LEUPHANA
UNIVERSITÄT LÜNEBURG

Förderung von Kernkompetenzen in der Lehramtsausbildung durch videobasiertes Microteaching und Peerfeedback

Altmann, Anna F.; Weber, Kira Elena; Prilop, Christopher Neil; Kleinknecht, Marc; Nückles, Matthias

Published in:
Lehrer. Bildung. Gestalten.

Publication date:
2019

Document Version
Verlags-PDF (auch: Version of Record)

Link to publication

Citation for pulished version (APA):
Altmann, A. F., Weber, K. E., Prilop, C. N., Kleinknecht, M., & Nückles, M. (2019). Förderung von Kernkompetenzen in der Lehramtsausbildung durch videobasiertes Microteaching und Peerfeedback. In T. Ehmke, P. Kuhl, & M. Pietsch (Hrsg.), *Lehrer. Bildung. Gestalten.: Beiträge zur empirischen Forschung in der Lehrerbildung* (S. 213-223). Beltz Juventa Verlag. https://content-select.com/de/portal/media/download_oa/9783779954811

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 03. Juli. 2025

Timo Ehmke | Poldi Kuhl |
Marcus Pietsch (Hrsg.)

Lehrer. Bildung. Gestalten.

Beiträge zur empirischen Forschung in der Lehrerbildung

BELTZ JUVENTA

Timo Ehmke | Poldi Kuhl | Marcus Pietsch (Hrsg.)
Lehrer. Bildung. Gestalten.

Timo Ehmke | Poldi Kuhl |
Marcus Pietsch (Hrsg.)

Lehrer.
Bildung.
Gestalten.

Beiträge zur empirischen Forschung
in der Lehrerbildung

BELTZ JUVENTA

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronische Systeme.

FSC
www.fsc.org
MIX
Papier aus verantwortungsvollen Quellen
FSC® C089473

Dieses Buch ist erhältlich als:
ISBN 978-3-7799-6177-2 Print
ISBN 978-3-7799-5481-1 E-Book (PDF) (Open Access)

1. Auflage 2019

© 2019 Beltz Juventa
in der Verlagsgruppe Beltz · Weinheim Basel
Werderstraße 10, 69469 Weinheim
Alle Rechte vorbehalten

Herstellung: Hannelore Molitor
Satz: Helmut Rohde, Euskirchen
Druck und Bindung: Beltz Grafische Betriebe, Bad Langensalza
Printed in Germany

Weitere Informationen zu unseren Autor_innen und Titeln finden Sie unter: www.beltz.de

Vorwort

In den letzten Jahren ist die Diskussion über die Qualität der Lehramtsausbildung an Universitäten zunehmend in das Zentrum wissenschaftlicher und bildungspolitischer Diskussionen gerückt. Hierzu haben insbesondere die „Qualitätsoffensive Lehrerbildung“, aber auch andere BMBF-Förderlinien entscheidend beigetragen. An vielen Universitäten wurde in der Folge nach Möglichkeiten gesucht, die Lehramtsausbildung zu verbessern und durch innovative didaktische Elemente eine stärkere Verzahnung von theoretischen Inhalten und praktischen Schulbezügen zu erreichen.

Entsprechend naheliegend war es, die AEPF-Tagung 2018 in Lüneburg unter das Motto „Lehrer. Bildung. Gestalten.“ zu stellen und auf diese Weise den Diskurs um die Ziele, Voraussetzungen, Prozesse und Wirkungen von Bildung zu fördern, insbesondere in der Lehrerbildung. Das hohe Interesse der AEPF-Community an der Thematik zeigte sich in insgesamt 23 Symposien, 124 Einzelbeträgen und 65 Postern.

Eine Auswahl der Beiträge, die sich explizit mit Fragen der Lehrerbildung beschäftigen, haben wir in diesem Konferenzband zusammengestellt. Wir haben diese vier Bereichen zugeordnet, in die jeweils Expertinnen und Experten mit einem einleitenden Überblickskapitel einführen. Die insgesamt 27 Beiträge widmen sich (1) den Einstellungen von (angehenden) Lehrkräften, (2) den Kompetenzen von (angehenden) Lehrkräften, (3) pädagogischen Interventionen und (4) den Bedingungen effektiven Unterrichts.

Wir danken allen Autorinnen und Autoren für ihre Beiträge und auch den Gutachterinnen und Gutachtern für ihre hilfreichen und konstruktiven Anmerkungen zu den eingesandten Manuskripten. Dazu, dass es uns gelungen ist, den Band bereits zur 84. AEPF in Münster zu veröffentlichen, haben Sie alle maßgeblich beigetragen!

Unser ausdrücklicher Dank gilt zudem zwei Personen: Nele Fischer, die vor und während der Tagung für einen reibungslosen und erfolgreichen Ablauf der AEPF in Lüneburg gesorgt hat. Herzlichen Dank auch an Kathrin Reese, die die Organisation dieses Konferenzbandes leistete.

Aber genug geschnackt: Nun wünschen wir Ihnen eine anregende Lektüre!

Lüneburg im Juli 2019
Timo Ehmke, Poldi Kuhl und Marcus Pietsch

Inhalt

Vorwort 5

I Einstellungen von (angehenden) Lehrkräften
Miriam M. Gebauer 11

Unter welchen Bedingungen entwickeln sich Überzeugungen im Praxissemester?
Eine Interviewstudie
Nicole Zaruba, Anna Gronostaj, Frederik Ahlgrimm und Miriam Vock 20

Entwicklung der Einstellungen Lehramtsstudierender zu Leistungsheterogenität
Der Einfluss von Lehrveranstaltungen
Sabine K. Lehmann-Grube, Andreas Hartinger und Markus Dresel 33

(Wie) wirken inklusionsorientierte Lehrveranstaltungen und Praxisanteile auf inklusionsbezogene Überzeugungen angehender Lehrkräfte?
Katja Scharenberg und Saskia Opalinski 44

Unterscheiden sich Studentinnen und Studenten des Grund- und Förderschullehramts in ihren Einstellungen und ihren Selbstwirksamkeitsüberzeugungen in Bezug auf den inklusiven Unterricht?
Frank Hellmich, Marwin Felix Loeper und Fabian Hoya 57

Welche Rolle spielen Einflussfaktoren vor und während des Studiums für die Überzeugungen von Lehramtsstudierenden zu schulischer Inklusion?
Ariane S. Willems und Karina Meyer 71

Einstellungen zur Leistungsheterogenität von Lehrkräften an integrativen und nicht-integrativen weiterführenden Schulen
Marcus Syring, Samuel Merk und Thorsten Bohl 85

Die Implementation digitaler Medien gestalten
Zur Bedeutung der Einstellungen von Fremdsprachenlehrkräften
Kerstin Drossel und Birgit Eickelmann 97

II Kompetenzen von (angehenden) Lehrkräften
Kai S. Cortina 110

Kompetenzerleben Lehramtsstudierender im Semesterpraktikum
Eva-Kristina Franz und Jana Groß Ophoff 116

Eine ‚gute' Erklärung für alle?!
Gruppenspezifische Unterschiede in der Beurteilung von Erklärqualität – erste Ergebnisse aus dem interdisziplinären Forschungsprojekt FALKE
Alfred Lindl, Lisa Gaier, Matthias Weich, Mario Frei, Christina Ehras, Maria Gastl-Pischetsrieder, Michael Elmer, Katharina Asen-Molz, Anna-Maria Ruck, Jana Heinze, Renate Murmann, Eileen Gunga und Simone Röhrl 128

Medienbezogene Kompetenzen von Lehrpersonen
Empirische Befunde und Perspektiven für die Lehrerausbildung
Ramona Lorenz, Manuela Endberg und Birgit Eickelmann 142

Entwicklung eines video- und textbasierten Instruments zur Messung kollegialer Feedbackkompetenz von Lehrkräften
Christopher Neil Prilop, Kira Elena Weber und Marc Kleinknecht 153

Learning by doing
Eine explorative Erhebung zur Förderung digitaler Kompetenzeinschätzungen durch die Verwendung digitaler Medien in der Hochschuldidaktik
Charlott Rubach, Rebecca Lazarides, Nina Brendel und Karsten Krauskopf 164

Zusammenhänge zwischen handlungsorientierten und thematischen Lerngelegenheiten und der DaZ-Kompetenz angehender Lehrkräfte
Tobias Schroedler und Ilse Stangen 176

Performanznahe und videobasierte Messung von DaZ-Kompetenz bei Lehrkräften
Skalierung und dimensionale Struktur des Testinstruments
Svenja Lemmrich, Sarah-Larissa Hecker, Stephanie Klein, Timo Ehmke, Anne Köker, Barbara Koch-Priewe und Udo Ohm 190

III Pädagogische Interventionen in der Lehrerbildung
Tobias Feldhoff 205

Förderung von Kernkompetenzen in der Lehramtsausbildung durch videobasiertes Microteaching und Peerfeedback
Anna F. Altmann, Kira Elena Weber, Christopher Neil Prilop, Marc Kleinknecht und Matthias Nückles 213

Veränderung der Selbstwirksamkeitserwartungen von Lehrkräften in einer Fortbildung zum Umgang mit Heterogenität im Unterricht
Claudia Schmaltz und Timo Leuders 224

Transfer vom Wissen zum Können
Potenzial videofallbasierten Lernens zum Aufbau handlungsleitender Kognitionen
Sabine Schöb 233

Lerngelegenheiten im Praxissemester aus Absolvierendensicht
Ergebnisse aus der VFL-Praxis-Studie
Gabriele Klewin, Anne Köker und Jan Christoph Störtländer 246

Unterrichtsbesuche am laufenden Band
Anforderungen und Belastungsempfinden im niedersächsischen Vorbereitungsdienst
Marike Feierabend und Frederike Bartels 261

IV Bedingungen effektiven Unterrichts
Marc Kleinknecht 272

Epistemisch-diagnostische Aktivitäten im Diagnoseprozess bei Lehrkräften im Mathematikunterricht
Angelika Wildgans-Lang, Andreas Obersteiner und Kristina Reiss 281

Welche Planungssituationen und Planungsmittel beschreiben Sonderpädagoginnen, Sonderpädagogen und Fachlehrkräfte, die im inklusiven Unterricht kooperieren?
Ausgewählte Ergebnisse einer Interviewstudie
Silvia Greiten, Franziska Müller, Matthias Trautmann und Daniel Mays 292

Wovon hängt ab, wie Lehrkräfte differenzieren?
Eine Re-Analyse von Fragebogen- und Video-Daten der DESI-Studie
Julia Jennek, Anna Gronostaj und Miriam Vock 303

Soziale Interaktionen in pädagogischen Beziehungen
Erfassung mit State Space Grids
Marion Scherzinger, Benjamin Roth und Alexander Wettstein 314

Unterrichtsqualität in MINT-Klassen
Befunde aus der quasi-experimentellen Studie ProBiNi
Sonja Nonte, Mareike Haas und Tobias C. Stubbe 325

Gerechtigkeit von Leistungsbewertung im Urteil von Oberstufenschülerinnen und -schülern
Carolin Körner, Sabine Gruehn und Sylvia Rahn 336

Schriftliches Argumentieren im Deutschunterricht
Entwicklung eines Kompetenztests für Lernende der Sekundarstufe I
Julian Kempf und Dorit Bosse 347

Verzeichnis der Autorinnen und Autoren 359

I Einstellungen von (angehenden) Lehrkräften

Miriam M. Gebauer

Einleitung

In der Empirischen Bildungsforschung liefern Untersuchungen zu Einstellungen, die (angehende) Lehrkräfte unter anderem in Bezug auf Merkmale von Schülerinnen und Schülern oder auf kontextuelle Determinanten haben, wichtige Ergebnisse, die für Theorie und Praxis von zentraler Bedeutung sein können (z. B. Skott 2015). Einstellungen werden als Urteile bezogen auf ein Objekt verstanden und sind relevant für intendiertes Verhalten und anschließende Handlungen (Ajzen 2001), aber auch für individuelles Wohlbefinden wie beispielsweise wahrgenommenes Belastungserleben (Jennings 2015). Infolgedessen werden Einstellungen von (angehenden) Lehrkräften im schulischen Dienst als wichtig erachtet, da sie von Bedeutung für die Gestaltung eines effektiveren Unterrichts (Stern/Staub 2002) oder für die gelungene Umsetzung von schulischen Entwicklungen sein können (z. B. Goldhaber/DeArmond/DeBurgomaster 2011).

Generell kann mit Blick auf die aktuelle Literatur folgende Forschungsperspektiven hinsichtlich der Erforschung von Einstellungen und Überzeugungen (angehender) Lehrkräfte festgestellt werden. Es finden sich Studien, die (a) sich hinsichtlich des Einstellungsobjektbezugs unterscheiden, die (b) Veränderung und Entwicklung von Einstellungen betrachten oder (c) Determinanten, Korrelate und Typisierung der Einstellungen von (angehenden) Lehrkräften untersuchen. Hinsichtlich des differenziellen Objektbezugs von Einstellungen können spezifische Merkmale von Schülerinnen und Schülern untersucht werden (z. B. Hachfeld/Schroeder/Anders/Hahn/Kunter 2012), epistemische Vorstellungen über lernen und unterrichten fokussiert werden (Staub/Stern 2002) oder schulische Entwicklungen und Implementierung von Bildungsreformen als Einstellungsobjekt betrachtet werden (z. B. Goldhaber et al. 2011). Studien zu Einstellungsveränderung untersuchen der Herausbildung von Einstellungen, zum Beispiel Ergebnisse zielgenauer universitärer Lehrangebote oder die Entwicklung im Rahmen des allgemeinen universitären für alle Studierenden zugänglichen Lehrangebots (z. B. Tait/Purdie 2000). Untersuchungen zu Einstel-

lungen von angehenden Lehrpersonen, die Determinanten und Zusammenhänge betrachten, fokussieren beispielsweise Merkmale auf Individualebene und dessen Bedeutung für die Einstellungen von (angehenden) Lehrkräften (z. B. Gebauer/McElvany 2017).

Komponenten und Veränderung von Einstellungen

Einstellungen können, nach einer sehr allgemeinen Definition von Eagly und Chaiken (1993) als „psychological tendency, expressed by evaluating a particular entity with some degree of favor or disfavor" (Eagly/Chaiken 1993, S. 1) beschrieben werden. Es liegen jedoch eine Reihe an Definitionen von Einstellungen vor (z. B. Fazio 2007; Visser/Mirabile 2004), die sich vor allem in ihrer Auffassung hinsichtlich der Kontinuität und Stabilität von Einstellungen über Zeit und Situationen hinweg unterscheiden. Diese unterschiedlichen theoretischen Auffassungen über zeitliche Kontinuität und situative Stabilität führen zu differenziellen Annahmen darüber, ob Einstellungsveränderungen möglich und gegebenenfalls sogar steuerbar sind sowie, ob implizite oder explizite Messverfahren genutzt werden müssen, um Einstellungen valide und reliabel zu erfassen (vgl. Bohner/Dickel 2011). Einstellungen können demnach Bohner und Dickel (2011) zufolge als objektevaluierende Assoziationen (Fazio 1995; 2007) verstanden werden, die langfristig erinnert und automatisch abgerufen (z. B. MODE Model; Fazio 2007) oder je nach situativem Kontext neu konstruiert werden (z. B. APE Model; Gawronski/Bodenhausen 2007).

Weitgehende Einigkeit besteht dahingehend, dass Einstellungen objektbezogene Bewertungen und Urteile sind, die in kognitive, affektive und behaviorale Komponenten unterschieden werden können (Fazio 2007; Maio/Haddock 2014). Dabei führen kognitive, affektiv oder behaviorale Assoziationen, Wahrnehmungen oder Erfahrungen in der Summe zu ablehnenden negativen sowie zustimmenden positiven Einstellungen, die sich wiederum in kognitive, affektive und behaviorale Reaktionen niederschlagen können (Eagly/Chaiken 1993). Überzeugungen können dementsprechend als die kognitive Einstellungskomponente eines Urteils betrachtet und sind Bewertungen eines Objekts aufgrund von Wissen und Erkenntnissen (Eagly/Chaiken 1993; Maio/Haddock 2010). Überzeugungen können in Abhängigkeit zu einem erwarteten Ergebnis stehen und zum Beispiel die Nützlichkeit aufgrund von Kenntnis über eine Handlung oder Verhaltensweise bewerten (Ajzen 2001). Die Entwicklung von Überzeugung ist unter anderem über konkrete Wissensvermittlung und Wissenserwerb möglich (Eagly/Chaiken 1993). Positive Emotionen wie Freude und negative Emotionen wie Antipathie oder Abneigung, die in Verbindung mit einem Objekt empfunden werden, zählen zur affektiven Einstellungskomponente. Das können zum Beispiel negative Empfindungen wie Angst vor fremden Perso-

nengruppen mit anderen kulturellen Wertvorstellungen und Verhaltensweisen sein. Die sogenannte Kontakthypothese (vgl. Allport 1954) beschreibt beispielsweise, wie negative Einstellungen in Bezug auf Personengruppen durch positive Erlebnisse, Freude und Spaß im Austausch mit der negativ bewerteten Personengruppe zu positiveren Emotionen führen (Kenworthy/Turner/Hewstone/Voci 2005). Zur behavioralen Komponente zählt eigenes Verhalten aufgrund von Einstellungen. Basierend auf Selbstwahrnehmungen vorangegangener Verhaltensweisen (Bem 1972) die aufgrund spezifischer Einstellungen gezeigt wurden, manifestieren Personen diese Einstellungen und formen neue Bewertungen (Maio/Haddock 2010).

Einstellungen von angehenden Lehrkräften

Einstellungen (und Überzeugungen) von angehenden Lehrkräften stehen in der empirischen Bildungsforschung zunehmend im Fokus (z. B. Fischer/Ehmke 2019; Syring/Tillmann/Weiß/Kiel 2018). Davon ausgehend, dass Einstellungen, neben der Kompetenzüberzeugung, wie die Selbstwirksamkeitsüberzeugung von Bedeutung für intendiertes und gezeigtes Verhalten ist (Ajzen 2001; Bandura 1997), wird angenommen, dass die Einstellungen von (angehenden) Lehrkräften in Bezug auf Lehren und Lernen ebenso relevant für die Gestaltung des Unterrichts sind. Lehrkräfte, die beispielsweise eher konstruktivistische lehr- und lerntheoretische Überzeugungen haben, bieten einen eher konstruktivistischen Unterricht an, der von Vorteil für die Lern- und Leistungsentwicklung von Schülerinnen und Schülern sein kann (Stern/Staub 2002). Des Weiteren sind, vor dem Hintergrund zunehmender Heterogenität der Schülerinnen- und Schülerschaft, Einstellungen in Bezug auf Schülerinnen- und Schülermerkmale vermehrt Untersuchungsgegenstand (z. B. Merk/Cramer/Dai/Bohl/Syring 2018) oder aber die Implementierung von Bildungsreformen, wie beispielsweise aktuell, schulische Entwicklungen im Zuge von Einführungen und Nutzung neuer technologischer Unterrichtsmedien (z. B. Lorenz 2018). Dabei kann davon ausgegangen werden, dass positivere Einstellungen in Bezug zu dem Einstellungsobjekt zu einem positiveren Umgang im Unterricht oder im schulischen Alltag führen können (Gebauer/McElvany 2017). In Abhängigkeit davon, welcher theoretische Rahmen zugrunde gelegt wird, kann eine zielgenaue Einstellungsveränderung durch entsprechende Maßnahmen erwartet werden (vgl. Bohner/Dickel 2011). Wissen über das Einstellungsobjekt oder positive affektive Erfahrungen im Umgang damit können zu positiveren Einstellungen führen (Eagly/Chaiken 1993), jedoch fehlen Erkenntnisse zur Strukturierung solcher Angebote sowie darüber, in welcher Abhängigkeit solche Angebote zu dem jeweiligen Einstellungsobjekt stehen. Darüber hinaus werden neben den Untersuchungen zu Einstellungen und Einstellungsobjekt sowie zu dessen Entwick-

lung auch Determinanten und Faktoren auf Seiten der (angehenden) Lehrkräfte untersucht, wie beispielsweise individuelle oder soziale Merkmale der (angehenden) Lehrkräfte, die die Ausprägung der Einstellungen oder die Herausbildungen moderieren oder mediieren (Merk et al. 2018). Systematische Untersuchungen zu relevanten Faktoren unter Berücksichtigung der Einstellungsobjekte stehen ebenfalls noch aus.

Aktuelle Forschungsbeiträge

Die Beiträge der AEPF Tagung in Lüneburg knüpfen an den aktuellen Forschungsstand an und liefern Untersuchungsbeiträge, die sich mit der zielgenauen Einstellungsveränderung und Herausbildungsverlauf über das reguläre Studium befassen oder den Blick auf Determinanten und Korrelate richten, die für die Einstellungsausprägung relevant sind. Zudem zeichnen sich die Beiträge durch ihre Vielfalt hinsichtlich der Einstellungsobjekte aus und greifen somit aktuelle Themen der Bildungsforschung auf (vgl. McElvany/Gebauer/Gräsel 2019), neben der Berücksichtigung der Leistungsheterogenität von Schülerinnen und Schüler ist das Unterrichten von inklusiven Klassen ein Objektbezug, der sich auf konkrete Schülerinnen und Schülermerkmale richtet. Generelle Vorstellungen über das Lehren und Lernen sowie Einstellungen in Bezug auf schulische Entwicklungsprozesse, zur Nutzung digitaler Medien im Unterricht sind weitere Objektbezüge, die an aktuelle Diskussionen und Reformen in der Bildungsforschung anschließen.

Veränderung und Entwicklung von Einstellungen

Die Untersuchungen von Einstellungsentwicklungen oder die Analyse zielgenauer Einstellungsveränderungen sind insbesondere im Rahmen der universitären Ausbildungsphase von angehenden Lehrkräften von Bedeutung. In dieser Ausbildungsphase erfolgt die fachliche und überfachliche Vorbereitung auf den späteren beruflichen Alltag, die auch im Rahmen von angeleiteten Praxisphasen (König/Tachtsoglou/Darge/Lünnemann 2014) stattfinden. Wenig ist bisher bekannt, wie Einstellungen während des Studiums entstehen oder herausgebildet werden können. Infolgedessen richten folgende Beiträge den Forschungsfokus auf ebendiese Phase und die Entwicklung oder Veränderungen von Einstellungen angehender Lehrkräfte in der ersten Ausbildungsphase.

Der Beitrag von Zaruba, Gronostaj, Ahlgrimm und Vock (1) betrachtet die Entwicklung von Überzeugungen in Bezug auf das Lehren und Lehren während des Praxissemesters. Es werden Überzeugungen untersucht, die in transmissive und konstruktivistische Vorstellungen unterschieden werden können und es wird gezeigt, unter welchen Bedingungen sich diese über den Verlauf des Pra-

xissemsters entwickeln. Mit leitfadengestützten Interviews untersuchen die Autorinnen und Autoren die selbstwahrgenommenen Veränderungen von Überzeugungen in Bezug auf das Lehren und Lernen. Annahme der Autorinnen und des Autors ist, dass die Studierenden aufgrund von Konvergenz- oder Divergenzerfahrungen eine Bestätigung der eigenen Überzeugungen erleben und diese stabil bleiben wenn aber keine konstruktive Auseinandersetzung stattfindet, verwerfen die Studierenden ihre Überzeugungen auch und erleben eine Art „Praxisschock". Letzteres erfolgt durch eine konstruktive Auseinandersetzung von praktischen Erlebnissen, die konträr zu den eigenen Überzeugungen stehen. Es konnten in dieser Studie stabile und adaptive Überzeugungsentwicklungen festgestellt werden, wobei die Überzeugungsadaption im Falle von Divergenzerfahrungen stattfand und eine geringfügige Anpassung oder komplette Revidierung der eigenen Überzeugungen stattfanden. Bemerkenswertes Ergebnis dieser Studie ist unter anderem, dass einige der befragten Studierenden, trotzdem sie das Lehren als Berufsziel anstreben, keine Vorstellungen über den Prozess der Wissensaneignung oder -vermittlung haben. Die Ergebnisse dieser Studie liefern vielversprechende Hinweise darauf, wie Praxiserfahrungen reflektiert, gedeutet und gegebenenfalls umgedeutet werden können und dies möglicherweise gezielt im Rahmen der begleiteten Praxisphasen, durch die Schaffung gemäßigter kognitiver Dissonanzerlebnisse, angewendet werden kann.

Lehmann-Grube, Hartinger und Dresel richten in ihrem Beitrag (2) den Blick auf die Herausbildung von Einstellungen während eines zielgenauen Studienangebots. Vor dem Hintergrund der diagnostischen Leistungen, die Lehrkräfte im beruflichen Alltag kontinuierlich erbringen müssen heben die Autorin und Autoren der Studie die Bedeutung von Einstellungen in Bezug auf Leistungsheterogenität hervor. Gleichzeitig betonen sie die Relevanz der universitären Ausbildungsphase für die Herausbildung von auf Leistungsheterogenität bezogenen Einstellungen und legen dar, wie angehende Lehrkräfte in entsprechend konzertierten Lehrveranstaltungen dahingehend sensibilisiert werden können, um großer Leistungsheterogenität von Schülerinnen und Schülern positiv zu begegnen. Anhand einer Prä-Postinterventionsstudie wird die Wirkung der Fokussierung auf Leistungsheterogenität im Rahmen von Lehrveranstaltungen untersucht, die durch Wissensvermittlung und Bereitstellung konkreter Handlungsmuster zum Umgang damit Einstellungen verändern sollen. Die Ergebnisse weisen auf ein Gelingen und positive Veränderungen in einigen Bereichen der differenziell erfassten Einstellungsfacetten hin. Beachtlich ist an dieser Studie neben dem Erkenntnisgewinn, dass Einstellungen positiv und zielgenau verändert werden können, die hohe Anzahl an Seminaren aus verschiedenen Disziplinen der Lehrerbildung, die im Rahmen dieser Studie die Maßnahme erfolgreich durchführten.

Scharenberg und Opalinski fokussieren in ihrem Beitrag (3) mit einem Prä-Post-Interventionsdesign die Veränderung von inklusionsbezogenen Überzeugungen, die durch ein entsprechendes Lehrangebot initiiert werden soll. Die Autorinnen stellen die Überprüfung eines Seminarkonzeptes zur Diskussion, in dem deklarativer Wissenserwerb von Studierenden mit schulpraktischen Erfahrungen verknüpft und angeleitet reflektiert wird. Die Auswertungen der Daten von rund einhundert Studierenden weisen darauf hin, dass, trotz der Annahme, dass Überzeugungen relativ stabil und schwer veränderbare Vorstellungen sind, eine zielgenaue Wissensvermittlung in Verbindung mit schulpraktischen Erfahrungen und angeleiteten Reflexionsprozessen zu positiveren inklusionsbezogenen Überzeugungen führen. Bemerkenswert sind die subgruppenspezifischen Ergebnisse der Studie, die differenzielle Zuwächse für weibliche und männliche Studierende in den unterschiedlichen Bildungsgängen zeigen und vielversprechende anschließende Forschungsdesiderate aufzeigen.

Determinanten, Korrelate und Typisierung

Disparitätsmerkmale, wie unter anderem das Geschlecht oder der studierte Bildungsgang, führen in einer Reihe an Untersuchungen zu Einstellungen zu differenziellen Ergebnissen für die untersuchten Subgruppen (z. B. Gebauer/McElvany 2017). Gründe dafür können sein, dass Frauen generell zu positiveren Einstellungen neigen oder ein spezieller studierter Bildungsgang zu unterschiedlichen Wissensbeständen führt, die zu unterschiedlichen Einstellungen führen. Studien, die weitere Determinanten von (angehenden) Lehrkräften untersuchen und differenzielle Ergebnisse berichten, sind die folgenden AEPF-Beiträge.

Hellmich, Loeper und Hoya untersuchen in ihrem Beitrag (4) die differenzielle Bedeutung des Studiengangs und der Studienphasen für die Einstellungen von angehenden Lehrkräften in der universitären Ausbildungsphase in Bezug auf einen aktuellen und bedeutungsvollen Einstellungsobjektbezug. Mit einem querschnittlichen Design und aussagekräftigen Stichprobenumfang liefern die vergleichenden Analysen wichtige Ergebnisse und Hinweise für anschließende Forschungen zu Berufswahlmotiven oder der Gestaltung der Lehrerinnen- und Lehrerbildungsforschung. Hervorzuheben ist bezüglich dieser Ergebnisse, dass Einstellungen bereits zu einem frühen Zeitpunkt des Studiums differenziell messbar sind und es scheint, dass Studierenden bereits in der Bachelorstudienphase der mögliche Mehrwert einer inklusiven Bildung für Schülerinnen und Schüler bewusst ist.

Willems und Meyer richten in ihrem Beitrag mit der Wahl des Einstellungsobjekts den Blick auf ein hochrelevantes Thema: die Inklusion von Schülerinnen und Schülern mit Behinderungen und sonderpädagogischen Förderbedarfen im regulären Schulbetrieb (5). Dabei untersuchen die Autorinnen die

Bedeutung von Studienwahlentscheidung, Kontakt zu Menschen mit Behinderungen, Fächer und Schulformwahl für proinklusive Einstellungen und subjektives Kompetenzerleben im Umgang mit Menschen mit Behinderung von angehenden Lehrkräften. Die für Formen der Behinderungen differenziell untersuchten Daten von Studierenden an mehreren Universitäten zeigen, dass Erfahrungen, die vor dem Studium stattfanden von Relevanz für positive auf Inklusion bezogene Einstellungen sind. Die Befunde dieser Studie unterstreichen die Bedeutung angeleiteter Praxisphasen während der universitären Ausbildungsphase in inklusiven schulischen Kontexten.

Syring, Merk und Bohl betrachten in ihrem Beitrag (6) abermals die Leistungsheterogenität von Schülerinnen und Schülern. Die Autoren legen die Bedeutung der Untersuchung von Einstellungen von Lehrkräften im schulischen Dienst in Bezug auf die Leistungsheterogenität von Schülerinnen und Schülern dar und betonen die Einschätzungen höher wahrgenommener Kosten und geringer zugeschriebenen Mehrwerts von Lehrkräften in Bezug auf diese Heterogenitätsform. Mit einer umfangreichen und aussagkräftigen Stichprobe untersuchen die Autoren die Einstellungen von Lehrkräften an Schulen mit äußerer Leistungsdifferenzierung, im Vergleich zu Lehrkräften an Gemeinschaftsschulen, bei denen mehrere Bildungsabschlüsse erreicht werden können. Grundsätzlich könnte bei Schulen mit äußerer Differenzierung eine höhere Leistungshomogenität der Schülerschaft erwartet werden. Mit anspruchsvollen Analysen, die unter Berücksichtigung von möglichen Korrelaten, einen Gruppenvergleich ermöglichen, zeigen die Ergebnisse, dass Lehrkräfte an Gemeinschaftsschulen, bei gleicher objektiver Leistungsheterogenität der Schülerschaft, positivere Einstellungen in Bezug auf den Umgang mit Leistungsheterogenen Schülergruppen aufweisen. Die Analysen unterstreichen indirekt, dass auch aufgrund schulstruktureller äußerer Differenzierungsmaßnahmen die Leistungsheterogenität keineswegs reduziert werden kann und (angehende) Lehrkräfte für diese Herausforderungen entsprechend vorbereitet werden müssen.

Drossel und Eickelmann richten den Forschungsfokus ebenfalls auf ein aktuell hochrelevantes Einstellungsobjekt: die Implementation digitaler Medien in unterrichtliche Lehr- und Lernprozesse (7). Die Autorinnen untersuchen mit Daten der IEA-Studie ICILS 2013 die Einstellungen von Lehrkräften im schulischen Dienst in Bezug auf die Nutzung digitaler Medien. Mit anspruchsvollen international vergleichenden Analysen können spezifische Einstellungstypen identifiziert werden, die sich hinsichtlich ihres Nutzungsverhaltens von digitalen Medien im Fremdsprachenunterricht unterscheiden. Im europäischen Vergleich kann für das Nutzungsverhalten unter Berücksichtigung der Einstellungen von Fremdsprachlehrkräfte an deutschen Schulen eine eher geringere Nutzungshäufigkeit festgestellt werden. Die differenzielle Betrachtung des Nutzungsverhaltens von Lehrkräften unter Berücksichtigung ihrer Einstellungen liefern wichtige Hinweise für die erfolgreiche Implementierung von schuli-

schen Entwicklungen und kann als eine relevante Gelingensbedingung betrachtet werden.

Insgesamt greifen alle skizzierten Untersuchungen mit der Wahl der Einstellungsobjekte aktuelle Themen der Empirischen Bildungsforschung auf. Gleichzeitig tragen die Studien substanziell zum theoretischen Erkenntnisgewinn bei. Sie liefern Hinweise darauf, wie und unter welchen Bedingungen Einstellungen veränderbar sowie welche Determinanten die Einstellungen von (angehenden) Lehrkräften von Relevanz sind.

Literatur

Ajzen, I. (2001): Nature and operation of attitudes. Annual Review of Psychology, 52(1) 27–58.

Allport, G. W. (1954): The Nature of Prejudice. Reading, MA: Addison-Wesley.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York, NY: W. H. Freeman.

Bem, D. J. (1972): Self-perception theory. In Advances in Experimental Social Psychology (Vol. 6, pp. 1–62). Academic Press.

Bohner, G. & Dickel, N. (2011): Attitudes and attitude change. Annual Review of Psychology, 62, 391–417.

Eagly, A. H. & Chaiken, S. (1993): The Psychology of Attitudes. Harcourt Brace Jovanovich College Publishers.

Fazio, R. H. (1995): Attitudes as object-evaluation associations: Determinants, consequences, And correlates of attitude accessibility. Attitude Strength: Antecedents and consequences, 4 247–282.

Fazio, R. H. (2007): Attitudes as object-evaluation associations of varying strength. Social Cognition 25(5), 603–637.

Fischer, N. & Ehmke, T. (2019): Empirische Erfassung eines „messy constructs“: Überzeugungen angehender Lehrkräfte zu sprachlich-kultureller Heterogenität in Schule und Unterricht. Zeitschrift für Erziehungswissenschaft 1–23.

Gebauer, M. M. & McElvany, N. (2017): Zur Bedeutsamkeit unterrichtsbezogener heterogenitätsspezifischer Einstellungen angehender Lehrkräfte für zukünftiges Unterrichtsverhalten. Psychologie in Erziehung und Unterricht, 64(3), 163–180.

Goldhaber, D., DeArmond, M. & DeBurgomaster, S. (2011). Teacher attitudes about compensation reform: Implications for reform implementation. ILR Review, 64(3), 441–463.

Gawronski, B. & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. Social Cognition 25(5), 687–717.

Hachfeld, A., Schroeder, S., Anders, Y., Hahn, A. & Kunter, M. (2012): Multikulturelle Überzeugungen. Zeitschrift für Pädagogische Psychologie.

Jennings, P. A. (2015): Early childhood teachers' well-being, mindfulness, and self-compassion in relation to classroom quality and attitudes towards challenging students. Mindfulness, 6(4), 732–743.

Kenworthy, J. B., Turner, R. N., Hewstone, M. & Voci, A. (2005): Intergroup contact: When does it work, and why. On the Nature of Prejudice: Fifty years after Allport 278–292.

König, J., Tachtsoglou, S., Darge, K. & Lünnemann, M. (2014): Zur Nutzung von Praxis: Modellierung und Validierung lernprozessbezogener Tätigkeiten von angehenden Lehrkräften im Rahmen ihrer schulpraktischen Ausbildung. Zeitschrift für Bildungsforschung, 4(1), 3–22.

Lorenz, R. (2018): Ressourcen, Einstellungen und Lehrkraftbildung im Bereich Digitalisierung. Digitalisierung in der schulischen Bildung: Chancen und Herausforderungen, 53.

Maio, G. R. & Haddock, G. (2010): The Psychology of Attitudes and Attitude Change. London: Sage.

McElvany, N., Gebauer, M. M. & Gräsel, C. (Hrsg.). (2019): Rückblick, Vorausschau und aktuelle Ergebnisse der Empirischen Bildungsforschung in Deutschland. Themenheft des Journal for Educational Research Online 11 (1).

Merk, S., Cramer, C., Dai, N., Bohl, T. & Syring, M. (2018): Faktorielle Validität der Einstellungen von Lehrkräften zu heterogenen Lerngruppen. Journal for Educational Research Online 10(2), 34–53.

Skott, J. (2015): The promises, problems, and prospects of research on teachers' beliefs. In H. Fives & M. G. Gill International Handbook of Research on Teachers' Beliefs 1, 37–54.

Syring, M., Tillmann, T., Weiß, S. & Kiel, E. (2018): Empirische Arbeit: Positive Einstellung zur Inklusion-ablehnende Haltung zur Umsetzung in der Schule. Analyse des Widerspruchs durch Überprüfung eines aus der Heterogenitätsforschung adaptierten Messinstruments für die Inklusion an Lehramtsstudierenden. Psychologie in Erziehung und Unterricht, 65(3) 206–220.

Staub, F. C. & Stern, E. (2002): The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. Journal of Educational Psychology, 94(2), 344.

Tait, K. & Purdie, N. (2000): Attitudes toward disability: Teacher education for inclusive environments in an Australian university. International Journal of Disability, Development and Education, 47(1) 25–38.

Visser, P. S. & Mirabile, R. R. (2004): Attitudes in the social context: the impact of social network composition on individual-level attitude strength. Journal of Personality and Social psychology, 87(6), 779.

Unter welchen Bedingungen entwickeln sich Überzeugungen im Praxissemester?

Eine Interviewstudie

Nicole Zaruba, Anna Gronostaj, Frederik Ahlgrimm und Miriam Vock

Abstract

Überzeugungen zum Lehren und Lernen sind als Teil der professionellen Kompetenz von Lehrkräften in Praxisphasen im Lehramtsstudium von Bedeutung. Anhand von Interviews mit $N = 16$ Studierenden haben wir die Entwicklung der Überzeugungen von Studierenden im Potsdamer Praxissemester im Zusammenhang mit ihren Praxiserfahrungen untersucht. Mittels qualitativer Inhaltsanalyse konnten wir drei Entwicklungswege identifizieren: Überzeugungen wurden zum ersten Mal relevant, blieben stabil oder veränderten sich. Ein wichtiger Faktor ist dabei der Umgang der Studierenden mit Praxiserfahrungen, die zu ihren Überzeugungen im Widerspruch stehen. Nach einem konstruktiven Umgang mit diesen Divergenzerfahrungen kam es zu einer leichten Adaption der Überzeugungen; ein Verwerfen der Überzeugungen führte zu einem Praxisschock. Überzeugungen blieben stabil, wenn die Praxiserfahrung Überzeugungen bestätigte oder wenn die praktische Umsetzung der Überzeugung noch nicht gelang, die Überzeugung aber eher abstrakt beibehalten wurde.

1 Einleitung

Praxisphasen im Lehramtsstudium werden von Studierenden als bedeutsame Lerngelegenheiten wahrgenommen (vgl. Boekhoff et al. 2008, S. 53). Ob und wie Praxisphasen aber tatsächlich zur Kompetenzentwicklung im Lehramtsstudium beitragen, ist bislang nur unzureichend erforscht (vgl. Hascher 2006, S. 130 ff.). Konzepte zur Lernbegleitung durch universitäre Begleitseminare und Mentorinnen und Mentoren in der Schule sollen den Studierenden Gelegenheit zur Praxisreflexion geben. Empirische Befunde zeigen jedoch, dass Praxisphasen nicht automatisch mit einer Kompetenzentwicklung einhergehen (vgl.

Rothland/Boecker 2015, S. 119), sondern sie auch die Gefahr unerwünschter Effekte bergen; z. B. dass Studierende eine problematische Praxis unreflektiert imitieren (vgl. Bach 2013, S. 123).

Überzeugungen zum Lehren und Lernen sind Teil der professionellen Handlungskompetenz von Lehrkräften (vgl. Baumert/Kunter 2006, S. 482) und entwickeln sich insbesondere in solchen Ausbildungsformaten, in denen schulpraktische Erfahrungen gesammelt werden können (vgl. Levin 2015, S. 259). In einer Vorläuferstudie hatten wir bereits gefunden, dass sich unterrichtsbezogene Überzeugungen von Studierenden im Praxissemester individuell unterschiedlich entwickeln und teilweise auch stabil bleiben (vgl. Zaruba et al. 2018, S. 232). Wie diese komplexen Entwicklungsprozesse (vgl. Fives/Buehl 2012, S. 486 f.) ablaufen, untersuchen wir im vorliegenden Beitrag. Im Folgenden stellen wir dar, welche Entwicklungsprozesse sich in Reflexionen Studierender zu ihren Überzeugungen identifizieren lassen und welche Erfahrungen den unterschiedlichen Entwicklungen zugrunde liegen.

2 Theoretischer Hintergrund

2.1 Überzeugungen von Lehrkräften zum Lehren und Lernen

Überzeugungen von Lehrkräften (*teacher beliefs*) sind individuelle mentale Konstrukte (vgl. Skott 2015, S. 18), die als Teil der professionellen Handlungskompetenz von Lehrerinnen und Lehrern angesehen werden, und „Zielvorstellungen [beschreiben], die präskriptive Richtungsweiser für Unterrichtsplanung und Unterrichtshandeln sind“ (Baumert/Kunter 2006, S. 482/497). Dabei beinhalten Überzeugungen immer subjektive und bewertende Elemente (Kunter/Pohlmann 2015, S. 267). Überzeugungen werden meist in Abgrenzung zu Wissen von Lehrkräften betrachtet. Unter Wissen werden „grundsätzlich erlernbare Fähigkeiten […] [verstanden], die zeitlich und situationsübergreifend relativ stabil […] und […] mithilfe testdiagnostischer Verfahren einer Erfassung zugänglich sind“ (König et al. 2017, o. S.). Im Unterschied zu Wissen müssen Überzeugungen nicht interindividuell verifizierbar sein, sondern nur dem subjektiven Rechtfertigungsglauben der Lehrkraft genügen.

Oft werden die Konstrukte nicht klar voneinander abgegrenzt; so werden Überzeugungen manchmal synonym zu Einstellungen oder subjektiven Theorien von Lehrkräften gebraucht (vgl. Fussangel 2008, S. 72). Der Begriff „Einstellungen“ fokussiert dabei die Bewertung eines Objektes, auf das sich die Einstellung bezieht (vgl. Kessels/Erbring/Heiermann 2014, S. 190), „subjektive Theorien“ beziehen sich auf die persönliche Wahrnehmung des Erlebens und Handelns einer Lehrkraft (vgl. Mandl/Huber 1983, S. 98). Einstellungen und subjektive Theorien können als einander überschneidende Konzepte verstan-

den werden, die unter den Begriff „Überzeugungen“ subsumiert werden können (vgl. Fussangel 2008, S. 72–78).

Im Lehramtsstudium verdient die Prägung von Überzeugungen besonderes Augenmerk, da die sich entwickelnden Überzeugungen später das unterrichtliche Handeln und die Wahrnehmung der Lehrkräfte beeinflussen können (vgl. Voss et al. 2011, S. 238 f.) sowie sich auf die Leistung von Schülerinnen und Schülern auswirken können (Staub/Stern 2002, S. 354).

Lehrkräfte haben Überzeugungen zu einer Vielzahl berufsrelevanter Gebiete, z. B. zur Fachdidaktik ihrer Fächer (vgl. z. B. Lavy/Shriki 2008, S. 259 ff.), oder fächerübergreifend über Lehr-Lern-Prozesse im Unterricht (vgl. Fives/Lacatena/Gerard 2015, S. 249 ff.). Diese unterrichtsbezogenen Überzeugungen werden theoretisch oft in zwei unterschiedliche Sichtweisen unterteilt: in eine konstruktivistische und eine transmissive Sicht auf Lehren und Lernen. Beide lassen sich auch empirisch finden (vgl. Fives/Lacatena/Gerard 2015, S. 253 ff.). Lehrkräfte mit konstruktivistischen Überzeugungen gehen davon aus, dass Lernen ein individueller Prozess ist, der auf einer aktiven Konstruktion von Wissen beruht (vgl. Richardson 1996, S. 105). Eine transmissive Sicht auf Lehren und Lernen hingegen betrachtet Schülerinnen und Schüler als passive Rezipientinnen bzw. Rezipienten eines klar definierten Wissenskanons (vgl. Voss et al. 2011, S. 238 f.). Im internationalen Vergleich zeigt sich, dass Primarstufenlehrkräfte in solchen Ländern stärker konstruktivistische Überzeugungen aufwiesen, in denen die Gesellschaft als individualistisch beschrieben wird (z. B. Deutschland und Schweiz) (vgl. Felbrich/Schmotz/Kaiser 2010, S. 324 f.). Empirisch gehen transmissive Lehrkräfteüberzeugungen mit einer weniger lernförderlichen Unterrichtsgestaltung einher (vgl. Dubberke et al. 2008, S. 201), konstruktivistische hingegen mit Unterrichtsqualität und dem Lernerfolg der Schülerinnen und Schüler (vgl. Voss et al. 2011, S. 250).

Die Konzepte Transmission und Konstruktivismus wurden von anderen Autorinnen und Autoren weiter ausdifferenziert und erweitert. Für den Naturwissenschaftsunterricht arbeitete Kleickmann (2008) beispielsweise fünf Überzeugungskonstrukte heraus, die sich den Überzeugungen über Lehr-Lern-Prozesse zuordnen lassen. Neben (1) *Transmission* beschrieb er vier weitere Facetten: (2) *Laisser-faire und Praktizismus*, (3) *Conceptual-Change-Lernen*, (4) *motiviertes und anwendungsbezogenes Lernen* sowie (5) *Entwicklung eigener Deutungen und Diskussion von Schülervorstellungen*. Die Überzeugungsfacetten 3–5 lassen sich einer konstruktivistischen Sicht auf Lehren und Lernen zuordnen. *Laisser-faire* und *Praktizismus* werden hingegen als ein falsch verstandener Konstruktivismus interpretiert (vgl. Mayer 2004, S. 17). Auch wenn sich die beschriebenen Überzeugungen prinzipiell konzeptuell trennen lassen, so ist das Überzeugungssystem von Lehrkräften meist komplex, mehrdimensional und auch inkongruent, sodass Lehrkräfte auch widersprüchliche Überzeugungen haben können (vgl. Fives/Buehl 2012, S. 486 ff.).

2.2 Die Entwicklung von Lehrerüberzeugungen in Praxisphasen

Überzeugungen von Lehrkräften sind während des Studiums und zu Beginn der praktischen Ausbildung noch eher unspezifische, wenig elaborierte Vorstellungen vom Lehren und Lernen, die stark von eigenen Erfahrungen als Schülerin bzw. Schüler geprägt sind (vgl. Fives/Lacatena/Gerard 2015, S. 257). Zentral für die Weiterentwicklung von Überzeugungen scheinen Felderfahrungen zu sein (vgl. Cook/Young 2004, S. 355 ff.).

Die empirischen Befunde zur Entwicklung von Überzeugungen in Praxisphasen zeichnen aber kein einheitliches Bild: Eine Langzeitstudie von Müller-Fohrbrodt/Cloetta/Dann (1978) zeigt, dass angehende Lehrkräfte in den 1970er Jahren mit zunehmender Praxiserfahrung ein rezeptiveres Verständnis von Lehren und Lernen entwickelten. Decker/Kunter/Voss (2015, S. 51 ff.) fanden hingegen bei Referendarinnen und Referendaren, dass diese durch intensive Reflexion stärkere konstruktivistische Überzeugungen entwickelten. Es ist ebenfalls plausibel, dass sich Praxisphasen gar nicht auf die Überzeugungen der Studierenden auswirken (vgl. Hascher 2006, S. 133). Zaruba et al. (2018, S. 232) fanden für Lehramtsstudierende im Praxissemester alle drei beschriebenen Verläufe: Veränderung hin zu stärker transmissiver Überzeugung, Veränderung hin zu stärker konstruktivistischer Überzeugung, keine Veränderung.

Auch Yerrick und Hoving (2003, S. 399 ff.) stellten fest, dass Studierende mit ihren Felderfahrungen in der Schule unterschiedlich umgehen und beschreiben zwei Entwicklungstypen: Studierende, die den transmissiven Lehrstil ihrer eigenen Schulzeit reproduzieren und die Veränderung ihrer Überzeugungen ablehnen, und solche, die durch Reflexion ihre Praxis und Überzeugungen adaptieren.

Das *Cognitive-Affective Model of Belief Change* von Gregoire (2003, S. 165) beschreibt, wie sich Überzeugungen von Lehrkräften verändern. Wie (angehende) Lehrkräfte mit Erfahrungen und Impulsen umgehen, die nicht ihrem aktuellen Überzeugungskanon entsprechen, wird in diesem Modell als zentraler Faktor – neben weiteren Kontextfaktoren wie Emotionen und Motivation – für die Entwicklung von Überzeugungen genannt (ebd.). Ausgehend von der Theorie der kognitiven Dissonanz (vgl. Festinger 1978/2012, S. 15 ff.) kann angenommen werden, dass Studierende, die ihre Erfahrungen in der Schulpraxis als konvergent mit ihren Überzeugungen bewerten, diese bestätigen und beibehalten. Von den Überzeugungen abweichende, also divergente Erfahrungen, könnten kognitive Dissonanzen auslösen und zu einer Veränderung der Überzeugungen führen (vgl. Gregoire 2003, S. 164). Wie Studierende im Praxissemester die Entwicklung ihrer Überzeugungen in Zusammenhang mit ihren Erfahrungen bewerten, wurde bislang noch nicht empirisch untersucht, weshalb wir in dieser Studie auf den Aspekt der kognitiven Dissonanz fokussieren.

In unserer Studie untersuchen wir, wie sich die Überzeugungen der Studierenden entwickeln und welche Rolle der Umgang mit Praxiserfahrungen spielt. Welche Entwicklungsprozesse berichten Studierende bei der Reflexion ihrer Überzeugungen zum Lehren und Lernen im Verlauf des Praxissemesters, und welche Rolle spielt dabei ihr Umgang mit Praxiserfahrungen?

3 Methodisches Vorgehen

3.1 Das Potsdamer Praxissemester

Seit dem Jahr 2008 absolvieren an der Universität Potsdam alle Lehramtsstudierenden im Masterstudium ein 16-wöchiges Schulpraktikum, das durch universitäre Seminare vor- und nachbereitet sowie begleitet wird, und in dem sie hospitieren und selbst unterrichten. In der Schule werden sie durch Mentorinnen und Mentoren im Unterricht begleitet.

3.2 Stichprobe und Studiendesign

$N = 16$ Studierende wurden im Sommer 2015 nach Abschluss ihres Praxissemesters interviewt. Alle meldeten sich freiwillig. Die meisten waren weiblich ($n = 11$) und studierten Lehramt für Gymnasien ($n = 11$), je eine Person studierte Grundschullehramt und Lehramt für die Sekundarstufe I. Drei machten keine Angabe. Die Interviews dauerten zwischen 55 und 98 Minuten.

Um die zugrundeliegenden Überzeugungen zu evozieren, wurden die Studierenden mittels Repertory-Grid-Technik (vgl. Kelly 1955, S. 152 ff.), einem Verfahren zur Ermittlung des persönlichen Konstruktsystems, befragt. Im Fokus standen dabei ihre Überzeugungen zu Motiviertem Lernen, Anwendungsbezogenem Lernen, Conceptual-Change Lernen, Transmission, Laisser-faire und Praktizismus. Dafür wurden zwei Vignetten erstellt, die eine eher konstruktivistisch eingestellte und eine transmissiv eingestellte Lehrperson beschreiben. Diesen Vignetten ordneten die Studierenden je zwei Lehrkräfte aus der eigenen Schulzeit zu oder Lehrkräfte, bei denen sie im Studium hospitiert hatten. Im Vergleich mit diesen Lehrkräften formulierten die Studierenden persönliche Konstrukte in den o. g. Überzeugungsfacetten. Z. B. wurden die Studierenden gefragt: „Wenn Sie Frau Meyer und Herrn Schmidt *(nach konstruktivistischer Vignette ausgewählt)* mit Herrn Albrecht *(transmissive Vignette)* vergleichen, welche Rolle spielte der Anwendungsbezug im Unterricht dieser Lehrkräfte?“ Die Antworten wurden nicht korrigiert, so dass auch naive oder fehlerhafte Konstrukte übernommen wurden. Anschließend wurde die Studentin bzw. der Student aufgefordert, eigene Überzeugungen zum persönlichen Konstrukt (z. B. Alltagsbezug) im Vergleich mit den Lehrkräften zu er-

läutern. Danach wurden die Studierenden zur Stabilität und Veränderung ihrer Überzeugungen befragt.

Die Interviews wurden im Anschluss nach dem Transkriptionssystem von Dresing und Pehl (2013, S. 17 ff.) und nach Mayring (2010, S. 62 ff.) inhaltsanalytisch darauf hin ausgewertet, welche Erfahrungen die angegebene Entwicklung beeinflusst haben. Dabei wurden zunächst in einem deduktiven Ansatz alle Textstellen identifiziert, in denen die Entwicklung der jeweiligen Überzeugungsfacette reflektiert wurde. Anschließend wurden Interpretationen der Praxiserfahrungen induktiv kodiert und geclustert. Zur Sicherung der Güte der Ergebnisse nahmen wir fallübergreifende Analysen von Erfahrungsstrukturen vor (vgl. Przyborski/Wohlrab-Sahr 2014, S. 25 ff.). Die kommunikative Validierung der Ergebnisse erfolgte durch *peer-checks* und *member-checks*. Für die Ergebnisdarstellung wurden die Namen der Studierenden pseudonymisiert.

4 Ergebnisse

4.1 Überzeugung wird zum ersten Mal relevant

Einige Studierende hatten vor dem Praxissemester noch keine Überzeugung zu dem besprochenen Thema: „Naja, irgendwie zu Beginn hatte ich nicht wirklich eine Einstellung, weil überhaupt keine Erfahrung“ (Jenny).

4.2 Überzeugungen bleiben stabil

Einige Studierende berichten, dass ihre Überzeugungen stabil geblieben waren. Hier lassen sich zwei unterschiedliche Entwicklungswege identifizieren (s. Abb. 1): Im ersten wird eine Umsetzung in der Praxis als schwierig bewertet; obwohl sie die praktische Umsetzung ihrer Überzeugungen noch nicht als gelungen einstuften, bleiben ihre Überzeugungen stabil: „Mein Anspruch hat sich nicht verändert […], verschiedene Methoden zielgerichtet einzusetzen. Ich habe es aber im Praxissemester oft nicht umsetzen können“ (Alma).

Im zweiten Entwicklungsweg findet sich aufgrund konvergenter Praxiserfahrungen eine Bestätigung der Überzeugungen. Jenny erklärt, dass für sie die Lehrerpersönlichkeit das Wichtigste für erfolgreiches Schülerlernen ist. Auf die Nachfrage, ob sich ihre Überzeugungen im Verlaufe des Praxissemesters verändert haben, antwortet sie: „Nein, höchstens bestätigt. Also am Anfang liest man ja nur irgendwelche Studien, man hat die Theorie so und so ist es und im Praxissemester hat sich das bestätigt […].“

4.3 Überzeugungsadaption nach Divergenzerfahrung

Konstruktiver Umgang: leichte Überzeugungsadaption

Wenn Studierende von einer Veränderung ihrer Überzeugungen berichten, dann sprechen sie meist von einer leichten Anpassung ihrer Überzeugungen hin zu einem „realistische[n] Bild“ (Jenny). Leo beschreibt die Veränderung seiner Überzeugung zum Thema offene Lernangebote und selbstgesteuertes Lernen:

> Das […] hat sich tatsächlich geändert, weil ich […] hätte am Anfang eher mich hier eingeordnet, also viel frei anwenden und so. Aber [… ich habe] halt festgestellt, dass man die Schüler leicht überfordert, wenn man sie halt immer nur in die freie Anwendung schickt und immer nur sagt: „Ok, jetzt seid mal kreativ und jetzt macht mal.“ Das ist auch nicht jedermanns Sache, das kann man nicht immer machen. Also, das ist auf jeden Fall wichtig […] aber ich glaube so ein Mittelweg ist da gar nicht verkehrt.

Bei diesen Studierenden lässt sich ein eher konstruktiver Umgang mit einer Divergenzerfahrung erkennen. Die Studentin bzw. der Student hat die Dissonanzerfahrung reflektiert und ihre bzw. seine Überzeugung in der Folge angepasst. In Leos Beispiel wird deutlich, dass die Divergenzerfahrung zu einer Adaption der Überzeugung führte. Die Überzeugung wurde nicht völlig infrage gestellt, sondern es wurde „ein Mittelweg“ gefunden.

