgang ergibt.

Betrachtet wird die Qualität der Passung der gemessenen Kompetenzerreichung als Stufe im Kompetenzstufenmodell zum Ergebnis der Prüfung als Note. Es ist die Hypothese zu prüfen, inwieweit ein Verfehlen der Mindeststandards beim VERA-Test einem Nicht-Bestehen der Prüfung entspricht. Dies ist die einzige Hypothese, welche sich ohne Interpretation aus den beschriebenen Anforderungen klar ergibt. Für weitere Hypothesen gibt es lediglich Anhaltspunkte: Die Erfüllung der Regelstandards (KSM III) geht im Allgemeinen einher mit einer Abschlussnote von 3 oder besser. Eine Prüfungsnote 1 erhält auf jeden Fall, wer die Regelstandards übererfüllt (KSM IV oder V).

(3) Abschließend wird diskutiert, inwieweit die intendierte Interpretation der Differenz zwischen Kompetenzerwartung und Kompetenzstand relevant für die Einschätzung der Validität ist.

Wenn eine Differenz zwischen Kompetenzerwartung und Kompetenzstand für Unterrichtsentwicklung als Rückkopplung in einem Regelkreislauf verwendet wird, welcher darauf abzielt, diese Differenz zu verringern, taugt diese festgestellte Differenz nur so lange gleichzeitig für eine Prognose, wie wir eine nicht funktionierende Steuerung unterstellen. Die identische Aussage ist valide nur entweder als Steuergröße in einem formativen Assessment oder als Prognose in einem summativen Verfahren.

Freitag
13. März

15:30 Uhr

HGB 40

Horz, Holger

„Forschungsprojekte in Anknopplung an Large-Scale-Assessments – Aktuelle Projekte und ihre Theorien, Methoden und zentralen Befunde“

Large-Scale-Assessments (LSAs) haben die Heterogenität der Schülerschaft in das Blickfeld der Öffentlichkeit gerückt (z.B. „PISA-Schock“ mit anschließender Diskussion um Kompetenzerwerbsunterschiede in Abhängigkeit von sozialer Herkunft). Sie bilden durch Bildungsmonitoringprozesse erhobene Kompetenzstände samt den schulischen Rahmenbedingungen ab. Durch die systematische Verknüpfung von Schülerleistungen mit Schüler- und Unterrichtsmerkmalen können wertvolle Erkenntnisse für Forschung, Politik und Praxis abgeleitet werden. Das interdisziplinäre Symposium (Erziehungswissenschaft, Psychologie, Fachdidaktiken) verknüpft vier Beiträge aus Projekten, die im Rahmen der BMBF-Initiative zur „Förderung von Forschungsvorhaben in Anknopplung an Large-Scale Assessments“ durchgeführt werden. Es werden theoretische Hintergründe, Methoden und zentrale Ergebnisse der Forschungsprojekte aus verschiedenen schulrelevanten Inhaltsgebieten berichtet und vergleichend diskutiert.

In dem ersten Beitrag untersuchen McElvany, Schwabe, Gebauer und Bos LSA-Lesekompetenztests daraufhin, inwieweit sich die Heterogenität der Schülerschaft (soziale Lage, Migrationshintergrund, Geschlecht, Schulform) auch durch relative Stärken und Schwächen in Testitems äußert, die als Unfairness von Aufgaben interpretiert werden können. Vorgestellt werden DIF-Analysen zu den Itemmerkmalen Antwortformat, Leseprozess und Textgattung unterschiedlicher Wellen der IGLU- und PISA-Tests in Deutschland.

Der zweite Beitrag von Greiff, Funke und Goldhammer fokussiert die Kompetenzentwicklung zum komplexen Problemlösen, die als wichtiger Bestandteil in die internationale PISA-Erhebung 2012 integriert war. Vorgestellt werden zentrale Ergebnisse in Bezug auf (1) die Weiterentwicklung und Verfeinerung von Messinstrumenten zur Erfassung der komplexen Problemlösefähigkeit, (2) die empirische Validierung im Sinne konvergenter und divergenter Validität und (3) die Rolle und Relevanz für curricular gebundene Kompetenzen am Beispiel der naturwissenschaftlichen Erkenntnisgewinnung.

Im dritten Beitrag von Horz, Goldhammer, Frey und Naumann wird die Entwicklung und Validierung eines computerbasierten, adaptiven und verhaltensnahen Kompetenztests zu ICT-Skills vorgestellt, der auch eine änderungssensitive Individualdiagnostik ermöglicht. Berichtet werden die Entwicklung des theoriebasierten Assessmentframework, der Aufgabenentwicklungsprozess bis hin zur quantitativen Überprüfung des Frameworks und der Analyse ausgewählter Zusammenhänge mit relevanten Drittvariablen.

Im vierten Beitrag thematisieren Trautwein, Körber, Schreiber und das HITCH-Projektteam die Messung historischer Kompetenzen mit weitgehend geschlossenen Aufgaben in Large-Scale Assessments. In dem Vortrag werden zentrale Befunde von großen Erhebungen aus den Jahren 2013 und 2014 vorgestellt, an denen Schüler/innen der 9. Jahrgangsstufe aller Schulformen in Deutschland, Schweiz und Österreich teilgenommen haben. Die Analysen prüfen, ob die theoretisch postulierten Dimensionen (historische Frage-, Methoden-, Orientierungs- und Sachkompetenz) empirisch modellierbare Subfacetten darstellen.

Abschließend diskutiert Leutner die Beiträge des Symposiums in Bezug auf den größeren Rahmen des Gesamtprogramms und dessen Intention. Das Symposium leistet somit einen wichtigen Beitrag, die Rolle von Large-Scale-Assessments in ihrer Funktion, ihrem Mehrwert und ihren Grenzen im Zusammenspiel von Bildungsmonitoring, Individual-Diagnostik und Interventionsforschung kritisch zu reflektieren.

Chair(s)

McElvany, Nele
Horz, Holger**„Forschungsprojekte in Anknopplung an Large-Scale-Assessments – Aktuelle Projekte und ihre Theorien, Methoden und zentralen Befunde“**

McElvany, Nele; Schwabe, Franziska; Gebauer, Miriam; Bos, Winfried

Leseleistungsdifferenzen ausgewählter Schülersubgruppen in Abhängigkeit von Itemmerkmalen bei IGLU und PISA - Projekt „Testfairness“

Ein zentrales Ergebnis nationaler und internationaler Large-Scale-Assessments (LSA) ist, dass der Kompetenzerwerb im deutschen Bildungssystem maßgeblich von schülerseitigen Merkmalen beeinflusst wird. Relevante Merkmale, für die absolute Leistungsunterschiede im Bereich Lesen sowohl am Ende der Grundschulzeit als auch am Ende der Pflichtschulzeit berichtet werden, sind die soziale Herkunft, die Familiensprache bzw. der familiäre Migrationshintergrund, das Geschlecht und die besuchte Schulform (u.a. Autorengruppe Bildungsberichterstattung, 2012; Prenzel et al., 2013). Für die Differenzen zwischen erworbenen Kompetenzen verschiedener Schülersubgruppen kann angenommen werden, dass sich diese nicht nur absolut von denen der jeweiligen Vergleichsgruppe unterscheiden, sondern dass auch differenzielle Leistungsprofile der Subgruppen bestehen (Haag et al., 2013; McElvany & Schwabe, 2013; Schwabe, McElvany, Trendtel, Gebauer & Bos, 2014). Um spezifische Stärken und Schwächen zu identifizieren, bietet sich die systematische Analyse der Effekte von bestimmten Itemmerkmalen an (zB. international: Taylor & Lee, 2012).

Vor diesem Hintergrund ist das übergeordnete Ziel dieses Beitrags, die systematische Analysen zu spezifischen Stärken und Schwächen im Bereich Lesen in Abhängigkeit von Schüler- und Itemmerkmalen, die im Rahmen eines drei-jährigen BMBF-geförderten Projekts entstanden sind, in einer Zusammenschau darzustellen. Konkret, werden die Fragen beantwortet, inwieweit spezifische Stärken und Schwächen für zentrale Schülersubgruppen nach (a) der sozialen Lage, (b) dem sprachlicher Hintergrund, (c) dem Geschlecht und (d) der Schulformzugehörigkeit in Abhängigkeit von den Itemmerkmalen (i) Antwortformat, (ii) Leseprozess und (iii) Textgattung bestehen.

Zur Untersuchung der Forschungsfragen wurden die deutschen Teilstichproben der IGLU-Erhebung 2001 und 2011 sowie der PISA-2009-Studie verwendet. Somit werden jeweils mehrere tausend Schülerinnen und Schüler der vierten Klassenstufe (IGLU) bzw. im Alter von 15 Jahren (PISA) für die Analysen herangezogen. Die Identifikation der Schülersubgruppen erfolgte anhand der Angaben im Schülerfragebogen, die Kategorisierung der Items anhand der Angaben im Datensatz. Als Analysestrategie für die Identifikation einer möglichen Abhängigkeit zwischen Itemperformanz und Schülersubgruppenzugehörigkeit bei eigentlich gleichen Lesekompetenzleistungen der Lernenden wurden Differential-Item-Functioning-Analysen, insbesondere item-subset-Berechnungen mit GLM-Modellen nach De Boeck et al. (2011) in R mit dem Paket TAM (Kiefer, Robitzsch & Wu, 2013) eingesetzt.

Zentrale Ergebnisse sind, dass hypothesenkonform bei Aufgaben mit offenem Antwortformat für Kinder sozial schwächerer Lage, Kinder mit Migrationshintergrund, Jungen und Kinder aus Schulformen ohne Abituroption eine spezifische Schwäche gefunden wurde. Dieses Befundmuster wurde durch ergänzende Analysen zur Rolle der Motivation vertiefend untersucht, die verdeutlichten, dass Kinder mit niedriger Lesemotivation spezifische Schwächen bei offenen Antwortformaten aufweisen. Für die Itemmerkmale eines geforderten anspruchsvollen Leseprozesses und der Textgattung Informativer Text konnten im Rahmen des Projekts bisher hingegen kein bzw. nur vereinzelt Effekte für das relative Abschneiden von Schülergruppen identifiziert werden.

Das Ergebnismuster wird mit Blick auf die Konstrukt-Relevanz der unterschiedlichen Itemmerkmale diskutiert. Der Beitrag liefert damit einen konkrete Hinweis dazu, welchen Grenzen die Interpretierbarkeit der Ergebnisse, die auf in heterogenen Stichproben gewonnen Daten aus Large-Scale Assessments basieren, unterliegt. Ausgehend von der überblicksartigen Zusammenschau der Ergebnisse aus unterschiedlichen Stichproben unterschiedlichen Alters werden mögliche Entwicklungsbedarfe hinsichtlich der Testkonstruktion und -interpretation unter Aspekten der Testfairness vorgeschlagen und mit Blick auf die Realisierbarkeit im LSA-Kontext diskutiert. Darüber hinaus werden aus den beobachteten subgruppenspezifischen Stärken und Schwächen in der Bearbeitung von unterschiedlichen Aufgabenstrukturen und -inhalten Implikationen zur gezielten Förderung abgeleitet.

Freitag
13. März

15:30 Uhr

HGB 50

Chair(s)

McElvany, Nele
Horz, Holger

„Forschungsprojekte in Anknüpfung an Large-Scale-Assessments –
Aktuelle Projekte und ihre Theorien, Methoden und zentralen Befunde“

Greiff, Samuel; Funke, Joachim; Goldhammer, Frank

Komplexes Problemlösen und seine Rolle in internationalen Large-Scale Untersuchungen: Ergebnisse eines Projektes zu fachübergreifenden Kompetenzen

Ausgehend von der zunehmenden wissenschaftlichen und gesellschaftlichen Relevanz fächerübergreifender Kompetenzen wurde die komplexe Problemlösefähigkeit als integraler Bestandteil in die internationale PISA Erhebung 2012 integriert: Über 40 der insgesamt etwa 70 teilnehmenden Länder wählten die Option einer computer-administrierten Erhebung der komplexen Problemlösekompetenz ihrer Schülerstichproben. Die von deutschen Forschern entwickelten Problemlöseverfahren basieren dabei auf Dörners Theorie der operativen Intelligenz und erlauben erstmalig eine psychometrisch angemessene Erfassung des Konstruktes der komplexen Problemlösekompetenz, die zugleich den Desiderata von internationalen Large-Scale Erhebungen Rechnung trägt.

In dem vorliegenden Beitrag wird ein Forschungsprojekt im Rahmen der BMBF-Initiative zur Förderung von Forschungsvorhaben in Anknüpfung an Large-Scale Assessments vorgestellt und in seinen Implikationen für zukünftige PISA-Zyklen im Speziellen und Large-Scale Assessments im Allgemeinen erörtert. Das Projekt gliederte sich in drei distinkte Arbeitsphasen, die wechselseitig aufeinander bezogen waren: (1) Weiterentwicklung und Verfeinerung von Messinstrumenten zur Erfassung der komplexen Problemlösefähigkeit unter Berücksichtigung der spezifischen Desiderata von Large-Scale Erhebungen; (2) Empirische Validierung im Sinne konvergenter und divergenter Validität der komplexen Problemlösefähigkeit zu anderen erziehungswissenschaftlich und psychologisch bedeutsamen Konstrukten; (3) Erforschung der Rolle und Relevanz der komplexen Problemlösefähigkeit für curricular gebundene Kompetenzen am Beispiel der naturwissenschaftliche Erkenntnisgewinnung.

Dieser Beitrag berichtet überblicksartig zentrale Ergebnisse des genannten Forschungsprojektes und zeigt dabei auf, welcher Mehrwert einerseits mit einer Inklusion fachübergreifender Kompetenzen in Studien wie PISA assoziiert ist und welche Grenzen andererseits mit einer solchen Erweiterung verbunden sind. Im Einzelnen (1) wird der in PISA 2012 verfolgte Messansatz mit klassischen Aufgaben zur komplexen Problemlösefähigkeit wie sie sich in der Forschung finden verglichen. Inhaltlich sinnvolle und psychometrisch notwendige Anpassungen an den Large-Scale Kontext werden kritisch vor dem Hintergrund des kognitionspsychologischen Theoriendstandes sowie des PISA Frameworks diskutiert und eingehend erörtert. Eine (2) Reihe von Validierungsstudien, einige davon direkt im Rahmen des Projektes durchgeführt, setzen komplexes Problemlösen in Beziehung zu anderen Konstrukten, die aus der psychologischen Forschung bekannt sind, wie etwa induktives Denken, Arbeitsgedächtnis, oder figural-numerische Fähigkeiten. Diese Studien belegen, dass die genannten fundamentalen kognitiven Fähigkeiten mit der komplexen Problemlösefähigkeit korreliert, keinesfalls aber identisch sind. Als bildungswissenschaftlich (3) zentral ist die Frage nach der Bedeutsamkeit der cross-curricularen Kompetenzen für spezifische Unterrichtsfächer anzusehen. Komplexes Problemlösen kann dabei als systematische Suche im Problemraum nach geeigneten Lösungsstrategien mit einer besonderen Affinität zu naturwissenschaftlichen Frage- und Problemstellungen verstanden werden. Vor diesem Hintergrund wurden Anforderungen für komplexes Problemlösen (fachübergreifend) und naturwissenschaftliche Erkenntnisgewinnung (fachbezogen), die sich als relevante schulische Inhalte in den deutschen Bildungsstandards wiederfinden, konzeptuell miteinander verglichen. Anschließend wurden komplexe Problemlöseaufgaben derart modifiziert, dass sie mit naturwissenschaftlichen Inhalten angereichert und somit als fachspezifische Problemstellungen verstanden werden können. In diesem Zusammenhang konnten erste Erkenntnisse über die wechselseitigen Beziehungen zwischen fachspezifischen (im Bereich der naturwissenschaftlichen Erkenntnisgewinnung) und fachübergreifenden Problemlösefähigkeiten gewonnen werden.

Neben einer kritische Reflektion des Problemlösekonzeptes aus der PISA 2012 Studie in enger Anlehnung an empirische Befunde aus der Kognitionspsychologie und der empirischen Bildungsforschung, wird als Ausblick der potentielle Nutzen von Prozessdaten im Rahmen computer-basierter Testungen (wie etwa beim komplexen Problemlösen) angedeutet. Spezifische Verhaltensweisen in der Testsituation, die in computer-generierten Logfiles gespeichert werden, erlauben es, das Blickfeld von einer rein deskriptiven Diagnostik individueller sowie nationaler Differenzen auf eine Diagnostik des Zustandekommens von Leistungsunterschieden (etwa durch verhaltensbezogene Analysen von Motivation, Strategiewissen, Explorationsverhalten, Selbstregulation) zu erweitern.

Freitag
13. März

15:30 Uhr

HGB 50

Horz, Holger; Frey, Andreas; Goldhammer, Frank; Naumann, Johannes; Engelhardt, Lena; Hartig, Katja; Kuchta, Kathrin; Wenzel, Franziska

CavE-ICT: Computerbasierte, adaptive und verhaltensnahe Erfassung von ICT-Skills und deren Potenziale zur Erfassung heterogenitätsreduzierender Interventionen.

In modernen Wissensgesellschaften hängt die gelingende soziale Teilhabe auch von Kompetenzen im Umgang mit Informations- und Kommunikationstechnologien (ICT-Literacy) ab (ETS, 2002). Entsprechend dieser Relevanz werden computerbezogene Kompetenzen in Large-Scale-Assessments (LSAs) wie PISA erhoben. Anhand der dabei erfassten Kompetenzen sind jedoch aufgrund des Populationsbezugs von LSAs keine reliablen Rückschlüsse auf die Ebene individueller Messergebnisse möglich. Deshalb konnte auch die Wirksamkeit heterogenitätsreduzierender Maßnahmen in diesem Bereich nicht anhand von LSA-Messungen dargestellt werden, was aber unter anderem aus testökonomischen Gründen wünschenswert scheint. Der vorliegende Beitrag fasst die dreijährige Arbeit eines Projekts zusammen, in dem ein computerbasierter, adaptiver und verhaltensnaher Test entwickelt wurde, der sowohl für LSA-Zwecke als auch zur Individualdiagnostik geeignet ist und die Erfassung interventionsbedingter Kompetenzzuwächse zulässt.

Im vorliegenden Beitrag wird zunächst die Entwicklung und Validierung eines Kompetenztests zur ICT-Literacy dargestellt. Der Beitrag setzt damit an, wie das Konstrukt ICT-Literacy theoretisch ausgearbeitet wurde, indem man frühere Konzeptionen der ICT-Literacy (z.B. Ferrari, 2012; van Deursen & van Dijk, 2009; International ICT Literacy Panel, 2002, Richter, Naumann & Horz, 2010) systematisch analysierte und Weiterentwicklungsbedarfe spezifizierte. So wurden z.B. kognitionspsychologische Aspekte der ICT-Literacy stärker differenziert und es erfolgte eine stärkere Berücksichtigung zentraler situationaler Faktoren. Auf Grundlage des theoretischen Frameworks wurde ein Assessmentframework entwickelt aus dem abgeleitet wurde, welche Art von Aufgaben die konstruktbestimmenden Fertigkeiten repräsentieren. Aus den Dimensionen und Kombinationen dieser Dimensionen ergaben sich Aufgabenklassen, die dann der Aufgabenspezifikation zu Grunde lagen.

Danach wird auf den Aufgabenentwicklungsprozess eingegangen, der exemplarisch dargestellt wird. Es wird gezeigt, wie eine verhaltensnahe Messung der ICT-Literacy durch computerbasierte Simulationsaufgaben erfolgt und wie die Simulationsaufgaben durch die Implementierung mit einer speziellen Software bis hin zur qualitativen Analyse in einem Cognitive Lab optimiert wurden. Zur Optimierung des Aufgabenkonstruktionsprozesses wurde weiterhin eine Simulationsstudie durchgeführt, um zu ermitteln, welche Größe der zu konstruierende Itempool haben muss, um für verschieden Testlängen hinreichend präzise Testwerte für Teilskalen gewährleisten zu können. Hier wurde weiterhin eruiert, welchen Einfluss (a) adaptives vs. nicht-adaptives Testen, (b) unterschiedlich große Itempools sowie (c) die Testlänge auf die Messpräzision haben.

Die quantitative Überprüfung des Frameworks erfolgte im Mai/Juni 2014 an einer Stichprobe von fast eintausend 15-jährigen Schülerinnen und Schüler in Rheinland-Pfalz und Baden-Württemberg. Nach der Skalierung des Tests mit dem Rasch-Modell wurden die Pilotierungsdaten genutzt, um sensu den PISA-Erhebungen weitere Merkmale der Schüler, die mittels Fragebögen erhoben wurden (Schulnoten, Hintergrundmerkmale, kognitive Grundfähigkeit etc.), mit den Daten der erfassten ICT-Literacy in Beziehung zu setzen. Die Ergebnisse der Pilotierung werden zusammen mit ausgewählten, relevanten Zusammenhängen der weiteren erhobenen Variablen vorgestellt.

In der abschließenden Diskussion sollen zwei Aspekte fokussiert werden: Zuerst wird der Projektstand dahingehend reflektiert, wie die LSA-Methodik mit Individual-Diagnostik verbunden werden kann. Ziel ist es zu zeigen, dass der generierte Itempool sowohl für individualdiagnostische Zwecke als auch für Large-Scale-Assessments einsetzbar ist. Dadurch kann anhand der Erfassung von ICT-Skills exemplarisch für andere Kompetenzbereiche gezeigt werden, wie die LSA-Methodik durch eine computerbasierte, adaptive und verhaltensnahe Erfassung von Kompetenzen auch eingesetzt werden kann, um individuelle Diagnostik zu ermöglichen und auch die Effektivität von heterogenitätsreduzierenden Interventionen abzubilden. Abschließend sollen die Probleme dieser vergleichsweise aufwändigen Kompetenzerfassung insbesondere unter der Perspektive diskutiert werden, wie solche Testentwicklungen tatsächlich anschlussfähig für laufende Large-Scale-Assessments wie PISA gemacht werden können und gleichzeitig der Zusatznutzen im Rahmen einer Individual-Diagnostik und Erfassung von Interventionsergebnissen erreicht werden kann.

Anm.: Dieses Forschungsprogramm wurde durch das Bundesministerium für Wirtschaft und Forschung gefördert im Rahmen des Programms „Forschungsvorhaben in Anknüpfung an Large Scale Assessments.“

Trautwein, Ulrich; Körber, Andreas; Schreiber, Waltraud

Messung historischer Kompetenzen in Large-Scale Assessments

Während in den USA im Rahmen von NAEP (National Assessment of Educational Progress) Large-Scale Testungen zur US-Geschichte und Weltgeschichte seit den 1960er Jahren durchgeführt werden, wird in der deutschsprachigen Fachdidaktik die Messung historischer Kompetenzen mit quantitativen Methoden – nach einem Vorlauf in den 1970er-Jahren (Borries, 1974) – erst seit der Ausdifferenzierung fachspezifischer Kompetenzmodelle in Folge des PISA-Schocks diskutiert (Barricelli, Gautschi & Körber, 2012; Körber, 2008). An den bisher vorliegenden Testinstrumenten wie NAEP oder den in Deutschland mit großem Medienecho diskutierten Schröder-Studien (Deutz-Schröder & Schröder, 2007; Schroeder, Deutz-Schroeder, Quasten & Schulze Heuling, 2012) wird die Fokussierung auf die Erfassung des historischen Wissens kritisiert (VanSledright, 2014; Borries, 2008; Pampel, 2008). Studien, die demgegenüber die Weinert'sche Kompetenzdefinition zugrunde legen, adressieren hingegen zumeist lediglich Teilfähigkeiten (z. B. Hartmann 2008; Hartmann, Sauer & Hasselhorn, 2009; Bertram, Wagner & Trautwein, 2013; Bertram, Wagner & Trautwein, 2014) bzw. die Verfügung über fachliche Konzepte (Seixas, 2008). Ein Großteil der empirischen Kompetenzforschung arbeitet zudem mit qualitativen Methoden, die für groß angelegte Untersuchungen nicht genutzt werden können (z.B. Wineburg, 1991, 1998, 2001; Rout, Britt, Mason und Perfetti, 1996; Baron, 2012; vgl. auch die Tagungsbände der „geschichtsdidaktik empirisch“: Hodel & Ziegler, 2008; Hodel & Ziegler, 2010; Hodel, Waldis & Ziegler, 2013). Bisher fehlt ein zeitökonomisch einsetzbares Messinstrument, welches eine reliable, zugleich aber auch die Einzeldimensionen übersteigende Erfassung historischer Kompetenzen ermöglicht.

In dem vom BMBF geförderten Verbundprojekt „Historical Thinking – Competencies in History“ (HiTCH) geht es um die Entwicklung eines Kompetenztests, der für den Einsatz in Large-Scale-Assessments gegen Ende der Sekundarstufe I in allen Schulstufen geeignet ist. Als heuristischer Rahmen bei der Itemgenerierung und Validierung des Tests wurde das in der Fachdidaktik Geschichte prominente Kompetenz-Strukturmodell historischen Denkens (Körber, Schreiber & Schöner, 2007) der sogenannten FUER-Gruppe zugrunde gelegt, da es (1) mit seinem Kompetenzbegriff basierend auf Weinert (2001) an Schulleistungsstudien direkt anschlussfähig ist, (2) relativ umfassend ist und damit auch Kompetenzdefinitionen anderer Modelle sowie Messkonzepte zu einzelnen Fähigkeiten, etwa aus der internationalen Forschung, integriert werden können (z.B. Second Order Concepts aus dem anglo-amerikanischen Raum, u.a. Lee, 1983; Seixas, 2008; Seixas & Morton, 2013), und (3) eine Differenzierung der Kompetenzbereiche vornimmt, die als Ausgangspunkt für die Entwicklung von Items dienen.

Im HiTCH-Projekt wird untersucht, ob sich mit weitgehend geschlossenen Aufgaben die historische Kompetenz reliabel erfassen lässt. Zudem wird geprüft, ob sich Hinweise darauf finden lassen, dass die im FUER-Modell postulierten Dimensionen (historische Frage-, Methoden-, Orientierungs- und Sachkompetenz) modellierbare Subfacetten darstellen. In dem Vortrag werden zentrale Befunde von Erhebungen aus den Jahren 2013 (N = 1.701) und 2014 (N = 1.265 sowie eine weitere, derzeit laufende Studie) vorgestellt, an denen Schüler/innen der 9. Jahrgangsstufe aller Schulformen in Deutschland, Schweiz und Österreich teilgenommen haben. Bei den Auswertungen kamen neben eindimensionalen IRT-Modellen auch konfirmatorische nested-factor Ansätze zur Anwendung. Die bislang vorliegenden Befunde weisen darauf hin, dass (1) historische Kompetenzen mit einem Multiple-Choice-Instrument, dessen Bearbeitungszeit mit zwischen 30 und 50 Minuten angesetzt wurde, reliabel gemessen werden können, dass (2) es möglicherweise eine mehrfaktorielle Struktur in der historischen Kompetenz gibt, die jedoch ggf. von einem Generalfaktor sowie itemspezifischen Effekten überlagert wird, und dass (3) die Beziehungen mit Drittvariablen im Sinne einer divergenten und konvergenten Validität insgesamt plausibel ausfallen.

Freitag
13. März

15:30 Uhr

HGB 50

Chair(s)

McElvany, Nele
Horz, Holger

„Forschungsprojekte in Anknüpfung an Large-Scale-Assessments –
Aktuelle Projekte und ihre Theorien, Methoden und zentralen Befunde“

Göllner, Richard

Sprachliche Komplexitätsmerkmale in Unterrichtsmaterialien, Testverfahren und Fragebogeninstrumenten

Sprache und sprachliche Bildung sind von zentraler Bedeutung für Bildungsprozesse. Sprache ist ein notwendiges Mittel, um neues Wissen und neue Fähigkeiten zu erwerben, Informationen zu übermitteln sowie abstrakte Ideen zu beschreiben. Im Rahmen dieses Symposiums werden daher verschiedene Aspekte zur Bedeutung sprachlicher Anforderungsmerkmale im Bildungskontext aus einer disziplinübergreifenden bzw. interdisziplinären Perspektive heraus beleuchtet. Es wird thematisiert, inwiefern sich Unterrichtsmaterialien entlang verschiedener linguistischer Merkmale charakterisieren lassen, inwiefern bildungssprachliche Komplexitätsmerkmale mit erhöhten Anforderungen für Schülerinnen und Schüler verschiedener Altersstufen einhergehen, inwieweit der familiäre Sprachhintergrund mit dem Textverstehen assoziiert ist und in welchem Ausmaß sprachliche Komplexitätsmerkmale auch für die Erfassung unterrichtlicher Prozesse relevant sind.

Hierzu werden im ersten Beitrag Ergebnisse zu linguistischen Analysen von Geographie-Sachbuchtexten vorgestellt. Ein Vergleich der Textmaterialien über Schulformen und Klassenstufen hinweg legt nahe, dass nicht von einer systematischen Konstruktion der Texte im Hinblick auf die sprachliche Komplexität ausgegangen werden kann (Beitrag 1). Anschließend werden in zwei Beiträgen bildungssprachliche Anforderungen beim Hörverstehen (Text- und Satzebene) näher untersucht. Es zeigt sich, dass bildungssprachliche Hörverstehentexte im Vergleich zu Alltagssprachlichen Texten höhere Anforderungen an Schülerinnen und Schüler der neunten Klassenstufe stellen. Differentielle Effekte bildungssprachlicher Texte auf die Hörverstehensleistungen von Schülerinnen und Schüler mit bzw. ohne Zuwanderungshintergrund sind nicht nachweisbar (Beitrag 2). Im dritten Beitrag zeigen sich für das Grammatikverständnis von deutschsprachigen und nicht deutschsprachigen Kindergartenkindern spezifische Leistungsunterschiede, die mittels syntaktischer Marker erklärt werden können (Beitrag 3). Abschließend wird im Rahmen eines vierten Beitrages die Bedeutung sprachlicher Anforderungsmerkmale für die Erfassung von Unterrichtsprozessen aus Schülersicht näher betrachtet. Hier sind insbesondere Nominalisierungen und die Wortanzahl mit der Messgenauigkeit von Fragebogenitems verbunden (Beitrag 4).