Praxisschock: Relativierung der Überzeugungen

Frieda lehnt mit Nachdruck das theoretische Wissen aus dem Studium ab und bewertet es als praxisfern. Sie wirkt deutlich desillusioniert:

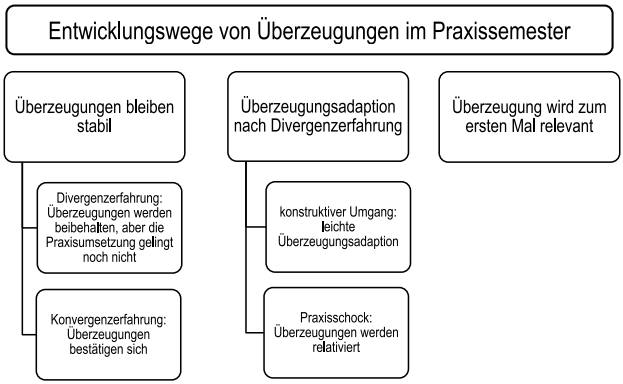
> […] dieses ganze Gerede […] über Selbststeuerung und Kompetenzentwicklung und diesem ganzen theoretischen Unterbau, den ich jetzt hatte […] da sind viele Worthülsen und Phrasen, die im Prinzip nichts mit dem Praxisalltag zu tun haben, super schwer umzusetzen sind und das ist eben diese gekünstelte Freiheit. [Die Schülerinnen und Schüler] würden […] garantiert nicht Lyrik machen oder zumindest nicht alle. […] das steht im Curriculum und ich muss das machen und […] das muss ich denen irgendwie beibringen und dabei tun wir so, als würden wir denen auch noch Kompetenzen beibringen irgendwie oder die darin fördern und das dann auch noch mit den unterschiedlichen Voraussetzungen.

Jenny begründet ihre Ernüchterung mit dem eigenen Anfangsidealismus: „Also, am Anfang […] dieser Idealismus, alle zu begeistern und immer alle mitzunehmen.“

Bei diesen Studierenden kann man auf der Grundlage der berichteten Desillusionierung eine Art „Praxisschock“ feststellen, also ein mehr oder weniger enttäuschtes Revidieren der eigenen progressiven Überzeugung.

Zusammenfassend ließen sich in den Interviews drei übergeordnete Entwicklungsmuster für die Überzeugungen Studierender im Praxissemester feststellen (s. Abb. 1). Entweder die Studierenden berichten, dass sie vor dem Praxissemester keine bewussten Überzeugungen zum Thema hatten und dass die Überzeugung erstmalig durch Erfahrungen im Praxissemester relevant wird. Eine weitere Variante ist, dass die Überzeugungen vorab vorhanden waren und diese im Verlauf des Praxissemesters stabil blieben. Eine dritte Möglichkeit ist, dass sich die Überzeugungen verändern. Stabile Überzeugungen berichten die Studierenden sowohl nach Konvergenz- als auch nach Divergenzerfahrungen. Im Fall der Divergenzerfahrungen wurde die Überzeugung in der Theorie beibehalten, in der Praxis wurde aber noch keine Umsetzungsmöglichkeit gefunden. Nach Divergenzerfahrungen kann es auch zu Überzeugungsadaptionen kommen. Wir fanden bei den Studierenden zwei unterschiedliche Wege, wie sie ihre Überzeugungen anpassten. Entweder adaptierten sie ihre Überzeugungen nur leicht und passten sie der von ihnen wahrgenommenen Realität an. Oder sie revidierten ihre Überzeugungen infolge einer Art Praxisschocks komplett, weil sie ihre ursprünglichen Überzeugungen als idealistisch oder praxisfern bewerteten.

Abbildung 1: Entwicklungswege von Überzeugungen im Praxissemester



5 Diskussion und Implikationen

Überzeugungen zum Lehren und Lernen sind komplexe mentale Konstrukte (vgl. Skott 2015, S. 18; Fives/Buehl 2012, S. 486 ff.), die sich je nach wahrgenommenen Lernangeboten im Praxissemester individuell entwickeln können

(vgl. Zaruba et al. 2018, S. 232). Eine wichtige Rolle in der Entwicklung von Überzeugungen scheint der Umgang mit solchen Erfahrungen zu spielen, die im Widerspruch zur bestehenden Überzeugung stehen (vgl. Gregoire 2003, S. 164). Wie aber Studierende im Praxissemester die Entwicklung ihrer Überzeugungen im Kontext ihrer Praxiserfahrungen bewerten, wurde bislang in der Forschung nicht betrachtet.

In den Interviews konnten wir drei Entwicklungswege von Überzeugungen feststellen. Entweder die Überzeugungen bleiben im Verlauf des Praxissemesters stabil oder sie werden nach Divergenzerfahrungen adaptiert. Eine weitere Möglichkeit ist, dass die Überzeugungen im Kontext des Praxissemesters erstmals relevant werden. Ausgehend vom *Cognitive-Affective Model of Belief Change* (vgl. Gregoire 2003, S. 164) kann angenommen werden, dass Studierende Impulse durch Praxiserfahrungen mit ihren Überzeugungen in Beziehung setzen und individuell unterschiedlich auf erlebte Divergenz und Konvergenz reagieren.

Viele Studierende haben zumindest eine vage Überzeugung zu allen angesprochenen Konstrukten. In einigen Fällen wurde aber deutlich, dass die Studierenden vor dem Praxissemester keine Vorstellungen zum Überzeugungskonstrukt hatten. Dieser Befund ist mit Blick auf die Forschungsmethodik relevant. Häufig werden Überzeugungen mittels Fragebögen erfasst (vgl. z. B. Decker/Kunter/Voss 2015, S. 49). Die Studierenden können aber nicht angeben, dass sie noch keine Überzeugung zu dem Themenbereich haben. Bei der zukünftigen Gestaltung von Fragebögen sollte dieser Befund berücksichtigt werden.

Ein weiteres zentrales Ergebnis ist, dass die Entwicklung der Überzeugungen im Praxissemester mit der Art des Umgangs mit Divergenzerfahrungen einhergeht. Wenn die Überzeugung noch nicht mit Praxiserfahrung verknüpft werden konnte, der Anspruch aber unverändert besteht, bleibt die Überzeugung stabil. Wenn Divergenzerfahrungen als Korrektur einer theoretischen Sicht zugunsten der Integration von Praxiserfahrungen verstanden werden, der grundsätzliche Anspruch aber fortbesteht, kommt es zu einer leichten Anpassung der Überzeugungen. Stehen sich aber Praxiserfahrung und Überzeugung aus Sicht der Studierenden diametral entgegen, lässt sich eine Desillusionierung im Sinne eines Praxisschocks und ein Rückfall in Richtung traditionellerer Überzeugungen beobachten. Werden Überzeugungen beibehalten, aber noch nicht in der Praxis erprobt, ist der Lernprozess noch „on-hold“ und eine Bewertung der Praxistauglichkeit der Überzeugung steht noch aus. Wenn es zu einer leichten Anpassung „an die Realität“ kommt, haben die Studierenden begonnen, Theorie und Praxis dialektisch zu betrachten (vgl. Cramer 2014, S. 353). Sie bewahren den Kern ihrer Überzeugung, gespeist durch den theoretischen Unterbau aus dem Studium und eigenen Schulerfahrungen, versuchen diese aber an das neue Praxiswissen anzupassen. Ausgehend von der Theorie

der kognitiven Dissonanz (vgl. Festinger 1978/2012, S. 15 ff.) kann man annehmen, dass die Studierenden nach Konvergenz ihrer Ansichten und der vorgefundenen Praxis streben. Eine Auflösung der Divergenzerfahrung zugunsten einer Integration beider Sichtweisen kann als konstruktive Auseinandersetzung mit Theorie und Praxis verstanden werden. Diese Auflösung findet beim Praxisschock nicht statt, sondern eher eine unkritische Imitation der Praxis.

Wie können Studierende in der Lernbegleitung besser bei der Auflösung von Divergenzerfahrungen unterstützt werden? Ein Ansatzpunkt wäre eine Schulung der Mentorinnen und Mentoren zum konstruktiven Umgang mit Divergenzerfahrung in der Unterrichtsnachbesprechung. Hier könnten auch Unterrichtskonzepte mit Blick auf das theoretische Vorwissen der Studierenden diskutiert werden. So könnten Studierende darin unterstützt werden, ihre Überzeugungen in der Praxis umzusetzen. Abhängig von der Überzeugung sollten aber auch Konvergenzerfahrungen Gegenstand einer kritischen Reflexion sein. Wenn sich nicht-lernförderliche Überzeugungen in der Praxis bestätigen, sollten Seminare, Hospitationen und schulische Lernbegleitung Lernanlässe durch gemäßigte kognitive Dissonanzen schaffen.

Dabei ist ein sensibles Vorgehen nötig, da ein Übermaß an unaufgelösten kognitiven Dissonanzen im Sinne eines Praxisschocks auch zu einem Zweifel an der eigenen Eignung als Lehrkraft führen könnte (vgl. Festinger 1978/2012, S. 15 ff.). Ein Berufseinstieg mit gefestigtem Selbstkonzept ist angesichts des Belastungsanspruchs an Lehrkräfte (vgl. Rothland 2007, S. 7 ff.) wünschenswert.

Unsere Studie beruht auf Selbstauskünften, die möglicherweise durch sozial erwünschte Aussagen von den tatsächlichen Erfahrungen abweichen. In weiteren Studien sollten daher auch implizite Überzeugungen untersucht werden. Anders als bei Zaruba et al. (2018, S. 232 f.) können wir keine Aussage über die Qualität der Entwicklung treffen. Da die Überzeugungsentwicklungen über alle Fälle und Überzeugungsfacetten hinweg analysiert wurden, kann nicht eingeschätzt werden, was z. B. die Entwicklung lernförderlicher Überzeugungen begünstigt. In einer Folgestudie könnte untersucht werden, welche Arten von Überzeugungen sich bestätigen und welche sich verändern.

Literatur

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

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft, H. 9, S. 469–520.

Boekhoff, Inga/Franke, Kerstin/Dietrich, Fabian/Arnold, Karl-Heinz (2008): „Effektivität universitärer Lehrerbildung in konsekutiven Studiengängen (EduLikS) unter besonderer Berücksichtigung Schulpraktischer Studien". www.uni-hildesheim.de/media/forschung/PK-Unterrichtsforschung/PDFs/2008_Druckversion_Netz_EduLikS_Bericht.pdf (Abfrage: 28.03.2019).

Cook, Paul/Young, Janet (2004): Face-to-face with children. In: Journal of Curriculum Studies, H. 36 (3), S. 341–360.

Cramer, Colin (2014): Theorie und Praxis in der Lehrerbildung: Bestimmung des Verhältnisses durch Synthese von theoretischen Zugängen, empirischen Befunden und Realisierungsformen. In: Die Deutsche Schule, H. 106 (4), S. 344–357.

Decker, Anna-Theresia/Kunter, Mareike/Voss, Thamar (2015): The relationship between quality of discourse during teacher induction classes and beginning teachers' beliefs. In: European Journal of Psychology of Education, H. 30 (1), S. 41–61.

Dresing, Thorsten/Pehl, Thorsten (2013): Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende. 5. Auflage. Marburg: Eigenverlag.

Dubberke, Thamar/Kunter, Mareike/McElvany, Nele/Brunner, Martin/Baumert, Jürgen (2008): Lerntheoretische Überzeugungen von Mathematiklehrkräften: Einflüsse auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern. In: Zeitschrift für Pädagogische Psychologie, H. 22 (34), S. 193–206.

Felbrich, Anja/Schmotz, Christiane/Kaiser, Gabriele (2010): Überzeugungen angehender Primarstufenlehrkräfte. In: Blömeke, Sigrid/Kaiser, Gabriele/Lehmann, Rainer (Hrsg.): TEDS-M 2008. Professionelle Kompetenz und Lernangelegenheiten angehender Primarstufenlehrkräfte im internationalen Vergleich. Münster: Waxmann, S. 297–326.

Festinger, Leon (1978/2012): Theorie der kognitiven Dissonanz. 2. Auflage, Faksimile-Reprint der Erstausgabe. Bern: Verlag Hans Huber.

Fives, Helenrose/Buehl, Michelle (2012): Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In: Harris, Karen R./Graham, Steve/Urdan, Tim (Hrsg.): APA Educational Psychology Handbook. Band 2: Individual differences and cultural and contextual factors. Washington: American Psychological Association, S. 471–499.

Fives, Helenrose/Lacatena, Natalie/Gerard, Laura (2015): Teachers' beliefs about teaching (and learning). In: Fives, Helenrose/Gregoire Gill, Michelle (Hrsg.): International handbook of research on teachers' beliefs. New York: Routledge. S. 249–265.

Fussangel, Kathrin (2008): Subjektive Theorien von Lehrkräften zur Kooperation. Eine Analyse der Zusammenarbeit von Lehrerinnen und Lehrern in Lerngemeinschaften. Dissertation. Wuppertal: Universität Wuppertal.

Gregoire, Michelle (2003): Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. In: Educational Psychology Review, H. 33, S. 147–179.

Hascher, Tina (2006): Veränderungen im Praktikum – Veränderungen durch das Praktikum. Eine empirische Untersuchung zur Wirkung von schulpraktischen Studien in der Lehrerbildung. In: Allemann-Ghionda, Cristina/Terhart, Ewald (Hrsg.): Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf. Weinheim und Basel: Beltz Juventa, S. 130–148.

Kelly, George (1955): The Psychology of Personal Constructs. New York: Norton.

Kessels, Ursula/Erbring, Saskia/Heiermann, Liesel (2014): Implizite Einstellungen von Lehramtsstudierenden zur Inklusion. In: Psychologie in Erziehung und Unterricht, H. 3, S. 189–202.

Kleickmann, Thilo (2008): Zusammenhänge fachspezifischer Vorstellungen von Grundschullehrkräften zum Lehren und Lernen mit Fortschritten von Schülerinnen und Schülern im konzeptuellen naturwissenschaftlichen Verständnis. Dissertation. Münster: Westfälische Wilhelms-Universität Münster.

König, Johannes/Doll, Jörg/Buchholtz, Nils/Förster, Sabrina/Kaspar, Kai/Rühl, Anna-Maria/Strauß, Sarah/Bremerich-Vos, Albert/Fladung, Ilka/Kaiser, Gabriele (2017): Pädagogisches Wissen versus fachdidaktisches Wissen? Struktur des professionellen Wissens bei angehenden Deutsch-, Englisch- und Mathematiklehrkräften im Studium. In: Zeitschrift für Erziehungswissenschaft. www.link.springer.com/content/pdf/10.1007%2Fs11618-017-0765-z.pdf (Abfrage: 03.06.2019).

Kunter, Mareike/Pohlmann, Britta (2015): Lehrer. In: Wild, Elke/Möller, Jens (Hrsg.): Pädagogische Psychologie. Berlin: Springer. S. 261–281.

Lavy, Ilana/Shriki, Atara (2008): Investigating changes in prospective teachers' views of a good teacher while engaging in computerized project-based learning. In: Journal of Mathematics Teacher Education, H. 11 (4), S. 259–284.

Levin, Barbara (2015): The development of teachers' beliefs. In: Fives, Helenrose/Gregoire Gill, Michelle (Hrsg.): International Handbook of Research on Teachers' Beliefs. New York: Routledge, S. 48–65.

Mandl, Heinz/Huber, Gunther (1983): Subjektive Theorien von Lehrern. In: Psychologie in Erziehung und Unterricht, H. 30, S. 98–112.

Mayer, Richard E. (2004): Should there be a three-strikes rule against pure discovery learning? In: American Psychologist, H. 59 (1), S. 14–19.

Mayring, Phillipp (2010): Qualitative Inhaltsanalyse: Grundlagen und Techniken. 11. Auflage. Weinheim: Beltz.

Müller-Fohrbrodt, Gisela/Cloetta, Bernhard/Dann, Hanns-Dietrich (1978): Der Praxisschock bei jungen Lehrern. Formen, Ursachen, Folgerungen. Stuttgart: Klett.

Przyborski, Aglaja/Wohlrab-Sahr, Monika (2014): Qualitative Sozialforschung: Ein Arbeitsbuch. 4. Auflage. München: Oldenbourg Verlag.

Richardson, Virginia (1996): The role of attitudes and beliefs in learning to teach. In: Sikula, John (Hrsg.): Handbook of research on teacher education. 2. Auflage. New York: Macmillan, S. 102–119.

Rothland, Martin (2007): Belastung und Beanspruchung im Lehrerberuf. In: Rothland, Martin (Hrsg.): Belastung und Beanspruchung im Lehrerberuf. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 7–10.

Rothland, Martin/Boecker, Sarah (2015): Viel hilft viel? Forschungsbefunde und -perspektiven zum Praxissemester in der Lehrerbildung. In: Lehrerbildung auf dem Prüfstand, H. 8 (2), S. 112–134.

Schraw, Gregory/Olafson, Lori (2015): Assessing teacher's beliefs. Challenges and solutions. In: Fives, Helenrose/Gregoire Gill, Michele: International handbook of research on teachers' beliefs. New York/London: Routledge, S. 87–105.

Skott, Jeppe (2015): The promises, problems, and prospects of research on teachers's beliefs. In: Fives, Helenrose/Gregoire Gill, Michelle (Hrsg.): International handbook of research on teachers's beliefs. New York/London: Routledge, S. 13–30.

Staub, Fritz/Stern, Elsbeth (2002): The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. In: Journal of Educational Psychology, H. 94 (2), S. 344–355.

Voss, Thamar/Kleickmann, Thilo/Kunter, Mareike/Hachfeld, Axinja (2011): Überzeugungen von Mathematiklehrkräften. In: Kunter, Mareike/Baumert, Jürgen/Blum, Werner/Klusmann, Uta/Krauss, Stefan/Neubrand, Michael: Professionelle Kompetenz von Lehrkräften. Münster: Waxmann, S. 235–257.

Yerrick, Randy/Hoving, Timothy (2003): One foot on the dock and one foot on the boat: Differences among preservice science teachers' interpretations of field-based science methods in culturally diverse contexts. In: Science Education, H. 87 (3), S. 390–418.

Zaruba, Nicole/Gronostaj, Anna/Kretschmann, Julia/Vock, Miriam (2018): Mehr Schüler*innenorientierung oder Praxisschock? Wie sich unterschiedliche Überzeugungen von Lehrkräften während des Praxissemesters entwickeln. In: Krüger, Anja/Radisch, Falk/ Willems, Ariane/Häcker, Thomas/Walm, Maik (Hrsg.): Empirische Bildungsforschung im Kontext von Schule und Lehrer*innenbildung. Bad Heilbrunn: Julius Klinkhardt, S. 223–235.

Entwicklung der Einstellungen Lehramtsstudierender zu Leistungsheterogenität

Der Einfluss von Lehrveranstaltungen[1]

Sabine K. Lehmann-Grube, Andreas Hartinger und Markus Dresel

Abstract

Einstellungen werden in den gängigen Modellen der Lehrerkompetenz als Teil der motivationalen Grundlagen professionellen Handelns gefasst. Im Hinblick auf den Umgang mit Leistungsheterogenität von Schülerinnen und Schülern gelten sie neben diagnostischen Kompetenzen als Prädiktor für eine differenzierende und individualisierende Unterrichtsgestaltung. Allerdings wurde bislang nicht untersucht, ob Einstellungen von Lehrkräften und Lehramtsstudierenden zu Leistungsheterogenität in Regelklassen durch spezifische Konzepte in der Lehrerbildung positiv beeinflusst werden können. (Eine Ausnahme bilden Studien, die sich auf Inklusion im engeren Sinne beziehen.) Modelle der Einstellungsveränderung legen nahe, dass durch eine systematische Auseinandersetzung mit dem Einstellungsgegenstand nachhaltige Veränderungen stattfinden können. In dieser Studie berichten wir darüber, wie sich die Einstellungen von Lehramtsstudierenden bezüglich der Leistungsheterogenität ihrer zukünftigen Schülerinnen und Schüler durch Lehrveranstaltungen, die diese Leistungsheterogenität zentral fokussierten, verändern. Die Ergebnisse zeigen positive Veränderungen im Vergleich zu Veranstaltungen, die Leistungsheterogenität kaum oder gar nicht fokussiert hatten.

1 Das diesem Bericht zugrundeliegende Vorhaben wurde im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA1509 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor(inn)en.

1 Problemstellung

Die Berücksichtigung heterogener Lernvoraussetzungen von Schülerinnen und Schülern (im Folgenden SuS) gilt als Qualitätsmerkmal professionellen Handelns von Lehrkräften (z. B. Strasser 2011; Tomlinson 2017). Dabei wird positiven Einstellungen gegenüber der Heterogenität zwischen SuS die motivationale Funktion zugeschrieben, sich mit den daraus ergebenden Herausforderungen auseinanderzusetzen und z. B. die unterschiedlichen Lernvoraussetzungen bei der Gestaltung des Unterrichts zu berücksichtigen (Gebauer/McElvany 2017; Schüle et al. 2016). Für die Lehrerbildung ergibt sich daraus der Auftrag, nicht nur Wissensgrundlagen für den Umgang mit Heterogenität zu vermitteln, sondern auch adäquate Einstellungen der Studierenden zu fördern. Zur Beeinflussung von Einstellungen lassen sich aus theoretischen Erklärungsmodellen der Einstellungsbildung und -veränderung zwar konkrete Maßnahmen ableiten; allerdings kann die Vielzahl empirisch belegter Einflussfaktoren, wie z. B. mitgebrachte Einstellungen, subjektive Relevanz und die komplexen Wechselwirkungen, in einem ökologischen Setting wie der Lehrerbildung allenfalls partiell kontrolliert werden (vgl. Bohner/Wänke 2006, S. 84; Petty/Cacioppo 1986, S. 183). Daher stellt sich die Frage, ob Faktoren und Bedingungen in Lehrveranstaltungen identifiziert werden können, die Einstellungen zu Schülerheterogenität messbar beeinflussen.

2 Theoretischer Hintergrund

Heterogenität hinsichtlich fachlicher Leistungsfähigkeit

Die fachliche Leistungsfähigkeit von SuS gilt als stärkster individueller Prädiktor von Schulleistungen (vgl. Schrader 2009). Zugleich sind fachliche Unterschiede zwischen den SuS, die als Streuung in Leistungsmaßen, aber auch darüber hinausgehend als spezifische Stärken und Schwächen der SuS gedacht werden können, ein Aspekt von Heterogenität, der durch die Qualität des Unterrichts gut berücksichtigt werden kann (Strasser 2011; Trautmann/Wischer 2011).

Unterrichtsqualität im Umgang mit heterogenen Lernvoraussetzungen der SuS bemisst sich v. a. nach dem Ausmaß adaptiver und differenzierender Unterrichtsgestaltung (Tomlinson 2017; Trautmann/Wischer 2011). Dabei zeigten einige Studien, dass positive Einstellungen zu Inklusion (Schüle et al. 2016) bzw. zu Leistungsheterogenität (Gebauer/McElvany 2017) als Prädiktoren für eine höhere Bereitschaft von Lehrkräften gesehen werden können, leistungsdifferenzierende Maßnahmen (z. B. durch entsprechende Aufgabenstellungen) einzusetzen.

Einstellungen

Einschlägige Definitionen des Konstrukts „Einstellungen“ beinhalten zwei Kernelemente: (1) die Bewertung und (2) das Objekt bzw. die Einheit. So definieren Bohner und Wänke (2006, S. 5) Einstellungen als „a summary evaluation of an object of thought“ und geben Beispiele für die bewertete Einheit, wie z. B. konkrete und abstrakte Objekte (Pizza, Freiheit der Rede), Personen (Milosevic) oder Gruppen (Ausländer) (fast identisch Eagly/Chaiken 1993, S. 1; Petty/Cacioppo 1986, S. 127). Die Heterogenität von SuS ist als relative Größe, die nach objektiven Maßstäben sowie in der subjektiven Wahrnehmung variieren kann (Strasser 2011; Trautmann/Wischer 2011), ein hoch abstraktes ‚Objekt‘, das gleichwohl kognitiv mit konkreten Schulklassen und Personen verknüpft sein kann.

Für die Erfassung von Einstellungen werden drei „response classes“ unterschieden: affektive Reaktionen, Verhaltenstendenzen und kognitive Repräsentationen (Bohner/Wänke 2006, S. 55 f.; Eagly/Chaiken 1993, S. 10 ff., 114 ff.). Ein bedeutsamer Strukturaspekt für die Untersuchung von Einstellungen zu Leistungsheterogenität ist die Konsistenz vs. Ambivalenz verschiedener Ziele und Inhalte, die in eine summarische Bewertung eingehen können (vgl. Bohner/Wänke 2006, S. 54 f.; Eagly/Chaiken 1993, S. 123 ff.). Ein Beispiel dafür wäre die Abwägung des Nutzens von Leistungsheterogenität für das Lernen der SuS gegenüber den Kosten (dem Aufwand und der damit verbundenen Belastung) für die Lehrkraft bei der Vorbereitung und Durchführung des Unterrichts.

An diesem Beispiel wird deutlich, dass Leistungsheterogenität mit subjektiven Überzeugungen für das Handeln im Hinblick auf die Leistungsentwicklung von SuS (als ein zentrales Ziel) und auf die eigenen Ressourcen verknüpft ist. Die Annahme der Theory of Planned Behavior (Ajzen 1991), dass Einstellungen zu Verhaltensoptionen durch Bewertung der Nutzen- und Kostenüberzeugungen gebildet werden (Erwartung x Wert), ist deshalb auch für Einstellungen zu Leistungsheterogenität sinnvoll. Weiterhin fließen in die Bildung individueller Einstellungen auch kollektive Bewertungen ein (vgl. Terry/Hogg 2000). So sind z. B. die Kategorien „Segregation“ vs. „Inklusion“ normativ mit verschiedenen Bewertungen von Leistungsheterogenität assoziiert (vgl. kritisch Trautmann/Wischer 2011). Solche normativen Einflüsse sollten in einer generellen Bewertungsfacette abgebildet werden, um ihre Rolle bei Einstellungsänderungen näher aufzuklären.

Einstellungen zu Leistungsheterogenität

Einstellungen zur Leistungsheterogenität von SuS im engeren Sinne wurden in verschiedenen Studien als Nutzen- und Kostenüberzeugungen erhoben.

Gebauer und McElvany (2011) berichteten aus einer Querschnittserhebung, dass die Überzeugungen zum Nutzen von Leistungsheterogenität die Intention vorhersagten, diese im Unterricht angemessen zu berücksichtigen; Kostenerwartungen zeigten dagegen keinen Zusammenhang. Hartwig et al. (2017) untersuchten die Zusammenhänge zwischen den Einstellungen von Lehrkräften und Lehramtsstudierenden und verschiedenen Kontextvariablen. Die Ergebnisse zeigten einen positiven Zusammenhang zwischen einer positiven Einschätzung der Ausbildung bezüglich des Sachwissens und der positiven Nutzenbewertung; zugleich fanden sich weniger negative Emotionen bezüglich Leistungsheterogenität. Wenn die Lehrkräfte und Studierenden ihre Ausbildung bezüglich des praktischen Umgangs mit Leistungsheterogenität als positiv bewertet hatten, dann ging dies mit geringeren Kostenbefürchtungen sowie weniger negativen Emotionen einher.

Die Befunde bieten wichtige Hinweise für die Bedeutsamkeit der Lehrerbildung in diesem Bereich. Für weitere Überlegungen zur Ausgestaltung der Lehre ist nun aber zu erfassen, inwieweit objektiv erfasste Merkmale der Ausbildung Effekte zeigen, die als Änderung der Einstellungen zu Leistungsheterogenität im Längsschnitt nachweisbar sind.

Änderung von Einstellungen

Die universitäre Phase der Lehrerbildung ist überwiegend durch Präsentation und Verarbeitung von Informationen, weniger durch praktische Erfahrungen mit SuS und Unterrichtsprozessen geprägt. Die Thematisierung der Unterschiedlichkeit von SuS sowie dazu passender individualisierender, schülerzentrierter und kooperativer Lernformen sind in diesem Kontext als Informationen über Leistungsheterogenität zu klassifizieren. Die Veränderung von Einstellungen durch derartige Informationen wird in Modellen der Persuasion (in Abgrenzung u. a. zur Kontakthypothese, vgl. Bohner/Wänke 2006) mit Prozessen der Informationsverarbeitung erklärt. Sowohl im Elaboration Likelihood Modell (ELM) (Petty/Cacioppo 1986) als auch im Heuristic Versus Systematic Processing Modell (HSM) (Chaiken 1980) werden zwei Wege bzw. Modi der Verarbeitung postuliert: zum einen die periphere automatische bzw. heuristische und zum anderen die zentrale kontrollierte bzw. systematische Verarbeitung. Beide Modelle sagen stärkere Einstellungsänderungen durch zentrale systematische Verarbeitung von Informationen über den Gegenstand vorher. Diese findet aber nur statt, wenn (1) die Fähigkeit der Person, (2) ihre Motivation und (3) die Zeit für die Verarbeitung gegeben sind. Die Qualität und Quantität der Argumente bzw. Informationen entfaltet (4) im Wechselspiel mit diesen Bedingungen bestimmte Wirkungen (vgl. Bohner/Wänke 2006, Kap. 7; Eagly/Chaiken 1993, Kap. 7).

Diese Bedingungen können in Lehrveranstaltungen realisiert werden, wenn z. B. Leistungsheterogenität zentral fokussiert wird: So können (1) Fähigkeiten zur Verarbeitung von Informationen über Leistungsheterogenität sukzessive aufgebaut werden, indem z. B. Kriterien zur Beurteilung von Unterrichtssituationen thematisiert werden. (2) Durch Gestaltungsmerkmale, wie z. B. den Einsatz von Unterrichtsvideos, kann die Motivation erhöht werden (vgl. Hellermann/Gold/Holodynski 2015). (3) Die Fokussierung gewährleistet die Zeit für die Verarbeitung. (4) Die Entwicklung neuer Lehrkonzepte verspricht darüber hinaus, dass aktuelle Evidenzen und didaktische Modelle behandelt werden (Qualität der Argumente).

Änderung von Einstellungen im Rahmen der Lehrerbildung

Empirische Studien mit belastbaren Evidenzen zur Änderung von Einstellungen durch Maßnahmen in der Lehrerbildung sind rar. Gao und Mager (2011) fanden in einem Kohortenvergleich von 168 Studierenden in einem Zertifikatsprogramm für inklusives Unterrichten die positivsten Einstellungen zur Inklusion bei Studierenden in der Mitte des Programms; das deutet auf eine diskontinuierliche Entwicklung hin. Kim (2011) verglich die Einstellungen zur Inklusion von 110 Teilnehmenden an zehn Lehrerbildungsprogrammen für „special education“ und fand positivere Einstellungen, wenn das Curriculum in die reguläre Lehrerbildung integriert war. Jones et al. (2009) entwickelten mathematikdidaktische Workshops, die auf die unterrichteten Klassenstufen der Lehrkräfte ausgerichtet waren, und berichteten positive Effekte im Prä-Post-Design (ohne Kontrollgruppe) auf professionelles Wissen und Einstellungen zu Mathematik. Eine Längsschnittuntersuchung, die mögliche Veränderungen von Einstellungen zur Leistungsheterogenität durch spezifische Maßnahmen erfasst und mit der Entwicklung in einer Kontrollgruppe vergleicht, ist uns bislang nicht bekannt.

Fragestellungen

Aus den theoretischen Überlegungen und dem Forschungsstand können folgende Fragestellungen abgeleitet werden:

(1) Beeinflusst die zentrale Fokussierung von Leistungsheterogenität in Lehrveranstaltungen die Entwicklung der Einstellungen zu diesem Heterogenitätsaspekt?
(2) Verändern sich die Einstellungen der Lehramtsstudierenden zu Leistungsheterogenität auch in Veranstaltungen, die diesen Heterogenitätsaspekt nur am Rande oder gar nicht fokussieren?

(3) Verändern sich die generelle Bewertung, die Nutzenbewertung und die Kostenbewertung in gleichem Maß oder zeigen sich differentielle Effekte?

3 Methode

Design und Stichprobe

Die Studie ist im klassischen Prä-Post-Interventionsdesign mit Kontrollbedingung gehalten. In 56 Lehrveranstaltungen (Wintersemester 2015/2016 bis Sommersemester 2017) füllten 583 Studierende auf freiwilliger Basis zu Beginn und am Ende des Semesters einen Fragebogen aus (Altersdurchschnitt = 22,5; SD = 3,3; Frauenanteil 71.1 %). Die Studierenden stammten aus vier Lehramtsstudiengängen (GS = 45.8 %, MS = 23.8 %, RS = 8.9 %, Gym = 16.0 %, keine Angabe = 5.4 %).

In 34 Lehrveranstaltungen wurde nach Angaben der Lehrenden die Leistungsheterogenität von SuS zentral fokussiert. So analysierten die Studierenden z. B. fachbezogene Präkonzepte von SuS, entwickelten Aufgabenstellungen für SuS verschiedener Leistungsfähigkeit oder Fördermaßnahmen für SuS mit besonderen Lernschwierigkeiten. In 22 Lehrveranstaltungen wurde Leistungsheterogenität nicht oder nur am Rande behandelt, dort wurden vorrangig andere Heterogenitätsaspekte oder allgemeine Grundlagen des Fachs fokussiert. In beiden Bedingungen wurden die Veranstaltungen von den jeweiligen Lehrenden anhand gemeinsamer Prinzipien zur Verknüpfung von Lern- und späterem Handlungsfeld konzipiert, u. a. mit einer ausgewiesenen Fallorientierung (vgl. Syring et al. 2013) und/oder dem Einsatz von Fremd- und Eigenvideos (vgl. Hellermann/Gold/Holodynski 2015). In beiden Bedingungen wurden einige Veranstaltungen in Tandemlehre gehalten mit einer Verzahnung von Fachwissenschaft, Fachdidaktik und Bildungswissenschaft.

Messinstrumente

Die Einstellungen wurden mit bipolaren Items (vgl. Osgood et al. 1971) als 6-stufiges Rating erfasst. Folgende drei Facetten wurden erhoben: „Generelle Valenz“ (drei Items), „Nutzen für SuS“ (vier Items) und „Kosten für die Lehrkraft“ (drei Items). Das Rating geschah immer mit Bezug zum Unterricht in einem Fach (Beispielitems: Für den Unterricht in meinem Fach „z. B. Deutsch“ halte ich die Unterschiedlichkeit der Schüler(-innen) in Bezug auf ihre fachliche Leistungsfähigkeit für … „schlecht vs. gut“ (generelle Valenz) … „hinderlich für das Lernen der Schüler(innen) vs. förderlich für das Lernen der Schüler(innen)“ (Nutzen für SuS) … „nicht belastend für die Lehrkraft vs. belastend für die Lehrkraft“ (Kosten für die Lehrkraft)). In der Pilotierung wurde die theoretisch

angenommene bi-faktorielle Struktur (1 Generalfaktor, 3 Subfaktoren, vgl. Reise/Moore/Haviland 2010) mit konfirmatorischen Faktorenanalysen bestätigt. Im aktuellen Datensatz ergab die Berechnung der internen Konsistenzen zu beiden Messzeitpunkten akzeptable bis gute Werte (α = .71 bis .86).

Analysen

Fehlende Werte (< 1 %) wurden je Messzeitpunkt mit dem Expectation Maximization Algorithmus (EM-Schätzer, vgl. Pigott 2001) geschätzt. Im Anschluss wurden die Skalenwerte gebildet (Gesamt, Valenz, Nutzen, Kosten) und am ersten Messzeitpunkt z-standardisiert. Der Heterogenitätsfokus (0 = *gar nicht/ am Rande*, 1 = *zentral*) wurde als Prädiktor auf der zweiten Ebene (Veranstaltungen) in die hierarchischen Mehr-Ebenen-Analysen (HLM 6) aufgenommen.

4 Ergebnisse

Auf der deskriptiven Ebene zeigte sich, dass die Studierenden schon zu Beginn der Veranstaltungen die Heterogenität der SuS hinsichtlich deren fachlicher Leistungsfähigkeit eher positiv und als eher nützlich für ihren zukünftigen Unterricht einschätzten (s. Tab. 1): Die Mittelwerte der Valenzskala und der Nutzenskala liegen über dem Skalenmittel von 3.5. Gleichzeitig schätzten sie auch die Kosten durch die Heterogenität als mittelhoch ein.

Tabelle 1: Mittelwerte und Standardabweichungen (in Klammern) der Einstellungsmaße je Heterogenitätsfokus und Messzeitpunkt

Leistungsheterogenität	*gar nicht / am Rande fokussiert* (N = 238, LV = 22)		*zentral fokussiert* (N = 345, LV = 34)	
	Beginn	Ende	Beginn	Ende
Gesamtskala	3.53 (.68)	3.64 (.67)	3.61 (.67)	3.79 (.72)
Generelle Valenz	3.88 (.90)	3.96 (.88)	3.89 (.92)	4.09 (.96)
Nutzen für SuS	3.76 (.89)	3.88 (.87)	3.91 (.87)	4.13 (.86)
Kosten für L	4.09 (.83)	3.98 (.79)	4.06 (.83)	3.95 (.90)

Anmerkungen: LV = Lehrveranstaltungen, SuS = Schülerinnen und Schüler, L = Lehrkraft; Werte der Kostenskala (hohe Kosten = negative Einstellung) wurden für die Bildung der Gesamtskala umgepolt.

Die hierarchischen Mehrebenenanalysen (s. Tab. 2) ergaben zunächst eine mittlere Stabilität der Einstellungen auf allen vier Skalen (Zeile 1, alle Koeffizienten ≥ .39, $p < .000$). *Zur Fragestellung 1*, ob die zentrale Fokussierung von Leistungsheterogenität in Lehrveranstaltungen die Entwicklung der Einstellungen beeinflusst, zeigen die Ergebnisse einen substantiellen positiven Effekt

(β = .18, $p < .05$). *Zur Fragestellung 2*, ob sich die Einstellungen auch in Veranstaltungen ohne (bzw. mit geringer) Fokussierung von Leistungsheterogenität verändern, ergaben die Analysen keinen Effekt, d. h. die Einstellungen veränderten sich in diesen Veranstaltungen nicht (Intercept = .06, $p > .05$).

Tabelle 2: Mehrebenen-Analyse der Prädiktoren der Einstellungsentwicklung in Lehrveranstaltungen von Semesterbeginn bis Semesterende

Leistungsheterogenität	Einstellungen zu Leistungsheterogenität (Ende)			
	Gesamtskala	Generelle Valenz	Nutzen für SuS	Kosten für L
Ebene 1 (Studierende)				
jeweilige Einstellungen zu Beginn	.49* (.04)	.39* (.05)	.44* (.04)	.43* (.04)
Ebene 2 (Veranstaltungen)				
zentral fokussiert	.18* (.07)	.14* (.09)	.24* (.08)	-.01 (.09)
nicht/am Rande fokussiert (Intercept)	.06 (.05)	.05 (.05)	.02 (.07)	-.07 (.06)
R^2	.24	.19	.24	.12

Anmerkungen: Level-1-Variablen z-standardisiert am ersten Messzeitpunkt. Level-2 dichotom (0, 1). Alle Prädiktoren unzentriert. Werte repräsentieren Regressionskoeffizienten, robuste Standardfehler (in Klammern) und Varianzaufklärung.
*$p < .05$

Zur Fragestellung 3, ob die Veränderungen der Facetten unterschiedlich ausfallen, zeigen die Ergebnisse, dass die Studierenden in Veranstaltungen, die die fachbezogene Leistungsheterogenität von SuS zentral fokussierten, insbesondere den Nutzen für das Lernen und die Entwicklung von SuS am Ende des Semesters positiver bewerteten (β = .24, $p < .05$), mit etwas geringerer Steigung auch die generelle Valenz (β = .14, $p < .05$). Die Bewertung der Kosten blieb dagegen unverändert. In den übrigen Veranstaltungen zeigten sich auf keiner Facette signifikante Veränderungen.

5 Diskussion

Das erste Ziel der Studie war, die Entwicklung der Einstellungen Lehramtsstudierender zur Heterogenität der fachlichen Leistungsfähigkeit von SuS in Lehrveranstaltungen mit verschiedenen Inhaltsschwerpunkten aufzuklären. Dafür wurden mit einem bereits validierten Instrument drei Facetten der Einstellungen (generelle Valenz, Nutzen für SuS, Kosten für die Lehrkraft) zu Beginn und am Ende von entsprechend konzipierten Lehrveranstaltungen erhoben. Die

Daten zeigen, dass der Ansatz, Leistungsheterogenität in Lehrveranstaltungen inhaltlich zentral zu fokussieren, bezogen auf die Einstellungen der Lehramtsstudierenden wirksam sein kann (Fragestellung 1), dass aber ohne die Fokussierung keine Veränderung stattfand (Fragestellung 2). Diese Ergebnisse stimmen mit Modellen und Befunden zur Einstellungsänderung überein, dass systematische und zentrale Verarbeitung von Informationen über Einstellungsgegenstände zu stärkeren und nachhaltigeren Einstellungsänderungen führen (vgl. Bohner/Wänke 2006; Eagly/Chaiken 1993). Auch Lipowsky und Rzejak (2017) stellen in ihrem Review wirksamer Bedingungen von Lehrerfortbildungen heraus, entscheidend sei u. a., „auf welchem Niveau, in welcher Tiefe […] die Beschäftigung mit dem Fortbildungsinhalt erfolgt" (S. 392) und weniger die Dauer an sich.

Ein weiteres Ziel der Studie war zu prüfen, ob sich die Einstellungen zu Leistungsheterogenität in allen drei erfassten Bewertungsfacetten (Valenz, Nutzen, Kosten) verändern (Fragestellung 3). Hier zeigen unsere Daten, dass sich durch die zentrale Fokussierung fachbezogener Leistungsheterogenität insbesondere die Nutzenbewertung der Studierenden positiv veränderte, in geringerem Maß die generelle Bewertung, die Kostenbewertung dagegen nicht. Das stimmt mit den Ergebnissen von Gebauer und McElvany (2017) überein, die einen Effekt der subjektiv wahrgenommenen Ausbildungsqualität bezüglich des Sachwissens über Leistungsheterogenität auf die Nutzenbewertung berichteten. Inwieweit die Studierenden die praxisorientierten Maßnahmen (Kooperation mit Schulen, Einsatz von Fremd- und Eigenvideos) als Ausbildungsqualität für die Praxis wahrgenommen haben, für die Gebauer und McElvany (2017) einen Effekt auf die Kostenbewertung fanden, muss hier offen bleiben. Angenommen werden kann aber, dass z. B. eigene und stellvertretende Erfahrungen mit geeigneten Maßnahmen zur angemessenen Berücksichtigung von Leistungsheterogenität (z. B. in analysierten Videos) die Nutzenbewertung positiv beeinflussen. Das würde für das Modell professioneller Entwicklung von Guskey (2002) sprechen, dass positive Erfahrungen mit konkreter Umsetzung neuer Maßnahmen zur Einstellungsänderung führen und nicht umgekehrt. Die unveränderte Kostenbewertung spricht andrerseits dafür, dass die Studierenden durch den Fokus deutlich erkannt haben, welch großes pädagogisches Können und Engagement ein angemessener Umgang mit Leistungsheterogenität verlangt.

Die Befunde ermutigen dazu, in Lehrveranstaltungen nicht nur die Herausforderungen und damit die Kosten von Heterogenität (z. B. hinsichtlich der Unterrichtsvorbereitung) zu adressieren, sondern auch den erwartbaren Nutzen explizit zu machen. Die Umsetzung von Modellen der Informationsverarbeitung bei der inhaltlichen Schwerpunktsetzung in Lehrveranstaltungen könnte, wie hier berichtet, ein Ansatz dafür sein. Es ist zu vermuten, dass die hochschuldidaktischen Maßnahmen, die eine Auseinandersetzung mit authentischen Situationen ermöglichten (s. o.), zu den Effekten der inhaltlichen

Schwerpunktsetzung beigetragen haben. In Bezug auf die Kostenbewertungen nehmen wir an, dass sich diese erst in der konkreten Schulpraxis durch den Aufbau von Routinen, z. B. bei der Vorbereitung differenzierender Aufgabenstellungen, positiv verändern (können). Hierzu sind Forschungsbemühungen erforderlich, die über die erste Phase der Lehrerbildung hinausgehen.

Allerdings sind durch das Design dieser Studie mit individuell konzipierten Lehrveranstaltungen, die – im Rahmen der gemeinsamen hochschuldidaktischen Prinzipien – innerhalb der Bedingungen variieren, die Aussagen zu den Effekten auf die inhaltlichen Schwerpunktsetzungen beschränkt. Auch können keine Aussagen über differenzielle Effekte in unterschiedlichen Fächern gemacht werden. Zudem können wir aus unseren Befunden keine Aussagen zu spezifischen Effekten einzelner hochschuldidaktisch-methodischer Maßnahmen ableiten. Auch hier sind weitere Forschungsarbeiten erforderlich.

Literatur

Ajzen, Icak (1991): The theory of planned behavior. In: Organizational Behavior and Human Decision Processes 50, H. 2, S. 179–211.

Bohner, Gerd/Wänke, Michaela (2006): Attitudes and attitude change. Reprint. Hove: Psychology Press.

Chaiken, Shelly (1980): Heuristic versus systematic information processing and the use of source versus message cues in persuasion. In: Journal of Personality and Social psychology 39, H. 5, S. 752–766.

Eagly, Alice H./Chaiken, Shelly (1993): The psychology of attitudes. Fort Worth, TX: Harcourt Brace Jovanovich.

Gebauer, Miriam M./McElvany, Nele (2017): Zur Bedeutsamkeit unterrichtsbezogener heterogenitätsspezifischer Einstellungen angehender Lehrkräfte für intendiertes Unterrichtsverhalten. In: Psychologie in Erziehung und Unterricht 64, H. 3, S. 163–180.

Gao, Wei/Mager, Gerald (2011): Enhancing preservice teachers' sense of efficacy and attitudes toward school diversity through preparation: A case of one U.S. inclusive teacher education program. In: International Journal of Special Education 26, H. 2, S. 92–107.

Guskey, Thomas R. (2002): Professional development and teacher change. In: Teachers and Teaching: Theory and Practice 8, H. 3, S. 381–391.

Hartwig, Svenja J./Schwabe, Franziska/Gebauer, Miriam M./McElvany, Nele (2017): Wie beurteilen Lehrkräfte und Lehramtsstudierende Leistungsheterogenität? Ausprägungen, Zusammenhänge und Prädiktoren von Einstellungen und Motivation. In: Psychologie in Erziehung und Unterricht 64, H. 2, S. 94–108.

Hellermann, Christina/Gold, Bernadette/Holodynski, Manfred (2015): Förderung von Klassenführungsfähigkeiten im Lehramtsstudium. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 47, H. 2, S. 97–109.

Jones, Elizabeth/Hampton, Eric M./Brown, Elizabeth M./Leinenbach, Marylin T. (2009): Impacting teacher mathematical knowledge and attitudes with grade-appropriate methods. In: Professional Development in Education 35, H. 2, S. 279–283.

Kim, Ji-Ryun (2011): Influence of teacher preparation programmes on preservice teachers' attitudes toward inclusion. In: International Journal of Inclusive Education 15, H. 3, S. 355–377.

Lipowsky, Frank/Rzejak, Daniela (2017): Fortbildungen für Lehrkräfte wirksam gestalten – Erfolgversprechende Wege und Konzepte aus Sicht der empirischen Bildungsforschung. In: Bildung und Erziehung 70, H. 4, S. 379–399.

Osgood, Charles E./Suci, George J./Tannenbaum, Percy H. (1971): The measurement of meaning. 8th Print. Urbana, Ill: University Press.

Petty, Richard E./Cacioppo, John T. (1986): The Elaboration Likelihood Model of persuasion. In: Advances in Experimental Social Psychology 19, S. 123–205.

Pigott, Therese D. (2001): A review of methods for missing data. In: Educational Research and Evaluation 7, H. 4, S. 352–383.

Reise, Steven P./Moore, Tyler M./Haviland, Mark G. (2010): Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores. In: Journal of Personality Assessment 92, H. 6, S. 544–559.

Schrader, Friedrich-Wilhelm (2009): Kognitive Voraussetzungen. In: Arnold, Karl-Heinz (Hrsg.): Handbuch Unterricht. 2., aktual. Auflage Bad Heilbrunn: Klinkhardt. 425–429.

Schüle, Christoph/Schriek, Josina/Besa, Kris-Stephen/Arnold, Karl-Heinz (2016): Der Zusammenhang der Theorie des geplanten Verhaltens mit der selbstberichteten Individualisierungspraxis von Lehrpersonen. In: Empirische Sonderpädagogik 8, H. 2, S. 140–152.

Strasser, Josef (2011): Heterogenität als Hausforderung pädagogischer Professionalität – Skizze eines Forschungsprogramms. In: Journal of Social Science Education 10, H. 2, S. 14–28.

Syring, Marcus/Bohl, Thorsten/Kleinknecht, Marc/Kuntze, Sebastian/Rehm, Markus/Schneider, Jürgen (2013): Fallarbeit als Angebot – fallbasiertes Lernen als Nutzung. Empirische Ergebnisse zur kognitiven Belastung, Motivation und Emotionen bei der Arbeit mit Unterrichtsfällen. In: Zeitschrift für Pädagogik 62, H. 1, S. 86–108.

Terry, Deborah J./Hogg, Michael A. (Hrsg.) (2000): Attitudes, behavior, and social context. The role of norms and group membership. Mahwah, NJ: L. Erlbaum Associates.

Tomlinson, Carol Ann (2017): How to differentiate instruction in academically diverse classrooms. 3rd Edition. Alexandria, VA: ASCD.

Trautmann, Matthias/Wischer, Beate (2011): Heterogenität in der Schule. Eine kritische Einführung. Wiesbaden: VS.

(Wie) wirken inklusionsorientierte Lehrveranstaltungen und Praxisanteile auf inklusionsbezogene Überzeugungen angehender Lehrkräfte?[1]

Katja Scharenberg und Saskia Opalinski

Abstract

Im Kontext inklusiver Bildung werden gleichermaßen Fragen inklusionsorientierter Lehrerbildung und die Bedeutung inklusionsbezogener Überzeugungen bei (angehenden) Lehrkräften diskutiert. Zu letzteren liegen mittlerweile zahlreiche, wenn auch widersprüchliche, Befunde vor, die jedoch überwiegend aus querschnittlichen Untersuchungen stammen und damit keine Aussagen zur Entstehung bzw. Veränderung von Überzeugungen erlauben. Es gibt jedoch erste empirische Hinweise darauf, dass inklusionsorientierte Lehrveranstaltungen zu positiveren Überzeugungen zur schulischen Inklusion führen können. Im Beitrag wird anhand eines Prä-Post-Designs der Frage nachgegangen, inwiefern durch inklusionspädagogische Lehrveranstaltungen und konzeptuell verankerte Praxisanteile im bildungswissenschaftlichen Anteil im Lehramtsstudium inklusionsbezogene Überzeugungen von Studierenden entwickelt werden können. Anhand der Befunde und bezogen auf die Implementation von Praxisanteilen in der hochschulischen Lehrerbildung werden professionstheoretische und hochschuldidaktische Fragen inklusionsorientierter Lehrerbildung diskutiert.

1 Das diesem Artikel zugrundeliegende Vorhaben wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01JA1818B gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autorinnen.

1 Inklusionsorientierte Lehrerbildung

Nicht erst seit der Ratifizierung des Übereinkommens über die Rechte von Menschen mit Behinderungen (UN-BRK) (vgl. United Nations 2006) kommt es zur Umgestaltung des deutschen Schulsystems: Nach dem Ansatz der Inklusion wird dabei, in Anlehnung an die UNESCO (2003), einerseits das Lernen aller Kinder und Jugendlichen in den Blick genommen. Andererseits werden dabei besonders vulnerable bzw. als marginalisiert geltende Gruppen und Individuen fokussiert (vgl. Kiuppis/Hausstätter 2014). Die daraus resultierende Veränderung von Schule und Unterricht, um die Bedürfnisse aller Lernenden zu berücksichtigen und deren Barrieren für Lernen und Teilhabe zu minimieren, stellt für Schulen und die dort tätigen Lehrkräfte eine große Herausforderung dar. In diesem Kontext werden gleichermaßen die Bedeutung inklusionsbezogener Überzeugungen bei (angehenden) Lehrkräften und Fragen inklusionsorientierter Lehrerbildung diskutiert. Dies ist aktuell eine der zentralen Herausforderungen des deutschen Bildungssystems (vgl. McElvany et al. 2019), wobei derzeit Fragen der institutionellen, curricularen sowie hochschuldidaktischen Umsetzung und der Wirkungen inklusionsbezogener Lehrveranstaltungen im Vordergrund stehen (vgl. Greiten et al. 2017). Wenn davon ausgegangen wird, dass die verschiedenen Aspekte professioneller Kompetenz von Lehrkräften (vgl. Baumert/Kunter 2006) generell erlern- und veränderbar sind, kommt konkreten Lerngelegenheiten und deren Nutzung eine besondere Bedeutung zu (vgl. Voss 2019). Gleichzeitig stellt dieser Bereich derzeit noch ein Theorie- und Forschungsdesiderat dar (vgl. Greiten et al. 2017; Moser 2018).

2 Inklusionsbezogene Überzeugungen von Lehrkräften

Berufsbezogene Überzeugungen als Teilaspekt der professionellen Kompetenz von (angehenden) Lehrkräften sind „affektiv aufgeladene, eine Bewertungskomponente beinhaltende Vorstellungen über das Wesen und die Natur von Lehr-Lernprozessen, Lerninhalten, die Identität und Rolle von Lernenden und Lehrenden […] sowie den institutionellen und gesellschaftlichen Kontext von Bildung und Erziehung, welche für wahr oder wertvoll gehalten werden“ (Reusser/Pauli 2014, S. 642). Sie sind handlungsleitend bei der Anwendung von Wissen und wirken auf die Leistungen von Schülerinnen und Schülern (vgl. Baumert/Kunter 2006). Auch im Diskurs um inklusive Bildung werden sie, neben Wissen und Fähigkeiten, als zentral für eine erfolgreiche Implementation durch die Lehrkräfte herausgestellt (vgl. Heinrich/Urban/Werning 2013). Berufsbezogene Überzeugungen im Sinne eines „individuell verinnerlichten kollektiven Habitus“ (Fischer et al. 2014, S. 20) haben für die Lehrkräfte subjektiv geltenden, normativ-evaluativen Charakter, sind sich selbst erhaltend, robust gegen-

über wahrgenommenen Widersprüchen, lebensgeschichtlich geformt sowie zeitlich stabil und gelten somit als insgesamt schwer veränderlich (vgl. Oser/Blömeke 2012). Dass entsprechende Überzeugungen nicht erst mit Beginn der Professionalisierung als (angehende) Lehrkraft, sondern bereits in der eigenen Schulzeit herausgebildet werden (vgl. Reusser/Pauli 2014), ist im Kontext inklusiver Bildung von besonderer Bedeutung, da die Mehrheit der derzeit angehenden Lehrkräfte noch nicht über systematische Erfahrungen mit inklusiver Bildung aus der eigenen Schulzeit verfügt. Hier deutet sich eine besondere Herausforderung für eine inklusionsorientierte Lehrerbildung an, da nur selten an konkrete Erfahrungen und in diesem Kontext erworbenes Wissen angeknüpft werden kann.

Zu den berufsbezogenen Überzeugungen angehender Lehrkräfte im Kontext inklusiver Bildung liegen mittlerweile zahlreiche nationale und internationale Forschungsbefunde vor (vgl. u. a. Avramidis/Bayliss/Burden 2000; Bosse/Spörer 2014; Hecht/Niedermair/Feyerer 2016; Moser et al. 2014; Schneider 2019; Sharma/Nuttal 2014), die ungeachtet der teils widersprüchlichen Ergebnisse auf neutral-positive Überzeugungen zur schulischen Inklusion hinweisen. Als bedeutsame Einflussfaktoren konnten bisher u. a. der gewählte Lehramtsstudiengang sowie bisherige Erfahrungen mit schulischer Inklusion identifiziert werden. Bezüglich des Lehramtsstudiengangs (Primarstufe, Sekundarstufe I oder II, Förderschule) wurden Unterschiede zwischen den Studierenden der verschiedenen Lehrämter aufgezeigt, die auf positivere Einstellungen von Studierenden der Lehramtsstudiengänge Primarstufe oder Sonderpädagogik im Vergleich zu anderen Lehrämtern schließen lassen (vgl. Moser et al. 2014; Ruberg/Porsch 2017; Trumpa et al. 2014). Die beruflichen und privaten Erfahrungen mit Inklusion bzw. im Umgang mit Menschen mit Behinderungen und sonderpädagogischem Förderbedarf stehen in positivem Zusammenhang zu den inklusionsbezogenen Überzeugungen der (angehenden) Lehrkräfte (vgl. Avramidis/Bayliss/Burden 2000; Bosse/Spörer 2014; Hellmich/Görel 2014; Ruberg/Porsch 2017). Darüber hinaus haben sich in einzelnen Studien geschlechtsspezifische Unterschiede angedeutet: Wiederholt konnten für weibliche Lehrpersonen positivere inklusionsbezogene Überzeugungen gezeigt werden (vgl. de Boer 2012). Da die bisherigen Befunde jedoch überwiegend aus querschnittlichen Untersuchungen stammen, sind kaum Aussagen zur Entstehung bzw. Veränderung dieser Überzeugungen möglich.

Es deutet sich jedoch an, dass sich inklusionsbezogene Einstellungen angehender Lehrkräfte im Studienverlauf verändern, da Studierende höherer Fachsemester positivere inklusionsbezogene Einstellungen aufweisen als Studienanfängerinnen und -anfänger (vgl. Bosse/Spörer 2014; Hecht/Niedermair/Feyerer 2016; Schneider 2019). Darüber hinaus gibt es erste empirische Hinweise darauf, dass inklusionsorientierte Lehrveranstaltungen zu positiveren Überzeugungen bezogen auf Inklusion führen können (vgl. Lancaster/Bain 2007;

Schneider 2019). Weiterführend werden sowohl in der Lehrerbildung als auch hinsichtlich der Veränderung berufsbezogener Überzeugungen angehender Lehrkräfte situierte Lerngelegenheiten, im Sinne von (hochschul-)didaktischen Formaten verbunden mit expliziten Praxiserfahrungen, als bedeutsam und wirksam herausgestellt (vgl. Fischer et al. 2014; Moser 2018; Reusser/Pauli 2014; Van Laarhoven et al. 2007).

In Anlehnung an die Überlegungen Banduras (1995) zur Entstehung bzw. Veränderung von Selbstwirksamkeitserwartungen ist vorstellbar, dass sich eine Veränderung der inklusionsbezogenen Überzeugungen von angehenden Lehrkräften vor dem Hintergrund eigener sowie stellvertretender Bewältigungserfahrungen und deren Reflexion vollziehen kann. Darauf verweisen auch die oben berichteten Befunde zum Zusammenhang zwischen konkreten (positiv bewerteten) Erfahrungen im inklusiven Unterricht und den inklusionsbezogenen Einstellungen von Lehrkräften.

3 Fragestellungen

Basierend auf diesen theoretischen Überlegungen und dem skizzierten Forschungsstand geht unser Beitrag folgenden Forschungsfragen nach:

1. Inwiefern verändern sich inklusionsbezogene Überzeugungen von Studierenden durch inklusionspädagogische Lehrveranstaltungen mit konzeptuell verankerten Praxisanteilen im bildungswissenschaftlichen Anteil der Lehramtsstudiengänge?
2. Inwiefern sind dabei studiumsbezogene (Lehramt Primarstufe oder Sekundarstufe I) oder individuelle Merkmale der Studierenden (Geschlecht) von Bedeutung?

Erwartet wird eine positive Entwicklung der inklusionsbezogenen Überzeugungen, die auf einer intensiveren thematischen Auseinandersetzung im Rahmen konkreter Praxiserfahrungen und -einblicke beruht. Es werden weiterhin positivere Überzeugungen der Studierenden des Lehramts Primarstufe angenommen, da die Grundschule als prospektives Arbeitsfeld ein Lernort mit sehr heterogenen Lerngruppen ist (vgl. Scharenberg 2012) und als derzeit am weitesten fortgeschritten bei der Implementation inklusiver Bildung betrachtet werden kann (vgl. Hollenbach-Biele 2016).

4 Methodisches Vorgehen

Der vorliegenden Arbeit liegt ein Prä-Post-Design zugrunde. Die Datenerhebungen erfolgten über mehrere Semester (Sommer 2016 bis 2018) jeweils zu Beginn und Ende der Semester in vier einsemestrigen, jeweils zweistündigen inklusionspädagogischen Seminaren an der Pädagogischen Hochschule Freiburg, die im Wahlpflichtbereich des Studienangebots der Bildungswissenschaften verortet waren, wobei Inklusion als Querschnittsthematik im Lehramtsstudium verankert ist. Die Seminare bezogen sich inhaltlich auf Inklusion in Schule und Unterricht. Alle Seminare zeichneten sich durch eine konzeptuell in das Seminarkonzept eingebettete, mehrtägige Exkursion aus, bei der die Studierenden an Schulen hospitierten, die als *good practice*-Beispiele im Umgang mit der Heterogenität ihrer Schülerinnen und Schüler und der Realisierung inklusiver Bildung betrachtet werden können. Die Auswahl der Schulen erfolgte dabei einerseits auf Basis von Selbstbeschreibungen, z. B. auf den Homepages der Schulen, und andererseits anhand externer Beurteilungen, z. B. Preisträgerschulen des Deutschen Schulpreises oder des Jakob Muth-Preises. In den Seminaren wurden jeweils zu Beginn die theoretischen Grundlagen inklusiver Bildung erarbeitet sowie in die Methode der Beobachtung im pädagogischen Feld eingeführt. Daran anschließend fand die Exkursion statt. In der Nachbereitung der Exkursion wurde eine Reflexion der Studierenden hinsichtlich eigener Interpretationen und Haltungen angeregt, die durch die Seminarleitung moderiert wurde. Im weiteren Semesterverlauf schlossen sich Sitzungen zur grundlegenden Thematik der inklusiven Gestaltung des Bildungssystems sowie von Schule und Unterricht an. In der weiteren Seminararbeit erfolgte immer wieder der Rückgriff auf die Erfahrungen und Beobachtungen aus der Exkursion.

Die Überzeugungen der Studierenden (insgesamt $n = 95$) wurden mittels der Subskala ‚Inklusionsorientierung (IO)' (12 Items) ($\alpha_{T1} = .72$; $\alpha_{T2} = .77$) des ‚Beliefsinventar Lehrkräfte im Bereich schulischer Förderung (BILF)' (vgl. Moser et al. 2014) erfasst. Auf einer vierstufigen Likert-Skala (1 = *trifft nicht zu*, 2 = *trifft eher nicht zu*, 3 = *trifft eher zu*, 4 = *trifft zu*) wurden damit einerseits eine allgemeine positive Überzeugung hinsichtlich schulischer Inklusion (z. B.: *Inklusive Settings sind prinzipiell für alle Schüler*innen geeignet*) und andererseits eine Selektionsorientierung bzw. eher kritische Einstellung gegenüber schulischer Inklusion (z. B.: *In homogenen Schulklassen lernen die Schüler*innen effizienter als in heterogenen Gruppen*) abgebildet. Negativ formulierte Items wurden für die Auswertung umkodiert, so dass höhere Skalenwerte eine positivere Inklusionsorientierung anzeigten. Darüber hinaus umfasste der Fragebogen soziodemographische Merkmale wie das Geschlecht (1 = *männlich*, 2 = *weiblich*) und studienbezogene Merkmale wie z. B. den Studiengang (1 = *Lehramt Primarstufe*, 2 = *Lehramt Sekundarstufe I*). Weiterhin wurde, als

Indikator für die bisherigen beruflichen Erfahrungen, erfragt, inwiefern sich die Studierenden im Rahmen ihres Studiums bereits mit Inklusion auseinandergesetzt hatten (1 = *Ich habe noch keine Lehrveranstaltungen zum Thema „Inklusion“ besucht*, 2 = *Das Thema „Inklusion“ wurde in einzelnen Sitzungen von Lehrveranstaltungen kurz angesprochen*, 3 = *Das Thema „Inklusion“ war Schwerpunkt in einer oder mehreren Lehrveranstaltungssitzungen*, 4 = *Ich habe eine oder mehrere Lehrveranstaltungen zum Thema „Inklusion“ besucht*).

Mit Blick auf die Zusammensetzung der gesamten Stichprobe sowie der einbezogenen Seminarkohorten (s. Tab. 1) zeigte sich, dass etwa 40 % der Befragten einen Studiengang für das Lehramt Primarstufe studierten, die restlichen 60 % einen Studiengang für das Lehramt Sekundarstufe I. Die Mehrheit der Studierenden war weiblich. Über 80 % der befragten Studierenden gaben an, sich im Rahmen der bisher besuchten Lehrveranstaltungen noch nicht oder nur oberflächlich mit schulischer Inklusion beschäftigt zu haben. Knapp 5 % hatten bereits Lehrveranstaltungen zum Thema besucht, für weitere 10 % wurde die Inklusionsthematik in einzelnen Lehrveranstaltungssitzungen angesprochen.

Tabelle 1: Beschreibung der Stichproben

	SoSe 2016		SoSe 2017		WiSe 2017/18		SoSe 2018		Gesamt	
	n = 19		*n* = 28		*n* = 22		*n* = 27		*n* = 96	
Geschlecht[a]										
männlich	5.3	(1)	25.0	(7)	9.1	(2)	37.0	(10)	20.8	(20)
weiblich	94.7	(18)	75.0	(21)	90.9	(20)	63.0	(17)	79.2	(76)
Inklusionserfahrung[a,b]										
1	42.1	(8)	35.7	(10)	31.8	(7)	33.3	(9)	35.4	(34)
2	42.1	(8)	39.3	(11)	59.1	(13)	44.4	(12)	46.9	(45)
3	10.5	(2)	17.9	(5)	4.5	(1)	11.1	(3)	11.5	(11)
4	5.3	(1)	3.6	(1)	4.5	(1)	11.1	(3)	6.3	(6)
Studiengang[a]										
Primarstufe	78.9	(15)	39.3	(11)	9.1	(2)	37.0	(10)	39.6	(38)
Sekundarstufe I	21.1	(4)	60.7	(17)	90.9	(20)	63.0	(17)	60.4	(58)

[a] Angaben in Prozent (in Klammern: *n*).
[b] 1 = *Ich habe noch keine Lehrveranstaltungen zum Thema „Inklusion“ besucht*,
2 = *Das Thema „Inklusion“ wurde in einzelnen Sitzungen von Lehrveranstaltungen kurz angesprochen*,
3 = *Das Thema „Inklusion“ war Schwerpunkt in einer oder mehreren Lehrveranstaltungssitzungen*,
4 = *Ich habe eine oder mehrere Lehrveranstaltungen zum Thema „Inklusion“ besucht*.

Zur Prüfung von Effekten der Teilnahme an inklusionsorientierten Lehrveranstaltungen auf die Überzeugungen bei Studierenden wurden multivariate Varianzanalysen spezifiziert. Alle Analysen erfolgten mit dem Statistikprogramm SPSS 25 (IBM Corporation 2017). Die Bestimmung der Teststärke (1–β) erfolgte mit der Software G*Power 3.1.9.2 (vgl. Faul et al. 2009).

5 Ergebnisse

Im Vorfeld der multivariaten Analysen erfolgte getrennt für die beiden Messzeitpunkte eine Analyse fehlender Werte, da unvollständig ausgefüllte Fragebögen auf eine geringere Teilnahmemotivation der Probandinnen und Probanden hinweisen können. Der Anteil fehlender Werte zu T1 und T2 belief sich auf insgesamt nur 0.6 % der Werte, sodass fehlende Werte imputiert (*single imputation*) wurden. Da es zwischen Studierenden, die an beiden bzw. nur an einem Messzeitpunkt an der Befragung teilgenommen hatten, weder zu T1[2] noch zu T2[3] signifikante Unterschiede gab, erfolgten alle weiteren Analysen auf Basis der Fälle, für die längsschnittliche Daten zu zwei Messzeitpunkten vorlagen ($n = 66$).

Zunächst erfolgten querschnittliche Auswertungen zu T1 und T2 unter Berücksichtigung soziodemografischer und studienbezogener Variablen. Hinsichtlich der inklusionsbezogenen Überzeugungen der Studierenden lagen zu T1 keine signifikanten Unterschiede in Abhängigkeit der bisherigen Auseinandersetzung mit Inklusion vor (F (3,62) = 1.267, $p = .293$). Es zeigten sich weiterhin zu T1 keine signifikanten Unterschiede zwischen den verschiedenen Seminargruppen (F (3,61) = 0.741, $p = .532$). Deshalb erfolgten alle weiteren Analysen ohne Differenzierung nach dem Ausmaß der bisherigen Auseinandersetzung mit Inklusion oder nach der Seminargruppe.

Unterschiede zu Semesterbeginn und Semesterende

In den weiteren querschnittlichen Analysen konnten für beide Messzeitpunkte (Tab. 2, linke Spalte) keine signifikanten Unterschiede in Abhängigkeit von Geschlecht ($T = -0.23$, $p = .823$) oder Studiengang ($T = 1.34$, $p = .184$) ermittelt werden.

---

2 $n_{nur\ T1} = 20$, $n_{beide\ MZP} = 66$, $M_{nur\ T1} = 2.96$, $SD_{nur\ T1} = 0.33$, $M_{beide\ MZP} = 2.96$, $SD_{beide\ MZP} = 0.40$, $T = 0.31$, $p = .975$.

3 $n_{nur\ T2} = 9$, $n_{beide\ MZP} = 66$, $M_{nur\ T2} = 3.01$, $SD_{nur\ T2} = 0.21$, $M_{beide\ MZP} = 3.09$, $SD_{beide\ MZP} = 0.41$, $T = -0.61$, $p = .543$.

Tabelle 2: Inklusionsorientierung und ihre Veränderung im Längsschnitt nach Personenmerkmalen

		T1 [a]		T2 [a]		Veränderung T1 → T2			
	n	*M*	*(SD)*	*M*	*(SD)*	Δ	*p*	*d*	1–β
Gesamt	66	2.97	(0.41)	3.11	(0.42)	0.14	.002	0.41	0.91
Geschlecht									
männlich	15	2.95	(0.45)	3.09	(0.44)	0.14	.068		
weiblich	51	2.98	(0.40)	3.11	(0.42)	0.13	.009	0.36	0.72
Studiengang									
Primarstufe	30	3.04	(0.38)	3.21	(0.40)	0.17	.012	0.52	0.78
Sekundarstufe I	36	2.91	(0.43)	3.03	(0.43)	0.12	.056		

Mittlere Inklusionsorientierung (*M*) und Standardabweichung (*SD*) der Studierenden ($n = 66$) zu Semesterbeginn (T1) und Semesterende (T2).
[a] Zum jeweiligen Erhebungszeitpunkt (T1, T2) gab es im Querschnitt keine signifikanten Mittelwertunterschiede.

Veränderung im Semesterverlauf

Über den Semesterverlauf war insgesamt eine positive Veränderung der Inklusionsorientierung feststellbar (Tab. 2, rechte Spalte): Am Ende des Semesters wiesen die Studierenden eine signifikant höhere Inklusionsorientierung auf ($M_{T1} = 2.97$, $SD_{T1} = 0.41$; $M_{T2} = 3.11$, $SD_{T2} = 0.42$; $T = -3.26$, $p = .002$, $d = .41$, $1-\beta = 0.98$), wobei der Effekt nach Cohen (1988, S. 24) als kleiner Effekt interpretiert werden kann. Darüber hinaus fiel die Entwicklung in Abhängigkeit von Geschlecht bzw. Studiengang unterschiedlich aus: Eine signifikante Veränderung der Inklusionsorientierung war nur für die weiblichen Studierenden ($M_{T1} = 2.98$, $SD_{T1} = 0.40$; $M_{T2} = 3.11$, $SD_{T2} = 0.42$; $T = -2.70$, $p = .009$, $d = 0.36$, $1-\beta = 0.72$) sowie die Studierenden der Primarstufe ($M_{T1} = 3.04$, $SD_{T1} = 0.38$; $M_{T2} = 3.21$, $SD_{T2} = 0.40$; $T = -2.70$, $p = .012$, $d = 0.52$, $1-\beta = 0.78$) zu verzeichnen. Beide Effekte sind als kleine bis mittlere Effekte zu interpretieren (ebd.).

Tabelle 3: Allgemeines lineares Modell mit Messwiederholung zum Effekt von Seminarbesuch, Geschlecht und Studiengang

	Quadratsumme	*F*	*p*	partielles η^2	*f*
Seminarbesuch (Messwiederholung)	.477	8.012	.006	.114	0.36
Seminarbesuch * Geschlecht	.010	.165	.686	.003	
Seminarbesuch * Studiengang	.046	.779	.381	.012	
Seminarbesuch * Geschlecht * Studiengang	.031	.528	.470	.008	

Bei simultaner Berücksichtigung aller Variablen (Messwiederholung, Geschlecht, Studiengang) sowie ihrer jeweiligen Interaktionseffekte in einem Allgemeinen Linearen Modell (Tab. 3) ergab sich ein signifikanter Haupteffekt für die Messwiederholung auf die Inklusionsorientierung (F (1,62) = 8.012, p = .006, partielles η^2 = .047, f = 0.36) mit mittlerer Effektgröße (vgl. Cohen 1988, S. 287). Unter Kontrolle der jeweils anderen Variablen konnten dagegen keine signifikanten Interaktionseffekte mit dem Geschlecht oder Lehramtsstudiengang aufgezeigt werden.

6 Diskussion

Die vorliegenden Befunde zeigen, dass die Studierenden nach dem Besuch inklusionsorientierter Lehrveranstaltungen mit konzeptuell verankertem Theorie-Praxis-Transfer am Semesterende signifikant höhere Inklusionsorientierungen aufweisen. Unsere Ergebnisse sind damit übereinstimmend zum bisherigen Forschungsstand und bestätigen die Wirkung inklusionsorientierter Lehrveranstaltungen auf inklusionsbezogene Überzeugungen angehender Lehrkräfte (vgl. Kopp 2009; Schneider 2019; Sharma/Nuttal 2014; Sharma/Sokal 2015) sowie die Bedeutung einer Theorie-Praxis-Verknüpfung im Lehramtsstudium (vgl. Van Laarhoven et al. 2007).

Das vorgestellte, konsequent auf die Hospitationserfahrungen bezogene Seminarkonzept ermöglicht die Verknüpfung von theoriebasiertem, kognitivem Wissenserwerb und fallbezogener Reflexion der Hospitationserfahrungen (vgl. Schneider 2019; Seitz 2007). Der Annahme folgend, dass die als zeitlich stabil geltenden berufsbezogenen Überzeugungen angehender Lehrkräfte bereits während der eigenen Schulzeit herausgebildet werden, die dabei jedoch selbst bisher kaum systematische Erfahrungen mit inklusiver Bildung gemacht haben (vgl. Reusser/Pauli 2014), stellen derartige durch einen Theorie-Praxis-Transfer gekennzeichnete Lerngelegenheiten in der Hochschulausbildung einen wichtigen Baustein bei der Implementation inklusiver Bildung und Entwicklung einer inklusiven Professionalität angehender Lehrkräfte dar. Da eigene Bewältigungserfahrungen für angehende Lehrkräfte nur eingeschränkt möglich sind und nicht konsequent durch Lehrveranstaltungen begleitet werden, kommt den stellvertretenden Erfahrungen im Rahmen universitärer Lehrveranstaltungen durch eine konzeptuell verankerte Hospitation in gelingenden inklusiven Lehr-Lernsettings eine besondere Bedeutung zu.

Inwiefern die Veränderung der inklusionsbezogenen Einstellungen der Studierenden tatsächlich auf das besuchte Seminar mit konzeptionell verankerter Hospitation zurückzuführen ist, muss in weiterführenden Untersuchungen in einem quasi-experimentellen Design mit Kontrollgruppen bzw. im Vergleich verschiedener didaktischer Formate noch geprüft werden. Des Weiteren ist

einschränkend festzustellen, dass die Erhebung der Überzeugungen durch Selbstauskünfte zu Effekten sozialer Erwünschtheit führen kann (vgl. Lüke/Grosche 2017). Um diese identifizieren zu können, wird in weiteren Erhebungen eine Skala zur Erfassung der Tendenz zur sozialen Erwünschtheit eingesetzt werden.

Der vorliegende Beitrag untersuchte die Entwicklung der inklusionsbezogenen Überzeugungen von Studierenden im Rahmen von Seminaren mit konzeptionell verankertem Theorie-Praxis-Transfer und leistet damit einen bedeutsamen Beitrag im Rahmen der Forschung zu Effekten und Wirkungen der inklusionsorientierten Lehrerbildung. In weiteren Untersuchungen sollte weiterführend der Frage nachgegangen werden, welchen Beitrag verschiedene hochschuldidaktische Lehrformate (z. B. Vorlesung, Seminar mit oder ohne Theorie-Praxis-Transfer) zur Entwicklung einer inklusiven Professionalität angehender Lehrkräfte und der Veränderung ihrer berufsbezogenen Überzeugungen im Kontext Inklusion leisten können, um durch eine gezielte Weiterentwicklung der gewählten bzw. den jeweiligen Rahmenvorgaben entsprechenden hochschuldidaktischen Ansätze einen bestmöglichen Beitrag zur inklusiven Lehrerbildung an den Hochschulen leisten zu können. Schließlich bleibt auf Basis der vorliegenden Daten bisher offen, inwiefern angehende Lehrkräfte im Rahmen inklusionsorientierter Lehrveranstaltungen für die prospektive Tätigkeit in inklusiven Schulsettings erforderliche Kompetenzen erwerben (vgl. Greiten et al. 2017) und inwiefern sich die (veränderten) berufsbezogenen Überzeugungen auch im späteren beruflichen Handeln niederschlagen (vgl. Heinrich/Urban/Werning 2013).

Trotz der dargestellten Limitationen und weiterhin offenen Fragen kann mit dem vorliegenden Beitrag gezeigt werden, dass die oft als problematisch beschriebene Veränderung berufsbezogener Überzeugungen angehender Lehrkräfte im Kontext Inklusion im Rahmen universitärer Lehrveranstaltungen gelingen kann. Weitere Untersuchungen zur Bearbeitung der aufgezeigten, weiterhin bestehenden Forschungsdesiderata befinden sich derzeit in Vorbereitung.

Literatur

Avramidis, Elias/Bayliss, Phil/Burden, Robert (2000): Student teachers' attitudes towards the inclusion of children with special educational needs in the ordinary school. In: Teaching and Teacher Education 16, 3, 277–293.

Bandura, Albert (1995): Exercise of personal and collective efficacy in changing societies. In: Bandura, Albert (Hrsg.): Self-efficacy in changing societies. Cambridge: Cambridge University Press, S. 1–45.

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, 4, 469–520.

Bosse, Stefanie/Spörer, Nadine (2014): Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. In: Empirische Sonderpädagogik 6, 4, 279–299.

Cohen, Jacob (1988): Statistical power analysis for the behavioral sciences (2. Aufl.). Hillsdale, NJ: Erlbaum.

de Boer, Anke (2012): Inclusion: a question of attitudes? A study on those directly involved in the primary education of students with special educational needs and their social participation. Groningen: Stichting Kinderstudies.

Faul, Franz/Erdfelder, Edgar/Buchner, Axel/Lang, Albert-Georg (2009): Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. In: Behavior Research Methods 41, 4, 1149–1160.

Fischer, Christian/Kopmann, Henrike/Rott, David/Veber, Marcel/Zeinz, Horst (2014): Adaptive Lehrkompetenz und pädagogische Haltung. Lehrerbildung für eine inklusive Schule. In: Zierer, Klaus (Hrsg.): Jahrbuch für Allgemeine Didaktik 2014. Thementeil: Allgemeine Didaktik für eine inklusive Schule. Baltmannsweiler: Schneider Verlag Hohengehren, S. 16–34.

Greiten, Silvia/Geber, Georg/Gruhn, Annika/Köninger, Manuela (2017): Inklusion als Aufgabe für die Lehrerausbildung. Theoretische, institutionelle, curriculare und didaktische Herausforderungen für Hochschulen In: Greiten, Silvia/Geber, Georg/Gruhn, Annika/Köninger, Manuela (Hrsg.): Lehrerausbildung für Inklusion. Fragen und Konzepte zur Hochschulentwicklung. Münster: Waxmann, S. 14–36.

Hecht, Petra/Niedermair, Claudia/Feyerer, Ewald (2016): Einstellungen und inklusionsbezogene Selbstwirksamkeitsüberzeugungen von Lehramtsstudierenden und Lehrpersonen im Berufseinstieg. Messverfahren und Befunde aus einem Mixed-Methods-Design. In: Empirische Sonderpädagogik 8, 1, 86–102.

Heinrich, Martin/Urban, Michael/Werning, Rolf (2013): Grundlagen, Handlungsstrategien und Forschungsperspektiven für die Ausbildung und Professionalisierung von Fachkräften für inklusive Schulen. In: Döbert, Hans/Weishaupt, Horst (Hrsg.): Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen. Münster: Waxmann, S. 69–133.

Hellmich, Frank/Görel, Gamze (2014): Erklärungsfaktoren für Einstellungen von Lehrerinnen und Lehrern zum inklusiven Unterricht in der Grundschule. In: Zeitschrift für Bildungsforschung 4, 3, 227–240.

Hollenbach-Biele, Nicole (2016): Inklusion statt Förderschule? Zum Stand des gemeinsamen Unterrichts in Deutschland. In: Bertelsmann Stiftung (Hrsg.): Inklusion kann gelingen! Forschungsergebnisse und Beispiele guter schulischer Praxis. Gütersloh: Bertelsmann Stiftung, S. 11–33.

IBM Corporation (2017): IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corporation.

Kiuppis, Florian/Hausstätter, Rune S. (2014): Inclusive education for all, and especially for some? On different interpretations of who and what the "Salamanca process" concerns. In: Kiuppis, Florian/Hausstätter, Rune S. (Hrsg.): Inclusive education twenty years after Salamanca. New York: Peter Lang Publishing, S. 1–5.

Kopp, Bärbel (2009): Inklusive Überzeugungen und Selbstwirksamkeit im Umgang mit Heterogenität – Wie denken Studierende des Lehramts für Grundschulen? In: Empirische Sonderpädagogik 1, 1, 5–25.

Lancaster, Julie/Bain, Alan (2007): The design of inclusive education courses and the self-efficacy of preservice teacher education students. In: International Journal of Disability, Development and Education 54, 2, 245–256.

Lüke, Timo/Grosche, Michael (2017): What do I think about inclusive education? It depends on who is asking. Experimental evidence for a social desirability bias in attitudes towards inclusion. In: International Journal of Inclusive Education 22, 1, 38–53.

McElvany, Nele/Schwabe, Franziska/Bos, Wilfried/Holtappels, Heinz Günter (2019): Vorwort. In: McElvany, Nele/Schwabe, Franziska/Bos, Wilfried/Holtappels, Heinz Günter (Hrsg.): Lehrerbildung – Potentiale und Herausforderungen in den drei Phasen. Münster: Waxmann, S. 7–8.

Moser, Vera (2018): Lehrerinnen- und Lehrerbildung. In: Sturm, Tanja/Wagner-Willi, Monika (Hrsg.): Handbuch schulische Inklusion. Opladen: Verlag Barbara Budrich, S. 283–298.

Moser, Vera/Kuhl, Jan/Redlich, Hubertus/Schäfer, Lea (2014): Beliefs von Studierenden sonder- und grundschulpädagogischer Studiengänge. In: Zeitschrift für Erziehungswissenschaft 17, 4, 661–678.

Oser, Fritz/Blömeke, Sigrid (2012): Überzeugungen von Lehrpersonen. Einführung in den Thementeil. In: Zeitschrift für Pädagogik 58, 4, 415–421.

Reusser, Kurt/Pauli, Christine (2014): Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In: Terhart, Ewald/Bennewitz, Hedda/Rothland, Martin (Hrsg.): Handbuch der Forschung zum Lehrerberuf (2. Aufl.). Münster: Waxmann, S. 642–661.

Ruberg, Christiane/Porsch, Raphaela (2017): Einstellung von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion. In: Zeitschrift für Pädagogik 63, 4, 393–415.

Scharenberg, Katja (2012): Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie. Münster: Waxmann.

Schneider, Edina (2019): Welche Überzeugungen zeigen Lehramtsstudierende zur schulischen Inklusion? Quantitative Befunde aus der wissenschaftlichen Begleitstudie des Projektes KALEI. In: Klektau, Claudia/Schütz, Susanne/Fett, Anne Julia (Hrsg.): Heterogenitätssensibilität durch Fallarbeit fördern. Zum Stellenwert von Kasuistik und Inklusion in der Lehrer*innenbildung. Halle/Saale: Zentrum für Lehrer*innenbildung der Martin-Luther-Universität Halle-Wittenberg, S. 133–144.

Seitz, Simone (2007): Warum Unterrichtsentwicklung und Lehrer/innenprofessionalisierung zusammengehören In: Demmer-Dieckmann, Irene/Textor, Annette (Hrsg.): Integrationsforschung und Bildungspolitik im Dialog. Bad Heilbrunn: Klinkhardt, S. 127–135.

Sharma, Umesh/Nuttal, Anthony (2014): The impact of training on pre-service teacher attitudes, concerns, and efficacy towards inclusion. In: Asia-Pacific Journal of Teacher Education 44, 2, 142–155.

Sharma, Umesh/Sokal, Laura (2015): The impact of a teacher education course on pre-service teachers' beliefs about inclusion: an international comparison. In: Journal of Research in Special Educational Needs 15, 4, 276–284.

Trumpa, Silke/Janz, Frauke/Heyl, Vera/Seifried, Stefanie (2014): Einstellungen zu Inklusion bei Lehrkräften und Eltern – Eine schulartspezifische Analyse. In: Zeitschrift für Bildungsforschung 4, 3, 241–256.

UNESCO (2003): Overcoming exclusion through inclusive approaches in education. A challenge & a vision. Conceptual paper. http://unesdoc.unesco.org/images/0013/001347/134785e.pdf (Abfrage: 22.02.2018).

United Nations (2006): Convention on the rights of persons with disabilities. http://www.un.org/disabilities/documents/convention/convention_accessible_pdf.pdf (Abfrage: 22.02.2018).

Van Laarhoven, Toni R./Munk, Dennis D./Lynch, Kathleen/Bosma, Julie/Rouse, Joanne (2007): A model for preparing special and general education preservice teachers for inclusive education. In: Journal of Teacher Education 58, 5, 440–455.

Voss, Thamar (2019): Lehrkraftwissen und dessen Erwerb. Was muss eine Lehrkraft wissen und wo lernt sie es? In: McElvany, Nele/Schwabe, Franziska/Bos, Wilfried/Holtappels, Heinz Günter (Hrsg.): Lehrerbildung – Potentiale und Herausforderungen in den drei Phasen. Münster: Waxmann, S. 9–28.

Unterscheiden sich Studentinnen und Studenten des Grund- und Förderschullehramts in ihren Einstellungen und ihren Selbstwirksamkeitsüberzeugungen in Bezug auf den inklusiven Unterricht?

Frank Hellmich, Marwin Felix Loeper und Fabian Hoya

Abstract

Einstellungen von (angehenden) Grund- und Förderschullehrkräften zur Inklusion und ihre Selbstwirksamkeitsüberzeugungen werden als wichtige persönliche Ressourcen für die Durchführung und Gestaltung des inklusiven Unterrichts in der Grundschule erachtet. In unserer Studie wurden insgesamt $N = 606$ Studentinnen und Studenten des Lehramts für Grundschulen und des Lehramts für sonderpädagogische Förderung zu ihren Einstellungen zur Inklusion in Schule und Unterricht sowie zu ihren Selbstwirksamkeitsüberzeugungen in Hinblick auf die Gestaltung des eigenen inklusiven Unterrichts befragt. Die Ergebnisse verdeutlichen, dass die Studierenden des Lehramts für sonderpädagogische Förderung über signifikant positiver ausgeprägte Einstellungen zur schulischen Förderung im inklusiven Unterricht verfügen und sich unter individuellem und kollektivem Aspekt als signifikant selbstwirksamer im inklusiven Unterricht wahrnehmen als die Studierenden des Lehramts an Grundschulen. Die Studierenden unterscheiden sich allerdings nicht in ihren Einstellungen zur sozialen Teilhabe von Kindern mit und ohne sonderpädagogischem/n Förderbedarf im inklusiven Klassenzimmer.

1 Problemstellung

Mit der Ratifizierung der UN-Behindertenrechtskonvention durch die Bundesrepublik Deutschland gehen gegenwärtig Fragen nach einer geeigneten Aus-, Fort- und Weiterbildung von (angehenden) Lehrerinnen und Lehrern für den inklusiven Unterricht einher. Noch wenig beantwortet ist beispielsweise die Frage danach, wie und auf welche Weise Grund- und Förderschullehrkräfte

zukünftig gemeinsam in kooperativen Settings im inklusiven Klassenzimmer erfolgreich unterrichten können. Als wichtige Voraussetzungen hierfür werden nach Martínez (2003) persönliche Ressourcen von Lehrkräften wie beispielsweise geeignete Einstellungen der Lehrkräfte zur Inklusion sowie positive Selbstwirksamkeitsüberzeugungen in Bezug auf die Gestaltung des eigenen inklusiven Unterrichts aufgefasst. Während gegenwärtig Studien vorliegen, in denen belegt werden konnte, dass Förderschullehrkräfte über positivere Einstellungen zur Inklusion sowie über höhere Selbstwirksamkeitsüberzeugungen verfügen (z. B. Kuhl/Redlich/Schäfer 2014; Lamture/Gathoo 2017) ist noch nicht in hinreichender Weise betrachtet worden, ob und inwiefern diese Unterschiede bereits zu früheren Zeitpunkten, d. h. bereits im Studium oder während des Referendariats, beobachtet werden können. Vor diesem Hintergrund untersuchen wir im Rahmen unserer Studie, ob und inwiefern sich Studierende des Lehramts an Grundschulen und solche des Lehramts für sonderpädagogische Förderung hinsichtlich ihrer individuellen Ressourcen für die Gestaltung des inklusiven Unterrichts in der Grundschule (Einstellungen zur Inklusion, Selbstwirksamkeitsüberzeugungen in Hinblick auf die Gestaltung des eigenen inklusiven Unterrichts) unterscheiden.

2 Theoretischer Hintergrund

Folgt man Baumert und Kunter (2006), so werden neben fachlich-didaktischem und pädagogischem Wissen Überzeugungen und Motivation als wichtige Facetten der Professionalität von Lehrerinnen und Lehrern aufgefasst. Auch in Überlegungen zu erforderlichen Kompetenzen von Lehrkräften, die im inklusiven Unterricht in den einzelnen Jahrgangsstufen tätig sind, werden personelle Ressourcen wie Einstellungen zur Inklusion, Selbstwirksamkeitsüberzeugungen in Hinblick auf die Gestaltung des eigenen inklusiven Unterrichts oder motivationale Orientierungen in starkem Maße betont (Martínez 2003). Gerade vor dem Hintergrund, dass Lehrkräfte verschiedener Professionen – beispielsweise Grund- und Förderschullehrkräfte – kooperativ im inklusiven Klassenzimmer unterrichten, stellt sich die Frage nach der Kompatibilität der jeweiligen individuellen Ausgangslagen der einzelnen Akteurinnen und Akteure. In den vergangenen Jahren wurden Einstellungen von Lehrkräften zur Inklusion als Dreh- und Angelpunkte für das Gelingen von inklusivem Unterricht in zahlreichen Studien betrachtet (z. B. Hellmich/Görel 2014; Kuhl/Redlich/Schäfer 2014). Unter Einstellungen werden dabei nach Eagly und Chaiken (1993) psychologische Tendenzen verstanden, die sich in Form einer Annahme oder Ablehnung bei der Bewertung eines Einstellungsobjekts ausdrücken. In verschiedenen Studien konnte belegt werden, dass Förderschullehrkräfte über signifikant höhere Einstellungen zur Inklusion verfügen als Regelschullehrkräfte (Kuhl/Red-

lich/Schäfer 2014; Moser et al. 2014; Trumpa et al. 2014). Bei der Messung der Einstellungen zur Inklusion wurde in den Studien von Moser et al. (2014) sowie Trumpa et al. (2014) zum einen zwischen den Einstellungen von Regel- und Förderschullehrkräften zur Inklusion in Hinblick auf die soziale Teilhabe von Kindern und Jugendlichen und zum anderen zwischen ihren Einstellungen zur Inklusion in Bezug auf die schulische Förderung aller Schülerinnen und Schüler im inklusiven Unterricht unterschieden. Im Detail wurde dabei deutlich, dass die Förderschullehrkräfte in der Studie von Moser et al. (2014) auf beiden Fragebogenskalen signifikant höhere Einstellungen als die Regelschullehrkräfte zeigten. In der Studie von Trumpa et al. (2014) konnte nachgewiesen werden, dass die Förderschullehrkräfte zwar signifikant höhere Einstellungen zur schulischen Förderung im inklusiven Unterricht haben als die an der Studie beteiligten Regelschullehrkräfte, sie zeigten allerdings signifikant niedrigere Werte auf der Fragebogenskala zur Erfassung der Einstellungen zur sozialen Teilhabe im inklusiven Unterricht als die Regelschullehrkräfte. Die Befundlage kann diesbezüglich als nicht konsistent bezeichnet werden. Darüber hinaus liegen zwei Studien vor, bei denen Studierende des Grund- und Förderschullehramts zu ihren Einstellungen zur Inklusion befragt worden sind (McHatton/Parker 2013; Scholz/Rank 2016). In der Studie von McHatton und Parker (2013) konnte belegt werden, dass Studierende des Förderschullehramts über positivere Einstellungen zur Inklusion verfügten als die Studierenden des Grundschullehramts. Scholz und Rank (2016) erklären als ein Ergebnis ihrer Studie, dass die Studierenden des Grundschullehramts signifikant stärker als die Studierenden des Förderschullehramts der Überzeugung waren, dass Kinder mit sonderpädagogischem Förderbedarf besser an Förder- als an Regelschulen unterrichtet werden können.

Auch die Ergebnisse zu den Unterschieden in den Selbstwirksamkeitsüberzeugungen von (angehenden) Regel- und Förderschullehrkräften in Hinblick auf die Gestaltung des eigenen inklusiven Unterrichts fallen weitgehend einheitlich aus. Unter Selbstwirksamkeit wird dabei die Kompetenzüberzeugung verstanden, erforderliches Verhalten auch unter schwierigen Voraussetzungen ausführen zu können, um angestrebte Ziele zu erreichen (Bandura 1997). Einheitlich wird in den vorliegenden Untersuchungen deutlich, dass Förderschullehrkräfte über signifikant höhere Selbstwirksamkeitsüberzeugungen in Bezug auf das Unterrichten im inklusiven Klassenzimmer verfügen als Regelschullehrkräfte (Freytag 2001; Lamture/Gathoo 2017). Lamture und Gathoo (2017) konnten beispielsweise im Detail verdeutlichen, dass sich die an der Untersuchung beteiligten Förderschullehrkräfte signifikant selbstwirksamer fühlten, Entscheidungen im inklusiven Unterricht zu treffen oder die Eltern an der Gestaltung und Umsetzung inklusiven Unterrichts zu beteiligen als die befragten Regelschullehrkräfte. Bezüglich ihrer Selbstwirksamkeitsüberzeugungen, ein positives Schulklima erzeugen zu können und instruktionale bzw. diszipli-

narische Maßnahmen im inklusiven Unterricht anwenden zu können, wurden hingegen keine Unterschiede zwischen Regel- und Förderschullehrkräften in dieser Studie gefunden.

Unterschiede in den Selbstwirksamkeitsüberzeugungen von angehenden Lehrkräften des Regel- und des Förderschullehramts konnten ebenfalls in verschiedenen Studien nachgewiesen werden (Leyser/Zeiger/Romi 2011; Sharma/Shaukat/Furlonger 2015). So konnten Sharma, Shaukat und Furlonger (2015) aufzeigen, dass diejenigen angehenden Lehrkräfte mit einer sonderpädagogischen Ausbildung, Wissen über sonderpädagogische Förderbedarfe, Unterrichtserfahrungen und persönlichen Erfahrungen mit Menschen mit Behinderungen über höhere Selbstwirksamkeitsüberzeugungen zum inklusiven Unterricht verfügten als angehende Lehrkräfte ohne diese Voraussetzungen. Die Befunde von Özokcu (2017) geben in ähnlicher Weise Auskunft darüber, dass Lehrkräfte, die sonderpädagogische Fortbildungen besucht haben, mehr Erfahrungen im inklusiven Unterricht gesammelt haben und zuvor mit Menschen mit Behinderungen interagiert haben, signifikant höhere Selbstwirksamkeitsüberzeugungen aufwiesen als Lehrkräfte ohne derartige Erfahrungen.

Die Befundlage zu den Unterschieden in den Selbstwirksamkeitsüberzeugungen angehender Regel- und Förderschullehrkräfte ist allerdings nicht einheitlich. So liegen auch Studien vor, bei denen keine Unterschiede in den Selbstwirksamkeitsüberzeugungen von Regel- und Förderschullehrkräften in Hinblick auf den inklusiven Unterricht gefunden werden konnten (Demirtas 2018; Flores et al. 2014).

In den vorliegenden Untersuchungen wurden darüber hinaus lediglich die individuellen Selbstwirksamkeitsüberzeugungen der jeweiligen (angehenden) Lehrkräfte in den Blick genommen. Kollektive Selbstwirksamkeiten – verstanden als die Kompetenzüberzeugungen von (angehenden) Lehrkräften, im Team gemeinsam Ziele im inklusiven Unterricht erreichen zu können – sind noch nicht betrachtet worden.

3 Forschungshypothesen

Vor dem dargestellten theoretischen und empirischen Hintergrund nehmen wir erstens an – auch wenn die Befundlage nicht vollständig einheitlich ist –, dass Studierende des Förderschullehramts der sozialen Teilhabe von Kindern mit und ohne Förderbedarfe/n im inklusiven Unterricht und der schulischen Förderung aller Schülerinnen und Schüler im inklusiven Klassenzimmer eine signifikant höhere Bedeutung beimessen als Studierende des Grundschullehramts (z. B. Kuhl/Redlich/Schäfer 2014; McHatton/Parker 2013; Moser et al. 2014). In Anlehnung an die Befunde von Lamture und Gathoo (2017) sowie Sharma, Shaukat und Furlonger (2015) gehen wir zweitens davon aus, dass sich

Studierende des Förderschullehramts signifikant als selbstwirksamer in Hinblick auf die Herausforderungen des inklusiven Unterrichts einschätzen als Studierende des Grundschullehramts. Explorativ wird drittens vermutet, dass diese Unterschiede auch in Bezug auf die kollektiven Selbstwirksamkeitsüberzeugungen zugunsten der Studierenden des Förderschullehramts zu beobachten sind.

4 Methode

Stichprobe

An der Studie sind insgesamt $N = 606$ Studentinnen und Studenten des Lehramts an Grundschulen und des Lehramts für sonderpädagogische Förderung an der Universität Paderborn beteiligt gewesen, davon 526 Studentinnen und 80 Studenten. Zum Zeitpunkt der Erhebungen waren die Studierenden im Durchschnitt ca. 22 Jahre alt ($M = 21.61$; $SD = 2.66$; $Min = 17$; $Max = 35$). 366 Studierende sind für das Lehramt an Grundschulen (275 im Bachelorstudiengang, 91 im Masterstudiengang) eingeschrieben, wohingegen 240 Studierende im Studiengang Lehramt für sonderpädagogische Förderung (190 im Bachelorstudiengang, 50 im Masterstudiengang) studieren. Die Studierenden des Bachelorstudiengangs für das Lehramt an Grundschulen und für das Lehramt für sonderpädagogische Förderung unterscheiden sich hinsichtlich ihres durchschnittlichen Alters signifikant voneinander ($M = 20.56$, $SD = 2.35$ versus $M = 21.57$, $SD = 2.23$; $t = 4.62$; $df = 461$; $p \leq .001$). Zwischen den Studierenden des Masterstudiengangs für das Lehramt an Grundschulen und für das Lehramt für sonderpädagogische Förderung sind keine signifikanten Altersunterschiede vorhanden ($M = 24.00$, $SD = 2.51$ versus $M = 23.34$, $SD = 2.37$; $t = 1.52$; $df = 139$; $p = .13$).

Von den befragten Studierenden haben 243 Studierende (40.1 %) angegeben, noch keine Erfahrungen im Unterrichten inklusiver Klassen gesammelt zu haben, wohingegen bereits 361 Studierende (59.57 %) über Erfahrungen verfügen. Zwei Studierende haben sich in unserem Fragebogen nicht zu ihren bisherigen Erfahrungen mit inklusivem Unterricht geäußert.

Messinstrumente

Die an der Studie beteiligten Studentinnen und Studenten wurden gebeten, einen Fragebogen zu ihren Einstellungen zur Inklusion sowie zu ihren Selbstwirksamkeitsüberzeugungen in Hinblick auf die Gestaltung des eigenen inklusiven Unterrichts auszufüllen. Für die Erfassung der Einstellungen der Studierenden zur Inklusion wurden zwei Fragebogenskalen eingesetzt, die auf den

Arbeiten von Kunz, Luder und Moretti (2010) sowie Woll (2017) basieren: „Einstellungen zur sozialen Teilhabe in der inklusiven Schule“ (4 Items; Beispielitem: „Wenn Kinder mit sonderpädagogischem Förderbedarf in einer inklusiven Schulklasse sind, dann finden sie dort auch Freundinnen und Freunde.“; $M = 3.71$, $SD = 0.66$, $Min = 1.25$, $Max = 5.00$, $\alpha = .79$) und „Einstellungen zur schulischen Förderung im inklusiven Unterricht“ (4 Items; Beispielitem: „Kinder mit sonderpädagogischem Förderbedarf werden in einer inklusiven Schulklasse letztendlich besser gefördert.“; $M = 2.90$, $SD = 0.77$, $Min = 1.00$, $Max = 5.00$, $\alpha = .78$). Die individuellen Selbstwirksamkeitsüberzeugungen der Studierenden in Hinblick auf die Gestaltung des (zukünftigen) eigenen inklusiven Unterrichts wurden auf der Grundlage eines Fragebogeninstruments von Kopp (2009) sowie Schwab, Hellmich und Görel (2017) erhoben (8 Items; Beispielitem: „Ich traue mir zu, Unterricht so zu organisieren, dass auch Kinder mit sonderpädagogischem Förderbedarf in ihrem eigenen Lerntempo zum Ziel kommen können.“; $M = 3.47$, $SD = 0.67$, $Min = 1.00$, $Max = 5.00$, $\alpha = .90$). Die kollektiven Selbstwirksamkeitsüberzeugungen der Studierenden in Hinblick auf den inklusiven Unterricht wurden anhand einer Fragebogenskala mit insgesamt acht Fragebogenitems gemessen (Beispielitem: „Was meinen Sie? Können Sie die folgenden Tätigkeiten im inklusiven Unterricht mit Schülerinnen/Schülern mit und ohne sonderpädagogischem/n Förderbedarf besser allein oder in Kooperation mit einer anderen Lehrkraft realisieren? – Alle Schülerinnen und Schüler unabhängig von ihren Leistungen unterrichten.“; $M = 3.72$, $SD = 0.73$, $Min = 1.00$, $Max = 5.00$, $\alpha = .82$). Für die Erstellung dieser Fragebogenskala wurden Fragebogenbogenitems von Skaalvik und Skaalvik (2007) adaptiert. Die an der Studie beteiligten Studentinnen und Studenten beantworteten die einzelnen Fragebogenitems auf fünfstufigen Likert-Skalen (1 = - -, 2 = –, 3 = 0, 4 = +, 5 = ++; Skala „Kollektive Selbstwirksamkeit“: 1 = besser allein, 2 = eher besser allein, 3 = unentschieden, 4 = eher besser in Kooperation, 5 = besser in Kooperation).

Durchführung/Datenanalyse

Die Befragungen wurden in regulären Lehrveranstaltungen, die jeweils für Studentinnen und Studenten der Bachelor- und Master-Studiengänge „Lehramt an Grundschulen“ und „Sonderpädagogische Förderung“ angeboten werden, durchgeführt. Die Teilnahme an den Befragungen erfolgte auf freiwilliger Basis, wobei den Studierenden mitgeteilt wurde, dass ihnen durch eine Nicht-Teilnahme keine Nachteile entstehen würden. Die Befragungen dauerten nicht länger als 20 Minuten. Für die Überprüfung der Unterscheidungshypothesen wurde eine mehrfaktorielle Varianzanalyse in SPSS (Version 25) berechnet. Für die Ermittlung der Effektstärken wurde η^2 herangezogen: $\eta^2 = .01$ (kleiner Effekt), $\eta^2 = .06$ (mittlerer Effekt) und $\eta^2 = .14$ (großer Effekt) (Sedlmeier/Renkewitz 2013).

5 Ergebnisse

In den Tabellen 1 und 2 sind die deskriptiven und in Tabelle 3 die analytischen Befunde aus der berechneten mehrfaktoriellen Varianzanalyse zu den Unterschieden in den Einstellungen der Studierenden aus den Bachelor- und Masterstudiengängen des Lehramts an Grundschulen und des Lehramts für sonderpädagogische Förderung zur Inklusion in Hinblick auf die schulische Förderung und die soziale Teilhabe im inklusiven Unterricht dargestellt, ebenso wie Unterschiede zwischen diesen Studierendengruppen in Hinblick auf ihre individuellen und kollektiven Selbstwirksamkeitsüberzeugungen.

Tabelle 1: Unterschiede zwischen Bachelor- und Masterstudierenden des Lehramts an Grundschulen und des Lehramts für sonderpädagogische Förderung hinsichtlich ihrer Einstellungen zur Inklusion (Schulische Förderung, Soziale Teilhabe) und ihrer Selbstwirksamkeitsüberzeugungen (individuell, kollektiv)

	Grundschullehramt						Lehramt Sonderpädagogik					
	BA (N = 273)		MA (N = 90)		Gesamt (N = 363)		BA (N = 189)		MA (N = 48)		Gesamt (N = 237)	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Einstellungen zur Inklusion:												
– Schulische Förderung	2.70	0.76	2.67	0.64	2.69	0.73	3.21	0.75	3.30	0.63	3.23	0.72
– Soziale Teilhabe	3.66	0.65	3.80	0.61	3.70	0.64	3.74	0.71	3.74	0.62	3.71	0.66
Selbstwirksamkeit (individuell)	3.40	0.63	3.07	0.65	3.32	0.65	3.79	0.58	3.47	0.70	3.73	0.62
Selbstwirksamkeit (kollektiv)	3.44	0.70	3.94	0.66	3.56	0.72	3.89	0.68	4.19	0.65	3.95	0.69

Anmerkungen. Listenweiser Fallausschluss (N = 600); BA = Bachelor of Education; MA = Master of Education

Tabelle 2: Unterschiede zwischen Studierenden der Bachelor- und Masterstudiengänge (Lehramt an Grundschulen *und* Lehramt für sonderpädagogische Förderung) hinsichtlich ihrer Einstellungen zur Inklusion (Schulische Förderung, Soziale Teilhabe) und ihrer Selbstwirksamkeitsüberzeugungen (individuell, kollektiv)

	BA (N = 462)		MA (N = 138)	
Variable	*M*	*SD*	*M*	*SD*
Einstellungen zur Inklusion:				
– Schulische Förderung	2.91	0.79	2.89	0.70
– Soziale Teilhabe	3.69	0.67	3.78	0.61
Selbstwirksamkeit (individuell)	3.56	0.64	3.20	0.69
Selbstwirksamkeit (kollektiv)	3.62	0.73	4.03	0.67

Anmerkungen: Listenweiser Fallausschluss (N = 600); BA = Bachelor of Education; MA = Master of Education

Tabelle 3: Mehrfaktorielle Varianzanalyse der untersuchten Variablen in Abhängigkeit des Studiengangs (Grundschullehramt, Lehramt Sonderpädagogik) und der Studienphase (Bachelorstudiengang, Masterstudiengang)

	Gesamtmodell			Studiengang			Studienphase			Studiengang x Studienphase		
	F	*p*	η^{2*}	*F*	*p*	η^2	*F*	*p*	η^2	*F*	*p*	η^2
Multivariat[a]				32.68	≤.001	.18	14.31	≤.001	.09	1.22	.30	.01
Univariat[b]												
Einstellungen zur Inklusion:												
– Schulische Förderung	26.11	≤.001	.12	60.02	≤.001	.09	0.20	.66	.00	0.68	.41	.00
– Soziale Teilhabe	1.13	.34	.01	0.01	.93	.00	1.09	.30	.00	0.99	.32	.00
Selbstwirksamkeit (individuell)	30.47	≤.001	.13	39.51	≤.001	.06	27.23	≤.001	.04	0.00	.95	.00
Selbstwirksamkeit (kollektiv)	30.29	≤.001	.13	25.91	≤.001	.04	33.91	≤.001	.05	2.18	.14	.00

Anmerkungen. *) Korrigiertes η^2; [a]) Multivariat $df = 3, 596$; [b]) Univariat $df = 1, 596$; Pillai-Spur

Signifikante Haupteffekte können auf der Grundlage der berechneten mehrfaktoriellen Varianzanalyse für die beiden festen Variablen ‚Studiengang' und ‚Studienphase' berichtet werden, nicht aber für den Interaktionseffekt ‚Studiengang x Studienphase'.

Im Detail werden dabei signifikante Unterschiede zwischen den Studierenden des Lehramts an Grundschulen und denjenigen für das Lehramt für sonderpädagogische Förderung in Bezug auf ihre Einstellungen zur schulischen Förderung im inklusiven Unterricht sowie in Hinblick auf ihre Selbstwirksamkeitsüberzeugungen (individuell, kollektiv) zugunsten der Studierenden des Lehramts für sonderpädagogische Förderung deutlich. Die Studierenden des Lehramts an Grundschulen und diejenigen des Lehramts für sonderpädagogische Förderung unterscheiden sich allerdings nicht signifikant hinsichtlich ihrer Einstellungen zur sozialen Teilhabe im inklusiven Klassenzimmer voneinander.

Signifikante Unterschiede zwischen den an der Studie beteiligten Bachelor- und Masterstudierenden liegen in Bezug auf ihre Einstellungen zur schulischen Förderung und zur sozialen Teilhabe im inklusiven Unterricht nicht vor. Jedoch verfügen die Bachelorstudierenden über signifikant positivere individuelle Selbstwirksamkeitsüberzeugungen als die Masterstudierenden, wohingegen die Masterstudierenden signifikant positivere kollektive Selbstwirksamkeitsüberzeugungen haben als die Bachelorstudierenden.

6 Diskussion

Der Ausgangspunkt der vorliegenden Studie war die Frage danach, ob Unterschiede zwischen Studentinnen und Studenten des Lehramts an Grundschulen und solchen mit dem Studienschwerpunkt ‚Sonderpädagogische Förderung' in Bezug auf ihre Einstellungen zur Inklusion und ihre Selbstwirksamkeitsüberzeugungen für die Gestaltung des eigenen inklusiven Unterrichts vorliegen. Für die Beantwortung dieser Forschungsfrage wurden insgesamt $N = 606$ Studierende des Lehramts an Grundschulen und des Lehramts für sonderpädagogische Förderung anhand eines Fragebogens gebeten, Stellung zu ihren Einstellungen zur Inklusion sowie zu ihren Selbstwirksamkeitsüberzeugungen zu nehmen. Die Ergebnisse verdeutlichen – so wie zuvor (explorativ) angenommen – signifikante Unterschiede in den Einstellungen zur schulischen Förderung im inklusiven Unterricht sowie in den Selbstwirksamkeitsüberzeugungen (individuell, kollektiv) zugunsten der Studierenden des Lehramts für sonderpädagogische Förderung. Die Befunde sind damit weitgehend kompatibel mit denjenigen bereits vorliegender Studien (z. B. McHatton/Parker 2013; Leyser/Zeiger/Romi 2011; Sharma/Shaukat/Furlonger 2015). Zukünftig müsste untersucht werden, anhand welcher Variablen sich die Unterschiede in den Einstel-

lungen zur Inklusion und den Selbstwirksamkeitsüberzeugungen bei Grund- und Förderschullehrkräften erklären lassen. Möglicherweise sind diese Unterschiede zugunsten der Studierenden des Lehramts an Förderschulen auf ihre Berufswahlmotive zurückzuführen, da sie sich bewusst für das Unterrichten von Kindern und Jugendlichen mit sonderpädagogischem Förderbedarf entschieden haben. Diese kausalen Zusammenhänge könnten in weiterführenden Untersuchungen in den Blick genommen werden.

Entgegen den Erwartungen unterscheiden sich die Studierenden des Lehramts an Grundschulen und des Lehramts für sonderpädagogische Förderung nicht signifikant in Hinblick auf ihre Einstellungen zur sozialen Teilhabe von Kindern mit und ohne Förderbedarf im inklusiven Klassenzimmer. Dieser Befund geht nicht einher mit demjenigen von Moser et al. (2014), die Unterschiede zugunsten der an der Studie beteiligten Förderschullehrkräfte finden konnten, oder demjenigen von Trumpa et al. (2014), die wiederum Differenzen zugunsten der im Rahmen ihrer Untersuchung befragten Regelschullehrkräfte nachweisen konnten. In nachfolgenden Studien müsste der Frage, unter welchen Bedingungen (angehende) Grund- und Förderschullehrkräfte der sozialen Teilhabe von Kindern mit und ohne Förderbedarf positiv gegenüber aufgeschlossen sind, nochmals genauer nachgegangen werden. Möglicherweise resultieren die jeweils im Rahmen der einzelnen Studien gemessenen Einstellungen der (angehenden) Lehrkräfte aus der Qualität der Erfahrungen, die sie im inklusiven Unterricht sammeln konnten. In verschiedenen Studien konnte nachgewiesen werden, dass enge Zusammenhänge zwischen den Erfahrungen von (angehenden) Lehrkräften aus dem inklusiven Unterricht, ihren Einstellungen zur Inklusion sowie ihren Selbstwirksamkeitsüberzeugungen bestehen (z. B. Bosse/Spörer 2014; Hellmich/Görel/Schwab 2016). In weitführenden Untersuchungen könnten bei der Betrachtung von Unterschieden in den Einstellungen und Selbstwirksamkeitsüberzeugungen zwischen angehenden Grund- und Förderschullehrkräften ihre Erfahrungen aus dem inklusiven Unterricht kontrolliert werden. Dies stellt sicherlich auch eine Limitation der vorliegenden Studie dar.