Beitrag 1: Detmar Meurers, Karin Berendes, Sowmya Vajjala & Doreen Bryant: Leseanforderung in der Sekundarstufe: Ein Vergleich der linguistischen Komplexität von Schulbuchtexten

Beitrag 2: Alexandra Marx, Thorsten Roick, Birgit Heppt & Petra Stanat: Bildungssprachliches und Alltagssprachliches Hörverstehen in der Erst- und Zweitsprache

Beitrag 3: Karin Berendes, Wolfgang Wagner, Detmar Meurers & Ulrich Trautwein: Grammatikverständnis von Kindern unterschiedlicher sprachlicher und sozioökonomischer Herkunft

Beitrag 4: Richard Göllner, Wolfgang Wagner, Eckhard Klieme, Detmar Meurers & Ulrich Trautwein: Schülerbeurteilungen des Unterrichts: Die Bedeutung der Lesekompetenz und sprachlicher Anforderungsmerkmale für die Bearbeitung von Fragebogenitems

Freitag
13. März

15:30 Uhr

HGC 20

Chair(s)

Göllner, Richard
Berendes, Karin
Meurers, Detmar**Sprachliche Komplexitätsmerkmale in Unterrichtsmaterialien, Testverfahren und Fragebogeninstrumenten**

Meurers, Detmar; Berendes, Karin; Vajjala, Sowmya; Bryant, Doreen

Leseanforderungen in der Sekundarstufe: Ein Vergleich der linguistischen Komplexität von Schulbuchtexten

Lesen stellt die zentrale Kernkompetenz für die Aneignung von Wissen und somit letztendlich für schulischen Erfolg dar (Baumert et al., 2001). Die Entwicklung von Leseleistungen ist eng an die Art und Häufigkeit des Umgangs mit Texten und entsprechende Lerngelegenheiten gekoppelt. Hierbei spielt der Schulunterricht eine zentrale Rolle (Hurrelmann, 2009). Es gibt Belege dafür, dass verschiedene Schulformen – auch bei Kontrolle der Ausgangsleistungen der Schülerschaft – unterschiedliche Entwicklungsmöglichkeiten bieten (Neumann et al., 2007). Ein institutionell bedingter Schereneffekt zum Nachteil der Hauptschüler und zugunsten der Gymnasiasten wird diskutiert. Bezogen auf die Leseleistungen können neben strukturellen und prozessualen Merkmalen des Unterrichts, wie beispielsweise auf das Lesen verwendete Zeit, auch die Lesematerialien zur Identifikation differenzieller Wirkungsmechanismen zwischen benachteiligten und privilegierten Lernumwelten herangezogen werden.

Ziel des interdisziplinären Projekts „Reading Demands in Secondary School“ (Berendes, Vajjala, Meurers & Bryant, 2013) ist es, die linguistische Komplexität von Schulbüchern über Schulformen und Klassenstufen hinweg zu analysieren und zu vergleichen. Es soll der Frage nachgegangen werden, ob die Sprachanforderungen in den Schulbüchern systematisch variieren und die linguistische Komplexität der Texte mit jeder Klassenstufe zunimmt. Des Weiteren soll geklärt werden, ob die Schulbücher der Gymnasien generell sprachlich anspruchsvollere Texte enthalten als die Hauptschulbücher. Die zugrundeliegende Annahme war, dass die Textschwierigkeiten und damit zugleich die Leseanforderungen nicht konsequent und systematisch variieren und dass möglicherweise „within-schoolbook differences“ und „within-school type differences“ größer sind als „between-school differences“.

Zur Klärung der Fragestellung werden Texteinheiten aus Geographie-Schulbüchern der Hauptschule und des Gymnasiums mittels computerlinguistischer Verfahren (Hancke, Vajjala & Meurers, 2012) untersucht. 5.973 Texte aus 35 Schulbüchern für die Klassenstufen 5–10 gingen in die Untersuchung ein. Die Bücher wurden auf Basis der „Zulassungen für allgemein bildende Gymnasien“ (LS, 2013a) und der „Zulassungen für Haupt- und Werkrealschulen“ (LS, 2013b) zusammengestellt und stammen von vier verschiedenen Verlagen. Jeder Text wurde einer von fünf Kategorien (Text, Quelle, Übung, Definition, Sonstiges) mit unterschiedlichen Subkategorien zugeordnet. Für die Auswertung wird zunächst lediglich die Textart „erklärender Sachtext“ herangezogen, die den Großteil (49%) der Schulbuchtexte ausmacht. Ein breites Spektrum lexikalischer, morphologischer und syntaktischer Merkmalsausprägungen wird zwischen den Schulformen und Klassenstufen varianzanalytisch verglichen. Des Weiteren wird untersucht, inwiefern sich die vielfältigen sprachlichen Eigenschaften mit maschinellen Lernverfahren zur Lesbarkeitsklassifikation nutzen lassen.

In dem Beitrag werden erste Ergebnisse des Projekts präsentiert und deren Implikationen für die Leseanforderungen von Sekundarschülerinnen und -schülern sowie die Entwicklung von Schulbuchtexten diskutiert.

Freitag
13. März

15:30 Uhr

HGC 20

Chair(s)

Göllner, Richard
Berendes, Karin
Meurers, Detmar

Sprachliche Komplexitätsmerkmale in Unterrichtsmaterialien, Testverfahren
und Fragebogeninstrumenten

Marx, Alexandra; Roick, Thorsten; Heppt, Birgit; Stanat, Petra

Bildungssprachliches und alltagssprachliches Hörverstehen in der Erst- und Zweitsprache

Schulleistungsstudien dokumentieren, dass Schülerinnen und Schüler mit Zuwanderungshintergrund häufig geringere akademische Kompetenzen erwerben als Schülerinnen und Schüler ohne Zuwanderungshintergrund. Als eine mögliche Ursache wird diskutiert, dass bildungssprachliche Fähigkeiten den akademischen Erfolg von Schülerinnen und Schülern entscheidend beeinflussen und dass der Erwerb von Bildungssprache in einer Zweitsprache mit besonderen Schwierigkeiten verbunden ist (Bailey & Butler, 2003). Bildungssprache wird dabei als jenes sprachliche Register verstanden, das im schulischen und universitären Kontext genutzt wird, um neues Wissen und Fähigkeiten zu erwerben, Informationen zu übermitteln und abstrakte Ideen zu beschreiben (Chamot & O'Malley, 1994). Dieses Register ähnelt der Schriftsprache und ist häufig abstrakt, sprachlich komplex, hoch strukturiert und wenig kontextualisiert (Bailey & Butler, 2003; Schleppegrell, 2004). Im Unterschied zur Bildungssprache gilt die Alltagssprache, die Schülerinnen und Schüler im Rahmen ihrer Kommunikation im sozialen Alltag nutzen, als deutlich informeller und weniger komplex.

Empirische Befunde stützen die Annahme, dass bildungssprachliche Texte höhere Anforderungen an Lernende stellen als alltagssprachliche Texte. Studien zum Erwerb von Bildungssprache in einer Zweitsprache konnten allerdings nicht konsistent aufzeigen, dass bildungssprachliche Texte in einer Zweitsprache (L2) eine höhere Hürde darstellen als in einer Erstsprache (L1). Eine mögliche Ursache für das Ausbleiben differenzieller Effekte könnte daraus resultieren, dass ein Großteil der bisherigen Studien zu Effekten bildungssprachlicher Merkmale auf Grundschulinnen und -schüler fokussierte. In diesem Altersbereich ähneln bildungssprachliche Texte jedoch häufig noch der Alltagssprache. Es ist möglich, dass Unterschiede im bildungssprachlichen resp. alltagssprachlichen Verstehen geschriebener und gesprochener Texte erst in höheren Klassenstufen evident werden, wenn schulische Texte zunehmend auch bildungssprachliche Elemente enthalten (Chudaske, 2012).

Vor diesem Hintergrund untersucht die vorliegende Studie die Abgrenzbarkeit von alltags- resp. bildungssprachlichem Hörverstehen bei Jugendlichen mit Deutsch als Erst- bzw. Zweitsprache. Weiterhin wurde der Frage nachgegangen, ob bildungssprachliche Hörverstehenskompetenzen bei Jugendlichen in einer Zweitsprache disproportional niedriger ausgeprägt sind als in einer Erstsprache.

Im Kontext der Studie bearbeiteten 459 Schülerinnen und Schüler (NL1 = 173; NL2 = 286) der neunten Klassenstufe bildungssprachliche und alltagssprachliche Hörverstehentexte sowie einen Wortschatztest und einen Test der grammatikalischen Fähigkeiten. Zusätzlich wurde der sozioökonomische Hintergrund der Schülerinnen und Schüler erhoben und ging als Kontrollvariable in die Analysen ein. In einem ersten Schritt wurden Items, die Differential Item Functioning in Abhängigkeit vom Sprachhintergrund aufwiesen, aus den Analysen entfernt. Anschließend wurde die erste Forschungsfrage zur Dimensionalität der Hörverstehenskompetenzen mittels einer konfirmatorischen Faktorenanalyse überprüft. Zusätzlich wurde die Dimensionalität anhand der Korrelationen zu Wortschatz und Grammatikkenntnissen der Schülerinnen und Schüler als externen Validierungskriterien überprüft. Zur Beantwortung der zweiten Forschungsfrage wurden die Mittelwerte der beiden Sprachgruppe im bildungs- resp. alltagssprachlichen Hörverstehen im Rahmen eines latenten Mehrgruppenmodells miteinander verglichen.

In Hinblick auf die erste Forschungsfrage zeigen die Ergebnisse, dass ein Modell des Hörverstehens, das zwischen bildungs- und alltagssprachlichem Hörverstehen differenziert, besser auf die Daten passt als ein eindimensionales Hörverstehensmodell. Konträr zu theoretischen Annahmen zur Bildungssprache zeigten sich allerdings höhere Korrelationen zwischen alltagssprachlichem Hörverstehen und Grammatikkenntnissen der Schülerinnen und Schüler als zwischen bildungssprachlichem Hörverstehen und den Grammatikkenntnissen. Der Wortschatz war gleichermaßen stark mit bildungssprachlichem Hörverstehen wie mit alltagssprachlichem Hörverstehen assoziiert. Hinsichtlich der zweiten Forschungsfrage zeigten sich für alle Jugendlichen schwächere Leistungen im bildungssprachlichen Hörverstehen als im alltagssprachlichen Hörverstehen. Besonders stark ausgeprägte Leistungsnachteile der L2-Jugendlichen gegenüber den L1-Jugendlichen im Bereich des bildungssprachlichen Hörverstehens zeigten sich indes nicht.

Die vorliegende Studie liefert Hinweise darauf, dass sich bildungs- und alltagssprachliches Hörverstehen als Subfacetten des Hörverstehens trennen lassen. Eine Validierung anhand von Außenkriterien stützt die theoretischen Annahmen zum Konzept der Bildungssprache jedoch nicht. Zudem treten die erwarteten Leistungsnachteile der L2-Jugendlichen im Bereich des bildungssprachlichen Hörverstehens nicht zutage, so dass die Befunde insgesamt nur bedingt dafür sprechen, dass sich bildungssprachliches und alltagssprachliches Hörverstehen empirisch und theoretisch voneinander abgrenzen lassen.

Freitag
13. März

15:30 Uhr

HGC 20

Chair(s)

Göllner, Richard
Berendes, Karin
Meurers, Detmar**Sprachliche Komplexitätsmerkmale in Unterrichtsmaterialien, Testverfahren
und Fragebogeninstrumenten**

Berendes, Karin; Wagner, Wolfgang; Meurers, Detmar; Trautwein, Ulrich

Grammatikverständnis von Kindern unterschiedlicher sprachlicher und sozioökonomischer Herkunft

Alltagssprachlich-kommunikative Fähigkeiten im Allgemeinen und bildungssprachliche Kompetenzen im Besonderen sind eine wesentliche Voraussetzung für eine erfolgreiche Bildungsbeteiligung (Berendes, Dragon, Weinert, Heppt & Stanat, 2013). Als zentral für die Beherrschung von Bildungssprache wird die syntaktische Kompetenz angesehen (Dehn, 2011). Kommt es zu Einschränkungen der syntaktischen Kompetenzen, so kann sich dies negativ auf das schulische Lernen auswirken, da die sprachliche Erfassung von Aufgabenstellungen, Texten sowie Erklärungen erschwert ist. Eine vorschulische Ausweitung und Differenzierung der syntaktischen Kompetenzen ist somit eine wichtige Voraussetzung im Hinblick auf das Verständnis der im schulischen Unterricht verwendeten bildungssprachlichen Strukturen. Der Erwerb von komplexen syntaktischen Strukturen scheint insbesondere für Kinder nicht-deutscher Familiensprache und Kinder aus Familien mit geringem sozioökonomischen Hintergrund eine Hürde darzustellen (Weinert & Ebert, 2013).

Ausgehend von der enormen Bedeutung von syntaktischen Fähigkeiten für Bildungsprozesse wird in dem Beitrag untersucht, inwieweit Kinder verschiedener sprachlicher und sozioökonomischer Herkunft unterschiedliche Ergebnisse im Grammatikverständnis erzielen. Verglichen wurde hierbei, ob komplexe grammatische Strukturen eine besondere Hürde für Kinder nicht-deutscher Familiensprache und Kinder aus Familien mit niedrigem sozioökonomischem Status darstellen. Hierfür wurden die Grammatikleistungen von etwa fünfjährigen Kindergartenkindern des Nationalen Bildungspanels Differential Item Functioning-Analysen auf der Basis kreuzklassifizierter Mehrebenenmodelle mit zufälligen Effekten für Person und Item unterzogen. Die Gesamtstichprobe umfasste 2.996 Kindergartenkindern (49.7% weiblich) aus 279 Kindergärten und 720 Kindergartengruppen. Für die berichteten Ergebnisse wurde mit verschiedenen Teilstichproben gerechnet. In die Berechnungen zum Sprachhintergrund gingen die Daten von 2.111 Kindern ein, die Berechnungen zum sozioökonomischen Hintergrund basieren auf den Daten von 1.819 Kindern. Als Prädiktoren zur Vorhersage relativer itemspezifischer Leistungsdifferenzen wurden verschiedene syntaktische Marker herangezogen. Diese Marker wurden mittels computerlinguistischer Methoden berechnet (siehe Hancke, Vajjala & Meurers, 2012).

Für monolingual deutschsprachige Kinder zeigte sich im Hinblick auf den sozioökonomischen Hintergrund (hoher vs. niedriger sozioökonomischer Status der Eltern) eine nahezu perfekte Übereinstimmung der relativen Itemschwierigkeiten. Der Leistungsvergleich zwischen monolingual deutschsprachigen und nicht-deutschsprachigen Kindern zeigte dagegen teilweise deutliche gruppenspezifische Itemschwierigkeiten. Die Variabilität in den Itemschwierigkeiten bei den nicht-deutschsprachigen Kindern ließ sich lediglich zu 93% durch die Itemschwierigkeiten bei monolingual deutschsprachigen Kindern vorhersagen. Post-Hoc Analysen liefern Belege dafür, dass sich die spezifischen Itemschwierigkeiten bei nicht-deutschsprachigen Kindern durch Berücksichtigung syntaktischer Merkmale als zusätzliche Prädiktoren erklären lassen. Als oberflächliches Merkmal konnte beispielsweise eine Zunahme der Satzlänge als Prädiktor identifiziert werden. Die Bedeutung der Ergebnisse für die Erfassung von grammatischen und schulischen Leistungen wird diskutiert.

Diese Arbeit nutzt Daten der Kindergartenkohorte des Nationalen Bildungspanels (NEPS), doi: 10.5157/NEPS:SC2:2.0.0. Die Daten des NEPS wurden von 2008 bis 2013 als Teil des Rahmenprogramms zur Förderung der empirischen Bildungsforschung erhoben, welches vom Bundesministerium für Bildung und Forschung (BMBF) finanziert wurde. Seit 2014 wird das NEPS vom Leibniz-Institut für Bildungsverläufe e.V. (IfBi) an der Otto-Friedrich-Universität Bamberg in Kooperation mit einem deutschlandweiten Netzwerk weitergeführt.

Freitag
13. März

15:30 Uhr

HGC 20

Chair(s)

Göllner, Richard
Berendes, Karin
Meurers, Detmar**Sprachliche Komplexitätsmerkmale in Unterrichtsmaterialien, Testverfahren
und Fragebogeninstrumenten**

Göllner, Richard; Wagner, Wolfgang; Klieme, Eckhard; Trautwein, Ulrich

Schülerbeurteilungen des Unterrichts: Die Bedeutung der Lesekompetenz und sprachlicher Anforderungsmerkmale für die Bearbeitung von Fragebogenitems

Schülerbeurteilungen des Unterrichts sind seit jeher eine wichtige und häufig genutzte Datenquelle zur Erfassung des Unterrichtsgeschehens. Schülerangaben sind inzwischen fester Bestandteil verschiedener empirischer Forschungsrichtungen und finden seit mehreren Jahren auch im Bereich der Schulentwicklung sowie des Bildungsmonitorings ihre praktische und regelmäßige Anwendung (u.a. Clausen, 2002; Helmke, 2006; Pietsch, 2009).