Die von uns berichteten Befunde können Berücksichtigung bei der Ausbildung von angehenden Grund- und Förderschullehrkräften im Rahmen des Lehramtsstudiums (z. B. in Lehrveranstaltungen oder Praxisphasen) finden, deuten sie doch auf sehr unterschiedliche Eingangsvoraussetzungen dieser beiden Gruppen von Studierenden in Hinblick auf den inklusiven Unterricht hin.

Literatur

Bandura, Albert (1997): Self-efficacy: The exercise of control. New York: Freeman.

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520.

Bosse, Stefanie/Spörer, Nadine (2014): Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. In: Empirische Sonderpädagogik 6, H. 4, S. 279–299.

Demirtas, Vesile Y. (2018): A study on teacher candidates' self-efficacy, motivation and affection levels of children. In: Journal of Education and Training Studies 6, H. 12, S. 111–125.

Eagly, Alice E./Chaiken, Shelly (1993): The psychology of attitudes. Fort Worth, TX: Harcourt Brace Jovanovich.

Flores, Margaret M./Thornton, Jennifer/Franklin, Toni M./Strozier, Shaunita (2014): Elementary general and special education teachers' mathematic skills and efficacy. In: Journal of Research in Education 24, H. 1, S. 69–82.

Freytag, Cathy E. (2001): Teacher efficacy and inclusion: The impact of preservice experiences on beliefs. Paper presented at the 24th Annual Meeting of the Southwest Educational Research Association, February 2001. New Orleans: Louisiana.

Hellmich, Frank/Görel, Gamze (2014): Erklärungsfaktoren für Einstellungen von Lehrerinnen und Lehrern zum inklusiven Unterricht in der Grundschule. In: Zeitschrift für Bildungsforschung 4, H. 3, S. 227–240.

Hellmich, Frank/Görel, Gamze/Schwab, Susanne (2016): Einstellungen und Motivation von Lehramtsstudentinnen und -studenten in Bezug auf den inklusiven Unterricht in der Grundschule – Ein Vergleich zwischen Deutschland und Österreich. In: Empirische Sonderpädagogik 8, H. 1, S. 67–85.

Kopp, Bärbel (2009): Inklusive Überzeugung und Selbstwirksamkeit im Umgang mit Heterogenität. Wie denken Studierende des Lehramts für Grundschulen? In: Empirische Sonderpädagogik 1, H. 1, S. 5–25.

Kuhl, Jan/Redlich, Hubertus/Schäfer, Lea (2014): Einstellungen verschiedener Lehrergruppen gegenüber Menschen mit geistiger Behinderung. In: Zeitschrift für Bildungsforschung 4, H. 3, S. 271–287.

Kunz, André/Luder, Reto/Moretti, Marta (2010): Die Messung von Einstellungen zur Integration (EZI). In: Empirische Sonderpädagogik 2, H. 3, S. 83–94.

Lamture, Sadhana/Gathoo, Varsha S. (2017): Self-efficacy of general and resource teachers in education of children with disabilities in India. In: International Journal of Special Education 32, H. 4, S. 809–822.

Leyser, Yona/Zeiger, Tali/Romi, Shlomo (2011): Changes in self-efficacy of prospective special and general education teachers: Implication for inclusive education. In: International Journal of Disability, Development and Education 58, H. 3, S. 241–255.

Martínez, Rebecca S. (2003): Impact of a graduate class on attitudes toward inclusion, perceived teaching efficacy and knowledge about adapting instruction for children with disabilities in inclusive settings. In: Teacher Development 7, H. 3, S. 473–494.

McHatton, Patricia A./Parker, Audra (2013): Purposeful preparation: Longitudinally exploring inclusion attitudes of general and special education pre-service teachers. In: Teacher Education and Special Education 36, H. 3, S. 186–203.

Moser, Vera/Kuhl, Jan/Redlich, Hubertus/Schäfer, Lea (2014): Beliefs von Studierenden sonder- und grundschulpädagogischer Studiengänge. In: Zeitschrift für Erziehungswissenschaft 17, S. 661–678.

Özokcu, Osman (2017): Investigating the relationship between teachers' self-efficacy beliefs and efficacy for inclusion. In: European Journal of Special Education Research 2, H. 6, S. 234–252.

Scholz, Markus/Rank, Astrid (2016): Perspektive Inklusion. Inklusionsverständnis und Einstellungen zur integrativen Beschulung bei Studierenden des Grundschul- und Förderschullehramts. In: Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete 85, H. 1, S. 53–67.

Schwab, Susanne/Hellmich, Frank/Görel, Gamze (2017): Self-efficacy of prospective Austrian and German primary school teachers regarding the implementation of inclusive education. In: Journal of Research in Special Educational Needs 17, H. 3, S. 205–217.

Sedlmeier, Peter/Renkewitz, Frank (2013): Forschungsmethoden und Statistik für Psychologen und Sozialwissenschaftler. 2. Auflage. Hallbergmoos: Pearson.

Sharma, Umesh/Shaukat, Sadia/Furlonger, Brett (2015): Attitudes and self-efficacy of preservice teachers towards inclusion in Pakistan. In: Journal of Research in Special Educational Needs 15, H. 2, S. 97–105.

Skaalvik, Einar M./Skaalvik, Sidsel (2007): Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. In: Journal of Educational Psychology 99, H. 3, S. 611–625.

Trumpa, Silke/Janz, Frauke/Heyl, Vera/Seifried, Stefanie (2014): Einstellungen zu Inklusion bei Lehrkräften und Eltern – Eine schulartspezifische Analyse. In: Zeitschrift für Bildungsforschung 4, H. 3, S. 241–256.

Woll, Anke (2017): „Kontaktbedingungen zwischen Menschen mit und ohne Behinderung als Prädiktoren von Einstellungen zu Inklusion". https://opus.ph-heidelberg.de/frontdoor/index/index/docId/230 (Abfrage: 21.03.2019).

Welche Rolle spielen Einflussfaktoren vor und während des Studiums für die Überzeugungen von Lehramtsstudierenden zu schulischer Inklusion?

Ariane S. Willems und Karina Meyer

Abstract

Proinklusive Überzeugungen von Lehrkräften sind zentrale Erfolgsdeterminanten für die Umsetzung einer inklusiven Beschulung von Schülerinnen und Schülern an Regelschulen. Eine gezielte Förderung positiver Einstellungen und Kompetenzüberzeugungen sollte daher auch bereits in der universitären Lehramtsausbildung erfolgen. Aus Forschungsperspektive ist in diesem Zusammenhang allerdings die Frage, welchen Einfluss inklusionsbezogene Lernerfahrungen während des Studiums gegenüber entsprechenden Erfahrungen vor Beginn des Studiums auf proinklusive Überzeugungen haben, bislang kaum untersucht. Der vorliegende Beitrag widmet sich der Analyse dieser Zusammenhänge auf Basis von Daten des Projekts „ELBI – Einstellungen von Lehramtsstudierenden zu Behinderung und Inklusion". Die Ergebnisse zeigen, dass für proinklusive Einstellungen sowohl die Vorerfahrungen mit Menschen mit Behinderung als auch das pädagogische Interesse als Studienwahlmotiv eine größere Rolle spielen als der Besuch inklusionsbezogener Lehrveranstaltungen und Praktika während des Studiums. Letztere erweisen sich demgegenüber als prädiktiv für die subjektive Einschätzung inklusiver Kompetenzen der Studierenden. Diskutiert werden Implikationen für die Lehramtsausbildung insbesondere hinsichtlich der Fragen, an welchen Stellen und wie die Förderung inklusiver Kompetenzen und proinklusiver Einstellungen ansetzen kann.

1 Problemstellung

Durch die Öffnung der Regelschulen für alle Schülerinnen und Schüler infolge der Ratifizierung der UN-Behindertenrechtskonvention (2009) steht die Lehramtsausbildung vor der Aufgabe, Lehrkräfte für den gemeinsamen Unterricht

von Schülerinnen und Schülern mit und ohne Behinderung bzw. sonderpädagogischen Förderbedarf auszubilden. Gemäß der *Standards für die Lehrerbildung* (KMK 2014) sollen Studierende u. a. dazu befähigt werden, individuelle Lernausgangslagen von Schülerinnen und Schülern festzustellen und darauf abgestimmt spezielle Fördermaßnahmen im Unterricht zu implementieren.

Als Bestandteile inklusiver Kompetenzen gelten dabei explizit auch die Befürwortung von schulischer Inklusion – im Sinne positiver proinklusiver Einstellungen – sowie die subjektive Einschätzung, hierfür über die notwendigen Kompetenzen zu verfügen (Gebhardt et al. 2018; Watkins 2012). Die Frage, wovon diese Überzeugungen im Einzelnen abhängen, ist Gegenstand unterschiedlicher Forschungsarbeiten. Während dabei zur Bedeutsamkeit von ausbildungs- und berufsbezogenen Faktoren, wie z. B. dem Besuch spezieller Fortbildungsmaßnahmen, relativ gesicherte Befunde vorliegen (Pit-ten Cate et al. 2018), ist weitgehend ungeklärt, wie motivationale Orientierungen oder inklusionsbezogene Erfahrungen, die Studierende vor Aufnahme des Studiums gemacht haben, inklusive Überzeugungen beeinflussen bzw. in welcher Relation die Einflüsse von Vorerfahrungen und konkreten Lernerfahrungen im Studium stehen. Ziel des vorliegenden Beitrages ist es daher, auf der Basis von Daten des Projekts „Einstellungen von Lehramtsstudierenden zu Behinderung und Inklusion" (ELBI) die relative Vorhersagekraft von Merkmalen der Studieneingangsvoraussetzungen sowie studienbezogener Variablen für die Ausprägung proinklusiver Einstellungen und Kompetenzeinschätzungen von Lehramtsstudierenden zu untersuchen.

2 (Proinklusive) Einstellungen und Kompetenzüberzeugungen als Aspekte professioneller Lehrerkompetenz

In Übereinstimmung mit Modellen zur Beschreibung der Struktur und Entwicklung professioneller Lehrerkompetenzen (Baumert/Kunter 2006; Krauss/Bruckmaier 2014) werden auch für Inklusion Einstellungen und Kompetenz- bzw. Selbstwirksamkeitsüberzeugungen als zentraler Bestandteil professioneller Kompetenzen von Lehrkräften aufgefasst (Gebhardt et al. 2018). Zwar steht die empirische Erforschung konkreter Effekte proinklusiver Überzeugungen auf das inklusionsbezogene Lehrerhandeln selbst noch am Anfang (Kullmann et al. 2014), Befunde der allgemeinen Einstellungs- und Selbstwirksamkeitsforschung stützen aber die Annahme, dass Einstellungen und Kompetenzüberzeugungen zentrale Prädiktoren professionellen Handelns sind (Fishbein/Ajzen 2010; Voss et al. 2011; Warner/Schwarzer 2009).

Für die Entwicklung entsprechender Lehrerkompetenzen spielen neben der Bereitstellung und Nutzung (qualitätsvoller) Lerngelegenheiten auch Studieneingangsvoraussetzungen – wie u. a. Studienwahlmotive sowie pädagogische

Vorerfahrungen – eine entscheidende Rolle (Kunter et al. 2011; Blömeke 2004; Rothland 2014, 2015). In Bezug auf die Entwicklung von inklusionsbezogenen Überzeugungen von (angehenden) Lehrkräften zeigt sich zudem, dass sowohl bei Lehrkräften (Pit-ten Cate et al. 2018) als auch bei Lehramtsstudierenden (Bosse/Spörer 2014; Hellmich/Görel/Schwab 2016; McElvany et al. 2018; Ruberg/Porsch 2017) inklusionsbezogene Kompetenzüberzeugungen positiv mit proinklusiven Einstellungen zusammenhängen. Dabei variieren Einstellungen zu Inklusion je nach Behinderungsform: (Zukünftige) Lehrkräfte schätzen demnach die Inklusion von Schülerinnen und Schülern mit körperlichen Einschränkungen weniger kritisch ein als die Inklusion von Schülerinnen und Schülern mit kognitiven Einschränkungen und Verhaltensauffälligkeiten (de Boer/Pijl/Minnaert 2011; Gebhardt et al. 2011; Ruberg/Porsch 2017). Proinklusive Überzeugungen nehmen mit zunehmender Berufserfahrung – operationalisiert über die Anzahl an Berufsjahren als Lehrkraft – ab, werden allerdings durch praktische Unterrichtserfahrungen in inklusiven Settings positiv beeinflusst (de Boer/Pijl/Minnaert 2011; Bosse/Spörer 2014; Hellmich/Görel/Schwab 2016; Gebhardt et al. 2011; McElvany et al. 2018; Ruberg/Porsch 2017). Während die Teilnahme an inklusionsbezogenen Aus- und Fortbildungskursen proinklusive Einstellungen von Studierenden (Kopp 2009) und Lehrkräften (de Boer/Pijl/Minnaert 2011; McElvany et al. 2018; Pit-ten Cate et al. 2018; Sharma et al. 2006) begünstigen, sind die Effekte solcher Maßnahmen auf die Kompetenzüberzeugungen nicht eindeutig nachweisbar (Loreman/Sharma/Forlin 2013; McElvany et al. 2018). Zudem weisen Bosse und Spörer (2014) darauf hin, dass Studienanfängerinnen und Studienanfänger ihre Kompetenzen positiver einschätzen als Masterstudierende. Insgesamt scheint außerdem der persönliche Kontakt mit Menschen mit Behinderung proinklusive Einstellungen (Bosse/Spörer 2014; Sharma et al. 2006) und – zumindest teilweise – auch Kompetenzüberzeugungen (Loreman/Sharma/Forlin 2013) zu begünstigen. Inwieweit proinklusive Überzeugungen bereits von der eigenen schulischen Sozialisation in inklusiven Settings bedingt werden, ist umstritten: Sharma und Kollegen (2006) weisen hier auf einen positiven Effekt der eigenen inklusiven Beschulung hin, der allerdings in einer Studie im deutschsprachigen Raum nicht repliziert wird (Schwab 2014).

Studien, die sich mit der Frage von Unterschieden in proinklusiven Überzeugungen von Studierenden verschiedener Lehrämter befassen, zeigen, dass sowohl Studierende des Förderschullehramtes als auch sonderpädagogische Lehrkräfte positivere Einstellungen zu Inklusion und höhere Kompetenzüberzeugungen aufweisen als Studierende anderer Lehrämter (Ruberg/Porsch 2017; Trumpa et al. 2014). Allerdings ist die Befundlage hier ebenso wie hinsichtlich der Frage von Unterschieden zwischen Sekundar- und Grundschullehrkräften (McElvany et al. 2018; Trumpa et al. 2014) nicht eindeutig. Offen sind schließlich auch die Fragen, inwieweit sich proinklusive Überzeugungen zwischen

Studierenden unterschiedlicher Fachrichtungen bzw. unterschiedlicher Studienwahlmotive unterscheiden: Zur Abhängigkeit proinklusiver Überzeugungen von konkreten Unterrichtsfächern geben zurzeit ausschließlich theoretisch-konzeptuelle Überlegungen (Rolka/Albersmann 2017; Stroh 2015) sowie erste Fallstudien (Abels et al. 2018) Anlass zu der Vermutung, dass inklusiver Unterricht im Kontext von MINT-Fächern besondere Herausforderungen birgt, was die Ausbildung proinklusiver Überzeugungen erschweren könnte. In Bezug auf die Frage des Einflusses von Studienwahlmotiven erscheint schließlich auf Basis der Befunde, dass Vorteile von Inklusion zumeist im sozialen Lernen gesehen werden (Trumpa et al. 2014), die Annahme plausibel, dass ein hohes Maß an pädagogischem Interesse der Lehramtsstudierenden proinklusive Überzeugungen begünstigen könnte. Vor diesem Hintergrund kann auch vermutet werden, dass Gymnasialstudierende der schulischen Inklusion kritischer gegenüberstehen als andere Lehramtsstudierende, da ihnen in verschiedenen Studien ein vergleichsweise höheres fachliches und geringeres pädagogisches Interesse attestiert wird (z. B. Retelsdorf/Möller 2012; siehe auch McElvany et al. 2018; Laubner 2019).

3 Forschungsfragen, Datengrundlage und Methodik

Ausgehend von den dargestellten Forschungsbefunden analysieren wir auf der Grundlage der Daten des Projektes *„ELBI – Einstellungen von Lehramtsstudierenden zu Behinderung und Inklusion“* die relative Vorhersagekraft von Merkmalen der Studieneingangsvoraussetzungen sowie studienbezogener Variablen für die Ausprägung proinklusiver Einstellungen und subjektiver Kompetenzeinschätzungen von Lehramtsstudierenden. Dabei gehen wir im Einzelnen folgenden Fragen nach:

1. Inwiefern werden proinklusive Einstellungen und subjektiv eingeschätzte inklusive Kompetenzen von Lehramtsstudierenden durch
   a. die Studienwahlmotivation,
   b. die (inklusionsbezogenen) Vorerfahrungen mit Menschen mit Behinderung,
   c. die Fächer- und Schulformwahl sowie
   d. konkrete inklusionsbezogene Lernerfahrungen in der universitären Ausbildung vorhergesagt?
2. Welche relative Bedeutung kommt den einzelnen Prädiktoren für die Ausprägung proinklusiver Einstellungen und subjektiv eingeschätzter inklusiver Kompetenzen zu?

Stichprobe

An der querschnittlich angelegten Fragebogenstudie nahmen im Mai 2018 $N = 424$ Lehramtsstudierende niedersächsischer Universitäten teil (Alter $M = 23.51$, $SD = 3.89$; 75 % weiblich; 45 % Masterstudierende). Knapp die Hälfte (44 %) besuchte im Studium bereits mindestens eine Lehrveranstaltung zum Thema Inklusion. Zudem gibt die Hälfte (50 %) der Studierenden an, durch Schulpraktika oder andere schulpraktische Tätigkeiten bereits über Erfahrungen mit inklusivem Unterricht zu verfügen (vgl. Tab. 2).

Messinstrumente

Für die Messung von *proinklusiven Einstellungen* wurden auf Basis bestehender Instrumente (Gebhardt et al. 2011; Kunz/Luder/Moretti 2010; Bosse/Spörer 2014) Skalen entwickelt, die auf zweifache Weise differenzieren: zum einen zwischen den drei Behinderungsformen körperlich, geistig und psychisch und zum anderen zwischen den beiden Facetten „Vorteile von Inklusion für die inkludierten Schülerinnen und Schüler" bzw. „Vorteile von Inklusion für die Mitschülerinnen und Mitschüler" (vgl. hierzu auch Willems/Meyer/Frechen 2019). Die Subskalen bestehen für die drei Behinderungsformen und die beiden Facetten aus jeweils vier sechsstufig skalierten Items (vgl. Tab. 1). Für die Erfassung der *subjektiv wahrgenommenen Kompetenzüberzeugung* wurden in Anlehnung an die „Standards für die Lehrerbildung" (KMK 2014) zwei Subskalen entwickelt, die die Ausprägung des subjektiv wahrgenommenen theoretischen Wissens bzw. der selbst eingeschätzten Handlungskompetenz erfassen. Für beide Dimensionen werden in dem vorliegenden Beitrag Kurzskalen bestehend aus jeweils vier vierstufig skalierten Items verwendet, wobei jeweils ein Item aus den Kompetenzbereichen Unterrichten, Erziehen, Beurteilen und Innovieren stammt.

Tabelle 1: Übersicht und statistische Kennwerte der latenten Konstrukte (Prädiktoren und Kriteriumsvariablen)

Bereich	Skalen (je 4 Items)	Beispielitem	*M (SD)*	*α*
Studienwahlmotive ($N \geq 423$)	Pädagogisches Interesse[a]	Ich habe ein Lehramtsstudium gewählt, weil mir die Entwicklung von Persönlichkeiten ein echtes Anliegen ist	3.52 (0.54)	.85
	Fachliches Interesse[a]	Ich habe ein Lehramtsstudium gewählt, weil ich die Inhalte meiner Fächer interessant finde	3.51 (0.53)	.88
Proinklusive Einstellungen (jeweils drei Skalen; $N \geq 289$)	Vorteile für die inkludierten Schülerinnen und Schüler	Kinder mit geistiger/körperlicher/psychischer Behinderung lernen in Regelschulen wichtige Dinge, die sie in Förderschulen nicht lernen würden	3.18^{A} (1.11) / 4.27^{B} (1.04) / 3.92^{C} (1.05)	.83 / .78 / .83
	Vorteile für die Mitschülerinnen und Mitschüler	Wenn Kinder mit geistiger/körperlicher/psychischer Behinderung in der Klasse sind, hat das für die anderen Kinder Vorteile, trotz möglicher Schwierigkeiten	4.54^{D} (1.00) / 5.14^{E} (0.77) / 4.55^{D} (0.97)	.83 / .77 / .78
Selbst eingeschätzte inklusive Kompetenzen ($N \geq 404$)	Theoretisches Wissen[a]	Ich weiß, was bei der Planung von Unterrichtseinheiten in inklusiven Klassen beachtet werden muss	2.41^{A} (0.65)	.78
	Handlungspraxis[a]	Ich bin der Lage, Unterricht für inklusive Klassen zu planen und zu gestalten	2.09^{B} (0.65)	.83

Anmerkung: Likertskalen von [a]1 – trifft gar nicht zu bis 4 – trifft voll zu bzw. [b]1 – stimme gar nicht zu bis 6 – stimme voll zu. Mittelwerte des gleichen Bereichs mit unterschiedlichen Indices (A-E) unterscheiden sich signifikant ($p \leq .05$).

Zur Messung der *Studienwahlmotive* wurde eine leicht adaptierte und auf die beiden Dimensionen des pädagogischen bzw. des fachlichen Interesses gekürzte Variante des Instrumentes FEMOLA (Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums, Pohlmann/Möller 2010) eingesetzt (für eine Übersicht: vgl. Tab. 1). Beide Dimensionen wurden mit jeweils vier vierstufig skalierten Items erfasst. Die einzelnen Skalen wurden anhand konfirmatorischer Faktorenanalysen mit M*plus* 8.3 (Muthén/Muthén 2017) auf ihre Eindimensionalität hin geprüft und weisen befriedigende bis sehr gute Kennwerte auf (jeweils $CFI \geq .98$, $TLI \geq .94$, $RMSEA \leq .12$, $SRMR \leq .02$). Fehlende Werte wurden modellbasiert unter Verwendung des *Full-Information-Maximum-Likelihood*-Algorithmus geschätzt. Die Reliabiliäten aller Skalen sind mindestens zufriedenstellend ($\alpha \geq .77$). Zusätzlich wurden verschiedene mani-

feste Indikatoren zur Beschreibung weiterer Studieneingangsvoraussetzungen und ausbildungsbezogener Merkmale genutzt, die in Tabelle 2 dargestellt sind.

Tabelle 2: Übersicht und statistische Kennwerte der manifesten Indikatoren (Prädiktoren) ($N \geqslant 411$)

Bereich	Item	Antwortformat	Verteilung
Erfahrungen mit (Menschen mit) Behinderung vor dem Studium	Ausmaß vorherigen Kontakts	1 (sehr wenig) – 5 (sehr viel)	$M = 2.53$ $SD = 1.11$
	Bewertung vorherigen Kontakts	1 (sehr negativ) – 5 (sehr positiv)	$M = 3.58$ $SD = 0.85$
	Kontakt in Schulklasse	0 (nein) / 1 (ja)	37 % ja
Schulform- und Fächerwahl	Studium Gymnasiallehramt	0 (nein) / 1 (ja)	87 % ja
	Mindestens ein MINT-Unterrichtsfach	0 (nein) / 1 (ja)	32 % ja
Ausbildungsstand	Aktuelles Fachsemester	Absolute Zahl	$M = 6.65$ $SD = 2.99$
Inklusionsbezogene Lernerfahrungen im Studium	Absolvierte Veranstaltungen zu Inklusion	Absolute Anzahl	$M = 0.9$ $SD = 1.41$
	Praktische Erfahrungen mit inklusivem Unterricht	0 (nein) / 1 (ja)	50 % ja

Zur Ermittlung der relativen Vorhersagekraft einzelner Merkmale auf die Ausprägung proinklusiver Einstellungen und Kompetenzüberzeugungen wurden drei identisch aufgebaute – jedoch für die drei Behinderungsformen separate – Strukturgleichungsmodelle mit M*plus* 8.3 (Muthén/Muthén 2017) angepasst. In alle Modelle gingen das pädagogische und das fachliche Interesse der Studierenden sowie die dargestellten studienbezogenen und dem Studium vorgelagerte Indikatoren (vgl. Tab. 2) als Prädiktorvariablen ein. Die subjektiv wahrgenommene inklusive Kompetenz wird in allen drei Modellen als abhängige Variable modelliert. In Modell 1 werden zusätzlich als abhängige Variablen die zwei Einstellungsskalen zur Inklusion bei geistiger Behinderung aufgenommen, in Modell 2 die Skalen zu körperlicher Behinderung und in Modell 3 die Skalen zu psychischer Behinderung (vgl. Abb. 1).

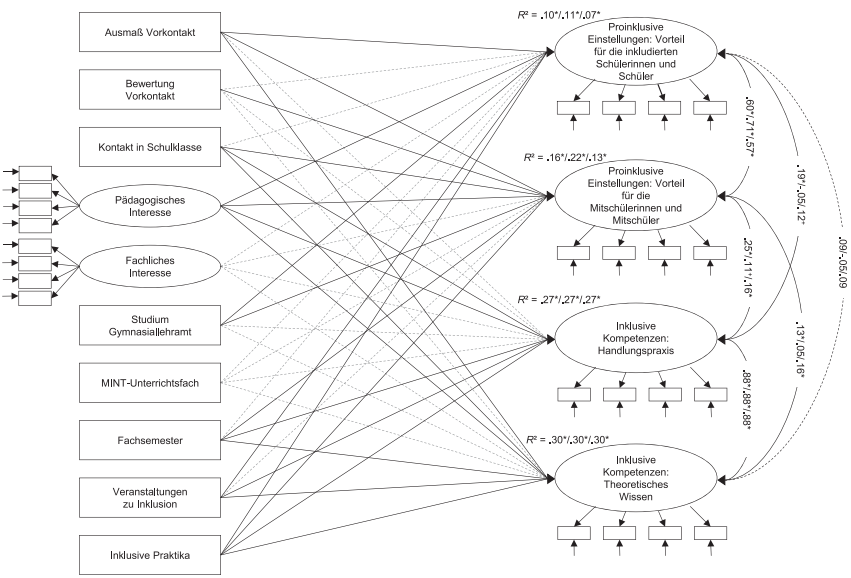
4 Ergebnisse

Die Analysen zur Beschreibung der abhängigen Variablen zeigen zunächst erwartungsgemäß, dass die proinklusiven Einstellungen von Studierenden am stärksten für den Bereich der Inklusion von Schülerinnen und Schülern mit körperlicher Behinderung ausgeprägt sind. Weniger positiv sind die Einstellun-

gen gegenüber der Inklusion von Schülerinnen und Schülern mit psychischer bzw. geistiger Behinderung. Die Unterschiede zwischen den Behinderungsformen sind dabei statistisch signifikant (vgl. Tab. 1). Unabhängig von der berücksichtigten Behinderungsform werden zudem die Auswirkungen für die Mitschülerinnen und Mitschüler signifikant positiver eingeschätzt als für die inkludierten Schülerinnen und Schüler selbst. Wie für diesen Ausbildungsabschnitt zu erwarten, schätzen Studierende ihre theoretischen inklusiven Kompetenzen signifikant höher als ihre praktischen Kompetenzen ein.

In Abbildung 1 und Tabelle 3 sind die Befunde der Strukturgleichungsmodelle dargestellt. Die Modelle weisen für alle drei Behinderungsformen gute bis sehr gute Modell-Fits auf. Durch die berücksichtigten Prädiktoren können insgesamt 27 % bzw. 30 % der Varianz in den Kompetenzüberzeugungen erklärt werden. Die Anteile erklärter Varianz bei den einstellungsbezogenen Dimensionen sind insgesamt geringer ($.07 \leq R^2 \leq .22$).

Abbildung 1: Strukturgleichungsmodelle zur Vorhersage von inklusiven Kompetenzeinschätzungen und Einstellungen zu Inklusion (Modell 1: geistige Behinderung/Modell 2: körperliche Behinderung/Modell 3: psychische Behinderung)



Anmerkung: Modell-Fit-Werte jeweils: $\chi^2/df \leq 1.73$, $CFI \geq .97$, $TLI \geq .97$, $RMSEA \leq .04$, $SRMR \leq .07$. Durchgezogene Pfeile markieren Pfadkoeffizienten bzw. latente Korrelationen (zwischen den abhängigen Variablen), die in mindestens einem der Modelle (geistig/körperlich/psychisch) auf dem Niveau von $p \leq .05$ signifikant werden.

Die berücksichtigten Prädiktoren unterscheiden sich in ihrer relativen Bedeutsamkeit deutlich je nach betrachteter abhängiger Variable bzw. je nach berücksichtigter Behinderungsform (vgl. Tab. 3): Während proinklusive Einstellungen insgesamt stärker durch unterschiedliche Studieneingangsvoraussetzungen er-

klärt werden können, werden die Kompetenzüberzeugungen eher durch konkrete Lernerfahrungen im Studium vorhergesagt. Das Ausmaß von Vorerfahrungen mit Menschen mit Behinderungen beeinflusst zwar die Kompetenzeinschätzung positiv, führt aber zu negativeren Bewertungen von Inklusion – v. a. in Bezug auf die wahrgenommenen Auswirkungen für die inkludierten Schülerinnen und Schüler. Dagegen begünstigt die positive Valenz von vorherigen Erfahrungen die positive Bewertung von Inklusion für die Mitschülerinnen und Mitschüler. Erfahrungen mit Behinderung in der eigenen Schulklasse zeigen kaum einen Effekt, erhöhen jedoch die inklusive Kompetenzeinschätzung.

Tabelle 3: Übersicht zu den Pfadkoeffizienten

	Vorteile von Inklusion bei geistiger/körperlicher/psychischer Behinderung		Inklusive Kompetenzen	
	für die inkludierten Schülerinnen und Schüler	für die Mitschülerinnen und Mitschüler	Theoretisches Wissen	Handlungspraxis
Ausmaß Vorkontakt	**-.23***/-.03./**-.14***	-.13^{+}/-.11^{+}/**-.16***	**.12***	**.11***
Bewertung Vorkontakt	.10/-.04/.13^{+}	**.29*/.23*/.24***	.06	.09^{+}
Kontakt in Schulklasse	-.00/-.01/.08	.03/**.13***/.05	**.13***	**.16***
Pädagogisches Interesse	**.20***/**.23***/.09	**.24*/.32*/.18***	**.11***	**.17***
Fachliches Interesse	-.08/.03/.04	.04/.05/.09	.04	.02
Gymnasiallehramt	.04/**.17***/.13^{+}	.04/**.17*/.14***	-.01	-.01
MINT-Unterrichtsfach	-.06/-.00/-.01	-.06/-.00/-.02	-.00	-.04
Fachsemester	-.07/.03/-.03	.07/**.14***/-.01	**.-14***	**-.14***
Veranstaltungen zu Inklusion	.12/**.17***/.13^{+}	.10/.11/.12	**.39***	**.35***
Inklusive Praktika	.01/**-.14***/.06	-.10/**-.15***/-.02	**.23***	**.19***

Anmerkung: * $p \leq .05$, $^{+}$ $p \leq .1$

Hinsichtlich der Studienwahlmotivation ergibt sich ein recht eindeutiges Befundmuster: Während das pädagogische Interesse positiv mit proinklusiven Einstellungen und Kompetenzeinschätzungen zusammenhängt, erweist sich das fachliche Interesse in keinem Fall als prädiktiv. Gleiches gilt für das Studium eines MINT-Unterrichtsfaches. Zwischen den verschiedenen Lehrämtern zeigen sich lediglich in den proinklusiven Einstellungen Unterschiede: Studierende für das Lehramt an Gymnasien weisen je nach Behinderungsform höhere proinklusive Einstellungen auf. Demgegenüber unterschieden sich die pro-

inklusiven Einstellungen von Studierenden höherer Fachsemester – unabhängig von der Anzahl an Veranstaltungen zu Inklusion, die sie bereits besucht haben, oder absolvierten inklusiven Praktika – in der Regel nicht von den Einstellungen Studierender niedriger Fachsemester, allerdings schätzen Studierende höherer Fachsemester ihre inklusiven Kompetenzen kritischer ein als Studierende niedrigerer Fachsemester. Studienbezogene Lernerfahrungen (Besuch von Lehrveranstaltungen bzw. Absolvieren von Praktika im inklusiven Unterricht) wirken sich nur auf die proinklusiven Einstellungen hinsichtlich der Inklusion von Schülerinnen und Schülern mit körperlicher Behinderung aus und dies in ambivalenter Form: Während der Besuch von Lehrveranstaltungen im Bereich Inklusion zu positiveren Einstellungen führt, führt das Absolvieren von Praktika im inklusiven Unterricht zu negativeren Einstellungen. Auf inklusionsbezogene Kompetenzeinschätzungen der Studierenden üben beide Aspekte dagegen einen positiven Einfluss aus, wobei der Effekt des Besuches von Lehrveranstaltungen im Bereich Inklusion nochmals stärker ist und sich beides zudem etwas positiver auf das theoretische Wissen als auf die selbst eingeschätzte praktische Handlungskompetenz auswirkt.

5 Diskussion und Ausblick

Ziel dieses Beitrags ist es, die Vorhersagekraft studienbezogener und dem Studium vorgelagerter Faktoren für die Ausprägung proinklusiver Einstellungen und Kompetenzüberzeugungen zu ermitteln und hieraus Implikationen für die Lehrerbildung abzuleiten. Als zentrales Ergebnis unserer Studie zeigt sich, dass subjektive proinklusive Einstellungen stärker mit Erfahrungen *vor* dem Studium sowie der konkreten Studienwahlmotivation zusammenhängen, während die Einschätzung inklusiver Kompetenzen zu bedeutsamen Anteilen durch inklusionsbezogene Lernerfahrungen *im* Studium zu erklären ist. Dass nur Vorerfahrungen mit Menschen mit Behinderung, die als positiv erlebt wurden, proinklusive Einstellungen begünstigen und inklusive Praktika teilweise einen gegenteiligen Effekt bewirken, verweist auf die Bedeutung der Qualität entsprechender Erfahrungen (vgl. Woll 2017). Gleichzeitig erscheint es möglich, dass die Zustimmungswerte zu Inklusion bei Studierenden mit Praxiserfahrung v. a. eine Bewertung dieser Erfahrung widerspiegeln (vgl. Gasterstädt/Urban 2016), was auf eine wichtige Limitation der Aussagekraft unserer Ergebnisse verweist, in denen weder spezifische Merkmale inklusiver Unterrichtserfahrung noch interindividuelle Unterschiede von Inklusions- und Behinderungsvorstellungen berücksichtigt werden. Zur Klärung dieser Fragen könnten Mixed-Methods-Ansätze beitragen, die Erkenntnisse qualitativer Zugänge zu subjektiven Perspektiven mit quantitativen Daten zu Einflussfaktoren auf proinklusive Überzeugungen verknüpfen (z. B. Hecht/Niedermair/Feyerer 2016).

Dass insbesondere Lehrveranstaltungen zu Inklusion sich als bedeutsam für die eigene inklusive Kompetenzeinschätzung zeigen, kann als Hinweis für die – zumindest subjektiv wahrgenommene – Wirksamkeit der Lehramtsausbildung interpretiert werden, wobei positive Selbstselektion den Effekt verzerren könnte, wenn vor allem Studierende mit hoher intrinsischer Motivation vermehrt Inklusionsveranstaltungen besuchen. Aufschluss hierüber könnten systematische Längsschnittstudien sowie Clusteranalysen zur Ermittlung interindividueller Zusammenhangsmuster (vgl. Schön/Stark/Stark 2018) ermöglichen, die über andere Stichprobenzusammensetzungen auch Unterschiede zwischen Lehramtsstudierenden unterschiedlicher Schulformen näher beleuchten könnte, als in unseren Analysen möglich war. Weiterer Klärung bedarf auch die Frage, inwieweit gefundene Differenzen zwischen Studierenden verschiedener Lehrämter replizierbar sind oder lediglich einen Stichprobeneffekt darstellen. Für detailliertere Aufschlüsse über die Ausbildung von inklusiven Kompetenzen ist für Folgestudien anzuregen, analog zur Einstellungsmessung auch hinsichtlich inklusiver Kompetenzen nach Behinderungsformen bzw. Förderschwerpunkten zu differenzieren und die diesbezügliche Selbsteinschätzung durch ein stärker objektivierbares Erhebungsmaß zu ergänzen.

Weniger eindeutig als bezüglich der eigenen Kompetenzeinschätzung fällt der Einfluss inklusionsbezogener Lerngelegenheiten im Studium auf proinklusive Einstellungen aus. Für die Gestaltung inklusionsbezogener Lehrveranstaltungen verweisen unsere Ergebnisse deshalb auf die Bedeutsamkeit der Förderung proinklusiver Einstellungen insbesondere hinsichtlich der Inklusion von Schülerinnen und Schülern mit geistigen Einschränkungen oder psychischen Erkrankungen. Insbesondere aber unterstreichen unsere Befunde die Annahme, dass nicht das Ausmaß, sondern die Valenz praktischer Erfahrungen mit Behinderung und Inklusion entscheidend für die Ausbildung proinklusiver Einstellungen ist. Praktika oder inklusive Lehrveranstaltungen benötigen entsprechend eine qualitätsvolle Begleitung, die ermöglicht, eigene Erfahrungen und möglicherweise hieraus begründete Vorbehalte gegen inklusive Beschulung kritisch zu reflektieren (Radhoff/Ruberg 2017). Darüber hinaus liefert die Bedeutsamkeit von pädagogischem Interesse für inklusive Überzeugungen ein Argument für die Stärkung des pädagogischen Profils der Lehramtsausbildung.

Literatur

Abels, Simone/Heidinger, Christine/Koliander, Brigitte/Plotz, Thomas (2018): Die Notwendigkeit der Verhandlung widersprüchlicher Anforderungen an das Lehren von Chemie an einer inklusiven Schule. In: Zeitschrift für interpretative Schul- und Unterrichtsforschung 7, H. 1, S. 135–151.

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520.

Blömeke, Sigrid (2004): Empirische Forschung zur Wirksamkeit der Lehrerbildung. In: Blömeke, Sigrid/Reinhold, Peter/Tulodziecki, Gerhard/Wildt, Johannes (Hrsg.): Handbuch Lehrerausbildung. Bad Heilbrunn: Klinkhardt. S. 59–91.

Bosse, Stefanie/Spörer, Nadine (2014): Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. In: Empirische Sonderpädagogik 6, H. 4, S. 279–299.

de Boer, Anke/Pijl, Sip Jan/Minnaert, Alexander (2011): Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. In: International Journal of Inclusive Education 15, H. 3, S. 331–353.

Fishbein, Martin/Ajzen, Icek (2010): Predicting and changing behavior. New York, Hove: Psychology Press.

Gasterstädt, Julia/Urban, Michael (2016): Einstellung zu Inklusion? Implikationen aus Sicht qualitativer Forschung im Kontext der Entwicklung inklusiver Schulen. In: Empirische Sonderpädagogik 8, H. 1, S. 54–66.

Gebhardt, Markus/Kuhl, Jan/Wittich, Claudia/Wember, Franz (2018): Inklusives Modell in der Lehramtsausbildung nach den Anforderungen der UN-BRK. In: Hußmann, Stephan/Welzel, Barbara (Hrsg.): DoProfiL – Das Dortmunder Profil für inklusionsorientierte Lehrerinnen- und Lehrerbildung. Münster: Waxmann. S. 279–292.

Gebhardt, Markus/Schwab, Susanne/Reicher, Hannelore/Ellmeier, Barbara/Gmeiner, Sonja/Rossmann, Peter/Gasteiger Klicpera, Barbara (2011): Einstellungen von LehrerInnen zur schulischen Integration von Kindern mit einem sonderpädagogischen Förderbedarf in Österreich. In: Empirische Sonderpädagogik 3, H. 4, S. 275–290.

Hecht, Petra/Niedermair, Claudia/Feyerer, Ewald (2016): Einstellungen und inklusionsbezogene Selbstwirksamkeitsüberzeugungen von Lehramtsstudierenden und Lehrpersonen im Berufseinstieg. In: Empirische Sonderpädagogik 8, H. 1, S. 86–102.

Hellmich, Frank/Görel, Gamze/Schwab, Susanne (2016): Einstellungen und Motivation von Lehramtsstudentinnen und -studenten in Bezug auf den inklusiven Unterricht in der Grundschule. In: Empirische Sonderpädagogik 8, H. 1, S. 67–85.

KMK (2014): Standards für die Lehrerbildung: Bildungswissenschaften. Bonn. KMK.

Kopp, Bärbel (2009): Inklusive Überzeugung und Selbstwirksamkeit im Umgang mit Heterogenität. In: Empirische Sonderpädagogik 1, H. 1, S. 5–25.

Krauss, Stefan/Bruckmaier, Georg (2014): Das Experten-Paradigma in der Forschung zum Lehrerberuf. In: Terhart, Ewald/Bennewitz, Hedda/Rothland, Martin (Hrsg.): Handbuch der Forschung zum Lehrerberuf. Münster: Waxmann. S. 241–261.

Kullmann, Harry/Lütje-Klose, Birgit/Textor, Annette/Schitow, Katharina (2014): Inklusiver Unterricht – (Auch) eine Frage der Einstellung! In: Schulpädagogik heute 5, H. 10, S. 1–14.

Kunter, Mareike/Kleickmann, Thilo/Klusmann, Uta/Richter, Dirk (2011): Die Entwicklung professioneller Kompetenz von Lehrkräften. In: Kunter, Mareike/Baumert, Jürgen/Blum, Werner/Neubrand, Michael (Hrsg.): Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann. S. 55–68.

Kunz, André/Luder, Reto/Moretti, Marta (2010): Die Messung von Einstellungen zur Integration (EZI). In: Empirische Sonderpädagogik 2, H. 3, S. 83–94.

Laubner, Marian (2019): Gymnasium – ein ‚Ausnahmefall' im Inklusionsdiskurs? In: Horn, Klaus-Peter/Rabenstein, Kerstin/Stubbe, Tobias C. (Hrsg.): Gymnasium und Inklusion. Manuskript im Erscheinen.

Loreman, Tim/Sharma, Umesh/Forlin, Chris (2013): Do pre-service teachers feel ready to teach in inclusive classrooms? In: Australian Journal of Teacher Education 38, H. 1, S. 27–44.

McElvany, Nele/Schwabe, Franziska/Hartwig, Svenja Janina/Igler, Jennifer (2018): Inklusive Lehr-Lern-Settings. Einstellungen und Motivation von Lehrkräften. In: Zeitschrift für Pädagogik 64, H. 6, S. 831–851.

Muthén, Linda K./Muthén, Bengt O. (2017): Mplus statistical analysis with latent variables: User's guide (version 8). Los Angeles, CA: Muthén & Muthén.

Pit-ten Cate, Ineke M./Markova, Mariya/Krischler, Mireille/Krolak-Schwerdt, Sabine (2018): Promoting inclusive education: The role of teachers' competence and attitudes. In: Insights into Learning Disabilities 15, H. 1, S. 49–63.

Pohlmann, Britta/Möller, Jens (2010): Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). In: Zeitschrift für Pädagogische Psychologie 24, H. 1, S. 73–84.

Radhoff, Melanie/Ruberg, Christiane (2017): Inklusion in der Lehrerausbildung. Aktuelle Strukturen und die Rolle der Universität. In: Greiten, Silvia/Geber, Georg/Gruhn, Annika/Köninger, Manuela (Hrsg.): Lehrerausbildung für Inklusion. Fragen und Konzepte zur Hochschulentwicklung. Münster: Waxmann. S. 64–76.

Retelsdorf, Jan/Möller, Jens (2012): Grundschule oder Gymnasium? Zur Motivation ein Lehramt zu studieren. In: Zeitschrift für Pädagogische Psychologie 26, H. 1, S. 5–17.

Rolka, Katrin/Albersmann, Natascha (2017): Inklusion im Mathematikunterricht der Sekundarstufe. Entwicklung, Erprobung und Evaluation eines Seminarkonzepts. In: Greiten, Silvia/Geber, Georg/Gruhn, Annika/Köninger, Manuela (Hrsg.): Lehrerausbildung für Inklusion. Fragen und Konzepte zur Hochschulentwicklung. Münster: Waxmann. S. 255–265.

Rothland, Martin (2014): Warum entscheiden sich Studierende für den Lehrerberuf? In: Terhart, Ewald/Bennewitz, Hedda/Rothland, Martin (Hrsg.): Handbuch der Forschung zum Lehrerberuf. 2., überarb. u. erw. Aufl. Münster: Waxmann. S. 349–385.

Rothland, Martin (2015): Die Bedeutung pädagogischer (Vor-)Erfahrungen von Lehramtsstudierenden – ein Mythos? In: Zeitschrift für Pädagogik 61, H. 2, S. 270–281.

Ruberg, Christiane/Porsch, Raphaela (2017): Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion. Ein systematisches Review deutschsprachiger Forschungsarbeiten. In: Zeitschrift für Pädagogik 63, H. 4, S. 393–456.

Schön, Michael/Stark, Lisa/Stark, Robin (2018): Einstellungen und Vorstellungen bezüglich Inklusion. Eine typenbildende Clusteranalyse. In: Bildungsforschung 15, H. 1, S. 1–21.

Schwab, Susanne (2014): Schulische Förderung und soziale Integration von Schülern mit sonderpädagogischem Förderbedarf in Integrationsklassen. In: Empirische Pädagogik 28, H. 3, S. 259–274.

Sharma, Umesh/Forlin, Chris/Loreman, Tim/Earle, Chris (2006): Pre-service teachers' attitudes, concerns and sentiments about inclusive education. In: International Journal of Special Education 21, H. 2, S. 80–93.

Stroh, Michael (2015): Inklusion im naturwissenschaftlichen Unterricht – Beschreibung eines Spannungsfeldes. In: Siedenbiedel, Catrin/Theurer, Caroline (Hrsg.): Inklusive Unterrichtspraxis und -entwicklung. Immenhausen: Prolog-Verlag. S. 110–124.

Trumpa, Silke/Janz, Frauke/Heyl, Vera/Seifried, Stefanie (2014): Einstellungen zu Inklusion bei Lehrkräften und Eltern – Eine schulartspezifische Analyse. In: zbf Zeitschrift für Bildungsforschung 4, H. 3, S. 241–256.

Voss, Thamar/Kleickmann, Thilo/Kunter, Mareike/Hachfeld, Axinja (2011): Überzeugungen von Mathematiklehrkräften. In: Kunter, Mareike/Baumert, Jürgen/Blum, Werner/Neubrand, Michael (Hrsg.): Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann. S. 235–257.

Warner, Lisa M./Schwarzer, Ralf (2009): Selbstwirksamkeit bei Lehrkräften. In: Zlatkin-Troitschanskaia, Olga/Beck, Klaus/Sembill, Detlef/Nickolaus, Reinhold: Mulder, Regina (Hrsg.): Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung. Weinheim: Beltz. S. 629–640.

Watkins, Amanda (2012): Inklusionsorientierte Lehrerbildung. Ein Profil für inklusive Lehrerinnen und Lehrer. Odense: European Agency for Development in Special Needs Education.

Willems, Ariane S./Meyer, Karina/Frechen, Moritz (2019). Skalendokumentation des Projekts EFBI – Beschreibung und Entwicklung von Einstellungen von FSJler*innen zu Behinderung und Inklusion. Göttingen: Göttinger Beiträge zur erziehungswissenschaftlichen Forschung. Manuskript in Vorbereitung.

Woll, Anke (2017): Kontaktbedingungen zwischen Menschen mit und ohne Behinderung als Prädiktoren von Einstellungen zu Inklusion. Dissertation, Universität Heidelberg.

Einstellungen zur Leistungsheterogenität von Lehrkräften an integrativen und nicht-integrativen weiterführenden Schulen

Marcus Syring, Samuel Merk und Thorsten Bohl

Abstract

Gemeinschaftsschulen als eine Form integrativer Schulen haben zum Ziel, Kinder und Jugendliche mit unterschiedlichen Begabungen zu fördern, womit sie explizit auf eine heterogene Schülerschaft abzielen. Um diesem Ziel gerecht zu werden, bedarf es Lehrkräfte, die sich produktiv u. a. mit Leistungsheterogenität auseinandersetzen. Ein wichtiger Prädiktor eines solchen Umgangs sind die Einstellungen von Lehrkräften, denn sie gelten als wahrnehmungs- und handlungsregulierend. Der Beitrag geht der Frage nach, inwiefern sich die Einstellungen zur Leistungsheterogenität bei Lehrkräften an Gemeinschaftsschulen von Lehrkräften anderer weiterführender, nicht-integrativer Schulen unterscheiden, insbesondere nach Kontrolle von personalen und institutionellen Kontextmerkmalen sowie der objektiven Leistungsheterogenität der Schülerschaft. In der vorliegenden Studie mit $N = 1455$ Lehrkräften wurden die Einstellungen mittels eines etablierten fünffaktoriellen Instruments (Mehrwert, Kosten, Motivation, Emotion, wahrgenommene Kompetenz) erhoben. Um Unterschiede in den Überzeugungen von potentieller Konfundierung zu bereinigen, wurde ein Matching-Verfahren durchgeführt. Die Ergebnisse der Analysen zeigten mittlere Effekte zugunsten der Lehrerinnen und Lehrer an Gemeinschaftsschulen für die Einschätzung des Mehrwerts und der intrinsischen Motivation.

1 Einleitung und Problemstellung

Ein wesentliches Ziel aktueller schulstruktureller Veränderungen stellt deutschlandweit die Verringerung sozialer Disparitäten dar (vgl. Wacker/Bohl 2016). Dieses Ziel soll durch die Einführung integrativer Schulformen, wie den Gemeinschaftsschulen, erreicht werden. Das führt – je nach Bundesland – zu einer weiteren Pluralisierung an Schulformen (z. B. in Baden-Württemberg)

oder zu einer Homogenisierung der Schullandschaft durch den Aufbau eines Zwei-Säulen-Systems (z. B. in Bremen, Saarland etc.; vgl. Tillmann 2012).

Gemeinschaftsschulen als eine Form integrativer Schulen haben zum Ziel, Kinder und Jugendliche mit unterschiedlichen Begabungen im Rahmen gemeinsamen Lernens zu fördern (vgl. Bohl/Wacker 2016b). So heißt es bspw. im Schulgesetz von Baden-Württemberg: „Die Gemeinschaftsschule vermittelt in einem gemeinsamen Bildungsgang Schülern der Sekundarstufe I je nach ihren individuellen Leistungsmöglichkeiten eine der Hauptschule, der Realschule oder dem Gymnasium entsprechende Bildung" (§ 8a, Abs. 1, BW SchG). Damit fokussieren Gemeinschaftsschulen explizit eine heterogene Schülerschaft, die vor allem auch durch ein hohes Maß an Leistungsheterogenität gekennzeichnet ist.

Um die genannten Ziele umzusetzen, bedarf es Lehrkräfte, die sich produktiv u. a. mit Leistungsheterogenität auseinandersetzen. Dazu benötigen sie nicht nur Wissen und Kompetenzen als kognitive Komponente, sondern auch entsprechende Überzeugungen sowie motivationale Orientierungen und selbstregulative Fähigkeiten (vgl. Baumert/Kunter 2006), denn Einstellungen und Überzeugungen gelten als wahrnehmungs- und handlungsregulierende Faktoren (vgl. Kunz/Luder/Moretti 2010).

Daraus ergibt sich die Frage, ob sich Lehrkräfte integrativer Schulformen von Lehrkräften anderer Schulformen in ihren Einstellungen zum Umgang mit Leistungsheterogenität unterscheiden.

2 Theoretischer Hintergrund

Einstellungen als Prädiktoren unterrichtlichen Handelns

Folgt man kompetenztheoretischen Modellen der Professionalität von Lehrkräften (vgl. Baumert/Kunter 2006), so kommen Überzeugungen, Einstellungen bzw. Orientierungen von Lehrkräften eine besondere Rolle für die Unterrichtspraxis zu. Forschungsbefunde liefern Evidenz für diese Bedeutung: Felbrich, Schmotz und Kaiser (2008) konnten etwa zeigen, dass Einstellungen wichtig für die Unterrichtsgestaltung und die Interaktionen in Lehr-Lernprozessen sind. Voss und Kolleg/innen (2011) wiesen nach, dass der Lernerfolg von Schülerinnen und Schülern ebenfalls durch die Einstellungen der Lehrkräfte, mediiert durch Unterrichtsplanung und Instruktion, prädiziert wird. Auch die Lehrergesundheits- bzw. Belastungsforschung, in der Einstellungen als personenbezogene Ressourcen zur Anforderungsbewältigung interpretiert werden (vgl. Pajares 1992), liefert Hinweise für die Bedeutung von Einstellungen von Lehrkräften: So wurde in unterschiedlichen Studien aufgezeigt, dass Heterogenität mit

einer Wahrnehmung als Herausforderung oder befürchteten Mehrbelastung einhergeht (vgl. z. B. Avramidis/Norwich 2002; Kiel/Syring/Weiß 2017).

Einstellungen können als „Vorstellungen und Annahmen von Lehrkräften über schul- und unterrichtsbezogene Phänomene und Prozesse mit einer bewertenden Komponente“ (Kunter/Pohlmann 2009, S. 267; vgl. auch Syring/Tillmann/Ivanova/Kiel 2019) definiert werden. Sie gelten als „wichtige handlungsregulierende Faktoren [, die] die Wahrnehmung“ (Kunz/Luder/Moretti 2010, S. 84; vgl. auch Pajares 1992) und das Verhalten (vgl. Ajzen/Fishbein 2000) beeinflussen.

In der Literatur besteht Einigkeit darüber, dass Einstellungen ein stabiles Konstrukt darstellen und verschiedene Dimensionen bzw. Komponenten umfassen (vgl. bspw. Drei-Komponenten-Modell: Werth/Meyer 2008). Häufig werden dabei eine kognitive, eine motivational-affektive und eine behaviorale Dimension unterschieden (vgl. Eagly/Chaiken 1993; Rosenberg/Hovland 1960).

Einstellungen von Lehrkräften zur Leistungsheterogenität

Einstellungen von Lehrkräften und Lehramtsstudierenden sind spezifisch für Objekte, Personen oder Ideen. Aus diesem Grund erscheint es bedeutsam, Heterogenität bspw. nicht als globale Kategorie zu betrachten, sondern je spezifische Heterogenitätsformen bzw. Differenzlinien (z. B. Leistungsheterogenität) zu betrachten. Dies belegen auch empirische Ergebnisse (vgl. z. B. Merk/Cramer/Dai/Bohl/Syring 2018).

Unter Leistungsheterogenität kann die unterschiedliche Leistungsausprägung trotz gleichen Alters und gleicher Bildungsstufe (vgl. z. B. Blossfeld et al. 2007) verstanden werden. Bisher liegen kaum empirische Ergebnisse zu den Einstellungen von Lehrkräften hinsichtlich Leistungsheterogenität vor. Befunde aus zwei eigenen Studien (vgl. Merk/Cramer/Dai/Bohl/Syring 2018; Syring/Tillmann/Ivanova/Kiel 2019) zeigen, dass Lehramtsstudierende

- Leistungsheterogenität im Vergleich zu anderen Heterogenitätsformen (wie z. B. ethnisch-kultureller Heterogenität) einen geringeren Mehrwert zuschreiben und damit auch höhere Kosten verbinden,
- motiviert sind, mit Leistungsheterogenität umgehen zu wollen und weder besonders positive noch besonders negative Emotionen mit Blick auf diesen Umgang verbinden,
- sich als kompetent für den Umgang mit Leistungsheterogenität empfinden,
- mit Migrationshintergrund signifikant geringere Kosten und negative Emotionen mit dem Unterrichten leistungsheterogener Klassen verbinden und sich dafür eine höhere Kompetenz und Motivation für den Umgang mit leistungsheterogenen Lerngruppen zuschreiben (vgl. Syring/Merk/Cramer/Topalak/Bohl 2019).

Für den empirischen Fokus auf Lehramtsstudierende wird die bestehende Forschung jedoch häufig kritisiert, da vermutet wird, dass sich die Einstellungen von im Dienst stehenden Lehrkräften nochmals von denen von Studierenden unterscheiden (vgl. Syring/Merk/Cramer/Topalak/Bohl 2019). Diesem Desiderat begegnet die vorliegende Studie, indem sie Einstellungen von im Dienst befindlichen Lehrkräften integrativer und nicht-integrativer Schulformen vergleicht.

3 Fragestellung

Aus dem genannten theoretischen Hintergrund ergibt sich die Frage, inwiefern sich Einstellungen zur Leistungsheterogenität bei Lehrkräften an Gemeinschaftsschulen (als integrative Schulform) von Lehrkräften anderer weiterführender, nicht-integrativer Schulen unterscheiden. Dabei interessiert insbesondere wie diese Unterschiede nach Kontrolle von personalen (z. B. Dienstjahre, Beschäftigungsumfang, Erschöpfung und Depression etc.) und institutionellen Kontextmerkmalen (Leitung der Schule, Kooperationsformen im Kollegium etc.) sowie der objektiven Leistungsheterogenität ausfallen.

4 Methodisches Vorgehen

Datenerhebung und Stichprobe

Die für die vorliegende Untersuchung genutzten Daten stammen aus der WissGem-Studie (Wissenschaftliche Begleitung der Einführung der Gemeinschaftsschule in Baden-Württemberg; vgl. Bohl/Wacker 2016a). Von April bis Juli 2014 wurden hier $N = 1.455$ Lehrkräfte an insgesamt 78 weiterführenden Schulen befragt. Die Stichprobe der Lehrkräfte an Gemeinschaftsschulen umfasst $n_{\text{Gem}} = 691$, die anderen Lehrkräfte verteilen sich auf Gymnasien, Realschulen und Werkrealschulen. Die Daten wurden mittels paper-pencil-Fragebogen vor Ort erhoben (für eine ausführliche Beschreibung und Dokumentation siehe: Batzel-Kremer et al. 2016).

Abhängige Variablen

Zur Messung der Einstellungen zur Leistungsheterogenität von Lehrkräften wurde ein etabliertes Instrument von Gebauer, McElvany und Klukas (2013) verwendet, welches sowohl die kognitive als auch die behaviorale und motivational-affektive Komponente von Einstellungen erfasst (siehe auch Merk/Cramer/Dai/Bohl/Syring 2018). Die Komponenten der Einstellungen bilden den

Mehrwert und die Kosten (kognitiv), die negativen Emotionen sowie die intrinsische Motivation (motivational-affektiv) wie auch die selbsteingeschätzte wahrgenommene Kompetenz (behavioral) ab (siehe Tabelle 1). Die Beantwortung erfolgte auf einer Likert-Skala (1 = stimmt gar nicht; 4 = stimmt genau). Eine konfirmatorische Faktorenanalyse mit τ-kongenerischen Messmodellen (vgl. Brown 2014) zeigte eine gute Anpassung der empirischen Kovarianzstruktur an die theoretisch erwartete ($\chi^2 = 40.190$, df = 18, Comparative Fit Index [CFI] = 0.928, Tucker-Lewis Index [TLI] = 0.936, Root Mean Square Error of Approximation [RMSEA] = 0.062, Standardized Root Mean Square Residual [SRMR] = .048).

Tabelle 1: Interne Konsistenz der Skalen zur Leistungsheterogenität

Dimension	Items	Beispielitem	McDonalds ω
Mehrwert	5	Schüler/innen profitieren durch Unterricht in leistungsheterogenen Lerngruppen hinsichtlich ihres schulischen Lernens.	.89
Kosten	5	Die leistungsbezogene Heterogenität in unseren Schulkassen erfordert von den Lehrkräften einen zusätzlichen Arbeitsaufwand.	.81
negative Emotionen	5	Im Hinblick auf das Unterrichten von leistungsheterogenen Klassen fühle ich mich angespannt bei dem Gedanken daran, die Kinder entsprechend ihren leistungsbezogenen Voraussetzungen zu unterrichten.	.86
Intrinsische Motivation	5	Die leistungsbezogene Heterogenität in unseren Schulklassen wird mir beim Unterrichten Spaß machen.	.92
Wahrgenommene Kompetenz	5	Im Hinblick auf das Unterrichten von leistungsheterogenen Klassen weiß ich, dass ich es schaffen werde, die erwarteten Lerninhalte zu vermitteln, wenn ich mich darum bemühe.	.81

Matching und statistische Analysen

Um Unterschiede in den Überzeugungen zur Heterogenität möglichst stark von potentieller Konfundierung zu bereinigen, wurde vor den eigentlichen Analysen ein Matching-Verfahren durchgeführt (vgl. Rosenbaum/Rubin 1983). Dabei wurden Lehrkräften, die an Gemeinschaftsschulen unterrichten, Lehrkräfte sonstiger weiterführender Schularten (Kontrollgruppe) zugeordnet, die diesen in den verfügbaren Kovariaten des reanalysierten Datensatzes möglichst ähnlich sind. Dabei wurde ein sogenanntes *Genetic Matching* eingesetzt (vgl. Diamond/Sekhon 2013), welches gegenüber üblicheren Verfahren (z. B. *Propensity*

Score Matching) den Vorteil hat, das Problem arbiträrer Modellspezifikationen zur Schätzung der Distanzmetrik zu umgehen.

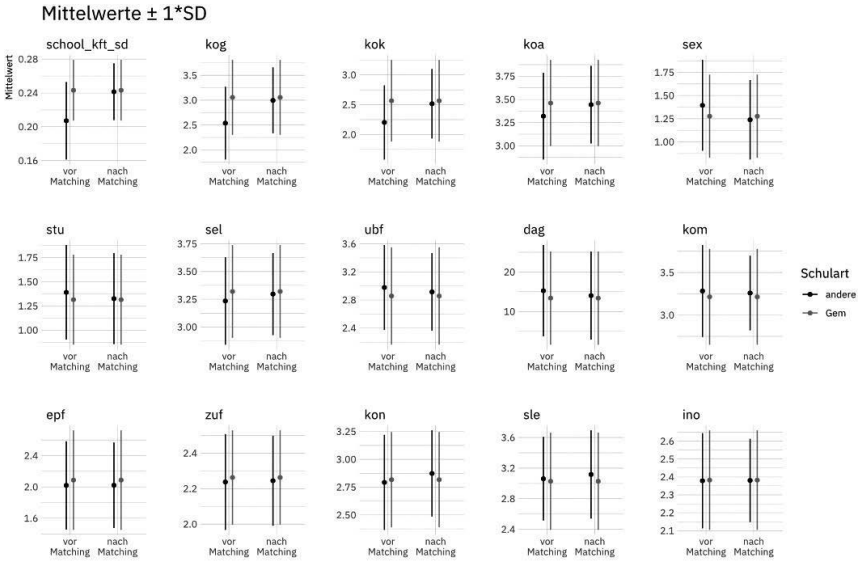
Da es sich bei der hier beschriebenen Studie um eine Sekundäranalyse eines bestehenden Datensatzes handelt, wurden in das Matching-Verfahren diejenigen Kovariaten einbezogen, die einerseits zum damaligen Zeitpunkt erhoben wurden und die andererseits von Relevanz für die Beschreibung von Lehrkräften an unterschiedlichen Schulformen erscheinen und ggf. die gefundenen Ergebnisse konfundieren würden. Die einbezogenen Kovariaten sind (siehe ausführlich, auch zur Messung: Batzel-Kremer et al. 2016): Geschlecht (sex), Beschäftigungsumfang (stu), Dienstjahre (dag), Selbstwirksamkeit (sel), Berufszufriedenheit (zuf), unterrichtsbezogene Führung (ubf), Erschöpfung und Depression (epf), Leitung der Schule (sle), Kooperationsformen des Kollegiums (Austausch: koa, gemeinsame Arbeitsorganisation: kog, Konstruktion und Reflexion: kok), Kommunikation im Kollegium (kom), Konsens im Kollegium (kon) sowie Innovationsbereitschaft des Kollegiums (ino). Eine Besonderheit stellt die Kovariate der objektiven Heterogenität (school_kft_sd) dar. Hierfür wurde für jede der 78 Schulen die Varianz des Kognitiven Fähigkeitstests (befragt wurden ca. 3.200 Schülerinnen und Schüler) berechnet. Diese wird als Maß der objektiven Heterogenität der Schule interpretiert. Vor dem Matching (siehe Abb. 1) unterschieden sich die Mittelwerte der Varianzen zwischen integrativen und nicht-integrativen Schulen um ca. eine Standardabweichung zu Gunsten eines höheren Mittelwertes bei den integrativen Schulen.

Da für das Matching ein Verfahren spezifiziert wurde, in dem Lehrkräfte der Kontrollgruppe (nicht-integrative Schulformen) mehrfach einer Lehrkraft der Treatmentgruppe (Gemeinschaftsschule) zugeordnet werden konnten, entstanden gewichtete Daten, die zudem eine hierarchische Struktur aufwiesen (Lehrkräfte in Schulen genestet). Diese Eigenschaften resultieren in Standardverfahren jedoch oft in inkorrekten Parameterschätzungen oder verzerrten Standardfehlern, was vorliegend durch entsprechend adjustierende Verfahren (Pseudo-Maximum-Likelihood-Punktschätzung; Varianzschätzung durch Taylor-Linearisierung; vgl. Bollen/Tueller/Oberski 2013) umgangen wurde.

5 Ergebnisse

Für eine hohe Qualität der Analysen ist von zentraler Wichtigkeit, die Balance der Kovariaten vor und nach dem genetischen Matching zu betrachten: Abbildung 1 stellt die arithmetischen Mittel und die Standardabweichungen entsprechend differenziert dar.

Abbildung 1: Arithmetische Mittel und Standardabweichungen (± 1) der in das Matching eingehenden Kovariaten. Subplots sortiert nach der Effektstärke der Mittelwertsunterschiede vor dem Matching (Cohen's d). Erläuterung der Abkürzungen siehe Methodenteil



Die Ergebnisse zur Frage der Unterschiede in den Einstellungen zur Leistungsheterogenität sind in Tabelle 2 deskriptiv dargestellt. Die Skalen „Kosten" und „negative Emotionen" wurden nicht rekodiert, d. h. das ein höherer Wert hier für höhere Kosten bzw. negativere Emotionen stehen. Abbildung 2 stellt das multiple-indicators multiple-causes (MIMIC; vgl. Jöreskog/Goldberger 1975) Modell dar, mithilfe dessen die Effektstärken geschätzt und inferenzstatistisch abgesichert wurden. Es zeigt sich, dass Lehrkräfte an Gemeinschaftsschulen sich von Lehrkräften nicht-integrativer weiterführender Schulen signifikant unterscheiden.

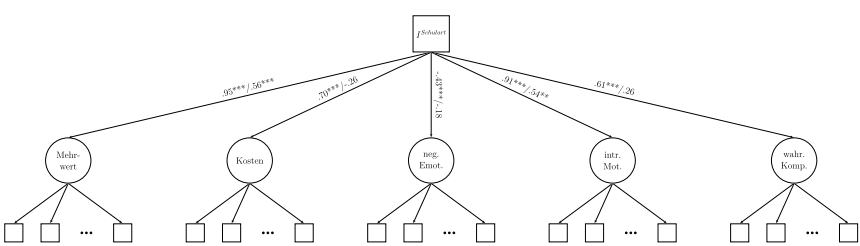
Tabelle 2: Mittelwerte und SD der Dimension der Einstellungen zur Leistungsheterogenität sowie y-standardisierte Effekte vor und nach dem Matching (siehe Abb. 2)

	nach dem Matching			**vor dem Matching**
Dimension	**Lehrkräfte an Gemeinschaftsschulen** M **(SD)**	**Lehrkräfte an nicht-integrativen Schulen** M **(SD)**	**y-stand. Effekt im MIMIC-Modell**	**y-stand. Effekt im MIMIC-Modell**
Mehrwert	3.18 (0.59)	2.85 (0.56)	0.56***	0.95***
Kosten	2.55 (0.62)	2.67 (0.53)	-0.26	-0.70***
negative Emotionen	2.36 (0.71)	2.49 (0.63)	-0.18	-0.43***
intrinsische Motivation	3.03 (0.65)	2.66 (0.58)	0.54**	0.91***
wahrgenommene Kompetenz	2.99 (0.49)	2.86 (0.41)	0.26	0.61***

Anmerkungen: ** $p < .01$. *** $p < .001$

Gemeinschaftsschullehrkräfte sehen einen höheren Mehrwert im Unterrichten leistungsheterogener Klassen und sind auch intrinsisch motivierter, mit Leistungsheterogenität umzugehen als in den Kovariaten vergleichbare andere Lehrkräfte. Beide Effekte können als in ihrer Größe moderat (vgl. Cohen 1988) eingeschätzt werden. Unterschiede in den negativen Emotionen aber auch in der wahrgenommenen Kompetenz zeigen keine statistische Signifikanz.

Abbildung 2: MIMIC Modell



Anmerkungen: y-standardisierte Pfadkoeffizienten vor/nach dem Matching. $I^{Schulart}$ = dummykodierte Indikatorvariable für die Schulart (= 1 falls Gemeinschaftsschullehrkraft, = 0 sonst); neg. Emot. = negative Emotionen; intr. Mot. = intrinsische Motivation; wahr. Komp. = wahrgenommene Kompetenz).
** p < .01. *** p < .001

6 Diskussion

Die Ergebnisse der Analysen des balancierten Datensatzes zeigten Effekte zugunsten der Lehrerinnen und Lehrer an Gemeinschaftsschulen: Für den eingeschätzten Mehrwert von Leistungsheterogenität und die intrinsische Motivation für den Umgang mit dieser ergaben sich signifikante Effekte mittlerer Größe zugunsten der Gemeinschaftsschullehrkräfte. Zusammenfassend kann also festgehalten werden, dass Gemeinschaftsschullehrkräfte Heterogenität stärker als eine positive Herausforderung verstehen, die den beruflichen Alltag bereichert, als vergleichbare Kolleginnen und Kollegen anderer Schularten.

Interessant erscheint dieser Befund auch vor dem Hintergrund des durchgeführten Matchings: Betrachtet man hier die Unterschiede in den Mittelwerten vor dem Matching, so ergeben sich mittlere bis große Effekte. Diese wären ohne die eingesetzte Zuordnung überinterpretiert worden. Andersherum kann man festhalten, dass trotz gleicher objektiver Heterogenität immer noch deutliche Unterschiede in den Einstellungen zu finden sind, vor allem was den Mehrwert auf kognitiver Ebene und die intrinsische Motivation auf motivational-affektiver Ebene angeht.

Welche Implikationen ergeben sich aus den gefundenen Ergebnissen für die Lehrerbildung? Die Ergebnisse lassen zwei Interpretationen zu: Einerseits scheinen sich Lehrkräfte mit günstigeren Einstellungen im Umgang mit Leistungsheterogenität bewusst für die Arbeit an integrativen Schulen zu entscheiden. Andererseits könnten sich die Einstellungen aufgrund des alltäglichen Umgangs mit einer heterogenen Schülerschaft bzw. einem auf Umgang mit Heterogenität ausgerichteten Unterrichts- und Schulkonzept positiv entwickelt haben. Mit Blick auf das einleitend erwähnte Ziel schulstruktureller Veränderungen, nämlich den Abbau sozialer Disparitäten (vgl. Wacker/Bohl 2016), stellen beide Interpretationen einen günstigen Umstand dar. Interpretiert man die Ergebnisse jedoch hinsichtlich nicht-integrativer Schulformen bei gleicher objektiver Heterogenität, so ergibt sich hier ein besonderer Bedarf in der Aus-, Fort- und Weiterbildung von Lehrkräften: In der ersten Phase der Lehrerbildung könnten Lehr-Lern-Gelegenheiten zur Reflexion der eigenen Einstellungen gerade bezüglich des späteren Einsatzortes angeboten werden. Hierbei kann auch die grundlegende Wahrnehmung von Heterogenität reflektiert werden. Für die zweite und insbesondere die dritte Phase der Lehrerbildung ergibt sich der Bedarf nach Fort- und Weiterbildung zum Erkennen von und Umgang mit Leistungsheterogenität. Bedeutsam erscheint zudem, die Wahrscheinlichkeit eines erfolgreichen Umgangs mit Heterogenität in der späteren Schulpraxis zu erhöhen, bspw. durch Vermittlung möglicher und wirksamer Unterrichtskonzepte oder durch Reflexion von Kooperations- und Schulentwicklungsmaßnahmen.

Bei der Interpretation der Ergebnisse ist einschränkend zu berücksichtigen, dass auch andere Kovariaten, die nicht in das Matching einbezogen wurden, einen Effekt auf die Einstellungen haben könnten bzw. im Zusammenhang mit ihnen stehen könnten. Sicherlich wäre es auch eleganter, die Werte der objektiven Heterogenität auf Klassen- und nicht nur auf Schulebene vorliegen zu haben, sodass sie noch besser den Lehrkräften zuzuordnen gewesen wären. Kritisch zu diskutieren ist auch der Aspekte, dass die Skala „wahrgenommene Kompetenz" durchaus auch dem Konstrukt der Selbstwirksamkeit zugeordnet werden könnte. Da hier jedoch ein bestehendes und etabliertes Instrument verwendet wurde, wurde auch einen Ausschluss der Skala verzichtet.

Abschließend soll darauf verwiesen werden, dass der Beitrag nur einen ersten Schritt einer langen Wirkungskette fokussiert hat: Den Unterschied in Einstellungen zwischen unterschiedlichen Gruppen von Lehrkräften. Empirisch weitestgehend noch ungeklärt bleibt die Frage, wie die Einstellungen im Unterrichtsverhalten wirksam werden und ob sich dies wiederum in den Outcomes der Schülerinnen und Schüler (z. B. Leistung) niederschlägt.

Literatur

Ajzen, Icek/Fishbein, Martin (2000): Attitudes and the attitude-behavior relation: reasoned and automatic processes. In: European Review of Social Psychology 11, H. 1, S. 1–33

Avramidis, Elias/Norwich, Brahm (2002): Teacher's attitudes towards integration/inclusion: a review of the literature. In: European Journal of Special Needs Education 17, S. 129–147

Batzel-Kremer, Andrea/Pietsch, Markus/Merk, Samuel/Bohl, Thorsten/Prinz, Eva/Schneider, Jürgen (2016): Die Gemeinschaftsschule in Baden-Württemberg: Situation und Sicht der Akteurinnen und Akteure. Propensity-Score-Analysen differenzieller Entwicklungsmilieus. In: Bohl, Thorsten/Wacker, Albrecht (Hrsg.): Die Einführung der Gemeinschaftsschule in Baden-Württemberg: Abschlussbericht der wissenschaftlichen Begleitforschung (WissGem). Münster u. a.: Waxmann, S. 47–72

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520

Blossfeld, Hans-Peter/Bos, Wilfried/Lenzen, Dieter/Müller-Böling, Detlef/Oelkers, Jürgen/Prenzel, Manfred/Wößmann, Ludger (2007): Bildungsgerechtigkeit. Jahresgutachten 2007. Wiesbaden: VS Verlag

Bohl, Thorsten/Wacker, Albrecht (2016a): Die Einführung der Gemeinschaftsschule in Baden-Württemberg. Abschlussbericht der wissenschaftlichen Begleitforschung (WissGem). Münster: Waxmann

Bohl, Thorsten/Wacker, Albrecht (2016b): Das Forschungsprojekt ‚Wissenschaftliche Begleitforschung Gemeinschaftsschulen in Baden-Württemberg (WissGem)'. Einordnung – Design – Vorgehensweise. In: Bohl, Thorsten/Wacker, Albrecht (Hrsg.): Die Einführung der Gemeinschaftsschule in Baden-Württemberg. Abschlussbericht der wissenschaftlichen Begleitforschung (WissGem). Münster: Waxmann, S. 11–26

Bollen, Kenneth/Tueller, Stephen/Oberski, Daniel (2013): Issues in the structural equation modeling of complex survey Data. In: 59th World Statistics Congress. Hong Kong.

Brown, Timothy A. (2014): Confirmatory factor analysis for applied research. New York: Guilford Press

Cohen, Jacob (1988): Statistical power analysis for the behavioral sciences. New York: Lawrence Erlbaum

Diamond, Alexis/Sekhon, Jasjeet S. (2013): Genetic matching for estimating causal effects: a general multivariate matching method for achieving balance in observational studies. Review of Economics and Statistics 95, H. 3, S. 932–945. doi:10.1162/REST_a_00318

Eagley, Alice H./Chaiken, Shelly (1993): The psychology of attitudes. Belmont: Thomson Wadsworth

Felbrich, Anja/Schmotz, Christiane/Kaiser, Gabriele (2008): Überzeugungen angehender Primarstufenlehrkräfte im internationalen Vergleich. In: Blömeke, Sigrid (Hrsg.): Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Münster u. a.: Waxmann, S. 297–325

Gebauer, Miriam M./McElvany, Nele/Klukas, Stephanie (2013): Einstellungen von Lehramtsanwärterinnen und Lehramtsanwärtern zum Umgang mit heterogenen Schülergruppen in Schule und Unterricht. In: McElvany, Nele/Gebauer, Miriam M./Bos, Wilfried/Holtappels, Günther (Hrsg.), Jahrbuch der Schulentwicklung. Weinheim, Basel: Beltz Juventa, S. 191–216

Jöreskog, Karl G./Goldberger, Arthur S. (1975): Estimation of a model with multiple indicators and multiple causes of a single latent variable. In: Journal of the American Statistical Association 70, H. 351a, S. 631–639

Kiel, Ewald/Syring, Marcus/Weiß, Sabine (2017): How can intercultural school development succeed? The perspective of teachers and teacher educators. In: Pedagogy, Culture and Society 25, H. 2, S. 243–261

Kunter, Mareike/Pohlmann, Britta (2009): Lehrer. In: Wild, Elke (Hrsg.): Pädagogische Psychologie. Berlin: Springer, S. 261–282

Kunz, André/Luder, Reto/Moretti, Marta (2010): Die Messung von Einstellungen zur Integration (EZI). Empirische Sonderpädagogik 2, H. 3, S. 83–94

Merk, Samuel/Cramer, Colin/Dai, Naien/Bohl, Thorsten/Syring Marcus (2018): Faktorielle Validität der Einstellungen von Lehrkräften zu heterogenen Lerngruppen. In: Journal for Educational Research Online (JERO) 10, H. (2), S. 34–53 (auch online unter www.j-e-r-o.com/index.php/jero/article/view/843 (Abfrage: 11.02.2019))

Pajares, M. Frank (1992): Teachers' beliefs and educational research: cleaning up a messy construct. In: Review of Educational Research 62, H. 3, S. 307–332

Rosenberg, Milton J./Hovland, Carl I. (1960): Cognitive, affective and behavioral components of attitudes. In: Hovland, Carl I./Rosenberg, M. J. (Hrsg.): Attitude organization and change. An analysis of consistency among attitude components. Bd. 3. New Haven: Yale University Press, S. 1–14

Rosenbaum, Paul R./Rubin, Donald B. (1983): The central role of the propensity score in observational studies for causal effects. In: Biometrika, 70, H. 1, S. 41–55

Schulgesetz für Baden-Württemberg (SchG). Verfügbar unter: http://www.landesrecht-bw.de/jportal/?quelle=jlink&query=SchulG+BW&psml=bsbawueprod.psml&max=true&aiz=true (Abfrage: 11.02.2019)

Syring, Marcus/Merk, Samuel/Cramer, Colin/Topalak, Cansu/Bohl, Thorsten (2019 im Druck): Der Migrationshintergrund Lehramtsstudierender als Prädiktor ihrer Einstellungen zu heterogenen Lerngruppen. In: Zeitschrift für Bildungsforschung

Syring, Marcus/Tillmann, Teresa/Ivanova, Alina/Kiel Ewald (2019): Passung von Transformationen am Gymnasium und Einstellungen von Gymnasiallehramtsstudierenden zur Heterogenität und Inklusion: Eine qualitativ-quantitative Untersuchung zu Wahrnehmung, Schulartunterschieden und Prädiktoren. In: Zeitschrift für Bildungsverwaltung 35, H. 1, S. 25–47

Tillman, Klaus-Jürgen (2012): Das Sekundarschulsystem auf dem Weg in die Zweigliedrigkeit: Historische Linien und aktuelle Verwirrungen. In: Pädagogik 64, H. 5, S. 8–12.

Voss, Thamar/Kleickmann, Thilo/Kunter, Mareike/Hachfeld, Axinja (2011): Überzeugungen von Mathematiklehrkräften. In: Baumert, Jürgen/Blum, Werner/Kunter, Mareike/Neubrand, Michael (Hrsg.): Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster u. a.: Waxmann, S. 235–257

Wacker, Albrecht/Bohl, Thorsten (2016): Schulsystem und Gemeinschaftsschule in Baden-Württemberg. In Bohl, Thorsten/Wacker, Albrecht (Hrsg.): Die Einführung der Gemeinschaftsschule in Baden-Württemberg. Abschlussbericht der wissenschaftlichen Begleitforschung (WissGem). Münster: Waxmann, S. 27–44

Werth, Lioba/Mayer, Jennifer (2008): Sozialpsychologie. Berlin: Spektrum.

Die Implementation digitaler Medien gestalten

Zur Bedeutung der Einstellungen von Fremdsprachenlehrkräften

Kerstin Drossel und Birgit Eickelmann

Abstract

Medienbezogene Einstellungen von Lehrkräften gelten als Bedingungsfaktor für die Implementation digitaler Medien in den Unterricht, wobei sich für den Fremdsprachenunterricht besondere Potenziale hinsichtlich des Einsatzes digitaler Medien ergeben. Da in bisherigen Untersuchungen lediglich Mittelwerte der Einstellungen aller Lehrkräfte betrachtet wurden, wird in diesem Beitrag eine Typologie der Einstellungen von Fremdsprachenlehrkräften zum unterrichtlichen Einsatz digitaler Medien ermittelt. Zudem werden die Nutzungshäufigkeit digitaler Medien im Unterricht sowie weitere Determinanten der Nutzung im Zusammenhang mit den identifizierten Typen untersucht. Die Identifikation der Typen erfolgt mithilfe einer Latent-Class-Analyse auf Basis der Daten der Lehrkräfte der IEA-Studie ICILS 2013 aus sechs europäischen Ländern. Dabei lassen sich vier Typen identifizieren: kritische ICT-Enthusiasten, ICT-Enthusiasten, ICT-Skeptiker und ICT-Skeptiker mit positiven Zügen. Mehrheitlich positive Einstellungstypen finden sich in allen Ländern mit Ausnahme von Deutschland. Erwartungskonform weisen die ICT-Enthusiasten zumeist die höchste Nutzungshäufigkeit auf, während ICT-Skeptiker selten digitale Medien im Unterricht einsetzen.

1 Problemstellung

Durch die Technisierung und Digitalisierung aller Lebens- und Arbeitsbereiche entstehen für Bildungssysteme Herausforderungen, da sich sowohl technische, pädagogische und didaktische Veränderungen ergeben, die den fachlichen und überfachlichen Kompetenzerwerb der Schülerinnen und Schüler unterstützen sollen (vgl. Eickelmann 2018; Heinen/Kerres 2017; KMK 2016). Dabei kommt

Lehrkräften sowohl als Kompetenzvermittlern als auch in Bezug auf die Umsetzung medienbezogener Innovationen eine wichtige Rolle zu. Dafür bedarf es des Einsatzes neuer Technologien in fachdidaktischen Lehr- und Lernszenarien, wobei sich insbesondere im Bereich der Fremdsprachen vielfältige Potenziale aufzeigen lassen, die sich vor allem auf das individuelle Lernen der Schülerinnen und Schüler als auch auf die Kommunikation und Kollaboration unter den Lernenden beziehen. Beispielsweise kann die Sprachkompetenz der Schülerinnen und Schüler durch den Zugang zur digitalen Welt, die eine große Bandbreite an fremdsprachlichen Ressourcen bietet (z. B. Texte, Filme, Audiodateien, Chats, Foren) und zahlreiche Möglichkeiten der sprachlichen Rezeption und Produktion offeriert, gefördert werden (vgl. Dausend 2017; Lopes 2018). Weiter sind authentische und lebensweltbezogene Sprachkontakte mit Jugendlichen anderer Nationen per Chat denkbar, aber auch der Einsatz von Programmen zum Vokabellernen, digitale Übungen zum Hörverständnis oder Wortschatzübungen zum individuellen Lernen (vgl. Biebighäuser/Zibelius/Schmidt 2012; Dausend 2017; Holmberg 2019). Damit der Einsatz digitaler Medien jedoch zu Entwicklungen von Kompetenzen der Schülerinnen und Schüler beitragen kann, ist eine Implementation in unterrichtliche Lehr- und Lernszenarien durch die Lehrkräfte erforderlich. Diese ist wiederum von den Einstellungen der Lehrkräfte gegenüber digitalen Medien abhängig (vgl. Eickelmann et al. 2014; Holmberg 2019; Lopes 2018; Tondeur et al. 2019). Der vorliegende Beitrag verfolgt in diesem Zusammenhang das Ziel, eine Forschungslücke bezüglich der Einstellungen von Fremdsprachenlehrkräften gegenüber digitalen Medien zu schließen, indem anhand von Sekundäranalysen mit den Daten europäischer Fremdsprachenlehrkräfte der IEA ICILS-2013-Studie (vgl. Fraillon et al. 2014) untersucht wird, ob sich Typen der medienbezogenen Einstellungen der Lehrkräfte identifizieren lassen und inwieweit diese Typen mit der unterrichtlichen Nutzung im Zusammenhang stehen. Bevor die eigene Untersuchung vorgestellt wird, werden die Einstellungen der Lehrkräfte als auch die Nutzung digitaler Medien in der Schulentwicklungs- und Schulqualitätsforschung verortet und im Hinblick auf den aktuellen Forschungsstand dargestellt. Der Beitrag schließt mit einer Diskussion, in der die im Rahmen der eigenen Untersuchung ermittelten Unterschiede und Gemeinsamkeiten als Ansatzpunkte für eine nachhaltige Verbesserung des mediengestützten Fremdsprachenunterrichts auf Basis von Grundüberzeugungen der Fremdsprachenlehrkräfte betrachtet werden.

Medienbezogene Einstellungen der Lehrkräfte

Ein zentraler Prädiktor des Verhaltens stellt die Einstellung zum Objekt dar (vgl. Eagly/Chaiken 1993). Diese ist mehrdimensional und umfasst mehrere qualitativ unterscheidbare Teilaspekte. Die Einstellung zum Objekt steht dabei am Anfang einer Ursache-Wirkungskette und nimmt somit eine zentrale Funktion für die Vorhersage der letztlichen Handlung ein. Im schulischen Kontext gelten Einstellungen als ein Baustein der professionellen Kompetenz von Lehrkräften, die zur Qualität von Unterricht beitragen (vgl. König 2016). In Schulentwicklungs- und Schulqualitätsmodellen sind Einstellungen der Lehrkräfte auf der Ebene der Voraussetzungen verortet (z. B. Eickelmann/Drossel angenommen; Fraillon et al. 2014; 2019). Die Einstellungen stehen damit in einem direkten Zusammenhang mit Faktoren auf der schulischen Prozessebene, die wiederum im Zusammenhang mit den Leistungsergebnissen der Schülerinnen und Schüler auf der Outputebene stehen.

Bei Betrachtung der medienbezogenen Einstellungen der Lehrkräfte im internationalen Vergleich wird deutlich, dass diese stark zwischen den Bildungssystemen variieren und Deutschland oftmals das Schlusslicht darstellt (vgl. Fraillon et al. 2014). Beispielsweise stimmten nur rund zwei Fünftel der Lehrkräfte in Deutschland zu, dass der Einsatz digitaler Medien im Unterricht zu einer Verbesserung der schulischen Leistungen der Schülerinnen und Schüler führen kann (vgl. ebd.). In Ländern wie Polen, Dänemark und Norwegen stimmten hingegen über 70 % der Lehrkräfte dieser Aussage zu. Zudem haben Eickelmann und Vennemann (2017) fünf medienbezogene Einstellungstypen von Lehrkräften in Deutschland, Norwegen und der Tschechischen Republik identifiziert. Für die Fremdsprachenlehrkräfte in Deutschland zeigte sich dabei, dass diese besonders häufig der Gruppe der ICT-Enthusiasten (36.5 %) und vergleichsweise zu geringen Anteilen den absoluten Zweiflern, die den Einsatz von digitalen Medien in der Schule ablehnen (23.3 %), zuzuordnen sind. Diese positive Sichtweise bestätigt sich auch mit einer Studie des BITKOM-Digitalverbunds (2011), in der dargelegt wurde, dass ein großer Anteil (67 %) der Fremdsprachenlehrenden über eine eher positive Einstellung zum Medieneinsatz im Unterricht verfügt.

Unterrichtliche Nutzung digitaler Medien

In Schulentwicklungs- und Schulqualitätsmodellen (z. B. Fraillon et al. 2014; Fraillon et al. 2019) lässt sich auch die Nutzung von Computern im Unterricht durch Lehrkräfte einordnen. Diese ist auf der unterrichtlichen Prozessebene

angesiedelt und steht somit in einem direkten Zusammenhang mit den medienbezogenen Einstellungen der Lehrkräfte.

Die Befundlage zur Nutzungshäufigkeit aus der IEA-Studie ICILS 2013 zeigt jedoch, dass lediglich etwa ein Drittel der Lehrkräfte in Deutschland angab, regelmäßig (mindestens wöchentlich) digitale Medien im Unterricht einzusetzen (vgl. Eickelmann/Schaumburg/Drossel/Lorenz 2014). Damit stellte Deutschland das Schlusslicht im internationalen Vergleich dar. In Bezug auf die hier untersuchten Bildungssysteme zeigt sich darüber hinaus, dass vor allem die Lehrkräfte in Dänemark (79.5 %), gefolgt von Norwegen (77.9 %), den Niederlanden (77.4 %) und der Tschechischen Republik (65.6 %) regelmäßig digitale Medien im Unterricht einsetzten. Neben Deutschland gaben in Polen weniger als die Hälfte der Lehrkräfte (41.5 %) an, Computer mindestens wöchentlich im Unterricht zu nutzen (vgl. ebd.).

3 Forschungsfragen

Bisher liegen keine international repräsentativen fachspezifischen Untersuchungen zu medienbezogenen Einstellungen der Lehrkräfte vor, die Variationen der Einstellungen gegenüber digitalen Medien berücksichtigen, wobei sich bezüglich des Fremdsprachenunterrichts besonderes Potenzial für den Einsatz digitaler Medien ergibt. Daher werden folgende Forschungsfragen formuliert:

1. Ist eine Typologie von Fremdsprachenlehrkräften in Deutschland und anderen ausgewählten europäischen Bildungssystemen auf Basis von deren Einstellungen zum unterrichtlichen Einsatz digitaler Medien identifizierbar?
2. Wie verbreitet sind die identifizierten Typen in den ausgewählten europäischen Bildungssystemen?
3. Inwieweit unterscheidet sich die regelmäßige Nutzung von Computern im Unterricht zwischen den Fremdsprachenlehrkräftetypen?

4 Methodisches Vorgehen

Stichprobe

Um die Forschungsfragen zu beantworten, werden Sekundäranalysen mit dem Datensatz der Lehrkräfte der IEA-Studie ICILS 2013 (vgl. Fraillon et al. 2014) durchgeführt. Neben Deutschland werden die Angaben der repräsentativ gezogenen Lehrkräfte der achten Jahrgangsstufe der europäischen Bildungssysteme in die Analysen eingehen, deren Schülerinnen und Schüler im Mittel über signifikant höhere computer- und informationsbezogene Kompetenzen verfügen

als die Achtklässlerinnen und Achtklässler in Deutschland (Dänemark, Niederlande, Norwegen, Polen und die Tschechische Republik), um sich aus deutscher Sicht an erfolgreicheren Bildungssystemen im Hinblick auf die medienbezogenen Kompetenzen zu orientieren. Im Sinne der Forschungsfragen werden lediglich diejenigen Lehrkräfte berücksichtigt, die eine Fremdsprache unterrichten. Die Gesamtstichprobe der Fremdsprachenlehrkräfte in den untersuchten Bildungssystemen umfasst 2.406 Fremdsprachenlehrkräfte, was einem Anteil von 27.6 % an der gesamten Stichprobe der Lehrkräfte entspricht (siehe Tabelle 1).

Tabelle 1: Verteilung der Fremdsprachenlehrkräfte (FSL) in den untersuchten Bildungssystemen

Bildungssystem	N_{GESAMT}	N_{FSL}	**Anteil FSL**
Dänemark	728	254	34.9 %
Deutschland	1.386	404	29.1 %
Niederlande	1.083	254	23.5 %
Norwegen	1.158	400	34.5 %
Polen	2.228	452	20.3 %
Tschechische Republik	2.126	642	30.2 %
Gesamt	**8.709**	**2.406**	**27.6 %**

Messinstrumente

Die Kriterien für die Itemauswahl der medienbezogenen Einstellungen der Lehrkräfte ergeben sich aus der fachdidaktischen Perspektive des Fremdsprachenunterrichts, der insbesondere interaktive und kooperative Lernformen für das Erlernen von Fremdsprachen in den Vordergrund rückt (vgl. Biebighäuser et al. 2012). So erfragen die sechs Items beispielsweise (vgl. Gerick et al. 2018, vgl. Abb. 1), inwiefern die Lehrkräfte die Einstellung haben, dass digitale Medien einen Einfluss auf die Zusammenarbeit zwischen den Schülerinnen und Schülern haben und somit eine effektivere oder eingeschränktere Kommunikation herbeigeführt wird. In Anlehnung an Fraillon et al. (2014; 2019) lassen sich die Einstellungen der Lehrkräfte in positive und negative Sichtweisen unterteilen, womit auch dem mehrdimensionalen Verständnis von Einstellungen Rechnung getragen wird. Die Antworten wurden dichotomisiert, indem die Kategorien ‚stimme zu' und ‚stimme eher zu' in ‚Zustimmung' und die Kategorien ‚stimme eher nicht zu' und ‚stimme nicht zu' in ‚Ablehnung' rekodiert wurden. Die Häufigkeit der unterrichtlichen Computernutzung der Lehrkräfte wird über eine fünfstufige Likert-Skala von ‚nie' bis zu ‚jeden Tag' erfasst. Dabei wird von einer regelmäßigen Nutzung gesprochen, wenn der Computer mindestens wöchentlich (bestehend aus den Antwortkategorien ‚jeden Tag' und

‚mindestens einmal in der Woche, aber nicht jeden Tag') im Unterricht eingesetzt wird.

Analysemethodik

Um die erste Forschungsfrage zu beantworten, werden mittels einer Latent-Class-Analyse (LCA) (vgl. Gollwitzer 2012) Lehrkräftetypen identifiziert (Mplus). Für die Beurteilung der Modellgüte werden das Akaike Information Criterion (AIC) und Bayesian Information Criterion (BIC) bzw. das sample size adjusted Bayesian Information Criterion (aBIC) als informationstheoretische Maße herangezogen (vgl. Kuha 2004). Das BIC hat dabei im Gegensatz zum AIC unter anderem den Vorteil, dass die Spezifikation konservativer erfolgt und weniger wahrscheinlich komplexe Modelllösungen bevorzugt werden (vgl. Wolf/Best 2010). Niedrige Werte stehen bei beiden Kriterien für eine gute Modellpassung. Das Entropie-Maß beschreibt die Güte der Klassifikation. Zudem werden die Ergebnisse des Vuong-Lo-Mendell-Rubin Tests (vgl. Lo/Mendell/Rubin 2001) berichtet, die darüber Aufschluss geben, ob das jeweils geschätzte Modell signifikant besser die Daten beschreibt, als ein Modell mit einer Klasse weniger.

Hinsichtlich des Umgangs mit fehlenden Werten wurden zunächst diejenigen Lehrkräfte aus den Analysen ausgeschlossen, die zu keiner der sechs medienbezogenen Einstellungen Angaben gemacht haben (N = 20), sodass letztlich die Informationen von 2.386 Lehrkräften in die LCA eingehen. Zudem wurde auf das Full-Information-Maximum-Likelihood-Verfahren zurückgegriffen (FIML, vgl. Lüdtke/Robitzsch/Trautwein/Köller 2007). Darüber hinaus wurde die Stratifizierung der Daten (Lehrkräfte in Schulen) durch den Analysetyp ‚Type = mixture complex' (vgl. Muthen/Satorra 1995) berücksichtigt, der bei der Berechnung der Standardfehler und Modellpassung die statistische Verzerrung (Designeffekt), die der Stichprobe geschuldet ist, berücksichtigt. Vor den eigentlichen Berechnungen wird zudem ein sogenanntes Senate Weight (SENWGTT) mit der Gesamtstichprobe berechnet, um sicherzustellen, dass trotz unterschiedlich großer Stichproben der jeweiligen Bildungssysteme jedes einzelne unter ihnen zu gleichen Teilen Fälle zur Gesamtstichprobe beiträgt (vgl. Martin 2005). In einem weiteren Schritt werden anhand deskriptiver Statistiken mit Hilfe des IDB-Analyzers (vgl. Rutkowski/Gonzalez/Joncas/von Davier 2010) die Zusammenhänge der identifizierten Einstellungstypen zur unterrichtlichen Computernutzung analysiert. Dabei wird die Gewichtungsvariable für Lehrpersonen einbezogen (vgl. Jung/Carstens 2015), um Stichprobenverzerrungen auszugleichen und valide Schätzungen in Bezug auf die untersuchte Population zu erhalten.

5 Ergebnisse

Forschungsfrage 1: Typologie der Einstellungen der Lehrkräfte

Im Hinblick auf die erste Fragestellung zeigt sich im Ergebnis der LCA anhand des Informationskriteriums BIC bzw. aBIC sowie dem Entropie-Wert, dass eine Vier-Gruppenlösung die Daten am besten beschreibt (siehe Tabelle 2). Die durchschnittliche korrekte Zuordnungswahrscheinlichkeit beträgt für den ersten Lehrkräftetyp 82 %, für den zweiten 80 %, für den dritten 82 % und für den vierten 69 %.

Tabelle 2: Modell-Fit-Werte für die LCA (N = 2.386)

Klassen	AIC	BIC	sample size adjusted BIC	Entropie	LMR-LR-Test
2	17004.619	17079.725	17038.421	0.636	0.2651
3	16810.810	16926.358	16862.813	0.614	0.4119
4	16659.744	16815.733	16729.948	0.612	0.5063
5	16642.968	16839.399	16839.399	0.652	0.7771

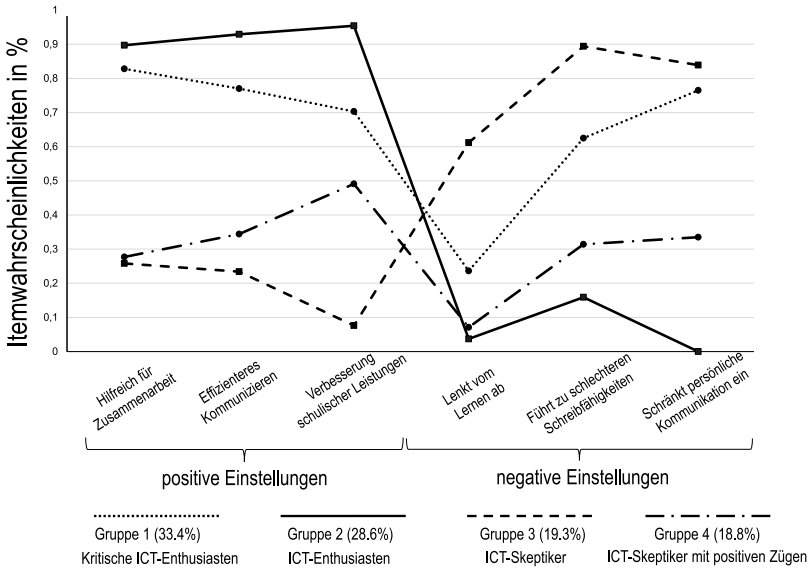
Der erste Lehrkräftetyp (vgl. Abb. 1) umfasst mit einem Anteil von einem Drittel (33.4 %) den größten Teil der Fremdsprachenlehrkräfte und ist durch einen vergleichsweise hohen Anteil positiver Einstellungen, aber auch zumeist einen hohen Anteil an kritischen Einstellungen gegenüber dem Einsatz neuer Medien in unterrichtliche Lehr- und Lernprozesse charakterisiert. Lehrpersonen, die dieser Gruppe zugeordnet werden können, nehmen statistisch beschrieben vor allem wahr, dass digitale Medien hilfreich für die Zusammenarbeit der Schülerinnen und Schüler sind (82.8 %). Ein hoher Anteil ist jedoch auch der Ansicht, dass der Einsatz digitaler Medien im Unterricht die persönliche Kommunikation einschränkt (76.5 %) sowie zu schlechteren Schreibfähigkeiten führt (62.5 %). Sie werden daher als kritische ICT-Enthusiasten bezeichnet.

In der zweiten Gruppe sind 28.6 % der Fremdsprachenlehrkräfte vertreten. Im Gegensatz zum ersten Typ sind die Lehrkräfte dieser Gruppe gegenüber dem Einsatz digitaler Medien durchweg positiv eingestellt. Die Lehrkräfte sehen vor allem Potenzial hinsichtlich einer Verbesserung der schulischen Leistungen (95.4 %) sowie einer effizienteren Kommunikation (92.9 %). Wie bei dem ersten Typ nehmen die Lehrkräfte auch mit einer hohen Wahrscheinlichkeit wahr, dass der Einsatz neuer Technologien zur Zusammenarbeit zwischen Schülerinnen und Schülern beiträgt (89.7 %). Hingegen ist die Wahrscheinlichkeit, dass Lehrkräfte dieser Gruppe den Einsatz kritisch einschätzen, sehr gering und beträgt beispielsweise für die Einstellung, dass der Medieneinsatz vom Lernen

ablenkt, lediglich 3.7 %. Lehrkräfte dieser Gruppe werden daher als ICT-Enthusiasten beschrieben.

Die dritte Gruppe umfasst etwa ein Fünftel der Lehrkräfte (19.3 %). Lehrkräfte dieser Gruppe weisen im Vergleich zu den anderen Typen die geringste Wahrscheinlichkeit auf, dass sie positiv gegenüber digitalen Medien eingestellt sind. Gering ist in dieser Gruppe vor allem die Wahrscheinlichkeit, dass Lehrkräfte der Ansicht sind, dass der Einsatz digitaler Medien die Leistungen der Schülerinnen und Schüler verbessert (7.6 %). Hingegen ist die Wahrscheinlichkeit, dass Lehrkräfte dieses Typs negative Einstellungen gegenüber dem Einsatz digitaler Medien in den Unterricht kommunizieren, vergleichsweise hoch. So liegt die Wahrscheinlichkeit, dass Lehrkräfte dieses Typs der Ansicht sind, dass der Medieneinsatz zu schlechteren Schreibfähigkeiten der Schülerinnen und Schüler führt, bei 89.4 %. Diese Gruppe wird aufgrund der geringen Wahrscheinlichkeiten für positive und der hohen Wahrscheinlichkeit für negative Einstellungen auch als ICT-Skeptiker bezeichnet.

Abbildung 1: Ergebnis der LCA hinsichtlich der Einstellungstypologie



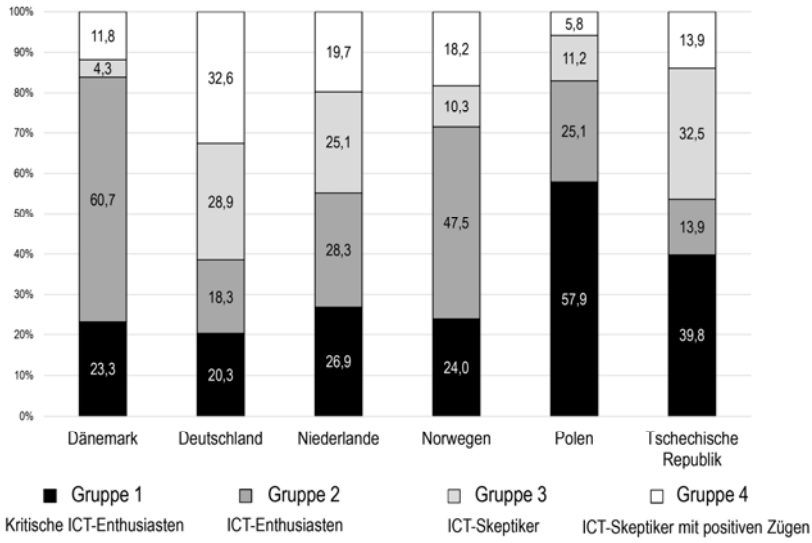
In der vierten Gruppe sind rund ein Fünftel der Fremdsprachenlehrkräfte vertreten (18.8 %). Lehrkräfte dieser Gruppe weisen ebenfalls mit einer geringen Wahrscheinlichkeit positive Einstellungen hinsichtlich des Einsatzes digitaler Medien auf, wobei die Wahrscheinlichkeit bei Lehrkräften des dritten Typs geringer ausfiel. So gibt ein höherer Anteil der Lehrkräfte der vierten Gruppe beispielsweise an, dass der Einsatz digitaler Medien den Schülerinnen und Schülern dabei hilft, zu lernen, mit anderen zusammenzuarbeiten (27.7 %) oder

zu einer effizienteren Kommunikation führt (34.4 %). Auffällig ist hingegen, dass der Anteil der Lehrkräfte mit negativen Einstellungen gegenüber dem Einsatz digitaler Medien ebenfalls vergleichsweise gering ausfällt. So liegt die Wahrscheinlichkeit, dass Lehrkräfte dieser Gruppe die Einstellung vertreten, dass digitale Medien die Schülerinnen und Schüler vom Lernen ablenken, bei lediglich 7.1 %. Die Gruppe wird daher als ICT-Skeptiker mit positiven Zügen charakterisiert.

Forschungsfrage 2: Verteilung der Typen der Einstellungen der Lehrkräfte auf die Bildungssysteme

Betrachtet man nun die Verteilung der Typen der Einstellungen der Lehrkräfte auf die ausgewählten Bildungssysteme, zeigt sich, dass die Anteile der Typen jeweils unterschiedlich auf die Systeme verteilt sind. Während in Polen der Großteil der Fremdsprachenlehrkräfte (57.9 %) dem ersten Typ (Kritische ICT-Enthusiasten) zuzuordnen ist und auch in der Tschechischen Republik vergleichsweise viele Lehrkräfte dieses Typs zu finden sind (39.8 %), ist diese Gruppe in Dänemark, Deutschland, den Niederlanden und Norwegen mit rund einem Fünftel vertreten. Die zweite Gruppe (ICT-Enthusiasten) findet sich anteilig vorwiegend in Dänemark (60.7 %). In Norwegen können noch knapp die Hälfte (47.5 %) der Lehrkräfte diesem Typen zugeordnet werden, in den Niederlanden (28.3 %) und Polen (25.1 %) beträgt der Anteil etwa ein Viertel. In den weiteren zwei Bildungssystemen liegt der Anteil des zweiten Typs jeweils unter einem Fünftel der Fremdsprachenlehrkräfte (Deutschland 18.3 %, Tschechische Republik 13.9 %). Hohe Anteile von ICT-Skeptikern (Gruppe 3) finden sich hauptsächlich in der Tschechischen Republik (32.5 %), in Deutschland (28.9 %) und den Niederlanden (25.1 %). In Polen (11.2 %), Norwegen (10.3 %) und Dänemark (4.3 %) sind jeweils ein Zehntel oder weniger der Fremdsprachenlehrkräfte dieser Gruppe zuzuordnen. Etwa ein Drittel der Fremdsprachenlehrkräfte in Deutschland (32.6 %) ist der vierten Gruppe (ICT-Skeptiker mit positiven Zügen) zuzuordnen. In den anderen fünf Bildungssystemen fällt dieser Anteil mit maximal einem Fünftel der Lehrkräfte jeweils deutlich geringer aus.

Abbildung 2: Verteilung der einstellungsbezogenen Fremdsprachenlehrkräftetypen im europäischen Vergleich



Forschungsfrage 3: Unterrichtliche Computernutzung nach einstellungsbezogenem Fremdsprachenlehrkräftetyp im europäischen Vergleich

Betrachtet man die mittlere Nutzungshäufigkeit von Computern durch Fremdsprachenlehrkräfte im Unterricht differenziert nach den identifizierten Fremdsprachenlehrkräftetypen (s. Tab. 3), ist ersichtlich, dass der Anteil der Lehrkräfte, der angibt, regelmäßig digitale Medien im Unterricht einzusetzen, zumeist in Gruppe drei (ICT-Skeptiker) am geringsten ausfällt. In Deutschland nutzen lediglich 13.8 % der Fremdsprachenlehrkräfte der dritten Gruppe regelmäßig (mindestens wöchentlich) digitale Medien im Unterricht. In Dänemark trifft dies etwa auf die Hälfte der Fremdsprachenlehrkräfte der Gruppe drei zu (49.9 %). In der Tschechischen Republik geben 56.8 %, in den Niederlanden 70.8 % und in Norwegen 65.6 % der Fremdsprachenlehrkräfte des dritten Typs an, regelmäßig digitale Medien im Unterricht zu nutzen. Auch in Polen (38.2 %) fällt der Anteil der regelmäßigen Nutzer und Nutzerinnen in der dritten Gruppe vergleichsweise gering aus. Sie ist jedoch höher als die dokumentierte Nutzungshäufigkeit bei Fremdsprachenlehrkräften der vierten Gruppe (25.7 %) (ICT-Skeptiker mit positiven Zügen). Hingegen sind es anteilig die meisten Fremdsprachenlehrkräfte der Gruppe zwei (ICT-Enthusiasten) in Dänemark (79.5 %), Deutschland (38.3 %), Norwegen (89.0 %) und der Tschechischen Republik (69.9 %), die regelmäßig Computer im Unterricht nutzen. In den Niederlanden trifft dies auf die Gruppe eins zu (92.4 %; kritische

ICT-Enthusiasten) und in Polen auf Fremdsprachenlehrkräfte der Gruppe drei (38.2 %; ICT-Skeptiker).

Tabelle 3: Unterrichtliche Computernutzung von Fremdsprachenlehrkräften nach einstellungsbezogenem Lehrkräftetyp im europäischen Vergleich (Kategorie: mind. wöchentlich in %; Standardfehler in Klammern)

	Gruppe 1: Kritische ICT-Enthusiasten	Gruppe 2: ICT-Enthusiasten	Gruppe 3: ICT-Skeptiker	Gruppe 4: ICT-Skeptiker mit pos. Zügen
Dänemark	73.1 (8.9)	79.5 (4.3)	49.9 (16.9)	76.4 (8.3)
Deutschland	20.8 (4.6)	38.3 (7.0)	13.8 (3.4)	24.6 (6.6)
Niederlande	92.4 (3.0)	79.7 (5.4)	70.8 (8.1)	82.9 (4.8)
Norwegen	77.6 (4.8)	89.0 (2.4)	65.6 (8.8)	82.4 (5.0)
Polen	37.4 (4.1)	37.4 (4.7)	38.2 (11.4)	25.7 (11.9)
Tschechische Republik	65.5 (3.5)	69.9 (5.8)	56.8 (4.6)	58.4 (7.0)

6 Diskussion

Die Ergebnisse verdeutlichen, dass nicht alle Fremdsprachenlehrkräfte digitale Medien als Gewinn für den Fremdsprachenunterricht, etwa in Bezug auf deren Rolle in der Zusammenarbeit der Schülerinnen und Schüler oder im Hinblick auf eine Leistungsverbesserung, sehen. Ein nicht zu vernachlässigender Teil sieht digitale Medien als potenzielles Risiko, beispielsweise in Bezug auf die Entwicklung von Schreibkompetenzen. Die Verteilung der Typen der Einstellungen der Lehrkräfte auf die ausgewählten Bildungssysteme variiert dabei stark. Beispielsweise sind die Fremdsprachenlehrkräfte in Deutschland besonders häufig den ICT-Skeptikern mit positiven Zügen sowie den ICT-Skeptikern zuzuordnen. In Bezug auf die Häufigkeit der regelmäßigen Computernutzung differenziert nach einstellungsbezogenem Fremdsprachenlehrkräftetyp können die Ergebnisse in den meisten Bildungssystemen als erwartungskonform bezeichnet werden. Der handlungsleitende Charakter von Einstellungen (vgl. Wilde/Kunter 2016) manifestiert sich hier durch eine häufigere Nutzung von digitalen Medien bei den ICT-Enthusiasten. Die geringste Computernutzung lässt sich hingegen in fast allen Bildungssystemen bei den ICT-Skeptikern beobachten.

Für künftige Untersuchungen könnte es zielführend sein, zu analysieren, inwieweit sich die identifizierten einstellungsbezogenen Fremdsprachentypen hinsichtlich soziodemografischer Merkmale unterscheiden und vor allem, inwieweit sich ein Zusammenhang mit digitalen Kompetenzen der Schülerinnen und Schüler zeigt. Darüber hinaus sind Kausalzusammenhänge aufgrund des

vorliegenden Querschnittsdesigns nicht abbildbar. Zielführend wären hier im Rahmen künftiger Forschungsprojekte Längsschnittstudien zur Aufdeckung kausaler Zusammenhänge sowie qualitative Befragungen einzelner Lehrkräfte zu Begründungslinien.