Trotz der häufigen Verwendung von Schülerbeurteilungen ist bislang jedoch vergleichsweise wenig über die psychometrische Qualität von Schülerangaben zum Unterricht bekannt. Gleichzeitig zeigt bereits eine oberflächliche Betrachtung vorhandener Instrumente, dass sich Fragebogenitems erheblich im Ausmaß ihrer sprachlichen Komplexität unterscheiden. Hierzu gehören u.a. Unterschiede im Bereich lexikalischer Merkmale (z.B. Anzahl der Nominalisierungen) als auch grammatikalischer Merkmale (z.B. durchschnittliche Satzlänge oder Anzahl von Nebensätzen) und linguistischer Merkmale (z.B. Mehrdeutigkeit einzelner Wörter). Einige dieser Sprachmerkmale konnte im Rahmen verwandter Untersuchungen als relevant für die Messgüte von Testaufgaben identifiziert werden. So fanden beispielsweise Shaftel, Belton-Kocher, Glasnapp und Poggio (2006) in einer Untersuchung von Schülerinnen und Schülern der vierten, siebten und zehnten Klasse, dass Präpositionalphrasen, mathematischer Fachwortschatz und mehrdeutige Wörter einen substantiellen Einfluss auf die Testleistung von Schülerinnen und Schülern ausüben. Effekte eines bildungssprachlich geprägten Wortschatzes als auch die Anzahl von Nominalphrasen erwiesen sich in einer neueren Untersuchung von Haag, Heppt, Stanat, Kuhl und Pant (2013) als prädiktiv.

Auf der Grundlage dieser Befunde soll im vorliegenden Beitrag die Bedeutung sprachlicher Merkmale für die Beurteilung des Unterrichts aus Schülersicht untersucht werden. Im Mittelpunkt der Analysen steht die Variabilität von Schülerangaben auf der Ebene zwischen Klassen und innerhalb von Klassen, welche im vorliegenden Beitrag als zentrale Kriterien betrachtet werden. Als Datengrundlage dienen Daten von Schülerinnen und Schülern der fünften Jahrgangsstufe (130 Klassen) der Studie TRAIN (Jonkman, Rose und Trautwein, 2013). Die Klassifikation potentieller Komplexitätsmerkmale erfolgte anhand der Klassifikationen von Hancke, Vajjala und Meurers (2012). Im Rahmen eines zweischrittigen Analyseverfahrens und der Anwendung eines latenten Mehrebenen-Dispersionsmodells wurden die beiden folgenden Fragestellungen untersucht: (a) inwieweit kann die Variabilität von Schülerangaben bei der Beurteilung von Einzelitems anhand der Leseleistung von Schülerinnen und Schülern prädiziert werden, (b) inwieweit können Unterschiede der Prädiktionsergebnisse zwischen Items durch sprachliche Anforderungsmerkmale erklärt werden. Erste Analyseergebnisse zeigen, dass die Variabilität von Schülerangaben signifikant mit der Leseleistung von Schülerinnen und Schülern in Verbindung steht. Entsprechende Zusammenhänge ergaben sich auf der Ebene innerhalb von Klassen und weniger stark auf der Ebene zwischen Klassen. Mit Blick auf die Bedeutung sprachlicher Merkmale scheinen Nominalisierungen und die Wortanzahl von Bedeutung zu sein.

Freitag
13. März

15:30 Uhr

HGC 20

Chair(s)

Göllner, Richard
Berendes, Karin
Meurers, Detmar

Sprachliche Komplexitätsmerkmale in Unterrichtsmaterialien, Testverfahren
und Fragebogeninstrumenten

Itzek-Greulich, Heike

Empirische Studien zur Beurteilung und Wirkung außerschulischer MINT-Lernorte

Außerschulische Lernorte, wie z. B. Schülerlabore, bieten gegenüber der Schule veränderte institutionelle Rahmenbedingungen. Diese zeigen sich bspw. in der Ausstattung (moderne Experimentiergeräte, Labor-Räume), dem Personal (oft wissenschaftliche Mitarbeiter), den Themen (vielfach forschungsorientiert) und dem Projektcharakter (mehrstündige Experimentiertage). Damit sind vielfältige Hoffnungen verbunden, wie z. B. einer Steigerung des Interesses, einer Stärkung des Fähigkeitsselbstkonzepts sowie des Gewinnens eines Einblicks in naturwissenschaftliche und technische Berufsfelder. Verschiedene Forschungsarbeiten widmen sich vor diesem Hintergrund der Wirksamkeit von Schülerlaboren. Ziel des Symposiums ist es, einen vielfältigen Blick auf Schülerlabore durch ausgewählte Studien zu werfen, um gemeinsame allgemeinere Aussagen der Forschung zu erkennen. Dabei steht die Heterogenität hinsichtlich Geschlecht, Geburtsland, Sprache und Eingangsvoraussetzungen im Fokus.

Die Beiträge des Symposiums unterscheiden sich in Bezug auf die Altersstufe (Primarstufe/Sekundarstufe), den Bildungskontext (Schweiz/Deutschland), die Einsatzorte (außerschulische Schülerlabore/mobile Lernlabore an Schulen), sowie in Bezug auf die Forschungsdisziplin (Didaktiker/ Erziehungswissenschaftler). Zudem beziehen sich die Lernlabore auf unterschiedliche Fachdomänen (Chemie/Physik/Natur, Mensch, Gesellschaft). Auch wird im Rahmen des Symposiums aufgezeigt, dass je nach Fragestellung unterschiedliche methodische Zugriffe (experimentelle Designs mit Pre-Post Erhebungen, qualitative Videoanalyse und quantitative Fragebogendaten) realisiert wurden.

Der erste Beitrag untersucht quantitativ, ob die Beurteilung des mobilen Lernlabors MobiLab, das Experimente zu naturwissenschaftlichen und technischen Themen in die Schweizer Primarschulen (4. bis 6. Klasse) bringt, durch personenbezogenen Faktoren beeinflusst wird. Dies ermöglicht Aussagen zur Frage, ob alle Schüler/innen gleichermaßen vom MobiLab profitieren, unabhängig ihres Geschlechts, ihres Geburtslandes und ihres Sprachgebrauchs zu Hause (Holmeier).

Der zweite Beitrag befasst sich mit den mittels Videografie beobachteten Unterschieden im Problemlöseprozess bei Schülerinnen und Schülern unter Berücksichtigung der individuellen Schülervoraussetzungen: Vorwissen, Zielorientierung und Interesse an Naturwissenschaft (Meier).

Der dritte Beitrag aus der Chemiedidaktik geht anhand einer randomisierten Interventionsstudie dreier Unterrichtsarrangements (außerschulisches Schülerlabor, herkömmlicher Schulunterricht und eine Kombination aus beidem) und einer Kontrollgruppe zur organischen Chemie auf situative Lernemotionen und situatives Interesse ein (Itzek et al.).

Die vorgestellte Studie am UniLab-Schülerlabor der Humboldt-Universität zu Berlin erforscht anhand einer Fragebogenstudie über mehrere Jahre, welche Unterschiede zwischen Mädchen und Jungen in der aktuellen Komponente des epistemischen Interesses nach dem Schülerlaborbesuchs auftreten (Priemer).

Freitag
13. März

15:30 Uhr

HGC 40

Chair(s)

Itzek-Greulich, Heike

Empirische Studien zur Beurteilung und Wirkung außerschulischer MINT-Lernorte

Holmeier, Monika

Wie unterschiedliche Schüler/innengruppen das mobile Lernlabor beurteilen.

Theoretischer Hintergrund: Während Schüler/innen am Ende der Grundschulzeit naturwissenschaftlichen Fragestellungen gegenüber aufgeschlossen sind (Wittwer, Saß & Prenzel 2008), zeigt sich im Verlauf der Schulzeit ein abnehmendes Interesse (Baumert & Köller, 1998; Hidi, 2000). Ein Faktor, der dieser Abnahme entgegenwirkt, ist ein kontextbezogener und anwendungsorientierter Unterricht (Hoffmann, Häußler & Lehrke, 1998; Parchmann & Gräsel, 2004), in dem die Schüler/innen Fragen aus ihrem Lebensweltkontext nachgehen und eigene Erfahrungen beim Experimentieren sammeln (Duit & Mikelskis-Seifert, 2007). Trotz dieser Ergebnisse zeigen die PISA-Befunde, dass Schüler/innen im naturwissenschaftlichen Unterricht selten Gelegenheit haben, eigenständig Experimente zu machen (PISA-Konsortium, 2006). Nicht zuletzt aufgrund dieser Befunde hat sich in den vergangenen Jahren eine Vielzahl an Lernlaboren etabliert, um Schüler/innen Erfahrungen mit Experimenten zu ermöglichen und das Interesse der Schüler/innen zu steigern (Brandt, Möller & Kohse-Höinghaus, 2008; Guderian, 2007; Pawek, 2009).

Das hier vorgestellte mobile Lernlabor MobiLab bringt über 130 Experimente zu acht naturwissenschaftlichen und technischen Themen in die Schweizer Primarschulen (4. bis 6. Klasse) und bietet den Kindern die Möglichkeit, während eines halben oder ganzen Tages eigenständig zu experimentieren und sich dadurch naturwissenschaftliche Phänomene selbst zu erschließen. Für die Experimente werden lediglich Alltagsgegenstände verwendet. Zudem fertigen die Kinder zu jedem der durchgeführten Experimente eigene Forschungsprotokolle an. Ziel ist es, die Kinder für naturwissenschaftliche und technische Fragen zu begeistern und sie zum Experimentieren zu inspirieren, um so die Freude an naturwissenschaftlichen Themen zu wecken.

Fragestellung: Im Beitrag wird untersucht, wie die Kinder das MobiLab beurteilen. Dabei wird geklärt, ob sich Beurteilungsunterschiede in Abhängigkeit des Geschlechts, des Geburtslandes und der gesprochenen Sprache zu Hause nachweisen lassen. Angenommen wird, dass die Beurteilung des MobiLabs positiv ausfällt. Im Sinne hoher Bildungsgerechtigkeit sollten das Geschlecht, das Geburtsland und die Sprache keinen Effekt auf die Beurteilung des MobiLabs haben.

Methode: Die Evaluation des MobiLabs startete im Herbst 2013. Bis Ende des Schuljahres 2014/2015 werden etwa 500 Schüler/innen zu drei Messzeitpunkten (direkt vor, direkt nach sowie 2.5 Monate nach dem MobiLab-Unterricht) mittels standardisierter Fragebogen befragt. Für den vorliegenden Beitrag wurden die Daten zum Messzeitpunkt 2 ausgewertet, der direkt nach dem MobiLab stattfand. Hier mussten die Schüler/innen das MobiLab anhand unterschiedlicher Skalen und Items beurteilen (Zeit, Anforderungsniveau, Spaß, Anzahl durchgeführter Experimente und Note). Die Reliabilitäten der Skalen sind ausreichend bis gut ($\alpha=.63-.88$; $N_{\min} = 331$).

Mittels deskriptiver Analysen wurde geprüft, wie die Kinder das MobiLab beurteilen. Anschließend wurden anhand von t-Tests für unabhängige Stichproben Unterschiede zwischen verschiedenen Schüler/innengruppen (Geschlecht, Geburtsland, Sprache) in Bezug auf die Beurteilungskriterien berechnet. Mittels Regressionen wurde abschließend geprüft, ob mögliche Unterschiede in der Beurteilung des MobiLabs auf die personenbezogenen Faktoren zurückzuführen sind oder diese sich verlieren, wenn das Interesse der Kinder am Experimentieren (Cronbachs $\alpha=.79$; $N=331$) und am Naturwissenschaftsunterricht ($\alpha=.80$; $N=334$) in das Modell aufgenommen werden.

Ergebnisse: Das MobiLab wird von den Kindern positiv beurteilt. Dabei zeigen sich Unterschiede in Bezug auf das Geburtsland und die Sprache. Kinder, die außerhalb der Schweiz geboren wurden, empfinden mehr Freude ($p<.05$), Kinder, die zu Hause wenig Deutsch sprechen, machen weniger Experimente ($p<.01$). Unter Kontrolle des Geschlechts, des Landes und der Sprache hat das Interesse am Unterricht und am Experimentieren einen signifikant positiven Einfluss auf die Note, die Freude und das Anforderungsniveau. Beim Aspekt Zeit und Anzahl der Experimente zeigt sich kein Effekt des Interesses, wohl aber der gesprochenen Sprache und des Geburtslandes. Kinder, die zu Hause wenig Deutsch sprechen oder außerhalb der Schweiz geboren sind, machen weniger Experimente und empfinden die Zeit zum Experimentieren eher als zu knapp.

Worauf die Ergebnisse zurückzuführen sind und wie das Konzept des MobiLab geändert werden sollte, damit im Sinne größerer Bildungsgerechtigkeit alle Kinder gleichermaßen vom MobiLab profitieren - unabhängig ihres Sprachgebrauchs und ihres Geburtslandes - wird zu diskutieren sein.

Freitag
13. März

15:30 Uhr

HGC 40

Chair(s)

Itzek-Greulich, Heike

Empirische Studien zur Beurteilung und Wirkung außerschulischer MINT-Lernorte

Meier, Angelika; Vogt, Franziska

Nutzung der Lernangebote in der Lernwerkstatt: Individuelle Lernwege beim naturwissenschaftlichen Problemlösen

Theoretischer Hintergrund: Die Studie wurde in einer Lernwerkstatt zu einem naturwissenschaftlichen Thema (Wetter und Klima) durchgeführt. Lernwerkstätten als außerschulische Lernorte bieten Lernumgebungen für eigenständiges, handlungsorientiertes und entdeckendes Lernen, wodurch eigene Lernwege und eine aktive Konstruktion von Wissen ermöglicht werden sollen (Schomaker, 2011). Unter konstruktivistischer Sichtweise wird Lernen als aktiver, problemlösender, selbstgesteuerter und vernetzender Aufbau von Wissen verstanden (Reusser, 2006). Für die Gestaltung von Lernumgebungen in den Lernwerkstätten heißt dies, dass Aufgaben gewählt werden, die den Lernenden Eigenaktivität ermöglichen, wobei insbesondere die kognitive Auseinandersetzung mit den Inhalten von Bedeutung ist (Mayer, 2004). Durch die Bearbeitung von problemorientierten Aufgabenstellungen erwerben die Lernenden sowohl Basiswissen über natürliche Phänomene, als auch Wissen über Denk- und Arbeitsweisen in den Naturwissenschaften und können durch die eigene Auseinandersetzung mit verschiedenen Inhalten eine positive Einstellung zu naturwissenschaftlichen Themen entwickeln (OECD, 2006; Prenzel & Seidel, 2008).

Fragestellung: Welche Unterschiede im Problemlöseprozess können bei Schülerinnen und Schülern mit unterschiedlichem Vorwissen beobachtet werden?

Methode: An der Studie nahmen 158 Schülerinnen und Schüler der 4. bis 6. Schulstufe teil, die zwischen 9 und 14 Jahre alt waren ($M = 11.2$ Jahre; $SD = 0.9$). Vor dem Besuch in der Lernwerkstatt wurden mit einem Fragebogen Daten zu den individuellen Voraussetzungen (Vorwissen, Zielorientierung und Interesse an Naturwissenschaft) erhoben. Während dem Besuch in der Lernwerkstatt wurden 20 Zweierteams nach dem Zufallsprinzip ausgewählt und bei der Bearbeitung einer problemorientierten Aufgabe videografiert. Als Datenbasis für die weiteren Auswertungen wurden die Transkripte der Videoaufzeichnungen von den Problemlösedialogen bei der Aufgabenbearbeitung verwendet. Zur Erfassung der Problemlöseprozesse wurde ein Kategoriensystem auf der Basis des Competent Problem Solving Model (Verschaffel et al., 1999) und des Experimentierzyklus (Frischknecht-Tobler & Labudde, 2010) entwickelt. Die Videoaufnahmen wurden in Episoden entsprechend den Veränderungen im Problemlöseprozess aufgeteilt (Jacobs et al., 2003). Bei der Codierung der Episoden wurde eine Beurteilerübereinstimmung von Cohens $\kappa = .91$ erreicht.

Ergebnisse: In der Videoanalyse wurden alle 427 Episoden aus den Videoaufzeichnungen je einer Kategorie zum Experimentieren als Problemlösen (Meier, in Vorb.) zugeordnet. Die Codierungen zeigen, dass mit dem Kategoriensystem die Teilschritte im Problemlösezyklus adäquat erfasst werden können. Aufgrund der Analyse der Problembearbeitung in den Teams wird ersichtlich, dass zwar alle Teilschritte beobachtet werden können, jedoch nur wenige Teams einen Problemlöseprozess durchlaufen, der als vollständig bezeichnet werden könnte. Der letzte Schritt im Problemlöseprozess, der der Überprüfung und Verifikation der Ergebnisse in Bezug auf die ursprüngliche Problemstellung dient, kann nur bei fünf von zwanzig Teams beobachtet werden. Ähnliche Befunde werden auch in Studien zum Problemlösen in Mathematik berichtet (z.B. Riebel, 2010; Rott, 2013; Schoenfeld, 1985; Verschaffel, Greer, & De Corte, 2000).

In der weiteren Auswertung wurden aufgrund der Punktzahl im Test zum Vorwissen ($M = 12.8$, $SD = 4.5$) diejenigen Teams ausgewählt, die im obersten Quartil (über 16 Punkte) liegen und verglichen mit Teams aus dem untersten Quartil (unter 10 Punkte). In allen Teams können Auslassungen und Rückschritte im Problemlöseprozess beobachtet werden. Hingegen zeigt der Vergleich zwischen den Teams mit hohem und denjenigen mit tiefem Vorwissen, dass nur in Teams mit geringem Vorwissen der Teilschritt zum Formulieren von Vermutungen nicht durchlaufen wird. Die Ergebnisse geben Hinweise darauf, wie die Lernprozesse der Schülerinnen und Schüler durch Lernbegleitung oder durch die Einbettung im Unterricht gefördert werden können.

Itzek-Greulich, Heike; Flunger, Barbara; Vollmer, Christian; Nagengast, Benjamin; Rehm, Markus; Trautwein, Ulrich
Emotionales Erleben in einem außerschulischen und innerschulischen MINT-Lehr-Lernarrangement

Theoretischer Hintergrund: Wenn auch in der Sekundarstufe I generell das Interesse an allen Fächern abnimmt (Daniels, 2008), verlieren doch Chemie und Physik in dieser Zeit am meisten an Attraktivität (Hoffmann, Lehrke, & Häußler, 1998; Woest, 1997). Untersuchungen wie TIMSS und PISA haben Hinweise geliefert, dass der naturwissenschaftliche Unterricht in Deutschland zu wenig problem- und anwendungsorientiert ist (Prenzel et al., 2005). Aufgrund dieser Ausgangslage wurden in den letzten Jahren verschiedene Ansätze entwickelt und erprobt. Dazu gehört auch der Einbezug außerschulischer Lernorte wie beispielsweise Schülerlabore (Engeln & Vorst, 2007). Ein wichtiges Ziel von Schülerlaboren ist es, die Begeisterung der Heranwachsenden für das naturwissenschaftliche Experimentieren sowie das Verständnis für die Naturwissenschaften zu steigern. Mittlerweile liegen in Deutschland mehrere Studien (Guderian & Priemer, 2008; Pawek, 2012) vor, die dem Besuch von Schülerlaboren eine positive Wirkung auf das Interesse und die Motivation von Schülerinnen und Schülern im naturwissenschaftlichen Unterricht zusprechen. Die in direktem Zusammenhang mit dem erlebten Unterricht stehenden lernrelevanten Emotionen fanden hierbei bisher kaum Berücksichtigung.

Fragstellung: Die vorliegende Studie untersucht, ob sich die mehrfach festgestellten Unterschiede im generellen Interesse (z.B. Glowinski & Bayrhuber, 2011) bezüglich außerschulischer Lernorte auch während des Schülerlaborbesuchs widerspiegeln. Hierbei wurden Lernemotionen und situatives Interesse während des Schülerlaborbesuchs berücksichtigt. Dies ist insbesondere von Relevanz, da positive Lernemotionen und situatives Interesse während Lernaktivitäten die Entwicklung von individuellem Interesse beeinflussen können (Hidi & Renninger, 2006; Krapp, 2002).

Methode: Die vorliegende Studie untersuchte die Wirkung verschiedener Lernarrangements (außerschulisches Schülerlabor mit und ohne Einbindung in den Unterricht, Schulunterricht, Kontrollgruppe) zum Unterrichtsthema „Starke Stärke“ (Itzek-Greulich et al., 2014) auf Lernemotionen und situatives Interesse von Schülerinnen und Schülern. Hierbei wurden Lernfreude, Langeweile, Ärger und Situatives Interesse (Pekrun, Goetz, Frenzel, Barchfeld, & Perry, 2011) erhoben. In der ersten Treatmentgruppe wurde die Unterrichtseinheit im Schülerlabor ohne Einbindung unterrichtet (SCOL only), in der zweiten wurde die Unterrichtseinheit nur im Schulunterricht behandelt (School only), in der dritten wurde der Schülerlaborbesuch in den Schulunterricht eingebunden (SCOL & School). Die Kontrollgruppe erhielt keinen Unterricht zum Thema „Starke Stärke“. Die randomisiert auf die Gruppen aufgeteilte Stichprobe (52 neunte Realschulklassen; 1415 Schüler und Schülerinnen) wurde mit Kurzfragebögen im Schuljahr 2013/14 während der Intervention nach dem Theorieteil (4 Schulstunden) und zudem nach dem Praxisteil (4 Schulstunden) befragt. Die Unterschiede der vier Gruppen in Lernemotionen und situativem Interesse wurden mit Mehrebenenregressionsanalysen geprüft. Mithilfe von Mplus wurde für jede der drei Lernemotionen (Lernfreude, Langeweile und Ärger) und das situative Interesse als abhängige Variable ein Regressionsmodell auf zwei Ebenen (Schüler- und Klassenebene) gerechnet. Als Referenzgruppe wurde die Kontrollgruppe gewählt.

Ergebnisse: Im Ergebnis zeigte sich in Bezug auf den Theorieteil der Intervention, dass Schüler und Schülerinnen der School-only Gruppe im Vergleich zur Kontrollgruppe weniger Langeweile und höheres situationales Interesse erleben. Lernende aller drei experimentellen Gruppen äußerten mehr Lernfreude während der Theoriephase als die Kontrollgruppe. In Bezug auf den Praxisteil berichteten Lernende des SCOL & School und SCOL only Treatments weniger Ärger im Vergleich zur Kontrollgruppe. In allen drei experimentellen Bedingungen wurde weniger Langeweile, höhere Freude und situationales Interesse als in der Kontrollgruppe berichtet. In einem weiteren Schritt wurden die Effekte des Geschlechts auf die abhängigen Variablen berücksichtigt. Hier zeigte sich, dass die Unterschiede zwischen den Treatments auch unter Kontrolle der Variable „Geschlecht“ gefunden werden können. Insgesamt zeigten sich in allen Lernarrangements hinweg stärkere Effekte während der Praxisphase der Intervention als während des Theorieteils. Das könnte darauf hindeuten, dass praktische Laborarbeit für Schülerinnen und Schüler motivierend ist, und dem Lernort dafür eventuell weniger starke Bedeutung zukommt. Jedoch zeigten sich auch differentielle Effekte der verschiedenen Lernarrangements in Bezug auf verschiedene abhängige Maße. Praktische Implikationen für die Organisation von Laborarbeit und Schülerlaborbesuchen werden diskutiert.

Freitag
13. März

15:30 Uhr

HGC 40

Chair(s)

Itzek-Greulich, Heike

Empirische Studien zur Beurteilung und Wirkung außerschulischer MINT-Lernorte

Priemer, Burkhard; Musold, Wiebke

Das Interesse von Mädchen und Jungen an physikalischen Themen nach dem Besuch eines Schülerlabors

Theoretischer Hintergrund: Verschiedene Studien haben gezeigt, dass Jungen ein höheres Interesse an naturwissenschaftlichen und insbesondere physikalischen Themen zeigen als Mädchen (z. B. Hoffmann, Häussler & Lehrke 1998, Holstermann & Bögeholz 2007, Schreiner & Sjøberg 2004, siehe auch Publikationen zu PISA 2012). Eine Ursache dafür kann im Image des Schulfaches Physik gesehen werden (Kessels 2002). Vor diesem Hintergrund stellt sich die Frage, ob sich bei außerschulischen Lernorten ebenfalls die gleichen Geschlechterunterschiede im Interesse zeigen.

Krapp (1992) unterscheidet zwei Formen des Interesses: das aktuelle und das dispositionale Interesse. Das aktuelle Interesse – auf das sich diese Untersuchung stützt – bezieht sich auf eine konkrete, vorübergehende Handlungssituation. Sowohl das aktuelle als auch das dispositionale Interesse sind durch drei Komponenten geprägt: die emotionale, die wertbezogene und die epistemische. Für den schulischen Bereich scheint insbesondere die epistemische Komponente bedeutsam zu sein. Diese steht auch im Fokus dieser Studie. Sie berücksichtigt die Bedürfnisse der Person, ihre Kompetenzen und Fähigkeiten bezüglich eines Interessengegenstandes zu erweitern.

Der Einfluss von Schülerlaboren auf die Interessensentwicklung von Schülerinnen und Schülern wurde in verschiedenen Studien untersucht (für einen Überblick siehe Guderian & Priemer 2008, Priemer und Pawek 2014). Als Schülerlabore können außerschulische Bildungseinrichtungen bezeichnet werden, in denen sich Schülerinnen und Schüler in projektähnlichen Unterrichtsformen mit naturwissenschaftlichen oder technischen Fragestellungen experimentell beschäftigen“ (Priemer und Lewalter 2009). Hauptanliegen dieser Studien war in der Regel der Nachweis des positiven Einflusses der Projekte der außerschulischen Lernorte auf das Interesse von Jugendlichen. In einigen dieser Studien wurde auch mituntersucht, ob es Unterschiede zwischen Mädchen und Jungen gibt (Engeln 2004, Pawek 2009). Berichtet werden keine bedeutsamen Unterschiede. Die zugrunde liegenden Stichproben der Studien waren aber z. T. recht klein, sodass diese Frage letztlich nicht hinreichend beantwortet werden konnte. Dies sollte die hier vorliegende Studie leisten.

Fragestellung: Gibt es Unterschiede zwischen Mädchen und Jungen in der aktuellen Komponente des epistemischen Interesses nach dem Besuch eines Schülerlabors?

Methode: Im UniLab-Schülerlabor der Humboldt-Universität zu Berlin wurden im Laufe von mehreren Jahren knapp N=7500 Schülerinnen und Schüler aus den Klassenstufen 4-13 (rund 50% Mädchen bzw. Jungen) zu ihrem aktuellen Interesse befragt. Die Skala der epistemischen Komponente des aktuellen Interesses besteht aus insgesamt fünf Einzelitems, die von Engeln (2004) übernommen wurde. Die Probanden wurden gebeten, den Fragebogen nach dem Schülerlaborprojekt auszufüllen (Bearbeitungszeit rund 4 min.).

Ergebnisse: Es ergaben sich hohe Werte im Interesse (MW=0.7, SD=0.2, N=7493 auf einer Skala von 0 bis 1). Die Untersuchung zeigte einen signifikanten Unterschied im Interesse zwischen Mädchen und Jungen zugunsten der Jungen, dieser hat allerdings eine sehr kleine Effektstärke ($t(7477)=3,37$, $p=.001$, $d=0,08$). Ein Besuch im Schülerlabor war also für beide Geschlechter gleichermaßen interessant. Neben diesem zentralen Befund werden im Vortrag weitere Ergebnisse z. B. bzgl. Klassenstufen, Mehrfachbesuchen und Projektthemen präsentiert.

Diese Studie weist zu der Frage nach Geschlechterunterschieden bzgl. des Interesses nach Schülerlaborprojekten erstmals eine hohe Stichprobengröße auf, unterliegt aber verschiedenen Einschränkungen. So wurde nur eine Komponente des aktuellen Interesses erhoben, nur ein Zeitpunkt – direkt nach dem Besuch des Schülerlabors – betrachtet und nur ein einziges Schülerlabor mit verschiedenen Projektthemen für verschiedene Klassenstufen untersucht. Dennoch scheint im Kontext der vorangegangenen Arbeiten die Aussage angemessen, dass Schülerlabore das Interesse von Mädchen und Jungen – im Gegensatz zu schulischem Unterricht – gleichermaßen fördern.

Dietrich, Julia

Motiviert für Physik et al.?! Individuelle und kontextuelle Faktoren für naturwissenschaftliche Motivation

Physik ist cool und Ingenieurinnen sexy? Mitnichten. Untersuchungen haben wiederholt belegt, dass naturwissenschaftliche Fächer und Berufe bei Jugendlichen auf wenig Begeisterung stoßen. So gewinnen die meisten westlichen Nationen nur schwer Nachwuchs für natur- und ingenieurwissenschaftliche Berufe und besonders Frauen sind unterrepräsentiert (OECD 2013). Dieses Symposium beleuchtet motivationale Ursachen für diese Tatsache. Es bringt Studien aus Psychologie und Erziehungswissenschaften mit quantitativen Fragebogen- und Tagebuchdaten und qualitativen Interviews zusammen. Deren Ergebnisse werden abschließend aus Sicht der Didaktik der Naturwissenschaften und Technik diskutiert.

Zunächst ermitteln Moeller und Kollegen in einer Tagebuchstudie die Charakteristika von Situationen mit optimaler Lernmotivation. Sie identifizieren Konstellationen situationaler Anforderungen und Ressourcen, die Schüler/innen typischerweise in ihrem Schulalltag erleben, und untersuchen, welche Konstellation stabilere Anteile fachspezifischen Selbstkonzepts und Interesses in Physik und Biologie vorhersagen.

Die Beiträge von Taskinen und Kollegen sowie Makarova und Kollegen verlegen sich auf die Analyse spezifischer Unterrichts- und Lehrermerkmale und deren Bedeutung für die Motivation in naturwissenschaftlichen Fächern. Taskinen et al. zeigen auf der Basis längsschnittlicher Fragebogendaten, dass Unterrichtsgestaltung und Enthusiasmus der Lehrperson für die Entwicklung von Motivation für das Fach Physik bedeutsam sind.

Makarova und Kollegen vertiefen die Untersuchung von motivationsförderlichen Unterrichtsmerkmalen in Chemie, Mathematik und Physik mit Fragebogen- und Interviewdaten. Besonderer Fokus ihrer Analyse ist das Zustandekommen geschlechtsuntypischer Studienwahlpräferenzen junger Frauen.

Dietrich und Kollegen befassen sich abschließend auf Basis von PISA Daten mit der Rolle der Eltern. Der Beitrag zeigt auf, dass positive elterliche Einstellungen gegenüber Naturwissenschaften und naturwissenschaftsbezogene Freizeitaktivitäten, die Eltern mit Kindern unternehmen, bedeutsame Faktoren für Interesse und Selbstkonzept in naturwissenschaftlichen Fächern und nicht zuletzt für Berufswahlintentionen sind.