Für die Praxis ergibt sich aus den hier gesammelten Erkenntnissen zunächst einmal ein besseres Verständnis der Einstellungen von Fremdsprachenlehrkräftetypen. Der europäische Vergleich hat gezeigt, dass in Deutschland besonders viele Fremdsprachenlehrkräfte dem Typ der ICT-Skeptiker zuzuordnen sind. Fortbildungen als elementarem Bestandteil von Lehrkräftebildung kann hier die Aufgabe übertragen werden, Fremdsprachenlehrkräften in Deutschland die Potenziale digitaler Medien für ihren Fachunterricht fachspezifisch zu verdeutlichen.

Literatur

Biebighäuser, Katrin/Zibelius, Marja/Schmidt, Torben (2012): Aufgaben 2.0. Konzepte, Materialien und Methoden für das Fremdsprachenlehren und -lernen mit digitalen Medien. Tübingen: Narr Verlag.

BITKOM (2011): Eine repräsentative Untersuchung zum Einsatz elektronischer Medien an Schulen aus Lehrersicht. Berlin: BITKOM.

Dausend, Henriette (2017): Tablets zur Förderung diskursiver Aushandlungsprozesse im Fremdsprachenunterricht. In: Bastian, Jasmin/Aufenager, Stefan (Hrsg.): Tablets in Schulen. Wiesbaden: Springer VS, S. 355–379.

Eagly, Alice H./Chaiken, Shelly (1993): The psychology of attitudes. Orlando, FL, US: Harcourt Brace Jovanovich College.

Eickelmann, Birgit (2018): Digitalisierung an Schulen. Eine Bestandsaufnahme. SchulVerwaltung Spezial, 20(4), 152–155.

Eickelmann, Birgit/Drossel, Kerstin (angenommen): Digitalisierung im deutschen Bildungssystem im Kontext des Schulreformdiskurses. In: Berkemeyer, Nils/Bos, Wilfried/Hermstein, Björn (Hrsg.): Schulreform – gestern, heute, morgen. Weinheim/Basel: Beltz Juventa.

Eickelmann, Birgit/Schaumburg, Heike/Drossel, Kerstin/Lorenz, Ramona (2014): Schulische Nutzung von neuen Technologien in Deutschland im internationalen Vergleich. In: Bos, Wilfried/Eickelmann, Birgit/Gerick, Julia/Goldhammer, Frank/Schaumburg, Heike/Schwippert, Knut/Senkbeil, Martin/Schulz-Zander, Renate/Wendt Heike (Hrsg.): ICILS 2013 – Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann, S. 129–154.

Eickelmann, Birgit/Vennemann, Mario (2017): Teachers' attitudes and beliefs regarding ICT in teaching and learning in European countries. European Educational Research Journal, S. 1–29. doi:10.1177/1474904117725899.

Fraillon, Julian/Ainley, John/Schulz, Wolfram/Duckworth, Daniel/Friedman, Tim (2019): IEA International Computer and Information Literacy Study 2018. Assessment Framework. Amsterdam: IEA.

Fraillon, Julian/Ainley, John/Schulz, Wolfram/Friedman, Tim/Gebhardt, Eveline (2014): Preparing for Life in a Digital Age. The IEA International Computer and Information Literacy Study International Report. Melbourne: Springer.

Gerick, Julia/Vennemann, Mario/Eickelmann, Birgit/Bos, Wilfried/Mews, Sina (2018): ICILS 2013 – Dokumentation der Erhebungsinstrumente der International Computer and Information Literacy Study. Münster: Waxmann.

Gollwitzer, Mario (2012): Latent-Class-Analysis. In: Moosbrugger, Helfried/Kelava, Augustin (Hrsg.): Testtheorie und Fragebogenkonstruktion. Berlin: Springer, S. 295–324.

Heinen, Richard/Kerres, Michael (2017): „Bildung in der digitalen Welt" als Herausforderung für Schule. DDS – Zeitschrift für Erziehungswissenschaft, Bildungspolitik und pädagogische Praxis. Themenschwerpunkt „Bildung in der Digitalen Welt", 2.

Holmberg, Jörgen (2019): Designing for added pedagogical value. A design-based research study of teachers' educational design with ICT. Stockholm: Department of Computer and Systems Sciences.

Jung, Michael/Carstens, Ralph (2015): ICILS 2013 User Guide for the International Database. Amsterdam: IEA.

KMK (2016): Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz. Berlin: Sekretariat der Kultusministerkonferenz.

König, Johannes (2016): Lehrerexpertise und Lehrerkompetenz. In: Rothland, Martin (Hrsg.): Beruf Lehrer/Lehrerin. Ein Studienbuch. Münster: Waxmann, S. 127–148.

Kuha, Jouni (2004): AIC and BIC: Comparisons of Assumptions and Performance. Sociological Methods and Research 33, H. 2, S. 188–229, doi:10.1177/0049124103262065.

Lo, Yungtai/Mendell, Nancy R./Rubin, Donald B. (2001): Testing the number of components in a normal mixture, Biometrika, 88(3), 767–778.

Lopes, António (2018): How Bold are Language Teachers? Comparative Analysis of the Data of a Transatlantic Survey on Technology-Mediated Task-Based Language Teaching. In: Lopes, António/Cecilia, Raúl Ruiz (Hrsg.) New Trends in Foreign Language Teaching: Methods, Evaluation and Innovation. Newcastle upon Tyne: Cambridge Scholars Publishing, S. 82–136.

Lüdtke, Oliver/Robitzsch, Alexander/Trautwein, Ulrich/Köller, Olaf (2007): Umgang mit fehlenden Werten in der psychologischen Forschung: Probleme und Lösungen. Psychologische Rundschau 58, H. 2, S. 103–117, doi:10.1026/0033-3042.58.2.103.

Martin, Michael O. (2005): TIMSS 2003 User Guide for the International Database. Chestnut Hill, MA: Boston College.

Muthen, Bengt O./Satorra, Albert (1995): Complex Sample Data in Structural Equation Modeling. Sociological Methodology 25, S. 267–316, doi:10.2307/271070.

Rutkowski, Leslie/Gonzalez, Eugenio/Joncas, Marc/von Davier, Matthias (2010): International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. Educational Researcher 39, H. 2, S. 142–151.

Tondeur, Jo/Scherer, Ronny/Baran, Evrim/Siddiq, Fazilat/Valtonen, Teemu/Sointu, Erkko (2019): Teacher educators as gatekeepers: Preparing the next generation of teachers for technology integration in education. British Journal of Educational Technology.

Wilde, Annett/Kunter, Mareike (2016): Überzeugungen von Lehrerinnen und Lehrern. In: Rothland, Martin (Hrsg.): Beruf Lehrer/Lehrerin. Ein Studienbuch. Münster: Waxmann, S. 299–316

Wolf, Christof/Best, Henning (2010): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden: VS Verlag für Sozialwissenschaften.

II Kompetenzen von (angehenden) Lehrkräften

Kai S. Cortina

Als Einleitung zu den in diesem Abschnitt zusammengefassten Beiträgen wird zunächst grob skizziert, welchen Herausforderungen Wissenschaftlerinnen und Wissenschaftler begegnen, die Ansätze und Methoden der empirischen Bildungsforschung nutzen, um belastbare Aussagen über die Güte der Lehrkräfteausbildung zu machen bzw. diese zu nutzen, um die Lehrkräfteausbildung zu verbessern. Vor dem Hintergrund, dass ein grundlegendes Spannungsverhältnis besteht zwischen dem Ziel, allgemeingültige Aussagen zu Qualitätsaspekten der Lehrkräfteausbildung zu machen, und der subjektiven Erfahrung von angehenden Lehrkräften über ihren eigenen Lernfortschritt, werden die Beiträge bezüglich ihrer Positionierung in diesem Spannungsfeld diskutiert.

Unterrichtsforschung und Ausbildung von Lehrkräften

Die empirische Unterrichtsforschung und die Aus- und Weiterbildung von Lehrkräften stehen traditionell in einem Spannungsverhältnis, das nicht selten zu Missverständnissen über die Stärken und Schwächen beider Seiten führt: Das primäre Ziel der Unterrichtsforschung ist es gemeinhin, verallgemeinerbare Aussagen über effektive Unterrichtsführung und Lernprozesse im Allgemeinen zu machen, wobei der bei den Lernenden erreichte fachspezifische Lernzuwachs in der überwältigen Mehrheit der Forschungsansätze die kritische Ergebnisvariable ist.

Auch wenn die so ermittelten allgemeinen Aussagen in vieler Hinsicht qualifiziert werden, z. B. nach Schulfach, Schulart, Zusammensetzung der Schülerschaft oder vielleicht sogar nach persönlichen Charakteristika der Lehrkräfte, bleiben es doch Aussagen, die hinsichtlich des effektiven Lernfortschritts der Schülerinnen und Schüler als generalisierbar verstanden werden, d. h. im Prinzip auf alle ähnlich charakterisierten Unterrichtssituationen zutrifft.

Lehrkräfte – zumal solchen mit langer Lehrerfahrung – reagieren sehr unterschiedlich auf dieses dominierende positivistische Wissenschaftsparadigma. Im Extremfall werden die so produzierten Erkenntnisse entweder als gesetzes-

artige Wahrheiten angesehen oder im entgegengesetzten Extrem als Behauptungen praxisferner ForscherInnen rundweg abgelehnt. Umgekehrt gilt, dass erfahrungsgesättigte Einsichten von Lehrkräften mit langjähriger Schulerfahrung von Forschenden entweder als nicht-generalisierbare Einzelerfahrungen belächelt werden oder als Beleg für die Krise der Programmatik in der eigenen Disziplin herhalten muss (vgl. Dammer 2015; Hartmann/Decristan/Klieme 2017). In der durchaus lebhaften Theorie-Praxis-Debatte in den Erziehungswissenschaften um die Rolle der Empirie (z. B. Helsper/Klieme/Terhart 2018; Krapp 2014) finden sich alle vier jeweils gegensätzlichen Positionen oft in abgestufter Kombination und sorgen so, zumal für diejenigen, die in den Beruf einsteigen, für ein eher schillerndes Bild dessen, was eigentlich vom Selbstverständnis her auch dazu angetreten war, jungen Lehrkräften den Weg in die Professionalität mit den Mitteln der Wissenschaft zu vereinfachen.

Den in diesem Abschnitt zusammengestellten Beiträgen ist gemein, dass sich ihr Forschungsschwerpunkt auf die jungen Lehrkräfte selbst richtet, d. h. deren Lernprozesse in den Vordergrund rücken und nicht diejenigen ihrer Schülerinnen und Schüler. Letztere sind nunmehr mittelbar tangiert, insofern man unterstellen muss, dass gelingende Lernprozesse angehender Lernkräfte sich auch günstig auf die Lernentwicklung der Lernenden auswirkt. Was so naheliegend und wenig bemerkenswert erscheinen mag, hat doch für die Auseinandersetzung um die Rolle der Empirie weitreichende Konsequenzen. Zum einen werden die jungen Lehrkräfte nicht von vornherein als Novizen den Experten (= erfahrene Lehrkräfte) als implizit defizitäre Gruppe gegenübergestellt, zum anderen tritt den Einsteigern die empirische Erziehungswissenschaft nicht mehr als Unternehmen objektiver Wahrheitsproduktion entgegen, deren Erkenntnisse man als „Unterrichtskunde“ lediglich absorbiert. Die empirische Forschung ist nun Dialogpartner, der auf Fragen, die im selbstgesteuerten Prozess der Verbesserung der eigenen Unterrichtspraxis aufgeworfen werden, tentative Antworten bereitstellen kann. Die Wissenschaft steht in dieser Rolle gleichberechtigt neben denjenigen Antworten, die eine junge Lehrkraft von einer Kollegin bekommt, die über langjährige Erfahrung verfügt, deren Urteil die Lehrkraft schätzt und deren Ratschlag sie akzeptieren kann, ohne dass dies impliziert, dass kategorial alle Lehrkräfte mit ähnlich langer Lernerfahrung diese Funktion erfüllen können. Unter der Hand wird so der Prozess der Aneignung von Professionalität selbst zum Gegenstand der wissenschaftlichen Betrachtung. Klassischen Dichotomisierungen wie „Laie vs. WissenschaftlerIn“ oder eben „Experte – Novize“ treten dann in den Hintergrund und erhöhen so die Akzeptanz dieser Forschung bei denjenigen, für die sie eigentlich nutzbringend sein sollte: Lehrkräfte in der Frühphase ihrer beruflichen Tätigkeit.

Im Beitrag von Franz und Groß Ophoff (1) wird dieses Paradigma besonders deutlich, weil hier die selbsteingeschätzten Kompetenzzuwächse junger Lehrkräfte in den Blick genommen werden. Dies ist insofern ein überzeugender Ansatz, als man wohl unterstellen darf, dass angehende Lehrkräfte sehr sensibel für die eigenen professionellen Herausforderungen sind und daher auch valide ihre positiven Entwicklungen wahrnehmen und somit auch selbst messen können. Es wird deutlich, wie differenziert der Effekt der schulpraktischen Studien wahrgenommen wird, und man kann darüber spekulieren, ob bestimmte strukturelle Änderungen in den schulpraktischen Studien die Lernwahrnehmung der Studierenden auch in anderen für wichtig erachteten Aspekten jenseits der konkreten Unterrichtsplanung steigern könnte.

Einen ebenfalls erfrischend anderen Bezugspunkt hat die FALKE Studie von Lindl, Gaier, Weich u. a. (2), die als *tertium comparationis* die Sprechakttheorie aus der Linguistik heranziehen, um die Qualität des Erklärens „sprechaktobjektiv" als eine fächerübergreifende Kompetenz berufsgruppenübergreifend zu analysieren. Hier darf man auf die zukünftigen Ergebnisse gespannt sein, gerade weil es ein vieldokumentiertes Phänomen in der empirischen Bildungsforschung betrifft, dass nämlich Einschätzungen der unterschiedlichen Berufsgruppen (hier: Schülerinnen, Studierende, Lehrkräfte, Fachdidaktiker) nur geringe Übereinstimmung aufweisen (Clausen 2002). Während dies traditionell so interpretiert wird, dass die unterschiedlichen Perspektiven unterschiedliche Facetten nur bedingt überlappender Konstrukte messen, wäre eine solche Interpretation durch die sprechaktanalytische Grundlegung nicht naheliegend. Der besondere Reiz der Studie liegt darin, dass subjektive Qualitätseinschätzungen vor der Folie einer universellen Theorie vorgenommen werden.

Auch Lorenz, Endberg und Eickelmann (3) setzen in ihrer eher deskriptiv angelegten Studie auf die Selbstreflexion der Lehrkräfte. In einer subjektiven Ist-Soll Analyse können die Autorinnen zeigen, dass zwar der wahrgenommene Anteil mediennutzungsrelevanter Anteile am Studium bei jüngeren Lehrkräften erwartbar zugenommen hat, diese dennoch deutliche Defizite in den ersten zwei Phasen der Ausbildung sehen. Dies wird dadurch unterstrichen, dass Defizite vor allem von solchen Lehrkräften beklagt werden, die moderne Medien tatsächlich im Unterricht einsetzen. Es ist bezeichnend, dass die Lehrkräfte die Ausbildung vor allem in der Referendariatsphase wünschen und deutlich weniger in der universitären Ausbildung sehen wollen. Das bestätigt die grundsätzliche Skepsis junger Lehrkräfte, sich Handlungskompetenzen als Lehrkraft im akademischen Setting aneignen zu wollen. Der Umgang mit Medien im Unterricht wird nicht als medienwissenschaftliches – also „kundliches" – Problem verstanden, das an der Universität gelehrt werden sollte, sondern primär als ein Erfahrungswissen, dass man sich am besten im direkten Unterrichtsvollzug im

Dialog mit kompetenten PartnerInnen für sich selbst „maßgeschneidert“ erwirbt. Im Unterschied zu vielen Aspekten der Fachdidaktik ist die Halbwertszeit von Wissen über und Einsatz von Medien im Unterricht besonders kurz.

Objektivierung des Subjekten

Der Beitrag von Prilop, Weber und Kleinknecht (4) stellt in der hier eingeführten Systematik einen Hybridansatz dar: Zwar geht es hier um das Kollegiale, nämlich das Feedback, das Lehrkräfte von Kolleginnen und Kollegen für ihre professionelle Weiterentwicklung bekommen, und somit nicht um die objektive Bestimmung richtigen oder falschen LehrerInnenhandels per se. Das Feedback wird aber dennoch klassisch als LehrerInnenkompetenz konzipiert und somit zugänglich zur Entwicklung psychometrisch validierbarer Messinstrumente. Der Beitrag ist insofern bemerkenswert, als es das Argument teilweise entkräftet, wonach bestimmte Aspekte einer „guten“ Lehrperson nicht messbar gemacht werden könnten und deshalb in der Debatte um Lehrkompetenz vernachlässigt werden. Es wäre zu wünschen, dass die Validität des Messinstruments auch jenseits von Expertenurteilen gezeigt werden kann. Wie in vielen anderen Bereichen, wo Kompetenzen konzeptionalisiert und mit geeigneten Methoden gemessen werden können, stellt sich gerade bei Aspekten der Lehrkraftkompetenzen unmittelbar die Frage, wie diese erworben werden und, damit zusammenhängend, wie man angehende Lehrkräfte effektiv dazu verhilft, professionelles Niveau zu erreichen, hier also konkret, mit welchem kollegialen Feedback über das kollegiale Feedback man Rückmelderoutinen für den Schulalltag optimiert. Hier finden sich im Beitrag durchaus wichtige Hinweise.

Rubach, Lazarides, Brendel und Krauskopf (5) nutzen in ihrer Studie ein experimentelles Design, das in der Forschung zu Lehrkraftausbildung häufiger zum Zuge kommen sollte, weil bei gut begründeter Wahl der Experimentalbedingungen direkte praktische Konsequenzen ableitbar sind. In ihrer Studie nehmen sie das Thema von Lorenz, Endberg und Eickelmann (3) wieder auf, nämlich die Frage, ob und wie jungen Lehrkräften Digitalisierung/Mediennutzung bereits an der Universität vermittelt werden sollte. Sie können zunächst zeigen, dass die selbsteingeschätzte Medienkompetenz in der Kontrollgruppe, die kein spezifisches Training erhalten hatte, am geringsten war. Die Unterschiede der didaktischen Herangehensweise in den beiden Interventionsgruppen machen – vielleicht überraschend – deutlich, dass es nicht immer optimal ist, die Lernposition der Schülerinnen und Schüler einzunehmen. Gerade bei der Mediennutzung ist man aus dieser Warte eher geneigt, den unmittelbar motivierenden Aspekt für die Lernenden im Unterricht zu sehen, der aber durch Routinisierung des Medieneinsatzes oft schnell verfliegt. Aus einer fachdidaktischen Perspektive (Lehrkraftperspektive) wird man notwendiger-

weise skeptischer und in Hinblick auf die Nutzung von Medien selektiver, weil im Hintergrund das pädagogische Gesamtkonzept einer Unterrichtseinheit steht. Werden digitale Medien bereits in der Konzeption von Unterrichtseinheiten integriert, zeigen sich stärkere Zuwächse in der selbsteingeschätzten Kompetenz in Hinblick auf die unterrichtliche Nutzung digitaler Medien im Unterricht. Dies bestätigt meine persönliche Erfahrung, wonach angehende Lehrkräfte, wenn sie digitale Medien in eine bestehende Unterrichtseinheit einbauen, diese oft „verschlimmbessern": Zwar sind die Schüler hoch motiviert, weil sie z. B. ihre Handys nutzen dürfen, doch die in der Regel als Spiele konzipierten Programme (z. B. Kahoot, Menti) bringen schnell unter der Hand kompetitive Elemente in den Unterrichtsfluss, die nicht selten didaktisch kontraproduktiv sind. Lehrkräfte geraten leicht in ein Fahrwasser, wo es nur noch um die richtigen Antworten geht, obwohl es pädagogisch wichtiger wäre, den falschen Antworten nachzugehen, um Missverständnisse produktiv aufzuarbeiten.

Aufbauend auf der psychometrisch soliden Entwicklung des DaZKom-Testinstruments (paper-pencil-Format), das die generelle Kompetenz misst, Lernende mit Deutsch-als-Zweitsprache zu unterrichten, gehen Schroedler und Stangen (6) der Frage nach, inwieweit diese Kompetenz bei angehenden Lehrkräften mit dem Nutzen von Lerngelegenheiten in der universitäre Ausbildung zusammenhängen. Es zeigt sich, dass die Nutzung thematisch gezielt angebotener Lehrveranstaltungen auch im Längsschnitt mit dem Kompetenzzuwachs korreliert. Trotz – oder gerade weil – diese Befunde spezifisch für die Situation der Lehrkräfteausbildung an der Uni Hamburg im Rahmen der ProfaLe (Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen) Initiative sind, lassen sich doch optimistische Schlussfolgerungen ziehen: Das Bemühen, DaZ zum übercurricularen Gegenstand der universitären Ausbildung zu machen, schlägt positiv im Kompetenzzuwachs zu Buche und ermutigt, solche Veranstaltungen ähnlich wie die des Hamburger Modellversuchs auch an anderen Standorten anzubieten. Dass die Befundlage in der Unterscheidung zwischen thematischen und handlungsorientierten Lerngelegenheiten früheren Studien widerspricht, mag irritierend sein und Anlass zu weiterer Forschung geben, ist aber in Hinblick auf die praktischen Schlussfolgerungen kaum mehr als ein Wermutstropfen. Entscheidend ist, dass in diesem Zusammenhang die universitäre Ausbildung nachweislich positive Effekte hat und das Thema folglich im Kanon der Ausbildungselemente für die Lehrkräfteausbildung einen festen Platz haben sollte.

Befunde aus Studien mit dem DaZKom-Test (wie u.a. bei Schroedler und Stangen) sowie Forschung zur Professionalisierung von Lehrkräften geben Anlass dazu, die Kompetenzmessung im Sinne der Performanzorientierung weiter zu denken. Lemmrich, Hecker, Klein, Ehmke, Köker, Koch-Priewe und Ohm (7) nutzen in ihrer Studie ein videobasiertes Testinstrument, um handlungsnahe Kompetenz von Lehrkräften im Bereich DaZ zu erfassen. Ziel ist es,

die Lehrkräfteaus- und Weiterbildung den Bedarfen im Hinblick auf sprachlich sensiblen Unterricht anzupassen. Die Studie präsentiert erste Ergebnisse aus dem Testentwicklungsprozess.

Ausblick

Die Akzeptanz der empirischen Schul- und Unterrichtsforschung hängt nicht nur davon ab, wie „angewandt" oder „nützlich" sie für angehende Lehrkräfte in der Ausbildungsphase ist, sondern auch davon, wieweit sie bereit ist, sich epistemologisch breiter aufzustellen und positivistische Positionen zu relativieren. Wie die Beiträge in diesem Abschnitt zeigen, sind die methodischen Herangehensweisen der empirischen Bildungsforschung erkenntnisbringend einsetzbar, auch wenn es primär um die Selbstreflexion von Lehrkräften geht (1 und 3), Linguistik als Tertiärwissenschaft hinzugezogen wird (2) oder experimentelle Ansätze mit Kompetenzselbsteinschätzung kombiniert werden (5). Es geht darum, Wirksamkeit von Maßnahmen zu dokumentieren und zu verstehen (6 und 7), die dazu anregen können, die Lehrkräfteausbildung zu verbessern – nicht durch wissenschaftliche Präskription, sondern im durch wissenschaftliche Erkenntnisse bereicherten Dialog von Lehrkräften, Ausbildenden und Forschenden.

Literatur

Clausen, Marten (2002): Unterrichtsqualität: Eine Frage der Perspektive? Münster: Waxmann.

Dammer, K.-H. (2015): Vermessene Bildungsforschung. Wissenschaftsgeschichtliche Hintergründe zu einem neoliberalen Herrschaftsinstrument. Baltmannsweiler: Schneider Hohengehren.

Hartmann, U., Decristan, S. & Klieme, E. (2017): Unterricht als Feld evidenzbasierter Bildungspraxis? Herausforderungen und Potenziale für einen wechselseitigen Austausch von Wissenschaft und Schulpraxis. In: Baumert, J. & Tillmann, K.-J., Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker, 179–199. Berlin: Springer.

Helsper, W., Klieme, E., Terhart, E. (2018): Perspektiven zur Entwicklung der empirischen Bildungsforschung. In: Empirische Bildungsforschung im Kontext von Schule und Lehrer*innenbildung, herausgegeben von Krüger, A., Radisch, F., Willems, A. S., Häcker, T., Walm, M., 21–50. Bad Heilbrunn: Klinkhardt.

Krapp, A. (2014): Geschichte der Pädagogischen Psychologie. In T. Seidel & A. Krapp (Hrsg.), Pädagogische Psychologie (S. 347–353). Weinheim: Beltz.

Kompetenzerleben Lehramtsstudierender im Semesterpraktikum

Eva-Kristina Franz und Jana Groß Ophoff

Abstract

Die vorliegende Studie fokussiert, in welchen Bereichen Studierende in schulpraktischen Ausbildungsphasen Kompetenzzuwächse erleben und welchen Einfluss Berufswahlmotive auf dieses Erleben haben. Zur Erfassung der selbsteingeschätzten Kompetenzen wurde ein Fragebogen entwickelt, mit welchem im Sommersemester 2018 an zwei verschiedenen Hochschulen 117 Studierende befragt wurden. Im Ergebnis können für das Kompetenzerleben fünf Faktoren unterschieden werden: *Unterricht*, *Lernprozessförderung*, *Diagnostische Kompetenz*, *Reflexion des Unterrichts und der Lehrerrolle* sowie *Wertbindung*. Die Studierenden erleben in ihrem Semesterpraktikum v.a. in den unterrichts- und leistungsbeurteilungsbezogenen Kompetenzbereichen signifikante Kompetenzzuwächse. Erste Analysen der Daten deuten zudem auf Zusammenhänge zwischen der selbsteingeschätzten Kompetenz mit verschiedenen Studienwahlmotiven hin.

1 Problemstellung

Die universitäre Lehrerbildung hat das Ziel, Studierende möglichst optimal auf deren zukünftige berufliche Tätigkeit vorzubereiten. Dabei kann das Unterrichten als zentrale Aufgabe angenommen werden (z. B. Rutsch/Rehm/Seidenfuß/Dörfler 2018). Um dieser gerecht zu werden, benötigen Studierende ein gewisses Maß an professioneller Kompetenz. Dieser entspricht nach Baumert und Kunter (2006) nicht ausschließlich Wissen. Auch Überzeugungen, Werthaltungen, selbstregulative Fähigkeiten und motivationale Orientierungen tragen konstituierend zur professionellen Kompetenz bei. Schulpraktische Studien ermöglichen Studierenden, pädagogische und didaktische Situationen im institutionellen Rahmen der Schule zu erleben und mitzugestalten. Sie werden als eine der formalen Lerngelegenheiten angesehen, mittels derer sich Professionelle Kompetenz entwickeln lässt (Voss/Kunina-Habenicht/Hoehne/Kunter

2015). Daher kommt ihnen in Modellen der Lehrerprofessionalität eine wichtige Bedeutung zu und ihre Wirksamkeit, meist operationalisiert über den Kompetenzzuwachs im Langzeitpraktikum, wird vermehrt analysiert (z. B. Gröschner/Hascher 2018). Alles in Allem spricht die Befundlage dafür, dass Studierende Kompetenzzuwächse während des Praktikums erleben, wenngleich dies nicht automatisch zu geschehen scheint (Gröschner/Schmitt 2010), sondern insbesondere durch die individuelle Nutzung von Lerngelegenheiten sowie durch individuelle Merkmale beeinflusst wird (König/Rothland 2018). Die im Folgenden dargestellte Studie nutzt ein Selbstreportverfahren zu Beginn und zu Ende des Integrierten Semesterpraktikums an Schulen, um herauszuarbeiten, in welchen Bereichen Studierende im Langzeitpraktikum Kompetenzzuwächse erleben sowie welchen Effekt motivationale Orientierungen wie Berufswahlmotive auf die wahrgenommene Kompetenzentwicklung haben.

2 Studierende in Praxisphasen – ein Angebots-Nutzungs-Modell

Die Integration schulpraktischer Studien bereits in die universitäre Phase der Lehrerbildung zielt darauf ab, verschiedene Wissensbereiche enger zu verzahnen (Verloop/van Driel/Meijer 2001; Rutsch et al. 2018). Handlungsleitend ist dabei die modellhafte Vorstellung, dass Studierende professionelle Kompetenzen entwickeln, indem sie einerseits Lerngelegenheiten an der Hochschule für sich nutzen (Kunter et al. 2011) und darüber hinaus eigene Erfahrungen im Unterrichten im Kontext schulpraktischer Studien machen (Korthagen 2010). In dem Wissen, dass Praxisphasen komplexe, vielschichtige Lerngelegenheiten darstellen, werden diese in den vergangenen Jahren immer weniger als „einheitliches Treatment“ betrachtet (Bock/Hany/Protzel 2017, S. 7), sondern es werden einzelne Komponenten schulpraktischer Studien getrennt untersucht (z. B. Cramer/Horn/Schweitzer 2009). Dabei setzt sich nach und nach ein Angebots-Nutzungs-Modell durch, welches Einflussfaktoren und Wirkgrößen auf das Praktikum an Schulen modelliert (Hascher/Kittinger 2014). In diesem werden neben der Qualität der Lehr-Lern-Prozesse im Praktikum an sich, u. a. auch die professionelle Kompetenz der Lehrenden an Schule und Universität, die individuellen Voraussetzungen und sozialen Hintergründe der Studierenden als Faktoren expliziert, welche den Nutzen des Angebots *schulpraktisches Studium* beeinflussen können.

Die im vorliegenden Beitrag fokussierten Berufswahlmotive können neben Fragen der *teacher beliefs* oder der berufsbezogenen Selbstwirksamkeitshypothese Prädiktoren für die Qualität der Nutzung universitärer und schulpraktischer Lernangebote darstellen (Rauin 2007).

3 Studien- und Berufswahlmotive

Die Motive, mit der Studierende ihr Studium aufnehmen, gelten als wesentliche Voraussetzung, dieses auch erfolgreich zu beenden (Rauin 2007; König/Rothland 2013; Rothland 2014). Insbesondere intrinsischen Motiven kommt eine hohe Bedeutung für adaptives und funktionales Verhalten in Lern- und Arbeitskontexten zu (Eccles/Wigfield 2002): entsprechend intrinsisch motivierte Lehramtsstudierende nutzen schulpraktische Lerngelegenheiten intensiver und zeigen ausgeprägtere Veränderungen in der Kompetenzselbsteinschätzung (Cramer 2016). Unter intrinsischen Motiven werden nicht nur das fachbezogene Interesse, sondern insbesondere auch das Interesse an der abwechslungsreichen und interessanten Arbeit mit Kindern und Jugendlichen, der Vermittlung von Wissen oder der Möglichkeit, einen wichtigen gesellschaftlichen Beitrag zu leisten, rubriziert. Zu den extrinsischen Berufswahlmotiven werden die Vereinbarkeit von Familie und Beruf, ein hoher Freizeitanteil, die Arbeitsplatzsicherheit und ein verhältnismäßig hohes Gehalt gezählt (Eberle/Pollack 2006).

Der aktuelle Forschungsstand legt die Vermutung nahe, dass eine intrinsische Motivlage zu einer intensiveren Nutzung der sich in den schulpraktischen Studien ergebenden Lerngelegenheiten führt (Hascher/Kittinger 2014). Besonders dem Interesse an der Arbeit mit Kindern kommt in diesem Kontext eine hohe Bedeutung zu. Erste Ergebnisse einer Studie von Biermann et al. (2018) stützen die Hypothese, dass auch die selbsteingeschätzte unterrichtliche Kompetenz im Schulpraktikum mit intrinsischen Motivationsaspekten zusammenhängt. Ob die Berufswahlmotive auch auf weitere Kompetenzbereiche einen Einfluss haben, stellt bis dato ein Forschungsdesiderat dar.

4 Fragestellung

In diesem Beitrag wird daher den Fragen nachgegangen, in welchen Bereichen (1) Studierende im Langzeitpraktikum Kompetenzzuwächse erleben und welchen Effekt Berufswahlmotive (2) auf die wahrgenommene Kompetenzentwicklung haben. Dabei ist davon auszugehen, dass sich intrinsische Motive, insbesondere das *Pädagogische Interesse*, positiv auf das Kompetenzerleben auswirken (Rauin 2007; Franz, in Druck), wohingegen extrinsische Berufswahlmotive sich tendenziell eher negativ auf das Kompetenzerleben auswirken könnten (ebd.).

5 Methodisches Vorgehen

Die Ergebnisse stützen sich auf Fragebogendaten von rund 400 Lehramtsstudierenden zwischen dem Sommersemester 2014 und 2018 an zwei baden-württembergischen Pädagogischen Hochschulen. Beide Hochschulen bilden im Schwerpunkt Studierende für das Lehramt Primarstufe und Sekundarstufe 1 aus. Die Stichprobe besteht zu ungefähr gleichen Teilen aus den beiden Studierendengruppen. Diese wurden in bildungswissenschaftlichen Begleitseminaren zu Beginn und gegen Ende des Praktikums darum gebeten, ihre eigenen unterrichtsbezogenen Kompetenzen einzuschätzen.

Ab dem Sommersemester 2018 wurden zusätzlich Berufswahlmotive erfasst (Pohlmann/Möller 2010). Die in dieser Kohorte befragten Studierenden sind ausnahmslos Studierende der Grundschulpädagogik.

Generell war die Teilnahme an der Befragung freiwillig, die Ergebnisse wurden während bzw. direkt im Anschluss an die Seminare an die Studierenden zurückgemeldet. Um die Datensätze von beiden Erhebungszeitunkten zuordnen zu können, zugleich aber eine anonymisierte Auswertung zu gewährleisten, wurde ein Personencode vergeben.

Instrumente

Auf Basis der konzeptionell als Qualifikationsziele für schulpraktische Studien fokussierten Kompetenzbereiche, *Unterrichten*, *Beurteilen*, *Erziehen* und *Innovieren* wurde ein Fragebogen zur Selbsteinschätzung der Kompetenzen entwickelt. Die Studierenden hatten darin die Möglichkeit, das Niveau ihrer Kompetenzen auf einer vierstufigen Likert-Skala einzuschätzen (4 = auf hohem Niveau, 3 = auf mittlerem Niveau, 2 = auf niedrigem Niveau, aber mit Potenzial, 1 = in unzureichender Weise).

Explorative Strukturgleichungsmodelle ergaben hinsichtlich der Dimensionalität der selbsteingeschätzten Kompetenzen fünf inhaltlich getrennte Dimensionen ($\chi 2/df = 455.782/302 = 1.51$; RMSEA = .045; CFI = .952; vgl. dazu Groß Ophoff/Franz/Immel, in Druck): *Unterrichtsplanung und -gestaltung* (F1), *Lernprozessförderung* (F2), *Diagnostische Kompetenz* (F3), *Reflexion von Unterricht und Lehrerrolle* (F4) und *Wertbindung* (F5). Eine Übersicht beispielhafter Items findet sich in Tabelle 1.

F1 bündelt Items, die sich auf die Planung und Gestaltung von Unterricht (vgl. Tabelle 1) beziehen. F2 versammelt Items, die sich auf die Fähigkeit zur Lernprozessförderung als Unterrichtselement beziehen, womit auch eine konstruktive Nutzung von Leistungsrückmeldungen in Verbindung gebracht wird. F3 lässt sich als Einschätzung der eigenen Diagnostischen Kompetenz beschreiben. F4 umfasst wiederum verschiedene Items, welche Fähigkeiten zur Reflexion des Unterrichts und der Lehrerrolle thematisieren und F5 wird schließlich durch Items repräsentiert, welche weniger Wissen und Können umfassen, sondern eine Wertbindung (sog. value commitment, vgl. Baumert/Kunter 2006) zum Ausdruck bringen (vgl. dazu Groß Ophoff/Franz/Immel, in Druck).

Neben deskriptiven Informationen (z. B. Alter, Geschlecht, Anzahl Fachsemester) wurden zum Ende des Praktikums (t2) die Studienwahlmotive anhand des Fragebogens zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA; Pohlmann/Möller 2010) erfasst. Das Instrument besteht aus 33 Items. Darin werden die Studierenden gebeten, für verschiedene Studienwahlmotive anzugeben, wie sehr dies für sie zutreffen (1 = „trifft gar nicht zu“ bis 5 = „trifft völlig zu“). Pohlmann und Möller (2010) unterschieden dabei die Subskalen (1) *Nützlichkeit*, (2) *Pädagogisches Interesse*, (3) *Fähigkeitsüberzeugung*, (4) *Soziale Einflüsse*, (5) *Geringe Schwierigkeit* und (6) *Fachliches Interesse*, wobei das *Pädagogische Interesse*, das *Fachliche Interesse* und die *Fähigkeitsüberzeugung* als intrinsische Dimensionen, die *Nützlichkeit*, die *Sozialen Einflüsse* und die *Geringe Schwierigkeit* als extrinsische Dimensionen aufgefasst werden.

Datenerhebung und Stichprobe

Im Sommersemester 2018 konnten insgesamt 117 Studierende (PH Freiburg: 59 %; PH Heidelberg: 41 %) in neun Begleitseminaren befragt werden[1]. An beiden Standorten war die Mehrheit der Studierenden weiblichen Geschlechts. Außerdem waren die Studierenden während des ISP im Durchschnitt 23 Jahre alt und befanden sich – wie curricular zu erwarten – mehrheitlich im fünften Semester. In beiden Merkmalen unterscheiden sich die Studierenden nicht signifikant ($p < .05$, zweiseitige Testung) zwischen den beiden Standorten.

1 Wir danken an dieser Stelle den Kolleginnen und Kollegen der Pädagogischen Hochschulen Freiburg und Heidelberg für die Bereitschaft, uns in ihre Seminargruppen einzuladen, sowie Caroline Immel für die sorgfältige Dateneingabe.

Statistische Analysen

Zur Berechnung der vorgestellten deskriptiven und inferenzstatistischen Analysen wurden die Analysesoftware SPSS 23 (2015) und Mplus (Version 6, Muthén/Muthén 2010) eingesetzt.

Für die Ermittlung der Skalenkennwerte für beide Instrumente wurden jeweils der Durchschnitt aus den Items gebildet, die zu dem jeweiligen Faktor als zugehörig identifiziert wurden. Für den Vergleich zwischen den Kompetenzeinschätzungen zu t1 und t2 wurden auf Basis der Skalenmittelwerte t-Tests für abhängige Stichproben berechnet. Hierbei wurden fehlende Werte paarweise ausgeschlossen. Um die Bedeutsamkeit der aufgefundenen Unterschiede beurteilen zu können, wird ergänzend die Effektstärke d berichtet, die gemäß Dunlap, Cortina, Vaslow und Burke (1996; vgl. Borenstein 2009) ermittelt wurde. Dabei kann ein $d \geq .80$ als großer Effekt, ein $d \geq .50$ als mittlerer Effekt und ein $d \geq .20$ als kleiner Effekt gewertet werden (Cohen 1992).

Angesichts der kleinen Stichprobe (und damit verbundener Schätzprobleme bei latenten Pfadanalysen, z. B. Schönbrodt/Perugini 2013) wurde anhand der Skalenmittelwerte ein manifestes Pfadmodell spezifiziert, in dem die Differenz zwischen den Kompetenzselbsteinschätzungen zu Beginn und am Ende des ISP (t2-t1) durch Berufswahlmotive vorhergesagt wird. Aus diesen Analysen wurden Fälle (N = 19) ausgeschlossen, die an einem der beiden Messzeitpunkte gefehlt haben. Weitere einzelne fehlende Werte wurden beibehalten, da das in der Analyse-Software Mplus implementierte Schätzverfahren (Full-Information-Maximum-Likelihood, vgl. Muthén/Kaplan/Hollis 1987) selbst bei Verletzung von Verteilungsvoraussetzungen mit fehlenden Werten umgehen kann.

6 Ergebnisse

Tabelle 1 und Tabelle 2 zeigen, dass die Studierenden in allen Kompetenzbereichen statistisch signifikante Kompetenzzuwächse erleben. Außerdem weisen die Skalen sowohl zu t1 als auch zu t2 mindestens akzeptable Reliabilitäten auf. Die Zuwächse lassen sich für die Skala *Unterrichtsplanung und -gestaltung* als großen Effekt (Tabelle 3) und für die Skalen *Lernprozessförderung* und *Diagnostische Kompetenz* als mittlere Effekte einstufen. Für die Skalen *Reflexion von Unterricht und Lehrerrolle* sowie *Wertbindung* erweisen sich dagegen die Veränderungen nur als klein.

Tabelle 1: Ergebnisse des Fragebogens zur Selbsteinschätzung der Kompetenz im Integrierten Semesterpraktikum

		t1		t2	
Subskalen	Beispielitem	*M (SD)*	*α*	*M (SD)*	*α*
F1: Unterrichts-planung und -gestaltung	Ich kann Unterricht zielgerichtet und systematisch planen, strukturieren und organisieren.	2.79 (0.49)	.77	3.27 (0.38)	.76
F2: Lernprozess-förderung	Ich fördere handlungsorientiertes und problemlösendes Lernen.	2.62 (0.59)	.78	2.93 (0.43)	.70
F3: Diagnostische Kompetenz	Ich kann unterschiedliche Lernvoraussetzungen und Lernschwierigkeiten erkennen.	2.65 (0.52)	.73	2.97 (0.43)	.80
F4: Reflexion von Unterricht und Lehrer-rolle	… kann unterrichtliche Interaktionen analysieren und reflektieren.	3.35 (0.48)	.68	3.57 (0.36)	.70
F5: Wertbindung	… zeige Engagement und Motivation für den Beruf.	3.60 (0.36)	.68	3.70 (0.28)	.69

Anmerkung: Die Berechnung der Mittelwerte und Standardabweichungen beruht auf der Stichprobe im Sommersemester 2018 (N_{min} = 101). Die Skalen-Reliabilitäten wurden auf Basis der Stichproben zwischen Sommersemester 2014 und Sommersemester 2018 ermittelt (N_{min} = 114, vgl. Groß Ophoff, Franz und Immel, in Druck)

Tabelle 2: Ergebnisse der t-Tests für abhängige Stichproben

	Differenz (t_2-t_1): M (SD)	r (t_1t_2)	t	df	d
F1: Unterrichtsplanung und -gestaltung	**0.50 (0.41)**	**.61**	12.196	97	1.09
F2: Lernprozessförderung	**0.35 (0.58)**	**.36**	5.982	94	0.69
F3: Diagnostische Kompetenz	**0.33 (0.50)**	**.50**	6.35	95	0.65
F4: Reflexion von Unterricht und Lehrerrolle	**0.20 (0.41)**	**.53**	4.788	95	0.47
F5: Wertbindung	**0.09 (0.37)**	**.36**	2.342	95	0.27

Anmerkungen: Abkürzungen: M = Mittelwert; SD = Standardabweichung; r = Korrelation zwischen Einschätzung zu t1 und t2; df = Freiheitsgrade; d = Effektstärke. Signifikante Ergebnisse ($p < .05$, zweiseitige Testung) sind fett hervorgehoben. N_{min} = 95

Tabelle 3 zeigt die Ergebnisse für das FEMOLA-Instrument (Pohlmann und Möller 2010) zum Ende des Semesterpraktikums. Hier weisen alle Subskalen gute bis sehr gute Reliabilitäten auf. Auffallend sind der sehr hohe Mittelwert und die niedrige Streuung im Teilbereich *Pädagogisches Interesse* und der geringe Mittelwert im Bereich *Geringe Schwierigkeit*.

Tabelle 3: Ergebnisse des Fragebogens zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA; Pohlmann und Möller 2010)

Subskalen	Beispielitem	M (SD)	α
M1: Nützlichkeit	Ich habe das Lehramtsstudium gewählt, weil ich als Lehrer/in durch die flexible Arbeitszeit soziale Kontakte pflegen kann. (1.1)	2.85 (0.83)	0.86
M2: Pädagogisches Interesse	… es für mich wichtig ist, einen Beitrag zur Ausbildung von Kindern und Jugendlichen zu leisten. (2.6)	4.59 (0.41)	0.75
M3: Fähigkeitsüberzeugung	… ich fachliche Inhalte interessant vermitteln kann. (3.2)	3.90 (0.59)	0.70
M4: Soziale Einflüsse	… mir von Freunden und Bekannten zum Lehramtsstudium geraten wurde. (4.5)	2.40 (0.96)	0.81
M5: Geringe Schwierigkeit	… das Studium nicht so anstrengend ist. (5.2)	1.55 (0.68)	0.89
M6: Fachliches Interesse	… ich mich gern mit den Inhalten meiner Fächer beschäftige. (6.3)	3.70 (0.70)	0.80

Anmerkung: $N_{min} = 100$

Tabelle 4: Ergebnisse für das Pfadmodell mit den Studienwahlmotiven (FEMOLA) als Prädiktoren der Veränderungen in der selbsteingeschätzten Kompetenzen (t2-t1) während des Integrierten Semesterpraktikums

	F1: Unterrichtsplanung und -gestaltung	F2: Lernprozessförderung	F3: Diagnostische Kompetenz	F4: Reflexion Unterricht/ Lehrerrolle	F5: Wertbindung
M1: Nützlichkeit	0.06	**0.29**	**0.25**	-0.01	-0.11
M2: Pädagogisches Interesse	0.10	0.10	**0.25**	**-0.24**	-0.15
M3: Fähigkeitsüberzeugung	-0.06	-0.16	0.02	0.16	0.01
M4: Soziale Einflüsse	0.00	-0.11	-0.00	0.15	**0.30**
M5: Geringe Schwierigkeit	-0.06	-0.02	-0.11	-0.01	-0.05
M6: Fachliches Interesse	-0.07	**0.26**	-0.10	-0.11	-0.01

Anmerkung: Signifikante Regressionskoeffizienten ($p < .05$, zweiseitige Testung) sind fett hervorgehoben

In Tabelle 4 sind die Ergebnisse für das Pfadmodell mit den Studienwahlmotiven als Prädiktoren der Veränderungen in den Kompetenzselbsteinschätzungen (Differenzwert t2-t1) wiedergegeben. Während demnach der Faktor *Unterrichtsplanung und -gestaltung* durch keine der berücksichtigten Studienwahlmotive vorhergesagt werden kann, erweisen sich für Veränderungen in Kompetenzen in der *Lernprozessförderung* nicht nur das *fachliche Interesse*, sondern auch die *Nützlichkeit* – ein extrinsisches Motiv – als signifikante Prädiktoren. Auch für die Veränderungen in der *Diagnostischen Kompetenz* erweist sich letzterer Faktor als bedeutsam. Zusätzlich zeigt sich hierauf ein positiver Effekt des *Pädagogischen Interesses*. Schließlich erweist sich das Studienwahlmotiv *Soziale Einflüsse* als positiver Prädiktor der *Wertbindung*.

7 Diskussion

Die Vergleiche zwischen den beiden Erhebungszeitpunkten verdeutlichen, dass Studierende das Praxissemester als förderlichen Lernraum erleben und Kompetenzzuwächse wahrnehmen. Dabei wird der eigene Lernzuwachs nach Angaben der Studierenden im Bereich der *Unterrichtsplanung und -gestaltung* als besonders deutlich erlebt. Erst mit einigem Abstand folgen weitere Bereiche wie die Fähigkeit zur *Lernprozessförderung* der *diagnostischen Kompetenz*.

Anders als bei Rauin (2007) zeigen die hier untersuchten Studierenden eine ausgeprägte intrinsische Studienwahlmotivation: der Wert für das *Pädagogische Interesse* als ein bedeutendes intrinsisches Berufswahlmotiv liegt im Mittel bei 4.59; der geringe Mittelwert von 1.55 für die *Geringe Schwierigkeit des Lehramtsstudiums* als extrinsisches Berufswahlmotiv stützt diese Annahme (Rauin 2007). Der Deckeneffekt im *Pädagogischen Interesse* könnte auf der ausschließlich aus Grundschulpädagog/innen bestehenden Kohorte in dieser Studie basieren. Für Studierende im Lehramt Sekundarstufe wäre dagegen ein deutlich geringerer Mittelwert zu erwarten (vgl. Franz, in Druck). Um eine höhere Generalisierbarkeit zu erreichen, sollten bei weiteren Erhebungen möglichst alle Schulformen vertreten sein. In diesem Zusammenhang wäre außerdem zu klären (und längsschnittlich zu untersuchen), inwiefern die Berufswahlmotivation ein stabiles Merkmal ist bzw. sich während des Semesterpraktikums verändert.

Angesichts des Deckeneffekts im *Pädagogischen Interesse* sowie der vergleichsweise geringen Streuung können auch Regressionseffekte nicht ausgeschlossen werden. Die signifikanten Pfade auf die selbsteingeschätzten Veränderungen in der *Diagnostischen Kompetenz* und der *Reflexionskompetenz* sollten daher zurückhaltend interpretiert werden. Auch die eher niedrigen Reliabilitätswerte für die Faktoren *Reflexion* und *Wertbindung* in Verbindung mit der vergleichsweise geringen Stichprobengröße in der vorgestellten Studie sprechen

für eine vorsichtige Interpretation dieser Pfadkoeffizienten (vgl. Kretzschmar/Gignac 2019; Schönbrodt/Perugini).

Neben der Untersuchung von deutlich größeren Stichproben bietet es sich ferner an ergänzend weitere Merkmale zu erheben wie z. B. motivationale Zielorientierungen, die Dickhäuser, Butler & Tönjes (2007) in Zusammenhang stellen mit der der Einstellungen gegenüber Hilfesuche und darüber vermittelt mit beruflichem Belastungserleben (z. B. Gröschner/Schmitt 2008). Im umgekehrten Sinne scheint in diesem Zusammenhang für die Studienzufriedenheit insbesondere die Nutzung der verfügbaren Lerngelegenheiten von (vermittelnder) Bedeutung zu sein (Biermann et al. 2018).

Mit Blick auf die Ergebnisse der Pfadanalysen erscheint es auf den ersten Blick zunächst unstimmig, dass eine ausgeprägtere extrinsische Studienwahlmotivation (*Soziale Einflüsse*) das Erleben von Kompetenzzuwächsen in der *Wertbindung* begünstigt. Eventuell stellt aber gerade das Semesterpraktikum für diese Studierendengruppe einen geeigneten Lernraum dar, dieses aus der Berufswahlmotivation potentiell resultierende Defizit zu bearbeiten. Hier ergeben sich weiterführende Fragen, die durch weitere – beispielsweise qualitative – Studien genauer beleuchtet werden sollten. Bis dato lag der Fokus der wissenschaftlichen Auseinandersetzung auf dem (im Verlauf des Praktikums) veränderten Kompetenz*erleben* der Studierenden. Zur weiteren Konstruktvalidierung sollte die Übereinstimmung zwischen Kompetenzselbsteinschätzungen und leistungsbezogenen Gutachten der Ausbildungsberater und -beraterinnen untersucht (vgl. Bodensohn/Schneider 2008) und überprüft werden, inwiefern letztere mit der Berufswahlmotivation hypothesenkonform korrelieren.

Literatur

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

Biermann, A./Grassmé, I./Gläser-Zikuda, M./Brünken, R. (2018). Effektive Nutzung von Schulpraktika: Ein Mediationsmodell zwischen individuellen Merkmalen, der Nutzung von Lerngelegenheiten und selbst eingeschätzter Kompetenz von Lehramtsstudierenden. Lehrerbildung auf dem Prüfstand, 12 (2), S. 222–239.

Bock, T./Hany, E. A./Protzel, M. (2017): Forschungsberichte aus der Erfurt School of Education. Lerngelegenheiten und Lerngewinne im Komplexen Schulpraktikum 2015–2017. www.uni-erfurt.de/ese/downloadbereich/veroeffentlichungen/ (26.3.2019).

Bodensohn, R./Schneider, C. (2008). Was nützen Praktika? Evaluation der Block-Praktika im Lehramterträge und offene Fragen nach sechs Jahren. Empirische Pädagogik, 22(3), 274–304.

Borenstein, M. (2009): Introduction to meta-analysis. Chichester: Wiley.

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155.

Cramer, C. (2016): Berufswahl Lehramt: Wer entscheidet sich warum? In M. Rothland (Hrsg.), Beruf Lehrer/Lehrerin. Ein Studienbuch. Münster: Waxmann. S. 261–276.

Cramer, C./Horn, K.-P./Schweitzer, F. (2009). Zur Bedeutsamkeit von Ausbildungskomponenten des Lehramtsstudiums im Urteil von Erstsemestern. Erste Ergebnisse der Studie „Entwicklung Lehramtsstudierender im Kontext institutioneller Rahmenbedingungen" (ELKiR). Zeitschrift für Pädagogik, 55(5), S. 761–780.

Dickhäuser, O./Butler, R./Tönjes, B. (2007). Das zeigt doch nur, dass ich's nicht kann. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 39(3), S. 120–126.

Dunlap, W. P./Cortina, J. M./Vaslow, J. B./Burke, M. J. (1996): Meta-analysis of experiments with matched groups or repeated measures designs. Psychological Methods, 1(2), S. 170–177

Eberle, T./Pollack, G. (2006): Studien- und Berufswahlmotivation von Lehramtsstudierenden. Paradigma. Beiträge aus Forschung und Lehre aus dem Zentrum für Lehrerbildung und Fachdidaktik. 1 (1), S. 19–36. Universität Passau.

Eccles, J. S./Wigfield, A. (2002): Motivational Beliefs, Values, and Goals. Annual Review of Psychology, 53, S. 109–132.

Franz, E.-K. (in Druck): Didaktische Adaptivität erwerben – Beiträge der Lehrerbildung zur Professionalisierung von Grundschullehrer*innen. In: Donie, Ch. et al. (Hrsg.) Grundschulpädagogik zwischen Wissenschaft und Transfer. Wiesbaden: Springer VS.

Gröschner, A./Hascher, T. (2018): Praxisphasen in der Lehrer/innen/bildung. In: Gläser-Zikuda, M./Harring, M./Rohlfs, C. (Hrsg.): Handbuch Schulpädagogik. Münster: Waxmann

Gröschner, A./Schmitt, C. (2010). Wirkt, was wir bewegen? – Ansätze zur Untersuchung der Qualität universitärer Praxisphasen im Kontext der Reform der Lehrerbildung. Erziehungswissenschaft, 21 (40), S. 89–97.

Gröschner, A./Schmitt, C. (2008). „Fit für das Studium?"-Studien- und Berufswahlmotive, Belastungserfahrungen und Kompetenzerwartungen am Beginn der Lehramtsausbildung. Empirische Befunde der wissenschaftlichen Begleitforschung zum Praxissemester an der Universität Jena. Lehrerbildung auf dem Prüfstand, 1(2), S. 605–624.

Groß Ophoff, J./Franz, E.-Kr./Immel, C. (in Druck): Selbstwirksamkeit als Prädiktor des Kompetenzerlebens von Lehramtsstudierenden im Semesterpraktikum. Empirische Pädagogik. Themenheft „Selbstwirksamkeit von Lehrkräften".

Hascher, T./Kittinger, C. (2014): Learning processes in internships – analyses from a study using learning diaries. In: Arnold, K.-H./Gröschner, A./Hascher, T. (Hrsg.): Pedagogical Field Experiences in Teacher Education. Münster: Waxmann. S. 221–235.

König, J./Rothland, M. (2013). Motivationale Bedingungen der Kompetenzentwicklung in der Lehrerbildung. In: Gehrmann et al. (Hrsg.): Formation und Transformation der Lehrerbildung. Entwicklungstrends und Forschungsbefunde. Bad Heilbrunn: Klinkhardt. S. 88–103.

König, J./Rothland, M. (2018). Das Praxissemester in der Lehrerbildung: Stand der Forschung und zentrale Ergebnisse des Projekts Learning to Practice. In: König, J./Rothland, M./Schaper, N. (Hrsg.): Learning to Practice, Learning to Reflect? Ergebnisse aus der Längsschnittstudie LtP zur Nutzung und Wirkung des Praxissemesters in der Lehrerbildung. Wiesbaden: VS Springer. S. 1–62.

Korthagen, F. A. (2010). How teacher education can make a difference. Journal of Education for Teaching, 36(4), S. 407–423.

Kretzschmar, A./Gignac, G. E. (2019). At What Sample Size Do Latent Variable Correlations Stabilize?. Journal of Research in Personality, 80, S. 17–22.

Kunter, M./Baumert, J./Blum, W./Klusmann, U./Krauss, S./Neubrand, M. (2011): Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramm COACTIV. Münster: Waxmann.

Muthén, B. O./Kaplan, D./Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. Psychometrika, 52 (3), 431–462.

Muthén, L. K./Muthén, B. O. (2010): MPlus Version 6. Statistical Analysis with latent variables. Los Angeles, CA: Muthén & Muthén.

Pohlmann, B./Möller, J. (2010): Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). Zeitschrift für Pädagogische Psychologie, 24(1), S. 73–84.

Rauin, U. (2007): Im Studium wenig engagiert – im Beruf schnell überfordert. Studierverhalten und Karrieren im Lehrerberuf – Kann man Risiken schon im Studium prognostizieren? Forschung Frankfurt a. M., 3, S. 60–64.

Rothland, M. (2014). Warum entscheiden sich Studierende für den Lehrerberuf? Berufswahlmotive und berufsbezogene Überzeugungen von Lehramtsstudierenden. Handbuch der Forschung zum Lehrerberuf, 2, 349–385.

Rutsch, J./Rehm, M./Vogel, M./Seidenfuß, M./Dörfler, T. (Hrsg.) (2018): Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen. Wiesbaden: Springer.

Schönbrodt, F. D./Perugini, M. (2013). At what sample size do correlations stabilize?. Journal of Research in Personality, 47(5), 609–612.

Verloop, N./van Driel, J./Meijer, P. (2001): Teacher knowledge and the knowledge base of teaching. International Journal of Educational Research, 35, S. 441–461.

Voss, T./Kunina-Habenicht, O./Hoehne, V./Kunter, M. (2015): Stichwort Pädagogisches Wissen von Lehrkräften. Empirische Zugänge und Befunde. Zeitschrift für Erziehungswissenschaft, 18(2), S. 187–223.

Eine ‚gute' Erklärung für alle?!

Gruppenspezifische Unterschiede in der Beurteilung von Erklärqualität – erste Ergebnisse aus dem interdisziplinären Forschungsprojekt FALKE

Alfred Lindl, Lisa Gaier, Matthias Weich, Mario Frei, Christina Ehras, Maria Gastl-Pischetsrieder, Michael Elmer, Katharina Asen-Molz, Anna-Maria Ruck, Jana Heinze, Renate Murmann, Eileen Gunga und Simone Röhrl

Abstract

Im interdisziplinären Forschungsprojekt FALKE (*Fa*chspezifische *L*ehrer*k*ompetenzen im *E*rklären) kooperieren die Didaktiken Biologie, Chemie, Deutsch, Englisch, Evangelische Religion, Geschichte, Grundschulpädagogik, Mathematik, Musik sowie Physik und die Bildende Kunst/Ästhetische Erziehung, die Deutsche Sprachwissenschaft, die Sprechwissenschaft sowie die Methoden der Empirischen Bildungsforschung.[1] Gemeinsame Ziele sind eine Untersuchung domänenübergreifender und -spezifischer Merkmale qualitätsvollen Erklärens sowie eine Verbesserung der Lehramtsausbildung durch eine evidenzbasierte Konzeption universitärer Lehrangebote zum Erklären im Fachunterricht. Nach einer Skizze des Forschungsstands widmet sich dieser Beitrag der projektbezogenen Konzeptualisierung und Operationalisierung von Erklären und diesbezüglicher Qualitätsaspekte wie Strukturiertheit, Adressatenorientierung, sprachlicher Verständlichkeit oder Sprech- und Körperausdruck. In einem einheitlichen Studiendesign werden diese anhand eines Onlinefragebogens mit Video-

1 Als Teilprojekt der Initiative KOLEG (Kooperative Lehrerbildung Gestalten; Förderkennzeichen: 01JA1512) an der Universität Regensburg wird FALKE im Rahmen der gemeinsamen Qualitätsoffensive Lehrerbildung von Bund und Ländern durch Mittel des Bundesministeriums für Bildung und Forschung gefördert. Als Betreuer fungieren in obiger Fächerreihenfolge durch Kommata getrennt: Prof. Dr. Arne Dittmer, Prof. Dr. Oliver Tepner, Prof. Dr. Anita Schilcher, Prof. Dr. Petra Kirchhoff, Prof. Dr. Michael Fricke, Dr. Josef Memminger, Prof. Dr. Astrid Rank, Prof. Dr. Stefan Krauss, Prof. Dr. Bernhard Hofmann/Dr. Gabriele Puffer, Prof. Dr. Karsten Rincke, Prof. Dr. Birgit Eiglsperger, Prof. Dr. Christiane Thim-Mabrey, PD Dr. Wieland Kranich, Prof. Dr. Sven Hilbert.

vignetten unter Perspektiventriangulation von vier Berufsgruppen analysiert (Schülerinnen und Schüler, Studierende, Lehrkräfte sowie Didaktikerinnen und Didaktiker). Neben einer enormen Beurteilungsheterogenität indizieren erste Ergebnisse signifikante Zusammenhänge zwischen der wahrgenommenen Erklärqualität und den genannten Kriterien, deren Bedeutung abschließend kritisch diskutiert wird.

1 Heterogene Forschungsperspektiven auf Erklären

Theoretisch-hermeneutische Untersuchungen

Dass der Begriff Erklären eine semantische Polyvalenz und eine komplexe Sprechhandlung impliziert, spiegelt sich in seiner alltagssprachlichen Verwendungsvielfalt wie auch in der Mannigfaltigkeit diesbezüglicher Forschungsperspektiven wider. Hierzu zählen wissenschaftstheoretische (Hempel/Oppenheim 1948), philosophische und sozialwissenschaftliche Betrachtungen (Bartelborth 2007) ebenso wie sprachwissenschaftliche Definitionsansätze, die sich dem Begriff auf morphologischer Ebene nähern und Erklären als einen ein (Kenntnis-)Defizit ausgleichenden Klärungsvorgang mit potenzieller perlokutiver Wirkung erachten (Klein 2009). Eine solche Wissensasymmetrie unterstreichen Kontrastierungen mit sinnverwandten Sprechhandlungen wie Erläutern, Begründen oder Argumentieren zwar als Charakteristikum von Erklären (Kiel 1999; Findeisen 2017). Allerdings sind einige Merkmale dieser Sprechakte ebenfalls in Erklärungen wiederzufinden, sodass sich ihre kommunikativen Funktionen überlagern (Hohenstein 2006). Eine vergleichbar geringe Trennschärfe weisen Typologisierungen nach der Prozessstruktur auf, obwohl sie alternative Klassifizierungen in Was-, Wie- und Warum- oder interpretierende, deskriptive und begründende Erklärungen bieten, die auf den Unterrichtskontext anwendbar erscheinen (Klein 2009). Hiervon hebt sich Kiels (1999) theoretische Modellierung aus pädagogischer Sicht insofern ab, dass er Erklärhandlungen neben ihrer prozessualen Gestaltung zudem nach ihrem ansteigenden Partizipationsgrad der Rezipienten an der Interaktion differenziert und damit die Adressatenrolle betont.

Empirischer Forschungsstand

Die Untersuchung explanativer Sprechakte hat in unterschiedlichen Bereichen Tradition (z. B. Laien-Experten-Kommunikation: Wittwer/Renkl 2008), jedoch wird die Fähigkeit, erklären zu können, vor allem als genuine Kompetenz von Lehrkräften angesehen. Dies bringt nicht nur deren Berücksichtigung als zentrale Facette des fachdidaktischen Wissens in Shulmans (1986) einflussreicher

Taxonomie des Lehrerprofessionswissens zum Ausdruck. Vielmehr ist sie unter Rückgriff auf dessen Beschreibung in zahlreichen Professionswissensmodellen und -tests für (angehende) Lehrkräfte enthalten (vgl. Krauss et al. 2017). Sie erweist sich als Kernfacette des fachdidaktischen Wissens (Kunter et al. 2013) und als Form direkter Instruktion (Hattie 2015) als prädiktiv valide für die Unterrichtsqualität und den Leistungszuwachs von Schülerinnen und Schülern.

Trotz dieser Tragweite ist der aktuelle Forschungsstand der einzelnen Unterrichtsfächer, von denen manche traditionell eher distal zur empirischen Bildungsforschung stehen, recht heterogen. Im mathematisch-naturwissenschaftlichen Bereich widmen sich zwar einige empirische Studien Erklärhandlungen (z. B. Prediger/Erath 2014; Wörn 2014; Kulgemeyer/Tomczyszyn 2015). Abgesehen von Wirtschaft und Rechnungswesen (Schopf/Zwischenbrugger 2015; Findeisen 2017) dominieren in den gesellschaftswissenschaftlichen, sprachlichen und ästhetisch bildenden Disziplinen bisher allerdings vorwiegend normative Überlegungen oder hermeneutische Arbeiten zum Erklären. Qualitative oder quantitative Untersuchungen liegen nur vereinzelt vor (z. B. Rathausky 2010; Neumeister 2011).

Neben der Bedeutung von Erklärungen in Lehr-Lernsituationen verdeutlicht diese Übersicht den hohen Fragmentierungsgrad bisheriger Untersuchungsansätze sowie zahlreiche Forschungsdesiderate, und zwar insbesondere mit Blick auf das gesamte Fächerspektrum an Schulen. Dies wiegt einerseits schwer, weil domänenspezifische Erklärgegenstände sowie die jeweilige Fachkultur die Art und Weise von Erklärungen wesentlich beeinflussen und spezielle Herausforderungen schaffen (Leinhardt 2001). Andererseits erforderte dies im Rahmen des FALKE-Projekts zuerst eine diskursive Identifikation übereinstimmender Ausgangspunkte und Ziele bezüglich der Analyse wesentlicher Qualitätskriterien mündlicher Erklärungen, wie im Folgenden darzulegen ist.

2 Modellierung und Messung von Erklärqualität

Der Begriff des Erklärens und Merkmale guten Erklärens

Literaturbasierte projektinterne Diskussionen über Bedeutung und Kriterien guten Erklärens aus inter- wie intradisziplinärer Perspektive mit der Absicht ihrer theoriekonformen, einheitlichen Konzeptualisierung und Operationalisierung führten zu einer gemeinsamen Definition von Erklären, die an die einschlägigen Vorarbeiten von Kiel (1999), Leinhardt (2001), Wagner und Wörn (2011) sowie Prediger und Erath (2014) anknüpft.

Unterrichtliches Erklären *einer Lehrkraft* ist ein komplexer, vorbereiteter oder sich situativ ergebender sowie interaktiver *Kommunikationsprozess*, der auf eine *Fähig- bzw. Fertigkeitsvermittlung* sowie die *Initiierung eines Verstehensprozesses* bei Adressatinnen und Adressaten abzielt.

Damit beschränkt sich diese Begriffsbestimmung auf Erklärungen, wie sie im Unterricht vorkommen und von Lehrpersonen intentional gegeben werden, ohne den prinzipiell dialogischen Charakter dieser hier primär illokutiv aufgefassten Sprechhandlung zu negieren. Die Absicht einer Wissens- oder Kompetenzvermittlung wird zwar als wesenhaft für eine Erklärung angesehen, allerdings deren Anregungspotenzial und perlokutive Wirksamkeit hinsichtlich individueller Lernprozesse nur angenommen.

Unter Voraussetzung einiger grundlegender Eigenschaften guter Erklärungen wie eine hinreichende fachliche Korrektheit oder Verwendung geeigneter Visualisierungen (Findeisen 2017; Schilcher et al. 2017) wurden fünf Konstrukte als zentrale Qualitätsindikatoren herausgearbeitet. Von diesen sind die ersten vier fachübergreifend einheitlich, das fünfte jeweils gemäß disziplineigenen Fragestellungen modelliert:

- Als bedeutsames Merkmal gilt die *Strukturiertheit* in Bezug auf den inhaltlichen Aufbau und die innere Gliederung eines Erklärprozesses, für den unter anderem eine klare stoffliche Abgrenzung, ein logisch-stringentes Vorgehen oder eine Fokussierung auf Kerninhalte förderlich erscheinen (Kiel 1999; Helmke 2010; Wagner/Wörn 2011).
- Auch die *Adressatenorientierung* wird als wesentlicher Qualitätsaspekt erachtet (Wittwer/Renkl 2008; Helmke 2010), worunter die möglichst optimale Abstimmung einer Erklärung auf volitionale, motivationale und kognitive Charakteristika von Rezipienten zu verstehen ist (Findeisen 2017).
- Ein weiteres Kriterium stellt die *sprachliche Verständlichkeit* dar (Kulgemeyer/Schecker 2013). Sie schließt Komponenten des semantischen und syntaktischen Arrangements einer Erklärung wie Begriffsbewusstsein, Aufmerksamkeitsmarker, syntaktische Komplexität oder idiolektische Besonderheiten einzelner Sprecherinnen und Sprecher ein.
- Der *Sprech- und Körperausdruck* umfasst hingegen die Bereiche Artikulation sowie Phonation mit Merkmalen der Prosodie und Stimmqualität wie auch Mimik und Gestik (Neuber 2013).
- In jeder Disziplin werden zudem *fachspezifische Qualitätsaspekte* des Erklärens konzeptualisiert. Hierzu gehören beispielsweise in Deutsch die Erklärtiefe, in Geschichte die Bedeutung historischer Kausalitätsstrukturen oder in Musik optische und akustische Darbietungsformen.

Erklärqualität und deren Kriterien in der Wahrnehmung von vier Berufsgruppen

Um gemäß dem Angebots-Nutzungs-Modell für Lehr-Lernkontexte von Helmke (2010) Vermittler- und Adressatenblickwinkel auf Qualitätsaspekte einzubeziehen, ist eine Perspektiventriangulation kriteriengestützter Einschätzungen von Erklärungen zwischen gleichberechtigten Standpunkten differenter Gruppen vorzunehmen. Deren Auswahl berücksichtigt die Forschung zur professionellen Wahrnehmung von Lehrkräften, die als zentraler Bestandteil ihrer Expertise gilt (z. B. Bromme 1992; Blömeke et al. 2015). Während Lehramtsstudierende demnach bei der Beurteilung von Erklärungen eine Novizensicht innehaben, bringen berufserfahrende Lehrkräfte an der Unterrichtspraxis ausgerichtete, Didaktikerinnen und Didaktiker als universitäre Lehrerbildner theorie- und forschungsbasierte Blickwinkel ein. Diesen drei Berufsgruppen von (angehenden) Erklärenden stehen die Schülerinnen und Schüler als Rezipienten gegenüber, die mit ihren eher intuitiven, impliziten und rollenspezifisch selektiven Beurteilungskategorien insbesondere bezüglich der Adressatenspezifität Experten sind (Kunter/Baumert 2006; Scherzinger/Wettstein 2018).

Gemeinsame Forschungsziele

Vor dem aufgezeigten theoretischen Hintergrund wurden in FALKE unter anderem folgende Fragestellungen disziplinübergreifend formuliert und im Fächerverbund betrachtet.

(1) Wie wird die Qualität einer Erklärung von den verschiedenen Personengruppen eingeschätzt? Welche Unterschiede gibt es innerhalb derselben beziehungsweise zwischen den Gruppen?
(2) Welche Kriterien sind für diese Einschätzungen allgemein sowie insbesondere fachspezifisch relevant und welche Zusammenhänge bestehen jeweils?
(3) Welche Gemeinsamkeiten und Unterschiede existieren zwischen den Fächern?

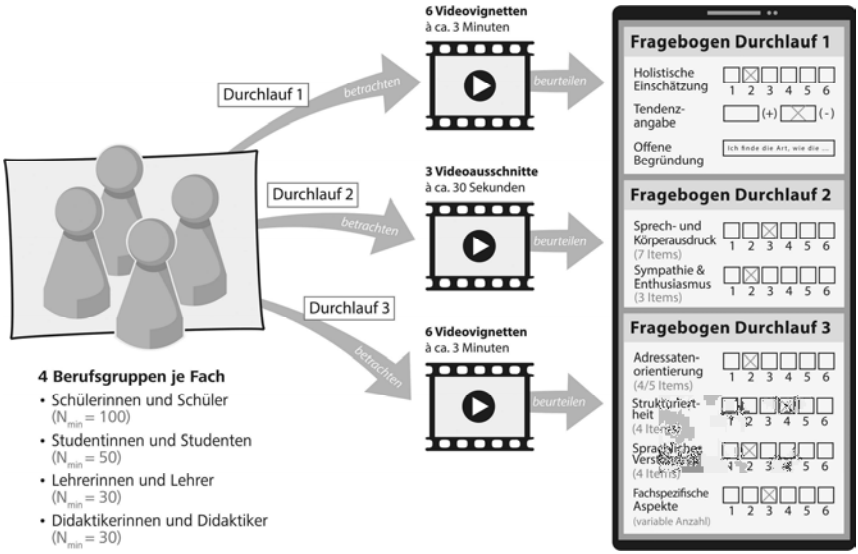
Im Folgenden sollen das hierfür verwendete Studiendesign und erste Ergebnisse dem begrenzten Umfang dieses Beitrags entsprechend überblicksartig präsentiert werden.

3 Einheitliches methodisches Vorgehen

Aufbau des Onlinefragebogens mit videographierten Erklärungen

Da sich zur Untersuchung der voranstehenden Forschungsfragen Videoformate als ökologisch möglichst valide Stimuli anbieten (Seidel/Thiel 2017), wurden in allen Fächern mindestens sechs je zirka dreiminütige, kontextreiche Videovignetten produziert.[2] Diese stellen verdichtete, inszenierte und typische unterrichtliche Erklärsituationen mit konkretem thematischen Bezug zum Schulcurriculum dar, beinhalten gezielte Variationen hinsichtlich fachheuristischer Aspekte und verzichten zur Komplexitätsreduktion sowie zur Aufmerksamkeitsfokussierung auf Interaktionsstrukturen. Zusammen mit pro Qualitätsmerkmal konstruierten Items wurden sie in einen computerbasierten Fragebogen mit systematisch variierter Präsentationsreihenfolge implementiert, der von den vier Berufsgruppen in drei Durchläufen zu bearbeiten ist (vgl. Abb. 1).

Abbildung 1: Fächerübergreifend identisches Studiendesign des FALKE-Projekts



Nach knappen einleitenden Bemerkungen zur Studie werden im ersten Durchgang demographische Angaben erhoben und im Anschluss an situierende Hinführungen alle Erklärungen in voller Länge dargeboten. Die Studienteilnehmerinnen und -teilnehmer geben jeweils ohne ablenkende Vorgabe spezifi-

2 Aufgrund fachspezifischer Gesichtspunkte wurden in Musik sieben und in Mathematik zwölf, in allen anderen Disziplinen sechs Erklärvideos erstellt.

scher Kriterien ein Globalurteil bezüglich der wahrgenommenen Erklärqualität auf einer sechsstufigen Notenskala ab, das sie mit einer positiven oder negativen Tendenz versehen und in einem Freitextfeld begründen können. Im zweiten Durchlauf werden drei ausgewählte Videos für 30 Sekunden angespielt, wobei nun die stimmlich-sprecherische und körperausdrucksbezogene Performanz und Persönlichkeitswirkung der/des Erklärenden anhand entsprechender Items (7 bzw. 3) auf geschlossenen sechsstufigen Antwortskalen einzuschätzen ist. Im dritten Durchgang werden alle Erklärvideos wiederum vollständig gezeigt und mithilfe der Items zur Adressatenorientierung (4/5), Strukturiertheit (4), sprachlichen Verständlichkeit (4) und zu fachspezifischen Aspekten bewertet (vgl. Abb. 1 und Tab. 1).

Stichprobe und psychometrische Gütekriterien

Zur Gesamtstichprobe von 2782 Personen[3] gehören 1328 Schülerinnen und Schüler verschiedener Schularten und Jahrgangsstufen in Bayern, die während regulärer Unterrichtsstunden unter Aufsicht eines geschulten Projektmitarbeiters teilnahmen. In universitären Aus- und Fortbildungsveranstaltungen bzw. nach Einladung via E-Mail beteiligten sich zudem 734 Lehramtsstudierende und 411 Lehrkräfte jeweils unterschiedlicher Schulformen sowie 309 Fachdidaktikerinnen und -didaktiker. Die Angehörigen aller Berufsgruppen bearbeiteten den Onlinefragebogen freiwillig und unentgeltlich.

Aufgrund computerbasierter Administration und geschlossener Antwortformate ist die Auswertungsobjektivität des Fragebogens gegeben und angesichts der theoriebasierten Konstruktion sowie Expertenbefragungen zu den Items in allen Fächern die inhaltliche und Augenscheinvalidität möglichst abgesichert. Die internen Konsistenzen der Skalen Adressatenorientierung, Strukturiertheit, Sprech- und Körperausdruck sowie sprachliche Verständlichkeit wurden mittels Cronbachs Alpha fachweise über alle Videos und Berufsgruppen hinweg bestimmt. Wie aus Tabelle 1 hervorgeht, sind die ermittelten Werte insgesamt zufriedenstellend, zumal hierbei konzeptionelle und inhaltliche Überlegungen vor strikten psychometrischen Ausschlusskriterien im Vordergrund stehen.

3 Die Erhebung war in Mathematik zum Publikationszeitpunkt noch nicht abgeschlossen, weshalb auf Angaben aus diesem Fach im Folgenden verzichtet wird.

Tabelle 1: Erklärqualitätskriterien, zugehörige Beispielitems sowie Itemanzahlen und interfachliche Spannweiten der internen Skalenkonsistenzen

Kriterium/ Skala	**Beispielitem**	**Item-anzahl**	**Cronbachs α**	
			Minimum (Fach)	Maximum (Fach)
Strukturiertheit	Die Erklärung hatte einen roten Faden.	4	.61 (Musik)	.82 (Geschichte)
Adressatenorientierung	Die Schülerinnen und Schüler konnten der Erklärung mit ihrem Wissen gut folgen.	4/5[+]	.47 (Biologie)	.75 (Physik)
sprachliche Verständlichkeit	Manche Sätze hat die Lehrkraft zu lang gemacht.	4	.44 (Religion)	.60 (Deutsch)
Sprech- und Körperausdruck	Die Lehrkraft hatte eine angemessene Sprechgeschwindigkeit.	7	.80 (Physik)	.89 (Musik)
fachspez. Aspekt (Geschichte)	Die Erklärung führte eine Denkweise vor, die auch für andere historische Erklärungen brauchbar ist.	variabel	---[++]	---[++]

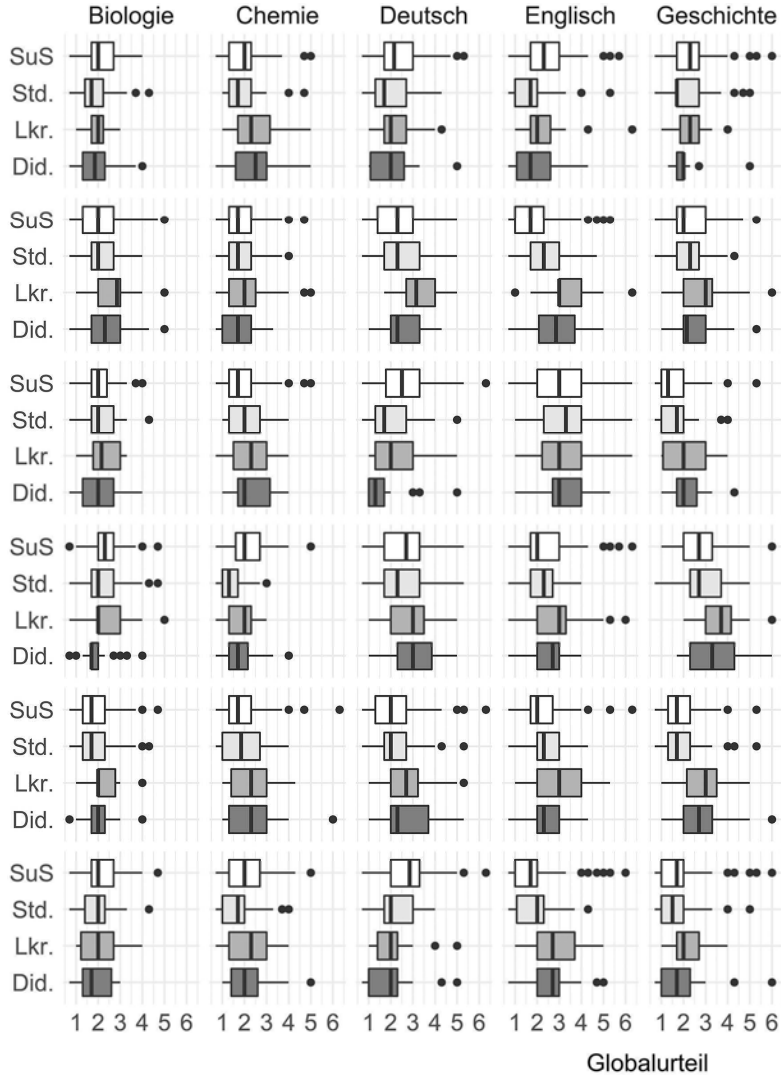
[+]: Bei Schülerinnen und Schülern mit vier, bei allen anderen Berufsgruppen mit fünf Items erhoben; [++]: Angesichts fachspezifischer Itemsätze sei hier auf Schilcher et al. (in Vorb.) verwiesen.

4 Erste Ergebnisse

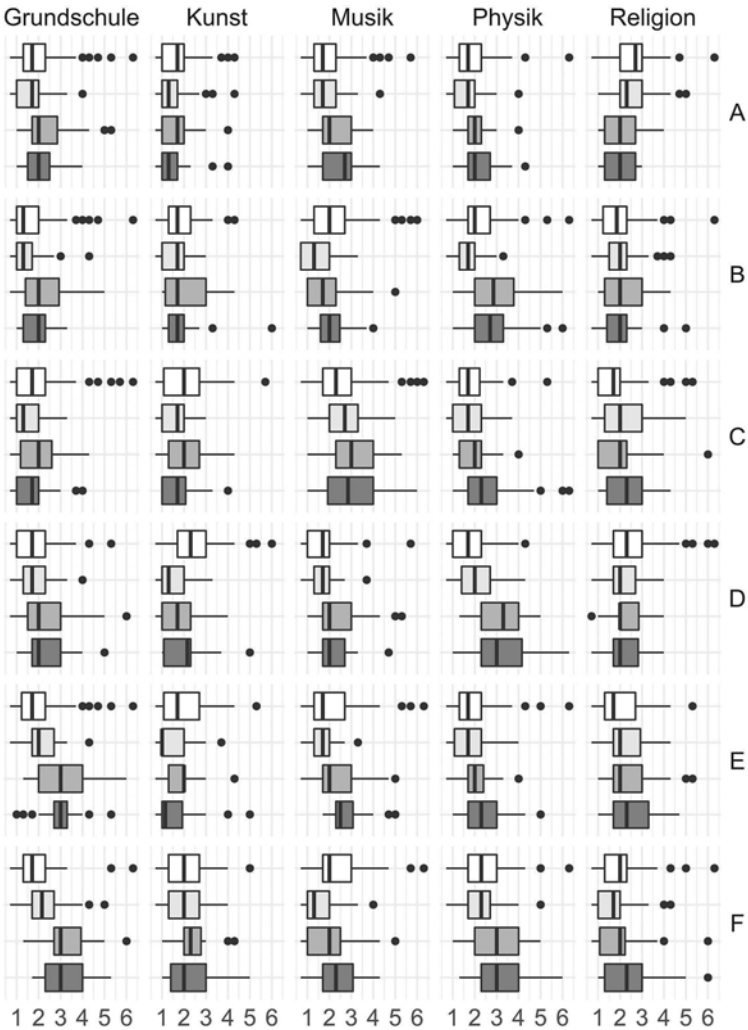
Beurteilungsheterogenität hinsichtlich der Erklärqualität

Eine rasche deskriptive Übersicht über die mittleren Tendenzen und Verteilungen der Globalbeurteilungen der Erklärungen (von „1+“ [0,7] bis „6–“ [6,3]) pro Fach, Video und Gruppe bietet Abbildung 2. Mit Blick auf die bezüglich der Globalurteile intra- wie interdisziplinär erzeugte Varianz fungieren die Videos als geeignete Stimuli für die intendierte Perspektiventriangulation (Abb. 2). Einerseits wird innerhalb derselben Berufsgruppe teils eine überraschende Einigkeit, teils eine enorme Streuung hinsichtlich der Globaleinschätzung erkennbar (z. B. Geschichte A Did. oder Physik B Std. vs. Englisch C SuS/Std./Lkr./Did. oder Musik C SuS/Lkr./Did.), wobei für die einzelnen Gruppen keine über Videos oder Fächer generalisierbaren Antworttendenzen vorliegen. Andererseits zeigen sich zwischen diesen abgesehen von wenigen einvernehmlichen Bewertungen (z. B. Kunst F oder Religion B) Unterschiede in Bezug auf die durchschnittliche Beurteilung der Erklärqualität, die in 47 der 60 Videos in entsprechenden einfaktoriellen Varianzanalysen signifikant sind.

Abbildung 2: Nach Fächern, Videos und Berufsgruppen differenzierte Darstellung der mittleren Tendenzen und der Streuungen hinsichtlich der globalen Beurteilung der Erklärungen



SuS: Schülerinnen und Schüler, Std.: Studierende, Lkr.: Lehrkräfte, Did.: Didaktikerinnen und Didaktiker;



A, B, C, D, E, F: fortlaufende Videonummerierung

Berufsgruppenspezifische Zusammenhänge zwischen Erklärqualität und -kriterien

In hierarchischen linearen Regressionsmodellen wird für jede Berufsgruppe die Globaleinschätzung durch die jeweiligen Mittelwerte der vier Skalen vorhergesagt (vgl. Tab. 2). Die nach Fächern, Videos und Personen geordnete Datenstruktur (Intraklassenkorrelation: 21 % bis 33 %) wird hierbei durch entsprechend variierende Modellkonstanten berücksichtigt.[4]

Tabelle 2: Hierarchische lineare Regressionsmodelle mit der abhängigen Variable Globalurteilunter Berücksichtigung der fach-, personen- und videospezifischen Datenstruktur

Berufsgruppe	*ICC*	Konstante	Strukturiertheit	Adressatenorientierung	Sprachliche Verständlichkeit	Sprech- und Körperausdruck	R^2 (marg.)	R^2 (kond.)
	%	*b*	*b*	*b*	*b*	*b*		
Schüler/-innen	25.7	.21*	.20**	.22**	.09**	.26**	.24	.37
Studierende	21.0	.14	.29**	.24**	-.02	.28**	.34	.47
Lehrkräfte	29.3	.51**	.24**	.37**	-.06	.18**	.31	.53
Didaktiker/-innen	33.0	.33*	.31**	.27**	-.01	.24**	.38	.51

ICC: Intraklassenkorrelation (in Prozent); *b*: unstandardisierte Regressionskoeffizienten; R^2: marginaler bzw. konditionaler Determinationskoeffizient; *: $p \leq .05$; **: $p \leq .01$

Laut Tabelle 2 stellen Strukturiertheit, Adressatenorientierung sowie Sprech- und Körperausdruck in allen Gruppen signifikante, nahezu gleichwertige Prädiktoren für die Beurteilung einer Erklärung dar, von denen sich einzig das deskriptiv größere Regressionsgewicht bezüglich der Adressatenorientierung bei Lehrkräften etwas abhebt. Die sprachliche Verständlichkeit besitzt erstaunlicherweise nur bei Schülerinnen und Schülern einen signifikanten Einfluss. Aufgrund ihrer Originalmetrik sind die unstandardisierten Regressionskoeffizienten intuitiv interpretierbar, sodass beispielsweise bei Lehrkräften eine im Mittel um eine Notenstufe bessere Bewertung der Skala Strukturiertheit mit einer Erhöhung der Globalbeurteilung um ca. eine Viertelnote einhergeht. Zudem ist die in den Modellen jeweils erzielte Varianzaufklärung von 24 % bis 38 % für die vier festen bzw. von 37 % bis 53 % für die festen und zufälligen Effekte als gut und als kriterialer Validitätsindikator zu erachten.

4 Aus Gründen der Modellvergleichbarkeit wird hier auf interdisziplinär variable Prädiktorensteigungen verzichtet, zumal diese nicht in allen Gruppen zu einer signifikanten Modellverbesserung führen (Schilcher et al. in Vorb.).

5 Diskussion und Ausblick

Die ersichtliche Beurteilungsheterogenität stimmt nicht nur mit Blick auf Prüfungssituationen, in denen mündliche Erklärleistungen von einer der betrachteten Gruppen bewertet werden, kritisch. Vielmehr unterstreicht diese, da man zumindest bei Expertinnen und Experten intuitiv von einer höheren Einigkeit ausgehen würde, die grundlegende Frage, an welchen Merkmalen sich Erklärqualität festmachen lässt. Erste Hinweise hierauf bieten die Ergebnisse der Regressionsanalysen, in denen Strukturiertheit, Adressatenorientierung sowie Sprech- und Körperausdruck in allen Berufsgruppen positiv und in ähnlicher Größenordnung mit der Erklärgüte assoziiert sind. Zugleich macht das bei Lehrkräften tendenziell hervortretende Regressionsgewicht der Adressatenorientierung jedenfalls ansatzweise auf gruppenindividuelle Bewertungsmaßstäbe aufmerksam, da diese die Qualität einer Erklärung eventuell durch entsprechende Ausbildungs- und Berufserfahrungen sensibilisiert insbesondere nach deren Adäquatheit für die Adressaten zu bemessen scheinen. Der verblüffend geringe bzw. fehlende Effekt sprachlicher Verständlichkeit sollte allerdings angesichts der breit gefächerten Operationalisierung dieses Aspekts vorläufig nur vorsichtig interpretiert, diese Skala theoretisch und empirisch reanalysiert und zu deren Bewertung unter anderem zusätzlich erfasste, offene Begründungen hinzugezogen werden. Nach erfolgter Evaluation und Überprüfung dienen die ermittelten Kriterien dabei weniger als einheitliche Beurteilungsmaßstäbe für Erklärungen, sondern die Resultate sollen langfristig zu einer evidenzbasierten Optimierung des Angebots und der strukturellen Nachhaltigkeit der Lehrerbildung beitragen. Dies betrifft vor allem die Einrichtung und Weiterentwicklung speziell konzipierter Seminare, die auf die Förderung fachunabhängiger wie -spezifischer Erklärkompetenzen angehender Lehrkräfte abzielen.

Literatur

Bartelborth, T. (2007): Erklären. Berlin: Walter de Gruyter.

Blömeke, S./Gustafsson, J.-E./Shavelson, R. J. (2015): Beyond Dichotomies. Competence Viewed as a Continuum. In: Zeitschrift für Psychologie 223, H. 1, S. 3–13.

Bromme, R. (1992): Der Lehrer als Experte. Zur Psychologie des professionellen Wissens. Bern: Huber.

Findeisen, S. (2017): Fachdidaktische Kompetenzen angehender Lehrpersonen: Eine Untersuchung zum Erklären im Rechnungswesen. Wiesbaden: Springer.

Hattie, J. (2015): Lernen sichtbar machen. 3. Auflage, übersetzt von Beywl, W./Zierer, K. Baltmannsweiler: Schneider.

Helmke, A. (2010): Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. 3. Auflage. Seelze-Velber: Klett.

Hempel, C./Oppenheim P. (1948): Studies in the Logic of Explanation. In: Philosophy of Science 15, H. 2, S. 135–175.

Hohenstein, C. (2006): Erklärendes Handeln im wissenschaftlichen Vortrag. Ein Vergleich des Deutschen mit dem Japanischen. München: Iudicium.

Kiel, E. (1999): Erklären als didaktisches Handeln. Würzburg: Ergon.

Klein, J. (2009): Erklären-Was, Erklären-Wie, Erklären-Warum, Typologie und Komplexität zentraler Akte der Welterschließung. In: Vogt, R. (Hrsg.): Erklären. Gesprächsanalytische und fachdidaktische Perspektiven. Tübingen: Stauffenburg, S. 25–36.

Krauss, S./Lindl, A./Schilcher, A./Fricke, M./Göhring, A./Hofmann, B./Kirchhoff, P./Mulder, R. H. (Hrsg.) (2017): FALKO – Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik. Münster: Waxmann.

Kulgemeyer, C./Schecker, H. (2013): Students explaining science – Assessment of science communication competence. In: Research in Science Education 43, H. 6, S. 2235–2256.

Kulgemeyer, C./Tomczyszyn, E. (2015): Physik erklären: Messung der Erklärensfähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation. In: Zeitschrift für Didaktik der Naturwissenschaften 21, S. 111–126.

Kunter, M./Baumert, J. (2006): Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. In: Learning Environments Research 9, H. 3, S. 231–251.

Kunter, M./Klusmann, U./Baumert, J./Richter, D./Voss, T./Hachfeld, A. (2013): Professional Competence of Teachers: Effects on Instructional Quality and Student Development. In: Journal of Educational Psychology 105, H. 3, S. 805–820.

Leinhardt, G. (2001): Instructional Explanations: A Commonplace for Teaching and Location for Contrast. In: Richardson, V. (Hrsg.): Handbook of research on teaching. Washington: American Educational Research Association, S. 333–357.

Neuber, B. (2013): Paraverbale und nonverbale Anteile der rhetorischen Kommunikation. In: Bose, I./Hirschfeld, U./Neuber, B./Stock, E. (Hrsg.): Einführung in die Sprechwissenschaft. Phonetik, Rhetorik, Sprechkunst. Tübingen: Narr, S. 134–140.

Neumeister, N. (2011): (Wie) Wird im Deutschunterricht erklärt? – Wissensvermittelnde Handlungen im Sprachunterricht der Sekundarstufe I. Ludwigsburg. phbl-opus.phlb.de/frontdoor/index/index/docId/32 (31.03.2019).

Prediger, S./Erath, K. (2014): Content, Interaction, or Both? Synthesizing Two German Traditions in a Video Study on Learning to Explain in Mathematics Classroom Microcultures. In: EURASIA Journal of Mathematics, Science & Technology Education 10, H. 4, S. 313–327.

Rathausky, A. (2010): Erklärprozesse im Fach Englisch. Eine qualitative Studie zur Vermittlung grammatischer Inhalte in der Sekundarstufe. Ludwigsburg. phbl-opus.phlb.de/frontdoor/index/index/docId/31 (31.03.2019).

Scherzinger, M./Wettstein, A. (2018): Classroom disruptions, the teacher-student relationship and classroom management from the perspective of teachers, students and external observers. A multimethod approach. In: Learning Environments Research 59, H. 3, S. 1–15.

Schilcher, A./Krauss, S./Lindl, A./Hilbert, S. (in Vorb.): FALKE – Fachspezifische Lehrerkompetenzen im Erklären: Untersuchungen zur Beurteilung und zu Kriterien unterrichtlicher Erklärqualität aus der Perspektive von 13 Fachbereichen. Weinheim/Basel: Beltz Juventa.

Schilcher, A./Krauss, S./Rincke, K./Hilbert, S. (2017): Ausblick – Aus FALKO wird FALKE. Fachspezifische Lehrerkompetenz im Erklären. In: Krauss et al. (2017), S. 153–199.

Schopf, C./Zwischenbrugger, A. (2015): Verständliche Erklärungen im Wirtschaftsunterricht: Eine Heuristik basierend auf dem Verständnis der Fachdidaktiker/-innen des Wiener Lehrstuhls für Wirtschaftspädagogik. In: Zeitschrift für ökonomische Bildung 3, S. 1–31.

Seidel, T./Thiel, F. (2017): Standards und Trends der videobasierten Lehr-Lernforschung. In: Zeitschrift für Erziehungswissenschaft 20, Supp. 1, S. 1–21.

Shulman, L. (1986): Those Who Understand: Knowledge Growth in Teaching. Educational Researcher 15, H. 2, S. 4–14.

Wagner, A./Wörn, C. (2011): Erklären lernen – Mathematik verstehen. Ein Praxisbuch mit Lernangeboten. Stuttgart: Klett.

Wittwer, J./Renkl, A. (2008): Why Instructional Explanations Often Do Not Work: A Framework for Understanding the Effectiveness of Instructional Explanations. Educational Psychologist 43, H. 1, S. 49–64.

Wörn, C. (2014): Unterrichtliche Erklärsituationen: Eine empirische Studie zum Lehrerhandeln und zur Kommunikation in Mathematikunterricht der Sekundarstufe I. Hamburg: Dr. Kovac.

Medienbezogene Kompetenzen von Lehrpersonen

Empirische Befunde und Perspektiven für die Lehrerausbildung

Ramona Lorenz, Manuela Endberg und Birgit Eickelmann

Abstract

Die Herausforderung, Schülerinnen und Schülern Medienbildung für das Leben in der digitalisierten Welt zu vermitteln, ist als eine zentrale Aufgabe der Schule anzusehen. Die Frage, inwiefern Lehrpersonen dafür ausgebildet sind und zukünftig werden sollten, wird in diesem Beitrag fokussiert. Daher wird auf Grundlage der repräsentativen Lehrerbefragung ($N = 1218$) im Rahmen der Studie „Schule digital – der Länderindikator 2017" untersucht, welche Lerngelegenheiten Lehrpersonen bezüglich wichtiger Aspekte schulischer Medienbildung in den ersten beiden Phasen der Lehrerbildung hatten. Kontrastierend wird ebenfalls der Frage nachgegangen, ob Lehrpersonen eine stärkere Vorbereitung auf Aufgaben der Medienbildung in der Lehramtsausbildung als notwendig erachten. Die Ergebnisse zeigen, dass nur geringe Anteile der Lehrpersonen medienbezogene Lerngelegenheiten in der Lehrerausbildung bekunden und gleichzeitig hohe Anteile eine bessere Vorbereitung in der Lehrerausbildung fordern. Zusätzlich zeigen vertiefende Korrelationsanalysen Zusammenhänge insbesondere der Nutzungshäufigkeit digitaler Medien im Unterricht mit Forderungen einer veränderten Ausbildungspraxis. Implikationen für die Lehrerausbildung werden abschließend hervorgehoben.

1 Einleitung

Konsequenzen der Digitalisierung sind in allen Lebensbereichen spürbar und erreichen auch die Schule. Lehrpersonen benötigen professionelles Wissen, um „ihren Bildungs- und Erziehungsauftrag in einer ‚digitalen Welt' verantwortungsvoll erfüllen [zu können]" (KMK 2016, S. 24). Der Erwerb dieses Wissens erfolgt im Rahmen der Lehrerausbildung (universitäres Studium und Referen-

dariat) in der Fläche noch nicht systematisch (van Ackeren et al. 2019). Gleichzeitig werden Lehrpersonen in der Schule tagtäglich mit den Potenzialen, aber auch Herausforderungen der Digitalisierung konfrontiert. Der vorliegende Beitrag greift die Perspektive der Lehrpersonen auf und beleuchtet, inwiefern diese im Rahmen ihrer Lehrerausbildung bedeutsames Wissen zur Förderung der Medienbildung der Schülerinnen und Schüler erworben haben und inwieweit sie eine stärkere Vorbereitung in der Lehrerausbildung für notwendig erachten. Mit Rückgriff auf den Ansatz der medienpädagogischen Kompetenz (Blömeke 2000) sowie den europäischen Referenzrahmen „European Framework for the Digital Competence of Educators" (DigCompEdu; Redecker 2017) zur Beschreibung der Kompetenzen von Lehrpersonen im Kontext der Digitalisierung werden im vorliegenden Beitrag als wichtige Teilbereiche der Medienbildung die Aspekte der Medienerziehung (Grafe 2011) sowie der informatischen Grundbildung (CECE 2017) in den Blick genommen. Zudem wird die durch Lehrpersonen im Unterricht initiierte Förderung der computer- und informationsbezogenen Kompetenzen der Schülerinnen und Schüler (Bos et al. 2014) fokussiert. Die Forschung weist bisher nur wenige Befunde zur Umsetzung dieser drei Bereiche im Unterricht auf; ihr Zusammenhang mit der Lehrerausbildung wurde bislang nicht untersucht. Im Fokus des Beitrags stehen daher die Einschätzungen von Lehrpersonen hinsichtlich der Fragen, (1) inwiefern in der eigenen Ausbildung Kompetenzen zur Auseinandersetzung mit zentralen Herausforderungen des Unterrichtens in der digitalen Welt erworben wurden und (2) ob in der Lehrerausbildung aktuell genügend auf diese Herausforderungen vorbereitet wird. Darüber hinaus erfolgt (3) die Untersuchung des Zusammenhangs zwischen Aspekten der Mediennutzung im Unterricht und den Wünschen der Lehrpersonen nach einer geänderten Ausbildungspraxis.

2 Relevanz der Medienbildung und der medienbezogenen Kompetenzen von Lehrpersonen

Vor dem Hintergrund aktueller Entwicklungen ergeben sich für Lehrpersonen neue Aufgaben im Bereich der Medienbildung, die Kompetenzen der Schülerinnen und Schüler im Kontext der Digitalisierung zu fördern. Einerseits betrifft dies einen kompetenten, selbstbestimmten und sicheren Umgang mit digitalen Medien (computer- und informationsbezogene Kompetenzen), aber auch medienerzieherische Aufgaben sowie ein, im Sinne einer informatischen Grundbildung, grundlegendes Verständnis von Informatiksystemen.

Theoretisch lassen sich diese Bereiche bereits im Ansatz der medienpädagogischen Kompetenz (Blömeke 2000) verorten, in dem fünf Kompetenzbereiche medienpädagogischen Handelns von Lehrpersonen differenziert werden: die

eigene Medienkompetenz, die mediendidaktische Kompetenz, die medienerzieherische Kompetenz, das Wissen zur Mediensozialisation sowie das Wissen zur Schulentwicklung mit digitalen Medien. Mit dem DigCompEdu, der von der Europäischen Kommission als Referenzrahmen entwickelt wurde, liegt ein aktuelles Modell vor, mit dem die Erfassung der digitalen Kompetenzen von Lehrenden möglich ist (Redecker 2017). Dieses umfasst insgesamt 22 Teilbereiche differenziert in berufliche und pädagogische Kompetenzen von Lehrpersonen. Auch hier stellt die Förderung der digitalen Kompetenz der Lernenden die Zielkomponente dar, sodass sich beide vorgenannten theoretischen Modelle zur Verortung der vorliegenden Forschung und damit zur Einordnung der drei benannten Bereiche der Förderung medienbezogener Kompetenzen von Lernenden, der Medienerziehung sowie der informatischen Grundbildung eignen. Im Folgenden wird näher auf diese drei Bereiche eingegangen.

Förderung der Medienkompetenz der Lernenden

Der Erwerb von Medienkompetenz ist angesichts der Bedeutung eines kompetenten Umgangs mit digitalen Medien zukünftig Aufgabe aller Phasen der Lehrerbildung (KMK 2016). Hinsichtlich der schulischen Medienbildung benötigen Lehrpersonen eigene Medienkompetenz und fachliche Medienexpertise für einen didaktisch sinnvollen Einsatz digitaler Medien im Unterricht (van Ackeren et al. 2019; Goertz/Baeßler 2018; Tulodziecki/Herzig/Grafe 2019). Die Hoffnung, dass junge Lehrpersonen, die selbst in einer digitalisierten Welt aufgewachsen sind, bereits automatisch über diese Kompetenzen verfügen, erweist sich als unbegründet (Eickelmann 2018). Dementsprechend besteht die Herausforderung, (angehenden) Lehrpersonen professionelles Wissen über den sinnvollen, adaptiven Einsatz digitaler Medien zu vermitteln, um didaktische und methodische Potenziale digitaler Medien für Lehr-Lernprozesse erkennen und nutzen können.

Bisher fehlen jedoch übergreifende Standards, die erst durch die Verabschiedung der neuen, um digitalisierungsbezogene Aspekte erweiterten KMK-Standards für die Lehrerbildung vorliegen werden. Das bisher festzustellende Fehlen an verbindlichen Anteilen medienbezogener Lehrerausbildung verdeutlicht auch die Zustimmung von mehr als 80 Prozent der Lehrpersonen in Deutschland bezüglich der Notwendigkeit einer stärkeren Vorbereitung auf den Einsatz digitaler Medien im Unterricht sowie auf die Förderung der computerbezogenen Fähigkeiten der Schülerinnen und Schüler in der ersten und zweiten Phase der Lehrerausbildung (Eickelmann/Lorenz/Endberg 2016).

Medienerziehung

Neben der Förderung der Medienkompetenz spielt die Medienerziehung eine zentrale Rolle für den kritischen und reflektierten Umgang mit Medien und Medieninformationen. Medienerziehung umfasst in diesem Sinne ein humanes bzw. verantwortliches Handeln im Kontext der Mediennutzung und Mediengestaltung (Tulodziecki 2008). Im Folgenden wird unter Medienerziehung die Auseinandersetzung der Lehrpersonen mit Medien als Gegenstand bzw. Inhalt des Unterrichts mit dem Ziel der Förderung eines kompetenten und kritischen Umgangs der Schülerinnen und Schüler mit digitalen Medien verstanden.

Der Blick in den Forschungsstand weist für Deutschland nur wenige empirische Studien auf, die Aktivitäten von Lehrpersonen im Kontext der Medienerziehung untersuchen. Die Befunde des Länderindikators 2017 zeigten, dass etwa drei Viertel (72.6 %) der Lehrpersonen der Sekundarstufe I es als sinnvoll erachten, Medienerziehung als Bestandteil ihres Fachunterrichts zu integrieren (Grafe/Lorenz/Endberg 2017). Aufgrund der gesamtgesellschaftlichen Relevanz der Digitalisierung ist die Medienerziehung als eine zunehmend relevante Aufgabe von Lehrpersonen anzusehen. Unklar bleibt bisher jedoch, inwiefern die Vorbereitung auf medienerzieherisches Handeln in Studium und Referendariat aus Sicht der Lehrpersonen zur Erfüllung dieses Anspruchs beitragen kann.

Informatische Grundbildung

Informatische Grundbildung ist – wie auch die Medienerziehung – ein wesentlicher Bestandteil von Medienbildung. Im Folgenden wird unter informatischer Grundbildung das Wissen über Grundlagen des Aufbaus, der Funktionsweisen und der Anwendung von Hardware, Software und Netzwerktechnologien verstanden. Sie vereint sowohl technische Grundlagen als auch das Wissen über mögliche gesellschaftliche Auswirkungen und die Kenntnis von Grundzügen der Wirkprinzipien von Algorithmen sowie deren praktischer Umsetzung, welche immer nachdrücklicher verbindlich eingefordert werden (u. a. Diethelm/Brinda 2018). Ein Blick in den Forschungsstand verdeutlicht, dass die informatische Grundbildung in der Schule von Lehrpersonen in Deutschland zunehmend aktiv gefördert wird. So zeigte der Länderindikator 2017, dass bundesweit durchschnittlich fast zwei Fünftel (38.1 %) der Lehrpersonen die Kompetenzen im Bereich der informatischen Grundbildung fächerübergreifend fördern, was einem signifikanten Anstieg im Vergleich zum Vorjahr (14.8 %) entspricht (Vahrenhold/Lorenz/Eickelmann 2017). Zur Erklärung dieses Befundes kann vermutet werden, dass dieser Anstieg auf eine Ausweitung der informatischen Grundbildung in einzelnen Bundesländern zurückzuführen ist. Gleichzeitig ist davon auszugehen, dass die von Starruß und Timmermann (2011) dokumentierten stark variierenden Verpflichtungsgrade und Zeiträume

im Rahmen der schulischen Laufbahn, in denen (Wahl-)Kurse zur Informatik bzw. informatischen Grundbildung angeboten werden, weiterhin Bestand haben. Daher stellt sich auch für den Bereich der informatischen Grundbildung die Frage, inwiefern eine (stärkere) Vorbereitung angehender Lehrpersonen zur Vermittlung informatischer Grundbildung aus Sicht der Lehrpersonen notwendig erscheint.

3 Forschungsfragen

Ausgehend von der aufgezeigten Relevanz der Aspekte von Medienbildung stehen im vorliegenden Beitrag die folgenden Fragen im Vordergrund, die die Vorbereitung der Lehrpersonen auf die mit der Digitalisierung verbundenen Aufgaben in den Blick nehmen:

(1) Welches Wissen haben Lehrpersonen hinsichtlich zentraler Bereiche der Digitalisierung in der Schule (Einsatz digitaler Medien im Unterricht, Medienerziehung, informatische Grundbildung) in der Lehrerausbildung erworben?
(2) In welchem Maße halten Lehrpersonen eine Vorbereitung auf die vorgenannten Bereiche in der Lehrerausbildung für erforderlich?
(3) Inwiefern hängen (a) die schulische Nutzungshäufigkeit digitaler Medien durch Lehrpersonen, (b) die Förderung medienerzieherischer Aspekte im Unterricht sowie (c) die Förderung informatischer Grundbildung mit den Forderungen einer veränderten Lehrerausbildung zusammen?

4 Methodik

Zur Beantwortung der Fragestellungen werden die Daten der Studie „Schule digital – der Länderindikator“ (Laufzeit 01/2015 bis 03/2018) herangezogen, die durch die Deutsche Telekom Stiftung ermöglicht wurde. Für die im Rahmen des Beitrags durchgeführten Analysen werden die repräsentativen Daten des Länderindikators 2017 (Lorenz et al. 2017) herangezogen, die auf Angaben von $N = 1218$ Lehrpersonen aller Schulformen der Sekundarstufe I in Deutschland (außer Förderschulen) basieren. Erhoben wurden die Daten mittels computergestützter Face-to-Face-Interviews (CAPI), wobei der Fragebogen gemeinsam mit einem/r Interviewenden digital bearbeitet wurde. Die Stichprobe setzt sich zu gut drei Fünfteln (61.3 %) aus weiblichen Lehrpersonen zusammen. Fast ein Drittel (30.9 %) der befragten Lehrpersonen ist bis zu 39 Jahre alt, gut ein Viertel (26.4 %) ist zwischen 40 und 49 Jahre alt und 42.8 % der Lehrpersonen sind 50 Jahre alt oder älter.

Neben demographischen Angaben und Fragen hinsichtlich der Nutzungshäufigkeit digitaler Medien im Unterricht (fünfstufig erfasst von *täglich* bis *nie*) standen Angaben zur Lehrerausbildung im Fokus. Die Lehrkräfte wurden jeweils in Bezug auf die universitäre Lehrerausbildung und in Bezug auf das Referendariat befragt, ob sie gelernt haben a) wie digitale Medien als Mittel des Lehrens und Lernens im Unterricht eingesetzt, b) wie medienerzieherische Ziele erreicht und c) wie die informatische Grundbildung der Schülerinnen und Schüler gefördert werden kann (mittels fünfstufigem Antwortformat von *stimme zu* bis *stimme nicht zu*). Für diese Aspekte wurde ebenso erfasst, ob die Lehrpersonen es als sinnvoll erachten, dass auf diese in den ersten beiden Phasen der Lehrerbildung stärker vorbereitet werden sollte (ja/nein-Antwortformat). Zudem wurde erfragt, ob Lehrpersonen im Unterricht Aspekte der Medienerziehung oder der informatischen Grundbildung vermitteln (ja/nein-Antwortformat).

Die Forschungsfragen werden mittels deskriptiver Auswertungen und Korrelationsanalysen (Korrelation nach Pearson) bearbeitet, die mit SPSS 25 durchgeführt werden.

5 Ergebnisse

Die deskriptiven Auswertungen hinsichtlich der ersten Forschungsfrage nach Lerngelegenheiten im Rahmen der ersten beiden Phasen der Lehrerausbildung im Kontext der Digitalisierung weisen aus Sicht der Lehrpersonen auf eine unzureichende Vorbereitung auf medienbezogene Aufgaben in beiden Ausbildungsphasen hin (vgl. Abb. 1): Je rund vier Fünftel der Lehrpersonen geben an, weder in der ersten noch in der zweiten Phase der Lehrerausbildung gelernt zu haben, wie Medien lernförderlich im Unterricht eingesetzt, medienerzieherische Ziele erreicht werden können und die Förderung der informatischen Grundbildung umgesetzt werden kann.

Bezüglich der Forderungen nach stärkerer Vorbereitung auf die genannten Aspekte in der Lehrerausbildung (Forschungsfrage 2) zeigt sich ein nahezu gespiegeltes Zustimmungsmuster: Die überwiegende Mehrheit fordert eine stärkere Vorbereitung auf zentrale, im Zuge der Digitalisierung im Unterricht anfallende Aufgabenbereiche bereits in der Lehrerausbildung (vgl. Abb. 2).