Freitag
13. März

15:30 Uhr

HGC 50

Chair(s)

Taskinen, Päivi
Dietrich, Julia

Motiviert für Physik et al.?! Individuelle und kontextuelle Faktoren für naturwissenschaftliche Motivation

Moeller, Julia; Dietrich, Julia; Judy Spencer, Justina; Schneider, Barbara; Lavonen, Jari; Salmela-Aro, Katariina

Zwischen Herausforderung und Überforderung: Situationales Engagement im Unterricht

Theoretischer Hintergrund: Engagement ist ein in letzter Zeit stark beforschtes Konstrukt zur Beschreibung von Lernmotivation. Viele Studien betonen die situationsspezifischen, veränderlichen Aspekte von Schul-Engagement, aber wenige untersuchen sie empirisch (Fredricks & McColskey, 2012). Diese Studie untersucht situationale Merkmale von Engagement und bezieht sich dabei auf die in der Engagementforschung etablierten Anforderungs-Ressourcen-Modelle (Bakker & Demerouti, 2007; Salmela-Aro & Upadaya, 2014) sowie auf Theorien des situationsspezifischen Erlebens von Herausforderungen und Fähigkeiten, nämlich die Flowtheorie (Csikszentmihalyi & Schneider, 2000) sowie die Theory of Challenge and Threat States in Athletes (TCTSA; Jones et al., 2009). Vorherigen Studien zu diesen Theorien zufolge beeinflussen vor allem die Konstellationen erlebter Herausforderungen und Ressourcen die Entwicklung von Lernmotivation und Engagement (Tuominen-Soini & Salmela-Aro, 2014).

Fragestellung: (1) Welche Konstellationen (Engagementprofile) situationaler Anforderungen und Ressourcen erleben Schüler/innen typischerweise in ihrem Schulalltag (und außerschulischen Alltag), und wie oft? (2) Wie hängt die Auftretenswahrscheinlichkeit bestimmter Profile mit Merkmalen der erlebten Situationen, Kontexte, und mit Personenmerkmalen zusammen? (3) In welchem Ausmaß sagen situationale Engagementprofile stabile Aspekte von Lernmotivation in den naturwissenschaftlichen Fächern Biologie und Physik vorher? Aus den oben genannten Theorien wurden jeweils spezifische Hypothesen abgeleitet, die aus Platzgründen nicht hier sondern im Vortrag dargestellt werden.

Methode: Die Stichprobe umfasste 193 Finnische und 274 US-Amerikanische Schüler/innen der Sekundarstufe II (9.-12. Klasse). Experience Sampling Method (ESM) Daten wurden mit einer Smartphone Application erhoben. Die Schüler/innen wurden für die Dauer der Erhebung mit Smartphones ausgestattet und füllten wiederholt Kurzfragebögen am Smartphone aus, sobald dieses ein Tonsignal abgab. Erfragt wurden die Herausforderungen, Ressourcen und Interesse in der aktuellen Situation, sowie einige Validierungsmaße (z.B. Affekt). Zusätzlich wurden nach der einwöchigen ESM Erhebung einmalig Fragebögen zur Lernmotivation in den Fächern Biologie und Physik vorgelegt. Diese Fragebögen enthielten etablierte Skalen zur Erfassung von Schulengagement (Salmela-Aro & Upadaya, 2012), Schulburnout (Salmela-Aro Kiuru, Leskinen, & Nurmi, 2009), Selbstkonzept und Fachinteresse (OECD, 2007).

Die Auswertungsverfahren folgten einem personenorientierten Ansatz und erlaubten eine Trennung in situationsspezifische und personenspezifische Effekte. Mit latenten Profilanalysen der ESM Daten identifizierten wir situationale Engagementprofile, also typische Konstellationen von situationalen Herausforderungen, Fähigkeitserleben und Interesse. Konfigurationsfrequenzanalysen und Mehrebenenmodelle (Situationen genestet in Individuen) untersuchten anschließend Zusammenhänge zwischen dem Auftreten eines Engagementprofils und anderen Situations-, Kontext- und Personenmerkmalen.

Ergebnisse: Die latenten Profilanalysen ergaben sechs Profile typischerweise erlebter Situationen: (1) Sehr hohe Werten in allen drei Indikatoren Herausforderungen, Fähigkeiten, und Interesse („Hohes Engagement“, 10%), (2) niedrige Herausforderungs- und hohe Fähigkeitswerte („Entspanntes Interesse“, 38%), (3) hohe Herausforderungs- und niedrige Fähigkeitswerte („Überforderung“, 6%), (4) niedrige Werte in allen drei Indikatoren („Apathie“, 12%), sowie zwei Profile, deren Merkmalskonstellation zwei bereits genannten Profilen glich, die jedoch deutlich niedrigere Werte auf den Indikatorvariablen aufwiesen („Moderates Engagement“, 18% und „Moderat entspanntes Interesse“, 15%).

Die inhaltliche Interpretation der Cluster erfolgte theoriegeleitet und wurde empirisch validiert durch die theoriekonformen erwarteten Zusammenhänge dieser Erlebnisse mit anderen Merkmalen der Situationen, Kontexte, und Personen: Situationen Hohen Engagements gingen beispielsweise mit sehr hoher Konzentration, Energetisierung und Aktivierung einher und konnten durch hohe Schulengagement-Werte im Fragebogen validiert werden. Dahingegen waren überfordernde Situationen von starken negativen Emotionen gekennzeichnet und ihr häufiges Erleben spiegelte sich in erhöhten Schulburnout-Werten im Fragebogen wieder. Apathische Situationen gingen mit niedrigen Werten in allen emotionalen und motivationalen Variablen einher, positiven wie negativen.

Die fachspezifische Lernmotivation (Selbstkonzept und Interesse) in den Fächern Biologie und Physik konnte schließlich durch die Häufigkeit vorhergesagt werden, mit der Schüler/innen bestimmte situationale Engagementprofile zeigten. Insgesamt lassen die Ergebnisse Rückschlüsse darauf zu, welche Konstellationen situationaler Anforderungen und Ressourcen optimales Schul-Engagement kennzeichnen.

Taskinen, Päivi; Spoden, Christian; Fischer, Hans E.

Bedeutung des Lehrerenthusiasmus, der Klassenführung und des Praxisbezugs des Unterrichts auf die Motivation von Schülerinnen und Schülern in Physik

Viele Jugendliche stufen Physik als uninteressant und schwierig ein. Als eine Ursache für die ungünstige motivationale Lage werden oft Merkmale des Unterrichts diskutiert. Bisher hat die Forschung vor allem die fachspezifische Gestaltung des Unterrichts fokussiert, und einige vorteilhafte Unterrichtspraktiken wie den Praxisbezug des Naturwissenschaftsunterrichts identifiziert (Bransford & Donovan, 2005; Hoffmann, 2002). Jedoch gibt es nur einzelne Studien, die weitere Merkmale des schulischen Unterrichts wie Lehrerenthusiasmus (z.B. Keller, Neumann, & Fischer, 2014) oder Klassenführung (z.B. Seidel & Prenzel, 2006) in einem Zusammenhang mit Physikmotivation bringen. Empirische Studien weisen generell darauf hin, dass der Enthusiasmus der Lehrkraft mit einer höheren Motivation in der Klasse einhergeht (McKinney, et al., 1984; Patrick, Turner, Meyer, & Midgley, 2003). Auch eine angemessene Klassenführung zeigt sich als relevant für eine positive Motivation (Kunter, et al. 2007). Es ist anzunehmen, dass diese Zusammenhänge sich auch in Physik zeigen. Der Unterricht kann jedoch auf Mädchen und Jungen womöglich unterschiedlich wirken, da das Ausgangsniveau der Physikmotivation und Erfahrungen von Mädchen und Jungen in Physik unterschiedlich sind.

Fragestellung: Das Ziel des Beitrags ist es, die Bedeutsamkeit von verschiedenen Unterrichtsmerkmalen für die Entwicklung der Physikmotivation von Neuntklässlern zu untersuchen. Dabei wird der Frage nachgegangen: (1) Welche Merkmale des Physikunterrichts sind für den Zuwachs der Motivation der Schülerinnen und Schüler von Relevanz? (2) Zeigen die Unterrichtsmerkmale unterschiedliche Effekte auf die Entwicklung der Physikmotivation von Mädchen und Jungen?

Methode.

Stichprobe: Es wurde eine Stichprobe von Schülerinnen und Schülern (N=1193; 50% weiblich) aus 47 deutschen Schulklassen untersucht, die im Rahmen des Forschungsprojekts „Quality of Instruction in Physics – comparing instruction and performance in Finland, Germany and Switzerland“ (QuIP) zu zwei Messzeitpunkten befragt wurden.

Instrumente: Die motivationalen Variablen Interesse, Selbstkonzept und instrumentelle Motivation wurden mit Hilfe von Kurzskaalen aus dem Programme for International Student Assessment (PISA) erfasst (Frey et al., 2009). Die dort ursprünglich in Bezug auf die Naturwissenschaften formulierten Items wurden für den Forschungskontext in Bezug auf das Unterrichtsfach Physik adaptiert. Die sechs-stufigen Likert-Skaalen erzielten zu Prä- ($\alpha(\text{Interesse})=.939$; $\alpha(\text{Selbstkonzept})=.943$; $\alpha(\text{instr. Motivation})=.925$) und zum Posttest ($\alpha(\text{Interesse})=.937$; $\alpha(\text{Selbstkonzept})=.943$; $\alpha(\text{instr. Motivation})=.924$) sehr gute Reliabilitäten.

Die Schüler wurden zusätzlich hinsichtlich verschiedener Merkmale ihres Unterrichts befragt, wobei in der hier vorliegenden Untersuchung der Lehrerenthusiasmus im Unterricht ($\alpha=.915$), die Klassenführung ($\alpha=.907$) und der Praxisbezug ($\alpha=.754$) analysiert wurden.

Vorgehen bei der Datenanalyse: Es wurden Veränderungen in den drei genannten motivationalen Dimensionen mit einem Modell latenter Veränderung (McArdle, 2009) geschätzt, um sie mit Hilfe der Einschätzungen zum Unterricht zu erklären.

Ergebnisse: Das Modell latenter Veränderung erzielte eine gute Anpassung an die Daten (RMSEA=.029; CFI=.995; TLI=.993; SRMR=.009). Veränderungen in den drei motivationalen Dimensionen lassen sich regressionsanalytisch durch die drei Unterrichtsmerkmale Lehrerenthusiasmus, Klassenführung und Praxisbezug signifikant erklären, wenn auch die Varianzaufklärung verhältnismäßig gering ausfällt ($R^2(\text{Interesse})=.04$; $R^2(\text{Selbstkonzept})=.11$; $R^2(\text{instr. Motivation})=.04$). Der stärkste positive Effekte konnte bei der Erklärung von Veränderungen im Interesse durch den eingeschätzten enthusiastischen Lehrstil gefunden werden ($\beta=0.402$, $p<.001$).

Unter Fixierung der Parameter des Messmodells unterscheiden sich die Regressionsparameter der Unterrichtsvariablen zwischen den Geschlechtern nicht bedeutsam (Modellanpassung gesamte Stichprobe: AIC=39696.53; BIC=40030.76; Zwei-Gruppenmodell: AIC=41481.48, BIC=41937.37).

Die Ergebnisse weisen insgesamt darauf hin, dass spezifische Unterrichtsmerkmale die Entwicklung von Motivation bei Jugendlichen beeinflussen. Neben der Unterrichtsgestaltung scheint auch Enthusiasmus als Merkmal der Lehrerpersönlichkeit für die Physikmotivation relevant zu sein. Dies ist plausibel, da Enthusiasmus von Lehrkräften zum einem mit aktiverem Lernen und höherer Schülerengagement einhergeht. Zum anderen ist es möglich, dass Lehrkräfte ihr Enthusiasmus an ihre Schüler übertragen (Patrick et al., 2003). Allerdings stellt sich für weitere Studien die Frage, worauf diese Wirkung zurückzuführen ist. Es kann gewiss nicht das Ziel sein, Lehrkräfte in Enthusiasmus zu lehren, sondern eher zu ermitteln, welche erlernbaren Unterrichtspraktiken eine enthusiastische Lehrkraft ausgleichen können.

Freitag
13. März

15:30 Uhr

HGC 50

Makarova, Elena; Aeschlimann, Belinda; Herzog, Walter

Mathematisch-naturwissenschaftlicher Unterricht auf der Sekundarstufe II - Ein multiperspektivischer Zugang

Ausgangspunkt für den vorliegenden Beitrag ist die anhaltende Unterrepräsentation von jungen Frauen in schulischen Leistungskursen und Studienrichtungen sowie in Berufsfeldern im Bereich Mathematik, Informatik, Naturwissenschaft und Technik (MINT) (vgl. OECD 2013; Solga und Pfafl 2009). Zwar haben sich Bildungsbeteiligung und Bildungsabschlüsse zwischen den Geschlechtern angeglichen, der Angleichung im Quantitativen steht aber eine Stagnation im Qualitativen gegenüber. Berufe im Sozial- und Gesundheitswesen werden vorwiegend von Frauen gewählt, während in der Mehrzahl der MINT-Berufen nach wie vor eine Dominanz der Männer festzustellen ist. Zugänge von Frauen in Ausbildungsberufen des MINT-Bereichs bilden daher weiterhin die Ausnahme, v.a. in Ländern wie Großbritannien, Frankreich, Österreich, Deutschland und der Schweiz (vgl. Gaupp et al. 2008; OECD 2013, S. 91-93).

Geschlechterunterschiede im mathematisch-naturwissenschaftlichen Unterricht (u.a. Blossfeld et al., 2009; OECD, 2013) lassen vermuten, dass die Geschlechtersegregation bei der Berufs- und Studienwahl, insbesondere die geringe Bereitschaft junger Frauen, naturwissenschaftlich-technische Berufe zu wählen, auch durch schulische Faktoren bedingt ist (Eccles, 1989; Prenzel et al., 2009; Willems, 2011).

Ausgehend davon fokussiert der beiliegende Beitrag die methodisch-didaktischen Formen der Unterrichtsgestaltung in den Fächern Chemie, Mathematik und Physik auf der Sekundarstufe II an Deutschschweizer Gymnasien. Der Beitrag verfolgt dabei einen multiperspektivischen Zugang, der einen Vergleich der Wahrnehmungsperspektiven von weiblichen und männlichen Jugendlichen einerseits, und eine Gegenüberstellung der Schülerinnen- und Schülerperspektive und der Sicht der Lehrpersonen andererseits, erlaubt. Zur Analyse der Forschungsfragen werden sowohl quantitative als auch qualitative Daten herangezogen. Die Datenbasis für den vorliegenden Beitrag stammt aus dem Forschungsprojekt „Geschlechtsuntypische Berufs- und Studienwahlen bei jungen Frauen“, das im Rahmen des Schweizerischen Nationalen Forschungsprogramms 60 „Gleichstellung der Geschlechter“ durchgeführt wurde. Die Studie wurde in 16 deutschsprachigen Kantonen der Schweiz durchgeführt und umfasste zwei Phasen. In einer ersten (quantitativen) Projektphase wurden Schülerinnen und Schüler (N = 3046) sowie deren Lehrpersonen (N = 123) mittels eines standardisierten Fragebogens zu ihrem Fachunterricht (Mathematik, Physik, Chemie) befragt (Makarova, Herzog, Ignaczewska & Vogt [Aeschlimann], 2012). In einer zweiten (qualitativen) Projektphase wurden mit Schülerinnen, die an der ersten Projektphase beteiligt waren und eine geschlechtsuntypische Studienwahlabsicht verfolgen (N = 16), Einzelinterviews durchgeführt (Makarova, Herzog, Vogt & Ignaczewska [Aeschlimann], 2012).

Ergebnisse zeigen, dass lehrerzentrierter Unterricht in den drei untersuchten Fächern dominiert, wobei sich auch fachkulturbedingte Differenzen in der methodisch-didaktischen Unterrichtsgestaltung eruieren lassen. Darüber hinaus belegen unsere Ergebnisse Geschlechterdifferenzen in der Wahrnehmung der Unterrichtsgestaltung, wonach der Unterricht in den mathematisch-naturwissenschaftlichen Fächern als abwechslungsreicher seitens männlicher Jugendlicher im Vergleich zu weiblichen Jugendlichen erlebt wird. Zudem zeigen unsere Analysen Diskrepanzen in der Wahrnehmung der Methodenvielfalt im mathematisch-naturwissenschaftlichen Unterricht aus der Schülerinnen- und Schülerperspektive und aus der Sicht der Lehrpersonen. Schliesslich schildern die qualitativen Daten inwiefern sich die Unterrichtsgestaltung in den Fächern Chemie, Mathematik und Physik für geschlechtsuntypische Studienwahlpräferenzen junger Frauen als bedeutsam erweist.

Freitag
13. März

15:30 Uhr

HGC 50

Chair(s)

Taskinen, Päivi
Dietrich, Julia

Dietrich, Julia; Taskinen, Päivi; Kracke, Bärbel

Die Rolle der Eltern für die Motivation Naturwissenschaftler/in werden zu wollen

Theoretischer Hintergrund: Naturwissenschaftsbezogene Karrieren stehen bei vielen Jugendlichen nicht an der ersten Stelle ihrer Berufswünsche. Das Elternhaus ist ein bisher wenig fokussierter Einflussfaktor der Überlegung von Jugendlichen, Naturwissenschaftler/in werden zu wollen. Eltern können die Lernmotivation (resp. Interessen und Fähigkeitsüberzeugungen) ihrer Kinder in Naturwissenschaften durch verschiedene Prozesse beeinflussen (Eccles, 1993; Lent, Brown, & Hackett, 2000). Beispielsweise können hier gemeinsame Aktivitäten (Bradley, Corwyn, Burchinal, McAdoo, & García Coll, 2001) oder familiäre Wertevermittlung (z. B. Gniewosz & Noack, 2012) eine wichtige Funktion haben. Die Lernmotivation ist wiederum entscheidend für die späteren Berufswahlentscheidungen: Jugendliche wählen Berufe, an deren Tätigkeiten sie interessiert sind und bei denen sie glauben, die beruflichen Anforderungen meistern zu können. Eltern sind zudem Rollenvorbilder und Informationsvermittler für spezielle Berufe, sodass sie die berufsspezifische Motivation von Jugendlichen beeinflussen können (Eccles, 1993; Krumboltz, 1994). Den Eltern kann dabei eine spezifische Funktion bei ihren Töchtern zukommen, da Mädchen generell weniger Erfahrung mit Naturwissenschaften haben (Alexander, Johnson, & Kelley, 2012) und Naturwissenschaften als männerdominiert wahrgenommen werden (Kessels & Hannover, 2007).

Fragestellungen: (1) Können naturwissenschaftsbezogene Aktivitäten im Kindesalter und elterliche Werte über Naturwissenschaften die kindliche Motivation in Naturwissenschaften (Interesse, Fähigkeitsselbstkonzept, naturwissenschaftliche Berufswahlintentionen) im Jugendalter erklären? (2) Gibt es geschlechtsspezifische Effekte des Elternhauses?

Methodik: Es wurden Neuntklässlerinnen und Neuntklässler aus PISA 2006 Deutschland untersucht (OECD, 2007). Um die Analysen auf die Jugendlichen zu fokussieren, die die kompetenzmäßigen Voraussetzungen für eine naturwissenschaftliche Karriere aufwiesen, umfasst die Analysestichprobe nur Jugendliche, die mindestens Kompetenzstufe V in den Leistungstests Naturwissenschaften erreichten (N = 1703). Deren Eltern machten Angaben zur eigenen Wertschätzung von Naturwissenschaften und naturwissenschaftlichen Aktivitäten in der Familie im Kindesalter.

Anhand eines Multigruppen-Strukturgleichungsmodells wurden Zusammenhänge der Elternvariablen mit den motivationalen Merkmalen der Jugendlichen ermittelt.

Ergebnisse: Die Ergebnisse zeigten positive Zusammenhänge naturwissenschaftlicher Aktivitäten mit den Eltern im Kindesalter zu Interesse ($\beta_{\text{Mädchen}} = .37$; $\beta_{\text{Jungen}} = .27$) und Fähigkeitsselbstkonzept im Jugendalter ($\beta_{\text{Mädchen}} = .27$; $\beta_{\text{Jungen}} = .17$). Die elterliche Wertschätzung von Naturwissenschaften zeigte kleinere Zusammenhänge mit Interesse ($\beta_{\text{Mädchen}} = .12$; $\beta_{\text{Jungen}} = .11$) und Fähigkeitsselbstkonzept ($\beta_{\text{Mädchen}} = .12$; $\beta_{\text{Jungen}} = .09$, n.s.), war jedoch signifikanter Prädiktor für naturwissenschaftliche Berufswahlintentionen.

Insgesamt weisen die Ergebnisse darauf hin, dass positive elterliche Einstellungen gegenüber Naturwissenschaften und naturwissenschaftsbezogene Freizeitaktivitäten im Kindesalter charakteristische Sozialisationsmerkmale sind, die entsprechende Fach- und Berufswahlmotive fördern können.

Freitag
13. März

15:30 Uhr

HGC 50

Chair(s)

Taskinen, Päivi
Dietrich, Julia

Motiviert für Physik et al.?! Individuelle und kontextuelle Faktoren für naturwissenschaftliche Motivation

Förtsch, Christian; Werner, Sonja; Dorfner, Tobias; von Kotzebue, Lena; Neuhaus, Birgit J.

Kognitive Aktivierung im Biologieunterricht – Wie werden situationales Interesse und Leistung von Lernenden beeinflusst?

Gemäß konstruktivistischen Lerntheorien ist die kognitive Aktivität von Schülerinnen und Schüler eine Voraussetzung zur Konstruktion von Wissen (Chi, 2009; Mayer, 2004, 2009). Neben Klassenmanagement und individueller Lernunterstützung wird auch kognitive Aktivierung als eine der drei Basisdimensionen von Mathematikunterricht angesehen (z.B. Kunter et al., 2013). Lipowsky et al. (2009) definieren für kognitive Aktivierung drei Schlüsselemente: kognitives Niveau von Schüleraktivitäten, Unterrichten von Konzepten und anregendes Unterrichtsgespräch. Allerdings finden sich in der Literatur nicht immer übereinstimmende Definitionen (vgl. u.a. Ergönenc et al., 2014; Klieme et al., 2006; Lipowsky et al., 2009; Neumann et al., 2012), die jedoch alle gemein haben, dass Schülerinnen und Schüler angeregt werden auf einer höheren Ebene zu denken und sich tiefer mit dem Inhalt auseinanderzusetzen. Der positive Einfluss von kognitiver Aktivierung auf die Leistung und das Fachinteresse von Schülerinnen und Schülern wurde bisher vor allem in Studien aus dem Bereich der Mathematik und Physik gezeigt (u.a. Ergönenc et al., 2014; Kunter et al., 2013; Lipowsky et al., 2009; Waldis et al., 2010). Es fehlen jedoch noch weitere Studien, die diese Ergebnisse festigen, als auch die Übertragbarkeit auf andere Fächer empirisch überprüfen. Die hier vorgestellte korrelative Teilstudie im Prä-Post-Design ist in das DFG-geförderte Projekt LerNT (von Kotzebue et al., 2014) eingebettet. Ziel ist die Analyse von kognitiver Aktivierung im Biologieunterricht und dessen Auswirkungen auf die Biologieleistung und das Biologieinteresse von Lernenden. Dazu wurden an verschiedenen bayerischen Gymnasien in der 6. Jahrgangsstufe zum Themenbereich Botanik Prozessdaten von 81 Biologieunterrichtsstunden (N = 28 Biologielehrkräfte; 3 Videos pro Lehrkraft) erhoben. Zusätzlich hierzu wurde mit einem Fragebogen das situationale Interesse (vgl. Vogt et al., 1999; N = 3 Items, $\alpha = 0,89$) der Lernenden (N = 741) nach jeder Stunde und am Ende des Schuljahres deren Leistung mithilfe eines standardisierten Leistungstests (freiwillige Lernstandserhebung in Natur-und-Technik; N = 16 Aufgaben, $\alpha = 0,67$) des Staatsinstituts für Schulqualität und Bildungsforschung München (ISB) gemessen. Die Analyse der Unterrichtsvideos erfolgte mittels eines, an die Biologie adaptierten, Ratingmanuals (N = 37 Items, $\alpha = 0,93$, $\kappa = 0,85$) (vgl. Rakoczy & Pauli, 2006; Vogelsang & Reinhold, 2013). Mehrebenenanalysen (HLM 7.01) zeigen nach Kontrolle von intrinsischer Motivation und Fachinteresse einen marginal signifikanten positiven Einfluss auf die Schülerleistung ($\beta = 0,19$, SE = 0,10, $p < 0,10$), sowie einen signifikanten positiven Einfluss auf das situationale Interesse der Schülerinnen und Schüler ($\beta = 0,24$, SE = 0,08, $p < 0,01$). Einschränkungen zeigen sich speziell in den Ergebnissen bezüglich der Biologieleistung. Gründe dafür sind die relativ kleine Stichprobe von 28 Lehrkräften und das Fehlen eines Prä-Leistungstest, welcher wegen der Nutzung des standardisierten Leistungstests nicht durchgeführt werden konnte. Gleichwohl liefern die Ergebnisse dieser Studie zusätzliche Hinweise auf die Bedeutung von kognitiver Aktivierung im Unterricht und auf die Übertragbarkeit dieses Unterrichtsqualitätsmerkmals auf das Fach Biologie.

Freitag
13. März

15:30 Uhr

GC 03/142

Chair(s)

Förtsch, Christian

Willems, Ariane S.

Profile des situationalen und individuellen Interesses von Schülerinnen und Schülern im Mathematikunterricht und ihr Zusammenhang zur Unterrichtsqualität

Heterogenität im schulischen Kontext bezieht sich nicht nur auf leistungs- oder herkunftsbezogene Merkmale, sondern umfasst gleichermaßen motivationale und interessenbezogene Dimensionen (Trautmann & Wischer, 2012). Die Fragen, wie schulische Interessen von SchülerInnen ausgeprägt sind, sich entwickeln und in welcher Weise Lehrkräfte durch unterrichtsbezogenes Handeln Interessen fördern können, sind daher auch Bestandteil der empirischen Bildungsforschung. Aktuelle Befunde zeigen, dass das Mathematikinteresse von SchülerInnen vergleichsweise gering ausgeprägt ist und im Laufe der Schulkarriere abnimmt (Baumert & Köller, 2000; Daniels, 2008; Köller et al., 2000; Kunter, 2005; Waldis, 2011; Willems, 2011). In der Schul- und Unterrichtsforschung wird aktuell vermehrt diskutiert, welche Bedingungen relevant für die Förderung von Interessen im Fachunterricht sind. Als eine Ursache wird dabei auch die Qualität unterrichtlicher Lehr-Lern-Prozesse hervorgehoben, wobei insbesondere die Bedeutsamkeit der Dimensionen Strukturiertheit/Zielklarheit, Schülerorientierung sowie kognitive Aktivierung herausgestellt wird (Klieme & Rakoczy, 2008; Waldis, 2011; Willems, 2011).

Den theoretischen Rahmen des Beitrags bilden zwei Ansätze der prozessorientierten Interessenforschung: Die Person-Gegenstandstheorie (Krapp, 2002) und das Vier-Phasen-Modell der Interessenentwicklung (Hidi & Renninger, 2006). Übereinstimmend wird in beiden Theorien zwischen zwei Entwicklungsstadien des Interesses unterschieden: dem situationalen und individuellen Interesse (Willems, 2011). Das individuelle Interesse wird als relativ dauerhaftes, stabiles Persönlichkeitsmerkmal aufgefasst, das in unterschiedlichen Situation zum Tragen kommt. Das situationale Interesse beschreibt einen Zustand des Interessiertseins, der in erster Linie durch situationsspezifische Merkmale einer konkreten (Lern-)Umgebung hervorgerufen wird und sich zu einem stabilisierten, situationsübergreifenden Interesse entwickeln kann (Hidi & Renninger, 2006). Auf Ebene des situationalen Interesses werden zudem zwei Komponenten unterschieden: Die Catch-Komponente, die sich auf das erste Auftreten eines situationalen Interesses bezieht und die Hold-Komponente, die eine relative Stabilisierung dieses Interesses innerhalb der Lernsituation beschreibt (Mitchell, 1993). Während das individuelle Interesse Gegenstand einiger Forschungsbemühungen ist, ist das situationale Interesse noch wenig untersucht (Minnaert, Boekaerts & Opdenakker, 2008; Tsai et al., 2008; Willems, 2011).

Ausgehend von diesen Überlegungen ist das Ziel des vorliegenden Beitrags, Interessen aus einer situationsspezifischen Perspektive des konkreten Unterrichts sowie aus Perspektive situationsübergreifender, fachlicher Interessen zu analysieren und die Frage zu beantworten, wie diese Interessendimensionen mit der Unterrichtsqualität aus Schülerperspektive zusammenhängen. Mittels des des personenzentrierten Analyseansatzes der Mehrebenen-Latent-Class-Analysen (Vermunt, 2008) wird folgenden Forschungsfragen nachgegangen: (1) Lassen sich auf Basis von Schülerangaben homogene Subgruppen identifizieren, die sich in charakteristischer Weise in ihrem Ausmaß der drei Interessendimensionen (Catch-Komponente, Hold-Komponente, individuelles Interesse) unterscheiden? (2) Unterscheiden sich diese Interessenprofiltypen in der Ausprägung der wahrgenommenen Unterrichtsqualität?

Die Grundlage der Analysen bilden Daten der SIGMA-Studie, im Rahmen derer N=951 SchülerInnen der 8. Jahrgangsstufe (Gymnasien) im Anschluss an eine Mathematikunterrichtsstunde zu ihrem situationalen (jeweils 6 Items zur Erfassung der Catch- und Hold-Komponente) und individuellem Interesse (Fachinteresse, 6 Items) befragt wurden. Darüber hinaus wurden unterschiedliche Unterrichtsqualitätsdimensionen (Strukturiertheit, Zielklarheit, inhaltliche Relevanz, kognitive Aktivierung, motivationsunterstützendes Verhalten der Lehrkraft) aus Schülersicht jeweils mit multiplen Indikatoren erfasst. Die Ergebnisse variablenzentrierter Analysen (CFA) zeigten zunächst, dass sich alle drei Interessendimensionen empirisch voneinander trennen lassen – wobei alle drei Dimensionen auf Stichprobenebene hoch positiv miteinander korrelierten (Willems, 2011). Die anschließenden Mehrebenen-Latent-Class-Analysen über die gesamten 18 Items der drei Interessendimensionen resultieren in einer stabilen Vier-Klassenlösung. Dabei zeigen sich zwei inhaltlich und schulpraktisch interessante Profilklassen: So lässt sich eine Schülergruppe identifizieren, die über unterdurchschnittliche Werte im Fachinteresse verfügen, gleichzeitig jedoch über überdurchschnittlich hohe Ausprägungen im situationalen Interesse. Außerdem lässt sich eine Profilklassen identifizieren, deren Schüler bei überdurchschnittlich hohen Ausprägungen des Fachinteresses durch ein gleichzeitig geringes Ausmaß des situationalen Interesses gekennzeichnet sind. Varianzanalytisch lassen sich auch charakteristische Mittelwertsunterschiede in der Wahrnehmung der Unterrichtsqualität in Abhängigkeit der Profilklassen nachweisen, die im Beitrag beschrieben und diskutiert werden. Die Analysen lassen den Schluss zu, dass spezifische Formen der Unterrichtsgestaltung dazu beitragen können, das situationale Interesse von Schülerinnen mit einem geringen Fachinteresse anregen zu können.

Freitag
13. März

15:30 Uhr

GC 03/142

Chair(s)
Förtsch, Christian

Herrmann, Christian; Seiler, Sara; Gerlach, Erin

Einfluss der Dimensionen Klassenführung und Schülerorientierung auf die Veränderung der Anstrengungsbereitschaft und des Fachinteresses im Schulfach Sport

Theoretischer Hintergrund: Die Unterrichtsqualität nimmt für den Lernerfolg von Schülern einen zentralen Stellenwert ein (Hattie, 2013). Dabei konnten drei Grunddimensionen der Unterrichtsqualität empirisch bestimmt werden (Klieme, Schümer & Knoll, 2001; Hamre & Pianta, 2010): Klassenführung, Schülerorientierung und kognitive Aktivierung. Zur Erfassung dieser Dimensionen haben sich neben Unterrichtsbeobachtungen vor allem die subjektiven Wahrnehmungen der Schüler als wichtige Informationsquelle erwiesen (Clausen, 2002). Für die Kernfächer (u. a. Mathematik, Deutsch) liegen bereits validierte Fragebogeninventare vor. Im vorliegenden Beitrag wird für die beiden primär motivationsfördernden Dimensionen Klassenführung und Schülerorientierung (Klieme & Rakoczy, 2008) eine fachspezifische Anpassung der Inventare auf den Sportunterricht vorgenommen. Die dritte Dimension kognitive Aktivierung muss zunächst außen vor gelassen werden, da vor dem Hintergrund der Spezifika des Sportunterrichts als Bewegungsfach eine Operationalisierung umfangreiche Adaptationen zur Folge hätte.

Fragestellung: Neben der Entwicklung einer fachspezifischen Testbatterie werden folgende Fragen bearbeitet: (1) Lassen sich die beiden Dimensionen der Unterrichtsqualität Klassenführung und Schülerorientierung im Sportunterricht abbilden? (2) Welchen Einfluss haben diese beiden Dimensionen auf die Veränderung des Fachinteresses und der Anstrengungsbereitschaft über ein Schuljahr?

Methode: In der von der Eidgenössische Sportkommission Schweiz geförderten IMPEQT-Studie wurden insgesamt 48 Klassen mit 1019 Schüler in der siebten (884 Schüler, 50.1% männlich; Alter: $M = 13.2$ Jahre; $SD = .63$) und darauf folgenden achten Jahrgangsstufe (767 Schüler, 47.5% männlich) zu acht Unterrichtsmerkmalen sowie zum Fachinteresse und zur Anstrengungsbereitschaft (adaptiert u.a. aus der PISA-, IGLU- und DESI-Studie) mittels Fragebogen erfasst. Die Berechnungen der konfirmatorischen Faktoranalysen und der Strukturgleichungsmodelle erfolgten in Mplus.

Ergebnisse: (1) Die achtfaktorielle Struktur der Unterrichtsmerkmale konnte in der konfirmatorischen Faktoranalysen mit guten Modellfits ($CFI = .94$; $RMSEA = .03$; $SRMR = .04$) bestätigt werden. Die acht Unterrichtsmerkmale wurden anschließend theoriegeleitet den beiden Dimensionen Klassenführung und Schülerorientierung zugeordnet und einer konfirmatorischen Faktoranalysen mit zweiter Ordnung unterzogen, welche ebenfalls über eine gute Modellanpassung verfügt ($CFI = .93$; $RMSEA = .03$; $SRMR = .05$). (2) In zwei erweiterten Strukturgleichungsmodellen zeigte sich, dass die Klassenführung ($CFI = .96$; $RMSEA = .04$; $SRMR = .10$) einen positiven signifikanten Einfluss auf die Veränderung des Fachinteresses ($\beta = .26$) sowie der Anstrengungsbereitschaft ($\beta = .35$) über ein Jahr hat. Die Schülerorientierung ($CFI = .95$; $RMSEA = .04$; $SRMR = .11$) hat zwar ebenfalls einen positiven signifikanten Einfluss auf die Veränderung des Fachinteresses und der Anstrengungsbereitschaft, dieser fällt jedoch deutlich geringer aus ($\beta = .08$).

Diskussion: Die entwickelte Testbatterie stellt eine Möglichkeit dar, die Unterrichtsqualität im Schulfach Sport aus Schülersicht in der Sekundarstufe zu erfassen. Die beiden Unterrichtsdimensionen Klassenführung und Schülerorientierung konnten im Sportunterricht empirisch identifiziert werden. Weiterhin zeigte sich, dass diese Dimensionen einen Einfluss auf die Veränderung der Motivation haben. Diese fällt im Fach Sport für die Klassenführung höher aus als für die Schülerorientierung. Die Entwicklung eines Instruments zur Erfassung der dritten und fachspezifisch akzentuierten Dimension kognitive Aktivierung steht noch aus und sollte Ziel zukünftiger Forschungsbemühungen sein.

Schütte, Kerstin; Köller, Olaf

Soziale Einflussfaktoren auf das Interesse an Mathematik

Jugendliche verbringen einen großen Teil ihrer Lebenszeit in der Schule sowie in Auseinandersetzung mit Schule. Schule ist also ein integraler Bestandteil ihres Lebens und ihre Rolle als Schülerin bzw. als Schüler ist klar definiert. Dennoch unterscheiden sich Jugendliche in ihrer schulbezogenen Identität (Oyserman, 2007). Neben dem Fähigkeitsselbstkonzept sind weitere Aspekte für das Handeln und Erleben von Schülerinnen und Schülern im akademischen Kontext relevant, die – um nicht als aversiv erlebt zu werden – ein stabiles, konsistentes Selbstkonzept ergeben sollten (Aronson, 1992). Dieser Beitrag fokussiert das Interesse an Mathematik, welches als begünstigende Bedingung zukünftiger Lernprozesse anzusehen ist. Das Interesse an einem Fach wird in dem Maße schwinden, in welchem ein Individuum negative Leistungsrückmeldung erfährt und das entsprechende Fähigkeitsselbstkonzept absinkt (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). Über diese Wahrnehmung der eigenen Fähigkeit hinaus hat die objektiv gemessene Leistung keinen Einfluss auf das Interesse (Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006). Mitschülerinnen und Mitschüler dienen aber nicht allein als Referenzrahmen, um über soziale Vergleiche zu einer Einschätzung der eigenen Kompetenz zu gelangen. Ebenso wie die Eltern transportieren Gleichaltrige durch ihr Handeln und gegebenenfalls durch direkte Äußerungen schul- bzw. fachbezogene soziale Normen (Altermatt & Pomerantz, 2003; Gniewosz & Noack, 2012). Die Wahrnehmung dieser Normen wirkt mutmaßlich ihrerseits auf das Interesse an Schule oder an einem spezifischen Fach.

Ein eher explorativer Anteil des Beitrags besteht im Vergleich unterschiedlicher Schulsysteme: Dem segregierten deutschen System wird die neunjährige finnische Gemeinschaftsschule gegenübergestellt. Die vergleichende Analyse gibt darüber Aufschluss, ob die betrachteten sozialen Einflussfaktoren in Lernumwelten, die bezogen auf die Kompetenzen der Schülerinnen und Schüler heterogen oder homogener zusammengesetzt sind, gleichermaßen wirksam sind. Die Hypothese, dass die wahrgenommenen Normen auf Seiten von Eltern und Freunden einen inkrementellen Beitrag zur Erklärung des Interesses an Mathematik leisten, wurde anhand von Daten aus der fünften Erhebung des Programme for International Student Assessment geprüft. Die $n = 5001$ Jugendlichen aus 230 Schulen in Deutschland und die $n = 8829$ Jugendlichen aus 311 Schulen in Finnland waren im Mittel 15,75 Jahre alt ($SD = 0,29$). Ihre Mathematikleistung wurde mit einem standardisierten Test erfasst. Ihr Interesse an Mathematik, ihr mathematisches Fähigkeitsselbstkonzept sowie die wahrgenommene mathematikbezogene Norm auf Seiten von Eltern und Freunden gaben die Schülerinnen und Schüler im Selbstbericht auf 4-stufigen Skalen an. Fehlende Werte wurden mit der Mplus-Software mehrfach imputiert ($m = 50$); die genestete Datenstruktur wurde in allen Analysen berücksichtigt. Die Prüfung von Messinvarianz ergab, dass die Messmodelle für die Substichproben aus Deutschland und Finnland vergleichbar waren. Im Mehrebenenmodell (doubly-latent model; Marsh et al., 2009) wurde die mittlere Mathematikkompetenz in der Schule auf Schulebene berücksichtigt. Wie jene der Eltern wurde die wahrgenommene Norm auf Seiten der Mitschülerinnen und -schüler nur auf Schülerebene modelliert, da sich die Skalen explizit auf Freunde, nicht auf die Klasse oder Schule bezog.

Das Befundmuster war in Deutschland und Finnland ähnlich. Beide wahrgenommenen mathematikbezogenen Normen leisteten einen inkrementellen Beitrag zur Vorhersage des Interesses an Mathematik. Unter Kontrolle des so genannten Fischteicheneffekts (negativer Effekt der mittleren Mathematikkompetenz in der Referenzgruppe auf das Fähigkeitsselbstkonzept) war der analoge kontextuelle Effekt auf das Interesse an Mathematik signifikant positiv. Die wahrgenommenen Normen waren miteinander sowie mit dem mathematischen Fähigkeitsselbstkonzept positiv korreliert. Insgesamt sprechen die Befunde also für eine konsistente schulbezogene Identität der Jugendlichen. Aufgrund der querschnittlichen Erhebung erlauben die vorliegenden Daten keine Aussage über die Wirkrichtung. Sie zeigen jedoch, dass Schülerinnen und Schüler mit geringen Kompetenzen und entsprechend geringem mathematischen Fähigkeitsselbstkonzept eher soziale Normen wahrnehmen – womöglich auch Kontexte aufsuchen, wo sie diese vermehrt finden – welche einem hohen Interesse und damit höheren Lerneranstrengungen entgegenstehen.

Freitag
13. März

15:30 Uhr

GC 03/142

Chair(s)

Förtsch, Christian

Kaiser, Christine

Das konzeptuelle Verständnis jugendlicher Lerner/innen von Demokratie

Sowohl in bildungswissenschaftlichen und lernpsychologischen als auch in politikdidaktischen Diskursen ist spätestens mit der Kompetenzdiskussion die möglichst umfangreiche Anhäufung von Faktenwissen als Ziel schulischer Lernprozesse in den Hintergrund gerückt. Ein Lernprozess wird nunmehr als erfolgreich angesehen, wenn er bei den Lernenden zum Aufbau eines umfassenden Verständnisses der grundlegenden Prinzipien eines Inhaltsbereichs führt, das bei der Interpretation neuer Wissensinhalte hilft und in Alltagssituationen anwendbar ist (Deutsches PISA-Konsortium, 2000; Klieme et al., 2006; Sander, 2007). Unterrichtsliche Lernprozesse können zu einem so verstandenen konzeptuellen Verständnis führen, wenn Lernende die Möglichkeit erhalten, neue Inhalte mit ihren bereits vorhandenen Vorstellungen in Beziehung zu setzen (vgl. Posner et al., 1982; Vosniadou, 1994). Sind die Vorverständnisse der Lernenden undifferenziert oder fehlerhaft, besteht ein grundlegendes Ziel von Unterricht in einer Ausdifferenzierung bzw. Veränderung der Konzepte. Conceptual Change-Ansätze beschäftigen sich mit der Frage, inwiefern und wodurch es im Zuge von Lernprozessen gelingt, vorhandene mentale Schemata durch neue Wissensbestände und Informationen auszudifferenzieren oder zu erneuern (vgl. Appleton, 1993; Chi & Roscoe, 2002; Treagust & Duit, 2008; Vosniadou, 1994).

Der lernpsychologische Ansatz des Conceptual Change wird in dieser Arbeit auf die politische Bildung übertragen. Um erfolgreiche politische Lernprozesse initiieren und unterstützen zu können, ist die Sichtbarmachung der subjektiven Konzepte der Lernenden zu Politik und Gesellschaft von grundlegender Bedeutung. Dies gilt besonders für das komplexe Themenfeld „Demokratie“. Empirische Studien bescheinigen Jugendlichen wiederholt ein mangelhaftes Verständnis der Funktionsweisen und Bedeutung demokratischer Systeme (vgl. Kerr et al., 2010).

An diesem Verständnis jugendlicher von Demokratie setzt die vorliegende Studie an: Sie ist als Vorstudie eines zweistufigen konsekutiven Forschungsdesigns zu verstehen (Creswell & Plano Clark, 2010; Kuckartz, 2014) und verfolgt das Ziel, Verständnisschwierigkeiten der Schüler/innen im Zusammenhang mit Demokratie herauszuarbeiten, die in wenig entwickelten oder fehlerhaften (Teil-)Konzepten zum Ausdruck kommen und die es im politischen Unterricht zu bearbeiten gilt. Die Studie dient der Beantwortung der Frage, wie sich die subjektiven Demokratiekonzepte Lernender beschreiben lassen.

Zur Überprüfung dieser Frage wurde ein explorativer qualitativer Zugang gewählt, der auf Gruppendiskussionen mit Schüler/innen basiert. Es wurden fünf Gruppendiskussionen mit je sechs Schüler/innen (N = 30) der 11. Klassenstufe aus Gymnasien in verschiedenen Bundesländern Österreichs geführt und anschließend vollständig transkribiert. Die Analyse der Gruppendiskussionen folgte zunächst der deduktiv-induktiven Kategorienbildung nach Mayring (2010). Die kategorisierten Sequenzen wurden anschließend paraphrasiert, mit Ankerbeispielen versehen und nach semantischen Ähnlichkeiten zu übergeordneten Teilkonzepten zusammengefasst. Die Interrater-Reliabilität, berechnet mittels Cohen's Kappa, erwies sich als zufriedenstellend.

Die Ergebnisse zeigen, dass die Demokratiekonzepte der befragten Schüler/innen durch vier thematisch breit gefasste Subkonzepte gekennzeichnet sind: 1) Mehrheitsentscheidungen bzw. Mitspracherecht, 2) Freiheit, 3) Gleichberechtigung und 4) Wohlstand. Jedes dieser Subkonzepte enthält sowohl elaborierte als auch weniger elaborierte bzw. fehlerhafte Teile. Im Subkonzept „Mehrheitsentscheidungen bzw. Mitspracherecht“ wird u.a. deutlich, dass das Prinzip der Mehrheitsentscheidungen häufig als Mehrheitsdiktatur missverstanden wird, bei der dem Willen der Mehrheit keine rechtlichen Grenzen gesetzt sind. Abgeleitet aus Alltagsbeobachtungen werden Regeln, die den Willen der Mehrheit binden, als undemokratische Einschränkungen wahrgenommen - ein Verständnis, das auch im Subkonzept „Freiheit“ zum Ausdruck kommt. Ein Grundproblem im Verständnis von Demokratie scheint demnach darin zu liegen, dass den Schüler/innen die Bedeutung von Regelungen und Gesetzen für das gesellschaftliche Zusammenleben nicht bewusst ist.

Daraus ergibt sich für den schulischen politischen Unterricht die Notwendigkeit, an diesen Missverständnissen anzusetzen und sie zu bearbeiten. Mithilfe der Ergebnisse der Gruppendiskussionen ist es nunmehr möglich, diejenigen Themenbereiche zu bestimmen, die mithilfe einer Intervention im politischen Unterricht bearbeitet werden sollten. Mithilfe eines quasiexperimentellen Prä-Post-Follow-Up Designs wird in einer Folgestudie der Frage nachgegangen, wie sich vorhandene Fehlkonzepte mit entsprechenden didaktisch-methodischen Mitteln im politischen Unterricht aufbrechen lassen.