Abbildung 1: Angaben der Lehrpersonen zu Lerngelegenheiten im Studium und Referendariat bezüglich Aspekten des Medieneinsatzes (Angaben in Prozent)

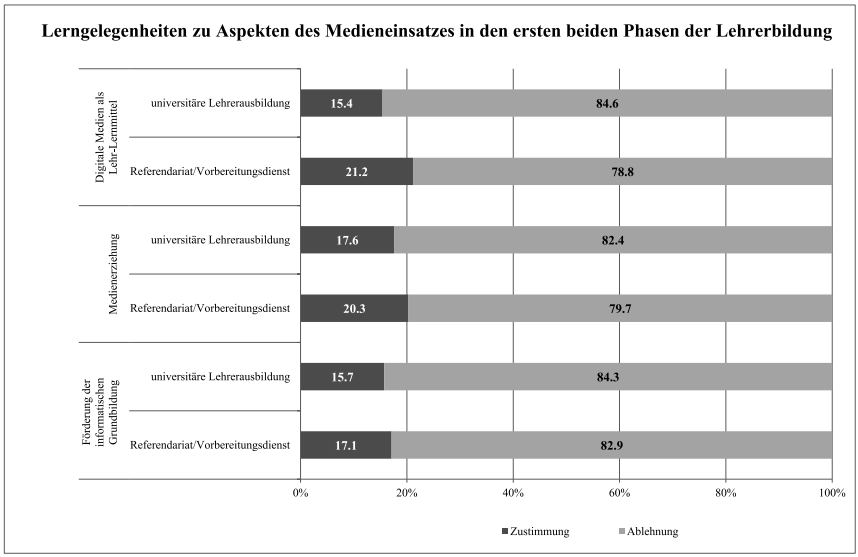
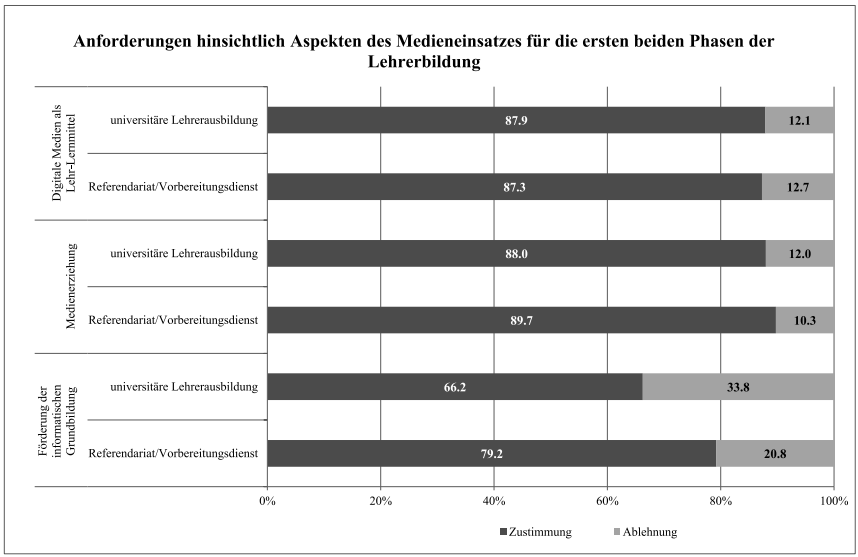


Abbildung 2: Anteile der Lehrpersonen, die der Forderung einer stärkeren Vorbereitung auf Aspekte des Medieneinsatzes durch Studium und Referendariat zustimmen (Angaben in Prozent)



Hinsichtlich der dritten Forschungsfrage werden Zusammenhänge zwischen den ersten beiden Ausbildungsphasen und dem unterrichtlichen Medieneinsatz untersucht. Dazu wird zunächst die Nutzungshäufigkeit digitaler Medien im Unterricht in Tabelle 1 deskriptiv abgebildet.

Tabelle 1: Nutzungshäufigkeit digitaler Medien im Unterricht (Anteile der Lehrpersonen in Prozent)

	Jeden Tag	Mindestens einmal in der Woche, aber nicht jeden Tag	Mindestens einmal im Monat, aber nicht jede Woche	Seltener als einmal im Monat	Nie
Nutzungshäufigkeit von Computern im Unterricht durch die Lehrpersonen	18.6 %	31.4 %	28.5 %	17.8 %	3.7 %

Außerdem wird der Zusammenhang zu im Unterricht aktiv vermittelten Kompetenzbereichen untersucht. Dabei geben drei Fünftel (60.7 %) der Lehrpersonen an, dass sie Kompetenzen im Unterricht vermitteln, die in den Bereich der Medienerziehung fallen. Für die informatische Grundbildung geben weniger als zwei Fünftel (38.1 %) der Lehrpersonen an, diese Kompetenzen zu vermitteln.

Die Ergebnisse der Korrelationsanalysen zum Zusammenhang der benannten Variablen (vgl. Tab. 2) implizieren, dass Lehrpersonen mit zunehmender Nutzungshäufigkeit digitaler Medien im Unterricht die Vorbereitung auf zentrale, im Zuge schulischer Digitalisierungsprozesse erforderliche Aufgabenbereiche einfordern; ein signifikanter Zusammenhang kann zu allen Bereichen und je beiden Ausbildungsphasen gezeigt werden.

Tabelle 2: Korrelationsanalysen zum Zusammenhang von unterrichtlichem Medieneinsatz und Forderungen einer verstärkten Vorbereitung in der Lehrerausbildung

	Einsatz digitaler Medien im Unterricht		**Medien-erziehung**		**informatische Grundbildung**	
	Universitäre Lehrer-ausbildung	Referen-dariat	Universitäre Lehrer-ausbildung	Referen-dariat	Universitäre Lehrer-ausbildung	Referen-dariat
Nutzungs-häufigkeit	.088**	.100**	.088**	.098**	.101**	.057*
Förderung mediener-zieherischer Aspekte	.070*	.056	.049	.015	.075*	.039
Förderung informa-tischer Grund-bildung	.018	-.021	-.007	-.052	.066*	.098**

* Die Korrelation ist auf dem Niveau von $p > 0.05$ (2-seitig) signifikant.
** Die Korrelation ist auf dem Niveau von $p > 0.01$ (2-seitig) signifikant.

Bezugnehmend auf die Förderung der informatischen Grundbildung im Unterricht bestehen signifikante Zusammenhänge zwischen den Förderungsaktivitäten und Forderungen nach besserer Vorbereitung auf die Vermittlung informatischer Grundbildung für die erste ($r = .066$) und die zweite ($r = .098$) Ausbildungsphase. Für den Bereich der Medienerziehung ergibt sich kein solch eindeutiges Bild.

6 Diskussion

Die Ergebnisse verdeutlichen, dass aus Sicht der Lehrpersonen in den ersten beiden Phasen der Lehrerbildung bisher noch nicht genug Gelegenheiten geboten werden, Aspekte der Medienerziehung, der informatischen Grundbildung und des lernförderlichen Einsatzes digitaler Medien zu erlernen und zu erproben. Diesbezüglich erscheinen aktuelle Entwicklungen wie die Überarbeitung der KMK-„Standards für die Lehrerbildung“ sowie die Entwicklung von Kompetenzrahmen für Lehrpersonen und die Lehrer(-aus-)bildung in den Bundesländern als ein Schritt in die richtige Richtung.

Die Befunde zeigen zudem, dass, je häufiger Lehrpersonen digitale Medien aktiv nutzen, diese zu größeren Anteilen und damit umso eher Veränderungen in der Lehrerausbildung fordern. Dies könnte dahingehend interpretiert werden, dass aufgrund fehlender Lerngelegenheiten im Rahmen der Lehrerausbildung das als fehlend wahrgenommene Wissen erst berufsbegleitend erworben werden muss. Dies könnte entsprechende Forderungen an eine veränderte Ausbildung – auch vor dem Hintergrund weiterhin steigender Herausforderungen in diesem Kontext – erklären. Perspektivisch scheint zunehmend wichtig zu sein, mehr und bessere Lerngelegenheiten in der Lehrerausbildung zu implementieren, die im Idealfall mit der dritten Phase (Fort- und Weiterbildung) im Sinne eines sukzessiven Kompetenzaufbaus verknüpft sind. Um Steuerungswissen und wissenschaftlich-analytisch abgesichertes Professionswissen zur Verfügung zu stellen, wäre die Wirksamkeit der aus Sicht der Lehrkräfte erforderlichen Veränderungen in der Lehrerausbildung empirisch zu überprüfen.

Bei der Interpretation der mit diesem Beitrag vorgelegten Ergebnisse ist zu berücksichtigen, dass bundeslandspezifische und hochschulstandortspezifische Konzepte der Lehrerausbildung unberücksichtigt bleiben und in zukünftigen Forschungsarbeiten angesichts der Vielfalt der Ansätze eine stärkere Beachtung finden sollten. Darüber hinaus liegen keine Angaben dazu vor, wie lange die eigene Lehrerausbildung zurückliegt, sodass dies nur annäherungsweise anhand des Alters der Lehrpersonen eingeschätzt werden kann. Perspektivisch scheint die systematische Vernetzung von theoretischen, empirischen und praxisorientierten Phasen angesichts der Vielfalt der Ansätze zukünftig möglicherweise

ebenso vielversprechend wie die Kooperation zwischen den verschiedenen Phasen, den lehrerbildenden Akteurinnen und Akteuren in den beteiligten Institutionen und an den Standorten. Diese Entwicklungen forschend zu begleiten wird in den nächsten Jahren vor dem Hintergrund des digitalen Wandels eine zunehmend wichtige Aufgabe für die empirische pädagogische Forschung sein.

Literatur

Ackeren van, Isabell/Aufenanger, Stefan/Eickelmann, Birgit/Friedrich, Steffen/Kammerl, Rudolf/Knopf, Julia/Mayrberger, Kerstin/Scheika, Heike/Scheiter, Katharina/Schiefner-Rohs, Mandy (2019): Digitalisierung in der Lehrerbildung. Herausforderungen, Entwicklungsfelder und Förderung von Gesamtkonzepten. In: Die Deutsche Schule 111, H. 1. www.dds.uni-hannover.de/fileadmin/schulentwicklungsforschung/DDS_Open_Access/DDS_1_2019_van_Ackeren_et_al.pdf (Abfrage: 05.05.2019).

Blömeke, Sigrid (2000): Medienpädagogische Kompetenz. Theoretische und empirische Fundierung eines zentralen Elements der Lehrerausbildung. München: kopaed.

Bos, Wilfried/Eickelmann, Birgit/Gerick, Julia/Goldhammer, Frank/Schaumburg, Heike/Schwippert, Knut/Senkbeil, Martin/Schulz-Zander, Renate/Wendt, Heike (Hrsg.) (2014): ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich. Münster: Waxmann.

CECE [The Committee on European Computing Education]. (2017): Informatics Education in Europe: Are We All In The Same Boat? doi: 10.1145/3106077

Diethelm, Ira/Brinda, Torsten (2018): Stellungnahme zu den Veränderungsvorschlägen zu den „Ländergemeinsamen inhaltlichen Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung“ (Fachstandards) der Kultusministerkonferenz auf Basis der KMK-Strategie zur „Bildung in der digitalen Welt“. fb-iad.gi.de/fileadmin/FB/IAD/Dokumente/gi-fbiad-stellungnahme-kmk-anforderungen-lehrerbildung-digitalisierung.pdf (Abfrage: 11.05.2019).

Eickelmann, Birgit (2018): Schulen und Lehrerbildung in der digitalen Welt – Thesen zwischen Hype und Zukunftsfähigkeit. In: Zeitschrift für Bildungsverwaltung 39, H. 2, S. 63–72.

Eickelmann, Birgit/Lorenz, Ramona/Endberg, Manuela (2016): Die Relevanz der Phasen der Lehrerausbildung hinsichtlich der Vermittlung didaktischer und methodischer Kompetenzen für den schulischen Einsatz digitaler Medien in Deutschland und im Bundesländervergleich. In: Bos, Wilfried/Lorenz, Ramona/Endberg, Manuela/Eickelmann, Birgit/Kammerl, Rudolf/Welling, Stefan (Hrsg.): Schule digital – der Länderindikator 2016. Kompetenzen von Lehrpersonen der Sekundarstufe I im Umgang mit digitalen Medien im Bundesländervergleich. Münster: Waxmann, S. 148–179.

Goertz, Lutz/Baeßler, Berit (2018): Überblicksstudie zum Thema Digitalisierung in der Lehrerbildung. Arbeitspapier Nr. 36. Berlin: Hochschulforum Digitalisierung.

Grafe, Silke (2011): ‚media literacy‘ und ‚media (literacy) education‘ in den USA: ein Brückenschlag über den Atlantik. In: Moser, Heinz/Grell, Petra/Niesyto, Horst (Hrsg.): Medienbildung und Medienkompetenz. Beiträge zu Schlüsselbegriffen der Medienpädagogik. München: kopaed, S. 59–80.

Grafe, Silke/Lorenz, Ramona/Endberg, Manuela (2017): Medienerziehung in der Sekundarstufe I im Bundesländervergleich. In: Lorenz, Ramona/Bos, Wilfried/Endberg, Manuela/Eickelmann, Birgit/Grafe, Silke/Vahrenhold, Jan (Hrsg.): Schule digital – der Länderindikator 2017. Schulische Medienbildung in der Sekundarstufe I mit besonderem Fokus auf MINT-Fächer im Bundesländervergleich und Trends von 2015 bis 2017. Münster: Waxmann, S. 178–204.

KMK [Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland]. (2016): Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz. www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2018/Digitalstrategie_2017_mit_Weiterbildung.pdf (Abfrage: 16.05.2019).

Lorenz, Ramona/Bos, Wilfried/Endberg, Manuela/Eickelmann, Birgit/Grafe, Silke/Vahrenhold, Jan (Hrsg.) (2017): Schule digital – der Länderindikator 2017. Schulische Medienbildung mit besonderem Fokus auf MINT-Fächer in der Sekundarstufe I im Bundesländervergleich und Trends von 2015 bis 2017. Münster: Waxmann.

Redecker, Christine (2017): European Framework for the Digital Competence of Educators. DigCompEdu. Luxembourg: Publications Office of the European Union.

Starruß, Isabelle/Timmermann, Bettina (2011): Informatische Bildung in Deutschland. Eine Analyse der informatischen Bildung an allgemeinbildenden Schulen auf der Basis der im Jahr 2010 gültigen Lehrpläne und Richtlinien. In: LOG IN 169/170, S. 49–59.

Tulodziecki, Gerhard (2008): Medienerziehung. In: Sander, Uwe/von Gross, Friederike/Hugger, Kai-Uwe (Hrsg.): Handbuch Medienpädagogik. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 110–115.

Tulodziecki, Gerhard/Herzig, Bardo/Grafe, Silke (2019): Medienbildung in Schule und Unterricht (2. Aufl.). Bad Heilbronn: utb.

Vahrenhold, Jan/Lorenz, Ramona/Eickelmann, Birgit (2017): Vermittlung informatischer Grundbildung im Unterricht der Sekundarstufe I im Bundesländervergleich. In Lorenz, Ramona/Bos, Wilfried/Endberg, Manuela/Eickelmann, Birgit/Grafe, Silke/Vahrenhold, Jan (Hrsg.): Schule digital – der Länderindikator 2017. Schulische Medienbildung in der Sekundarstufe I mit besonderem Fokus auf MINT-Fächer im Bundesländervergleich und Trends von 2015 bis 2017. Münster: Waxmann, S. 205–230.

Entwicklung eines video- und textbasierten Instruments zur Messung kollegialer Feedbackkompetenz von Lehrkräften

Christopher Neil Prilop, Kira Elena Weber
und Marc Kleinknecht

Abstract

Qualitativ hochwertiges kollegiales Feedback zwischen Lehrkräften fördert den Erwerb von professioneller Kompetenz. Für die Aus- und Fortbildung von Lehrkräften ist daher die Qualität von kollegialem Feedback sicherzustellen und zu erhöhen. Allerdings werden bislang keine Instrumente zur Messung von kollegialer Feedbackkompetenz in der Ausbildung von Lehrkräften eingesetzt. Es wurde ein video- und textbasiertes Instrument entwickelt und ein Kodierschema adaptiert, um Feedbackkommentare inhaltsanalytisch auf ihre Qualität hin zu überprüfen. Die Kombination von Video und Text als Stimulus gewährleistet hohe Authentizität, fokussierte Kontextualisierung und beugt kognitivem Overload vor.

Der Inhalt der video- und textbasierten Feedbacksituation und das Kodierschema wurde von Expertinnen und Experten validiert. Durch den Vergleich von Feedbackkommentaren von Bachelorstudierenden sowie Expertinnen und Experten konnte gezeigt werden, dass das Instrument Expertiseunterschiede abbilden kann. In Bezug auf die Reliabilität der Feedbackkodierungen ließen sich hohe Übereinstimmungen mit guten bis exzellenten ICCs feststellen.

Die Ergebnisse zeigen, dass mit Hilfe des video- und textbasierten Instruments und des Kodierschemas die kollegiale Feedbackkompetenz valide und reliabel gemessen werden kann.

1 Problemstellung

Um die Entwicklung der professionellen Kompetenz zu unterstützen, benötigen Lehrkräfte in allen Phasen der Ausbildung die Möglichkeit, qualitativ hochwertiges Feedback zu ihrem Unterricht von Kolleginnen und Kollegen oder Mentorinnen und Mentoren zu erhalten (Hammerness et al. 2005; Lu 2010). Feed-

backsitzungen können zur Reflexion über das eigene Unterrichtshandeln anregen und somit zu einer kritischen Überprüfung und Verbesserung der eigenen Unterrichtsentscheidungen führen (Tripp/Rich 2012). Zwar können Lehrkräfte ihre professionelle Kompetenz durch Praxiserfahrungen und eine kritische Selbstreflexion auch selbstständig weiterentwickeln, für substantielle Kompetenzzuwächse ist jedoch qualitativ hochwertiges Feedback notwendig (Ericsson/Krampe/Tesch-Römer 1993). Zahlreiche Studien (z. B. Allen et al. 2015; Tschannen-Moran/McMaster 2009; Weber et al. 2018) zeigen, dass Feedback und die gemeinsame Reflexion von unterrichtlichem Handeln sich positiv auf professionelle Kompetenz von Lehrkräften auswirken. Daher werden zunehmend Feedbackgelegenheiten in allen Phasen der Ausbildung von Lehrkräften implementiert (Kleinknecht/Gröschner 2016; Kraft/Blarzar/Hogan 2018).

Gleichzeitig zeigt sich, dass Feedback nicht zwingend zu Kompetenzzuwächsen führt. So führten Kluger und DeNisi (1996) eine Metastudie zu Feedbackinterventionen durch und konnten aufzeigen, dass ein Drittel aller Interventionen negative Effekte hatten. Eine effiziente Entwicklung der professionellen Kompetenz von Lehrkräften benötigt daher qualitativ hochwertiges kollegiales Feedback (Ericsson 2004). Um kollegiale Feedbackkompetenz zielgerichteter fördern zu können und dadurch Feedbackqualität zu steigern, bedarf es zunächst eines Instruments zur Messung von kollegialer Feedbackkompetenz.

2 Theoretischer Hintergrund

Studien zur Messung von Feedbackkompetenz fokussierten bislang Feedback zwischen Schülerinnen und Schülern (z. B. Gan/Hattie 2014) oder lehramtsferne Studienfächer (z. B. Gielen/De Wever 2015). Kollegiales Feedback in der Ausbildung von Lehrkräften wurde bislang kaum erforscht (z. B. Alqassab/Strijbos/Ufer 2018) und zudem nicht in Bezug zur professionellen Kompetenz von Lehrkräften gesetzt. Feedbackkompetenz wurde in den bisherigen Studien anhand des Feedbackinhalts bestimmt. Dabei stützten sich die Studien entweder auf die Klassifizierung von Hattie und Timperley (2007) oder Kriterien, die auf einem Artikel von Sluijsmans et al. (2002) beruhen. Beide Kategorisierungen ermitteln die Feedbackqualität anhand von evaluativen und tutorialen Feedbackmerkmalen (Narciss 2008).

Evaluatives Feedback kann den Lernenden summative Informationen über die Leistung (knowledge of performance), Bestätigungen über die Antwortrichtigkeit (knowledge of result) oder die richtige Aufgabenlösung (knowledge of correct response) bieten. Tutoriales Feedback stellt den Lernenden zusätzliche Informationen neben evaluativen Aspekten zur Verfügung, die für zukünftige Aufgaben genutzt werden können. Dies können unter anderem Beispiele für

Begriffe (knowledge about concepts), Ursachen von Fehlern (knowledge about mistakes), Lösungsstrategien (knowledge about how to proceed) oder metakognitive Fragen (knowledge on meta-cognition) sein (Narciss 2008). Tutoriales Feedback zieht tiefergehendes Lernen nach sich, da die Lernenden dazu befähigt werden, Kompetenzen selbst aufzubauen und ihr Handeln zu überprüfen und zu regulieren (Hattie/Timperley 2007). Im Gegensatz zu evaluativem Feedback fokussiert tutoriales Feedback nicht primär die vorliegende Aufgabe, sondern hilft den Lernenden prozess- und selbstregulationsbezogene Fähigkeiten zu entwickeln um zukünftige Aufgaben selbständig lösen zu können.

Sluijsmans, Brand-Gruwel und Van Merriënboer (2002) ermittelten, welche evaluativen und tutorialen Aspekte Feedback in der Ausbildung von Lehrkräften umfassen sollte, indem sie Merkmale eines qualitativ hochwertigen Feedbacks aus Expertinnen- und Expertenfeedback ableiteten. Das in ihrer Studie entstandene Kodiermanual wurde durch verschiedene Studien (z. B. Prins/Sluijsmans/Kirschner 2006; Gielen et al. 2010; Gielen/De Wever 2015) weiter ausgeführt. In den Studien wurde situatives Feedback anhand verschiedener Kategorien kodiert. Im Sinne des evaluativen Feedbacks wurde bewertet, inwiefern die Feedbackkommentare dem Kontext angemessen waren und die Leistungen anhand von Kriterien beurteilt wurden. Dies beinhaltete auch die Erklärung der Bewertung und das Hervorheben spezifischer Aspekte (Gielen et al. 2010). Als tutoriale Feedbackkomponenten wurde gewertet, ob in den Feedbackkommentaren Vorschläge zu alternativen Herangehensweisen präsentiert wurden (Gielen/De Wever 2015). Des Weiteren wurde geratet, inwiefern Fragen, die zur Reflexion anregen, Teil des Feedbacks waren, da diese zu einer höheren Aktivierung führen können (Nicol/Macfarlane-Dick 2006). Kodiert wurde auch die Gewichtung der Feedbacks, also ob positive, kritische oder positive und kritische Aspekte fokussiert wurden. Verschiedene Studien (Bandura/Cervon 1986; Kluger/DeNisi 1996) zeigen, dass die Gewichtung des Feedbacks zu verschiedenen Reaktionen führen kann, so dass eine Mischung aus positiven und kritischen Aspekten empfohlen wird (Gielen/De Wever 2015). Letztlich wurde auch bewertet, inwiefern die Feedbacks in Ich-Botschaften formuliert waren, da sich diese Eigenschaft auch in Feedbacks von Expertinnen und Experten zeigte (Prins/Sluijsmans/Kirschner 2006).

3 Fragestellungen

Da kollegiales Feedback eine entscheidende Rolle bei der Entwicklung von professioneller Kompetenz spielt (Hammerness et al. 2005), sollte kollegiale Feedbackkompetenz in der Ausbildung von Lehrkräften gefördert werden. Bislang fehlt jedoch die Möglichkeit zu evaluieren, inwiefern Interventionen die

kollegiale Feedbackkompetenz in Bezug auf unterrichtliches Handeln stärken. Daher bedarf es eines validen und reliablen Instruments, welches diese Kompetenz in einer realitätsnahen Situation misst. Die vorliegende Studie widmet sich dieser Forschungslücke anhand der folgenden Fragestellungen:

1. Inwiefern lässt sich mit einem video- und textbasierten Instrument kollegiale Feedbackkompetenz inhaltlich valide und reliabel messen?
2. Inwiefern lassen sich Kompetenzunterschiede mit dem video- und textbasierten Feedbackinstrument abbilden?

4 Methodisches Vorgehen

Um kollegiale Feedbackkompetenz zu messen, wurde ein video- und textbasiertes Instrument entwickelt. Instrumente, die Videos verwenden, erhöhen die Authentizität und spiegeln die Komplexität des Unterrichtsgeschehens wider (Borko/Whitcomb/Liston 2009). Videos können jedoch zu kognitivem Overload führen und benötigen Kontextinformationen (Derry/Sherin/Sherin 2014). Insbesondere für Studierende scheinen textbasierte Situationen Vorteile gegenüber ausschließlich videobasierten Situationen zu haben (Schneider et al. 2016). Um kognitivem Overload vorzubeugen, ausreichend Kontextinformationen zu bieten, aber gleichzeitig Authentizität zu gewährleisten, wurde daher ein video- und textbasiertes Instrument entwickelt.

Das entwickelte Instrument besteht aus einer video- und textbasierten Feedbacksituation zu der ein kurzes Feedback (max. 200 Wörter) verfasst werden soll. Die Feedbacksituation (siehe Abbildung 1) wurde basierend auf einer einminütigen Videosequenz, die als critical incident (Lievens/Motowidlo 2015) genutzt wurde, konstruiert. In der Videosequenz sollen die Schülerinnen und Schüler einer zweiten Klasse an verschiedenen Lernstationen versuchen herauszufinden, was mit Wasser passiert, wenn ein Objekt darin eingetaucht wird. Die Lehrerin überprüft beispielsweise die Ergebnisse, fordert die Schülerinnen und Schüler auf zu ihren Lernstationen zurückzukehren oder geht auf Unterrichtsstörungen ein. Die Videosequenz zeigt sowohl effektive als auch ineffektive Klassenführung. Neben der Videovignette werden grundlegende Informationen zur Lehrerin und zur Situation sowie eine Äußerung der Lehrerin zu ihrem Unterrichtshandeln als Text präsentiert. Die Äußerung der Lehrerin konzentrierte sich auf die Facette der Allgegenwärtigkeit im Bereich der Klassenführung. Dieser Bereich wurde gewählt, da Klassenführung als fachübergreifende Voraussetzung für erfolgreiches Unterrichten gilt (Kounin 1970; Wolff 2015).

Abbildung 1: Elemente des Feedbackinstruments

Videosequenz (1 min)
Stationenlernen zum Thema Schwimmen und Sinken

Information über die Lehrerin und die Situation
(z.B. *Eine Kollegin, mit der Sie sich gut verstehen, bittet Sie ihren Unterricht zu beobachten und ihr Feedback zur Klassenführung (Allgegenwärtigkeit) zu geben. / Diese Szene stammt aus einer zweiten Klasse. In der Unterrichtseinheit wird thematisiert, welche Gegenstände schwimmen oder sinken.*

Äußerung der Lehrerin
(z.B. *Manche Kinder sind natürlich etwas unkonzentriert, aber insgesamt kriege ich es echt gut hin die Schülerinnen und Schüler beim Arbeiten zu halten. Wie hast du das empfunden?)*

Feedbackkommentar
Wie würden Sie auf diese Äußerung in Bezug auf die Allgegenwärtigkeit eingehen?
Verfassen Sie bitte einen Feedbackkommentar.

Um den Inhalt der videobasierten Feedbacksituation zu validieren, bewerteten zwölf Expertinnen und Experten inwiefern die video- und textbasierte Feedbacksituation authentisch, interessant und typisch für eine kollegiale Feedbacksituation ist (Seidel/Thiel 2017). Die Gruppe der Expertinnen und Experten bestand aus der Leitung der Fach- und Kernseminarendes Studienseminars Lüneburg und Referentinnen und Referenten des Projekts „Kollegiale Unterrichtsreflexion" des Landesinstituts Hamburg, die in der Ausbildung von Referendarinnen und Referendaren oder in der Fortbildung von Schulkollegien eine hohe Anzahl an Feedbackgesprächen durchführen (Feedbackgespräche pro Jahr: $M = 77.67$, $SD = 45.61$, Min = 36, Max = 160). Die Expertinnen und Experten beurteilten die video- und textbasierte Feedbacksituation auf einer Vier-Punkte-Likert-Skala von *trifft nicht zu* (1) bis *trifft zu* (4). Werte, die sich im oberen Bereich von *trifft eher zu* (3) bis zu *trifft zu* (4) bewegen, werden als akzeptabel angesehen (z. B. Gold/Förster/Holodynski 2013).

Das Feedbackinstrument wurde mit 19 Bachelorstudierenden (84 % weiblich, $M_{\text{Alter}} = 22.58$, $SD_{\text{Alter}} = 1.71$, viertes Semester) und 20 Expertinnen und Experten (78 % weiblich, Feedbackgespräche pro Jahr: $M = 47.70$, $SD = 33.31$, Min = 10, Max = 120) erprobt, um zu ermitteln, inwiefern es Kompetenzunterschiede abbilden kann. Expertinnen und Experten waren hierbei die Leitungen der Fach- und Kernseminare aus Schleswig-Holstein.

Das in Bezug auf die video- und textbasierte Feedbacksituation verfasste Feedback wurde mit einer adaptierten Version eines von Prins et al. (2006) entwickelten Kodierschemas kodiert. Das Kodierschema wertet nicht die Anzahl verwendeter Kriterien oder bestimmter Fachbegriffe aus, sondern das Vorhandensein einer Reihe notwendiger Bestandteile (vgl. Gielen et al. 2010, S. 307). Jedes individuelle Feedback wurde als Analyseeinheit gewählt, da es die Semantik der Inferenz auf einer angemessenen Ebene erfasst (vgl. Chi 1996, S. 10).

Das Kodierschema erfasst die Feedbackqualität in Bezug auf das Klassenführungsmerkmal Allgegenwärtigkeit in sechs Kategorien und jeweils drei Qualitätsebenen (siehe Tabelle 1).

Tabelle 1: Kodierschema zur kollegialen Feedbackqualität in Bezug auf Klassenführung (Allgegenwärtigkeit)

Kategorie	**Gutes Feedback**		**Durchschnittliches Feedback**		**Minimales Feedback**	
Fokus	Feedback nimmt Bezug auf Allgegenwärtigkeit (inklusive Fachbegriffe)	2	Feedback nimmt Bezug auf Allgegenwärtigkeit (exklusive Fachbegriffe)	1	Feedback nimmt nicht Bezug auf Allgegenwärtigkeit	0
Konkretheit	Spezifische Situation wird dargelegt	2	Spezifische Phase wird dargelegt	1	Keine spezifische Situation oder Phase wird dargelegt	0
Alternativen	Alternativen werden dargelegt und erläutert	2	Alternativen werden dargelegt	1	Keine Alternativen werden dargelegt	0
Aktivierung	Aktivierende Fragen werden gestellt	2	Klärende Fragen werden gestellt	1	Keine Fragen werden gestellt	0
Valenz	Feedback ist ausgeglichen (positiv und kritisch)	2	Überwiegend positives Feedback	1	Überwiegend kritisches Feedback	0
Stil	Durchgehend in Ich-Botschaften formuliert	2	Teilweise in Ich-Botschaften formuliert	1	Keine Ich-Botschaften	0

Anmerkung: Adaptiert vom Feedback Quality Index (Prins et al. 2006)

Die Feedbackkommentare wurden von drei studentischen Hilfskräften kodiert. Die studentischen Hilfskräfte wurden von einem Mitglied des Forschungsteams im Kodieren geschult. Daraufhin wurden die Feedbackkommentare den studentischen Hilfskräften randomisiert zugeteilt. Jeder Feedbackkommentar wurde von zwei unabhängigen Kodierenden ausgewertet und die Reliabilität der Kodierung geprüft. Prozentuale Übereinstimmungen und Intra-Klassen-Korrelationen (ICC) zwischen den Kodierungen wurden berechnet. Um sicher-

zustellen, dass das Kodiermanual die Qualität von Feedback adäquat abbilden kann, wurde die Inhaltsvalidität der Kodierungen geprüft. Die erste Gruppe von Expertinnen und Experten wurde gebeten, zehn qualitativ unterschiedliche Feedbackkommentare in eine Rangordnung zu bringen. Im Anschluss wurden die Ränge, die durch die Kodierung und durch die Expertinnen und Experten gewonnen wurden, durch eine Spearman-Rangkorrelation verglichen.

5 Ergebnisse

Die Validierung der video- und textbasierten Feedbacksituation zeigte, dass die Expertinnen und Experten die Feedbacksituation authentisch, $M = 3.25$, $SD = 1.04$, interessant, $M = 2.90$, $SD = 1.10$, und typisch, $M = 3.00$, $SD = 0.85$, empfanden. Daher kann von einem akzeptablen Maß inhaltlicher Validität ausgegangen werden.

Die Kodierung der Feedbackkommentare resultierte in hohen prozentualen Übereinstimmungen und guten bis ausgezeichneten ICCs (siehe Tabelle 2). Die inhaltliche Validierung des Kodiermanuals zeigte, dass die Rangordnung der Expertinnen und Experten im Durchschnitt hoch positiv miteinander korrelierte, $r_S = 0.85$; $p < .01$. Es kann somit von einer angemessenen Intercoder-Reliabilität und inhaltlichen Validität des Kodierschemas ausgegangen werden.

Tabelle 2: Intercoder-Reliabilitäten: in prozentualer Übereinstimmung und Intra-Klassen-Korrelation (ICC)

Kategorie	Prozentuale Übereinstimmung	ICC
Fokus	87.2 %	0.80
Konkretheit	81.8 %	0.68
Alternativen	84.9 %	0.89
Aktivierung	91.6 %	0.93
Valenz	84.9 %	0.88
Stil	81.8 %	0.85

Die *t*-Tests für unabhängige Stichproben zeigten, dass die Bachelorstudierenden und Expertinnen und Experten sich signifikant und mit hoher Effektstärke in den Kategorien Konkretheit, Aktivierung und Stil unterschieden.

Tabelle 3: Mittelwerte (M), Standardabweichungen (SD) und Effektgrößen für die Kategorien kollegialer Feedbackkompetenz

	Studierende		Expertinnen und Experten					
	M	*SD*	*M*	*SD*	*Δ*	*t*	*df*	*p*
Fokus	1.11	0.57	1.00	0.00	0.11	0.81	37	.42
Konkretheit	**0.26**	**0.65**	**1.47**	**0.90**	**1.21**	**4.73**	**37**	**<.001**
Alternativen	0.74	0.93	1.21	0.92	0.47	1.58	37	.12
Aktivierung	**0.11**	**0.32**	**0.95**	**0.91**	**0.84**	**3.81**	**37**	**<.001**
Valenz	0.37	0.60	0.11	0.46	0.26	1.52	37	.14
Stil	**0.53**	**0.90**	**1.68**	**0.75**	**1.16**	**4.30**	**37**	**<.001**

Anmerkungen: Studierende: $n = 19$; Expertinnen und Experten: $n = 20$; *Kategorie*: M (min = 0, max = 2)

6 Diskussion

Bei der Auswertung der Feedbackqualität der Gruppe der Expertinnen und Experten zeigte sich, dass die Valenz des Feedbacks (positives vs. kritisches Feedback) größtenteils kritisch war. Dies deutet darauf hin, dass personale Faktoren wie beispielsweise die Selbstwirksamkeit des/der Feedbackempfangenden im Gestaltungsprozess einer genaueren Analyse bedürfen (Strijbos/Müller 2014). Es sollte daher untersucht werden, inwiefern die Feedbackgebenden personale Faktoren der Lehrkraft für die Valenz ihres Feedbacks heranzogen. Diese Faktoren finden sich zwar in verschiedenen Feedbackmodellen wieder, wurden bisher jedoch noch nicht näher untersucht. Des Weiteren zeigten sich keine Unterschiede zwischen den Gruppen in Bezug auf den angemessenen Fokus des Feedbacks. Daher sollte analysiert werden, inwiefern Expertinnen und Experten tiefergehend auf andere Aspekte neben der Allgegenwärtigkeit eingehen. Es ist zu erwarten, dass sich die Expertinnen und Experten in der vorliegenden Studie nicht allein auf diesen Bereich limitierten, sondern auch andere unterrichtsbezogene Aspekte thematisierten. Die Unterschiede in den Kategorien Alternativen und Konkretheit sind kongruent mit bisherigen Forschungsergebnissen aus dem Bereich der professionellen Wahrnehmung. Voraussetzung für beide Kategorien ist, dass zunächst kritische Unterrichtssituationen erkannt (noticing) und interpretiert (knowledge based reasoning) werden müssen. Forschungsarbeiten im Bereich der professionellen Wahrnehmung zeigen, dass Expertinnen und Experten eine höhere Zahl an suboptimalen Klassenführungssituationen wahrnehmen als Novizinnen und Novizen (Gold/Holodynski 2017). Inwiefern ein Zusammenhang zwischen diesen Kategorien und der professionellen Wahrnehmung besteht, sollte daher in zukünftigen Studien untersucht werden.

Durch die Auswertung der schriftlichen Feedbacks der Bachelorstudierenden zeigte sich in der vorliegenden Studie bereits in welchen Bereichen die

kollegiale Feedbackkompetenz in der universitären Ausbildung von Lehrkräften gefördert werden könnte. Neben mangelnder Konkretheit und Alternativen, unterscheiden sich die Studierenden insbesondere im Bereich der Aktivierung und des Stils von den Expertinnen und Experten. Um diesen Bereich stärker zu fördern, könnten beispielsweise im Schulpraktikum oder in Seminaren Prompts genutzt werden (Alqassab/Strijbos/Ufer 2017).

Wenngleich wir in der vorliegenden Studie erste Hinweise für die Validität des entwickelten Messinstruments erbringen konnten, handelt es sich bislang um eine eingeschränkte Validität. In weiteren Studien sollten entsprechend weitere Validierungsschritte (z. B. konfirmatorische Faktorenanalysen, konvergente und diskriminante Validierung) vorgenommen werden. Der Vergleich zwischen Bachelorstudierenden und Expertinnen und Experten deutet darauf hin, dass das Instrument Kompetenzunterschiede abbilden kann. Weiterführend sollte dennoch überprüft werden, inwiefern Kompetenzunterschiede zwischen praktizierenden Lehrkräften und Studierenden in verschiedenen Phasen ihrer Ausbildung mithilfe des Instruments abgebildet werden können. Insgesamt liefert das von uns entwickelte video- und textbasierte Instrument dennoch einen ersten Beitrag, um qualitativ hochwertiges kollegiales Feedback zwischen Lehrkräften zu messen und auf Grundlage dieser Messung effektiv zu fördern. Das video- und textbasierte Instrument und Kodierschema kann zudem leicht für Feedback zu anderen Bereichen der Unterrichtsqualität modifiziert werden.

Literatur

Allen, Joseph P./Hafen, Christopher A./Gregory, Anne C./Mikami, Amori Y./Pianta, Robert (2015): Enhancing secondary school instruction and student achievement: replication and extension of the My Teaching Partner-Secondary intervention. In: Journal of Research on Educational Effectiveness 8, S. 475–489.

Alqassab, Maryam/Strijbos, Jan-Willem/Ufer, Stefan (2017): Training peer-feedback skills on geometric construction tasks: role of domain knowledge and peer-feedback levels. In: European Journal of Psychology of Education 33, H. 1, S. 11–30.

Bandura, Albert/Cervone, Daniel (1986): Differential engagement of self-reactive influences in cognitive motivation. In: Organizational Behavior and Human Decision Processes 38, S. 92–113.

Borko, Hilda/Whitcomb, Jennie/Liston, Daniel (2009): Wicked problems and other thoughts on issues of technology and teacher learning. In: Journal of Teacher Education 60, H. 1, S. 3–7.

Chi, Michelene T. H. (1996): Quantifying qualitative analyses of verbal data: A practical guide. In: The Journal of the Learning Sciences 6, H. 3, S. 71–315.

Derry, Sharon J./Sherin, Miriam Gamoran/Sherin, Bruce L. (2014): Multimedia learning with video. In: R. Mayer (Hrsg.): The Cambridge handbook of multimedia learning. Cambridge, UK: Cambridge University Press, S. 785–812.

Ericsson, K. Anders (2004): Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. In: Academic Medicine 79, H. 10, S. 70–81.

Ericsson, K. Anders/Krampe, Ralf T./Tesch-Römer, Clemens (1993): The role of deliberate practice in the acquisition of expert performance. In: Psychological Review 100, H. 3, S. 363–406.

Gan, Mark J. S./Hattie, John (2014): Prompting secondary students' use of criteria, feedback specificity and feedback levels during an investigative task. In: Instructional Science 42, S. 861–878.

Gielen, Mario/De Wever, Bram (2015): Structuring peer assessment: Comparing the impact of the degree of structure on peer feedback content. In: Computers in Human Behavior 52, S. 315–325.

Gielen, Sarah/Peeters, Elien/Dochy, Filip/Onghena, Patrick/Stuyven, Katrien (2010): Improving the effectiveness of peer feedback for learning. In: Learning and Instruction 20, S. 304–315.

Gold, Bernadette/Förster, Stephan/Holodynski, Manfred (2013). Evaluation eines videobasierten Trainingsseminars zur Förderung der professionellen Wahrnehmung von Klassenführung im Grundschulunterricht. In: Zeitschrift für Pädagogische Psychologie, 27(3), 141–155.

Gold, Bernadette/Holodynski, Manfred (2017). Using digital video to measure the professional vision of elementary classroom management: Test validation and methodological challenges. In: Computers & Education 107, S. 13–30.

Hammerness, Karen/Darling-Hammond, Linda/Bransford, John/Berliner, David/Cochran-Smith, Marilyn/McDonald, Morva/Zeichner, Kenneth (2005): How teachers learn and develop. In: Darling-Hammond, Linda/Bransford, Joh/LePage, Pamela/Hammerness, Karen/Duffy, Helen (Hrsg.): Preparing Teachers for a Changing World: What teachers should learn and be able to do. San Francisco: Jossey-Bass, S. 358–389.

Hattie, John/Timperley, Helen (2007). The power of feedback. In: Review of Educational Research 77, H 1, S. 81–112.

Kleinknecht, Marc/Gröschner, Alexander (2016): Fostering preservice teachers' noticing with structured video feedback: Results of an online- and video-based intervention study. In: Teaching and Teacher Education 59, S. 45–56.

Kluger, Avraham N./DeNisi, Angelo (1996): The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin 119, H. 2, S. 254–284.

Kraft, Matthew A./Blazar, David/Hogan, Dylan (2018): The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. In: Review of Educational Research 88, H. 4, S. 547–588.

Kounin, Jacob S. (1970): Discipline and Group Management in Classrooms. New York: Holt, Rinehart & Winston.

Lievens, Filip/Motowidlo, Stephan J. (2016): Situational judgement tests: From measures of situational judgement to measures of general domain knowledge. In: Industrial and Organizational Psychology 9, H. 1, S. 3–22.

Lu, Hsiu-Lien (2010): Research on peer-coaching in preservice teacher education – A review of literature. In: Teaching and Teacher Education 26, S. 748–753.

Narciss, Susanne (2008): Feedback strategies for interactive learning tasks. In: Spector, J. Michael/Merrill, M. David/van Merrienboer, Jeroen J. G./Driscoll, Marcy P. (Hrsg.): Handbook of Research on Educational Communications and Technology. 3rd ed. Mahwah NJ: Lawrence Erlbaum Associates, S. 125–144.

Nicol, David J./Macfarlane-Dick, Debra (2006): Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. In: Studies in Higher Education 31, H. 2, S. 99–218.

Prins, Frans J./Sluijsmans, Dominique M. A./Kirschner, Paul (2006): Feedback for general practitioners in training: quality, styles, and preferences. In: Advances in Health Sciences Education 11, S. 289–303.

Schneider, Jürgen/Kleinknecht, Marc/Bohl, Thorsten/Kuntze, Sebastian/Rehm, Markus/Syring, Marcus (2016): Unterricht analysieren und reflektieren mit unterschiedlichen Fallmedien: Ist Video wirklich besser als Text? In: Unterrichtswissenschaft 44, H. 4, S. 474–490.

Seidel, Tina/Thiel, Felicitas (2017): Standards und Trends der videobasierten Lehr-Lernforschung. Zeitschrift für Erziehungswissenschaft, 20, H. 1, S. 1–21.

Sluijsmans, Dominique M. A./Brand-Gruwel, Saskia/van Merriënboer, Jeroen J. G. (2002): Peer assessment training in teacher education: effects on performance and perceptions. In: Assessment and Evaluation in Higher Education 27, S. 443–454.

Strijbos, Jan-Willem/Müller, Andreas (2014): Personale Faktoren im Feedbackprozess. In: Ditton, Hartmut/Müller, Andreas (Hrsg.): Feedback und Rückmeldungen: Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder. Münster: Waxmann, S. 87–134.

Tripp, Tony R./Rich, Peter J. (2012): The influence of video analysis on the process of teacher change. Teaching and Teacher Education 28, H. 5, S. 728–739.

Tschannen-Moran, Megan/McMaster, Peggy (2009): Sources of self-efficacy: Four professional development formats and their relationship to self-efficacy and implementation of a new teaching strategy. In: Elementary School Journal 110, H. 2, S. 228–245.

Weber, Kira E./Gold, Bernadette/Prilop, Christopher N./Kleinknecht, Marc (2018): Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online- and video-based self-reflection and feedback intervention. In: Teaching and Teacher Education 76, S. 39–49.

Wolff, Charlotte E. (2015): Revisiting "withitness": Differences in teachers' representations, perceptions, and interpretations of classroom management. Heerlen: Open University of the Netherlands.

Learning by doing

Eine explorative Erhebung zur Förderung digitaler Kompetenzeinschätzungen durch die Verwendung digitaler Medien in der Hochschuldidaktik

Charlott Rubach, Rebecca Lazarides, Nina Brendel und Karsten Krauskopf

Abstract

Die universitäre Lehrkräftebildung hat die Aufgabe, digitale Kompetenzen und Kompetenzeinschätzungen bei Lehramtsstudierenden zu fördern. Das Ziel der vorliegenden Studie ist vor diesem Hintergrund, zu untersuchen, inwieweit sich die digitalen Kompetenzeinschätzungen von Lehramtsstudierenden ($N_{T1} = 72$, $N_{T2} = 47$) abhängig von der Teilnahme an zwei Interventionsseminaren verändern. In Anlehnung an Forderungen der Kultusministerkonferenz wurden digitale Kompetenzen der Lehramtsstudierenden unter anderem in den Bereichen ‚Kommunizieren und Kooperieren', ‚Produzieren und Präsentieren' oder ‚Unterrichten und Implementieren' erfasst. Lehramtsstudierende des ersten Seminars nutzten digitale Medien zur Entwicklung eigener Unterrichtskonzepte in ihrer zukünftigen „Lehrkräfteperspektive". Die zweite Seminargruppe nutzte digitale Medien aus der „Perspektive von Lernenden", sodass digitale Medien unter anderem zur Bearbeitung von Aufgabenstellungen genutzt wurden. Die Ergebnisse zeigen, dass die Studierenden nach der Teilnahme am Seminar mit dem Konzept „Lehrkräfteperspektive" höhere digitale Kompetenzeinschätzungen berichten, insbesondere in den Bereichen ‚Kommunizieren und Kooperieren', ‚Problemlösen und Handeln' sowie ‚Unterrichten und Implementieren'. Lehramtsstudierende aus dem Seminar „Perspektive von Lernenden" berichten nach der Teilnahme am Ende des Semesters höhere Einschätzungen in den Bereichen ‚Kommunizieren und Kooperieren', ‚Schützen und sicher Agieren' als auch ‚Problemlösen und Handeln'.

1 Problemstellung

Die universitäre Lehrkräftebildung hat die Aufgabe, die Entwicklung digitaler Kompetenzen und Kompetenzeinschätzungen Lehramtsstudierender zu unterstützen (Herzig 2007). Das Ziel der Kompetenzförderung ist es, zukünftige Lehrpersonen auf eine effektive Förderung der digitalen Kompetenzen ihrer Schülerinnen und Schüler im Unterricht vorzubereiten (Bauer 2011; Herzig 2007). Jedoch nutzen Studierende des Lehramts national wie international digitale Medien im Vergleich zu Studierenden, die kein Lehramt studieren, weniger und schätzen sich als digital weniger kompetent ein (Bertelsmann Stiftung 2017; Farjon/Smits/Voogt 2019). Daraus lässt sich ableiten, dass in der universitären Lehrkräftebildung Handlungsbedarf besteht, da auch die Vorbereitung auf den Einsatz von Medien im Unterricht als unzureichend eingeschätzt wird (vgl. Eickelmann/Lorenz/Endberg 2016; Schiefner-Rohs 2015).

Eine Förderung (selbsteingeschätzter) digitaler Kompetenzen findet in der Lehrkräftebildung dann statt, wenn neben der Verwendung digitaler Medien durch Dozierende vor allem die Möglichkeit der selbstständigen Nutzung digitaler Medien möglich ist (Bastian/Aufenanger 2015; Bauer 2011; Tondeur et al. 2013). Vor diesem Hintergrund wurden zwei Seminare mit unterschiedlichen didaktischen Konzepten konzipiert. Gemeinsam war diesen Seminaren eine systematische Verwendung digitaler Medien im Sinne kognitiver Werkzeuge, d. h. Studierende konnten digitale Technologien nutzen und Lehr-Lernprozesse wurden mit Hilfe digitaler Medien gestaltet (Krauskopf/Zahn/Hesse 2015). Die daraus abgeleitete Fragestellung der vorliegenden Studie ist, welche Facetten digitaler Kompetenzeinschätzungen bei Lehramtsstudierenden im Rahmen dieser Seminare gefördert werden.

2 Theoretischer Hintergrund

Lehrkräftebildung und Digitalisierung

Angehende Lehrkräfte auf schul- und unterrichtsbezogene Digitalisierungsprozesse vorzubereiten gilt als aktuelle Herausforderung für die Bildungspraxis (Eickelmann et al. 2016). Der Beschluss der Kultusministerkonferenz zur Medienbildung in der Schule (2012) beschreibt, dass Lehrkräfte mit genutzten „Medien und Medientechnologien kompetent und didaktisch reflektiert umgehen“ (ebd. S. 7) sollen. Dem gegenüber konnte jedoch deutlich herausgestellt werden, dass die Vorbereitung auf den Einsatz digitaler Medien im Unterricht im Rahmen des Studiums von Lehrkräften als nicht ausreichend bewertet wird und Lehramtsstudierende geringe Kompetenzen bezüglich des Einsatzes digitaler Medien haben (Eickelmann/Lorenz/Endberg 2016; Farjon et al. 2019).

Weiterhin sind Lehramtsstudierende wenig motiviert in Bezug auf die Nutzung digitaler Technologien (Bertelsmann Stiftung 2017). Im Sinne eines kompetenzorientierten Verständnisses der Professionalisierung zukünftiger Lehrkräfte ist die universitäre Lehrkräftebildung gefordert, relevantes fachdidaktisches, bildungswissenschaftliches sowie fachliches Wissen zum effektiven Einsatz digitaler Medien zu vermitteln und digitale Kompetenzen und Kompetenzeinschätzungen bei ihren Studierenden zu steigern.

Digitale Kompetenzen von Lehramtsstudierenden

Digitale Kompetenzen beziehen sich auf Fähigkeiten, Wissen und die Motivation, um kompetent und mündig im digitalen Raum zu handeln (Eickelmann 2010). Die Wahrnehmung eigener Kompetenzen, die Voraussetzung des kompetenten Handelns im digitalen Raum (Klassen/Tze 2014), ist als digitale Kompetenzeinschätzung (synonym: selbsteingeschätzte digitale Kompetenz) zu bezeichnen. Die Erfassung digitaler Kompetenzeinschätzungen wird in der aktuellen Forschung sehr unterschiedlich vorgenommen. In mehreren nationalen wie internationalen Studien werden zumeist einzelne Fähigkeiten im Bereich ICT (Information, Communications and Technology) im Zusammenhang mit pädagogischen Anwendungsfeldern erfasst (vgl. Koh/Divaharan 2013). Harris und Kollegen (2017) verweisen in diesem Kontext jedoch darauf, dass eine inhaltsbezogene Erfassung digitaler Kompetenzen(-einschätzungen) von Lehramtsstudierenden bedeutsam ist. Als theoretische Grundlage der Erfassung solcher inhaltsbezogenen Kompetenzeinschätzungen kann das TPCK-Modell genutzt werden (Mishra/Koehler 2006), welches die Verknüpfung von technischem, pädagogischem und inhaltlichem Wissen für die sinnvolle Integration von Medien im Unterricht hervorhebt.

Mit dem Ziel der sinnvollen Förderung digitaler Kompetenzen bei Schülerinnen und Schülern definiert die KMK (2016) sechs konkrete Kompetenzbereiche, die Lehrkräfte im Rahmen schulischer Lehr-Lernprozesse bei Lernenden fördern sollen: ‚Suchen, Verarbeiten und Aufbewahren', ‚Kommunizieren und Kooperieren', ‚Produzieren und Präsentieren', ‚Schützen und sicher Agieren' sowie ‚Problemlösen und Handeln' und ‚Analysieren und Reflektieren'. Daran knüpfen Forderungen für Lehrkräfte und Lehramtsstudierende durch die Kultusministerkonferenz (2016, S. 24) an, da „Lehrkräfte […] selbst über allgemeine Medienkompetenz verfügen" sollten. Angenommen werden könnte, dass Lehrkräfte selbst hohe Fähigkeiten in den von der KMK (ebd.) definierten Kompetenzbereichen haben sollten und die Lehrkräftebildung in der Verantwortung ist, diese Kompetenzen zu fördern.

Zur Prüfung der Wirksamkeit der Lehrkräftebildung wurde in Anlehnung an die Strategie „Bildung in der digitalen Welt" der KMK (2016) von Rubach und Lazarides (under review) ein Messinstrument zur Erfassung digitaler

Kompetenzeinschätzungen bei Lehramtsstudierenden entwickelt – hierbei wurden sieben Bereiche digitaler Kompetenzeinschätzungen identifiziert: ‚Suchen und Verarbeiten', ‚Kommunizieren und Kooperieren', ‚Produzieren und Präsentieren', ‚Schützen und sicher Agieren', ‚Problemlösen und Handeln', ‚Analysieren und Reflektieren' sowie ‚Unterrichten und Implementieren'. Damit wird die vom TPCK-Modell beschriebene Relevanz grundlegender digitaler Kompetenzen aufgegriffen und im Sinne der Forderung der KMK (2016) erfasst.

Lehrformate zur Förderung digitaler Kompetenzeinschätzungen bei Lehramtsstudierenden

Die Förderung (selbsteingeschätzter) professioneller Kompetenzen im Umgang mit digitalen Technologien kann bei Studierenden des Lehramts durch verschiedene Lehrkonzepte gelingen (Krauskopf/Zahn/Hesse 2015; Tondeur et al. 2019). Einigkeit herrscht darüber, dass Lehramtsstudierende erstens grundlegende digitale Kompetenzen benötigen sowie zweitens im Rahmen der Ausbildung theoretische Grundlagen zur Relevanz von Medien in Lehr-Lern-Prozessen vermittelt werden sollten, damit sie Medien selbstständig nutzen und deren Bedeutung vor dem Hintergrund spezifischer Lehr-Lernziele kritisch reflektieren lernen (Bauer 2011; Brendel 2016; Voogt/McKenney 2017). Dabei wird von Tondeur und Kollegen (2019) darauf verwiesen, dass insbesondere Erfahrung in der Nutzung digitaler Medien hochrelevant sind, um Kompetenzen auszubilden.

Auf dieser Grundlage bestehen für die Lehrkräftebildung mindestens zwei Möglichkeiten, digitale Kompetenzen und Kompetenzeinschätzung der Studierenden zu fördern. Das Konzept der „Lehrkräfteperspektive" folgt Arbeiten von Bastian und Aufenanger (2015) sowie Bauer (2011), die die Relevanz des handlungsorientierten Lernens hervorheben: Digitale Medien werden von Studierenden selbstständig zur Entwicklung des eigenen Unterrichts genutzt, sie planen den Einsatz digitaler Medien und setzen sich aktiv mit Möglichkeiten der Nutzung auseinander. Andererseits können digitale Kompetenzeinschätzungen von Lehramtsstudierenden durch das Konzept der „Perspektive von Lernenden" gefördert werden, welches dem Konzept „Lernen am Modell" folgt: Dozierende der Lehrkräftebildung implementieren Medien didaktisch sinnvoll in Seminareinheiten als „Good-Practice-Beispiele" (Bastian/Aufenanger 2015; Kaufman 2014). Das bedeutet, Studierende nutzen digitale Medien im Rahmen eigener Lernprozesse, überprüfen die Relevanz digitaler Medien für Lehr-Lernprozesse und reflektieren eigene Erfahrungen.

3 Fragestellungen der vorliegenden Studie

Das Ziel der vorliegenden Studie ist es, zu untersuchen, inwieweit sich die digitalen Kompetenzeinschätzungen von Lehramtsstudierenden abhängig von der Teilnahme an den verschiedenen Seminaren verändern. Das Ziel beider Seminare war die Förderung digitaler Kompetenzeinschätzungen im Zuge der Nutzung digitaler Medien und die Reflexion der Nutzung für zukünftige Unterrichtsprozesse. Dabei nehmen wir an, dass in beiden Lehrformaten digitale Kompetenzen der Studierenden gefördert werden.

4 Methodisches Vorgehen

Seminarintervention

Für die vorliegende Studie wurden unter Berücksichtigung der Arbeiten von Bastian und Aufenanger (2015) sowie Bauer (2011) zwei Seminarformate entwickelt: Studierende der *ersten Interventionsgruppe* (IG_1, Lehrkräfteperspektive) erhielten die Möglichkeit, sich aktiv mit dem Einsatz von digitalen Medien für den eigenen Fachunterricht auseinanderzusetzen. In dieser Interventionsgruppe wurden im Rahmen eines fachdidaktischen Seminars im lehramtsbezogenen Studiengang empirische Ergebnisse zum Einsatz digitaler Medien im Unterricht besprochen sowie das Konzept des „Bring your own device" (BOYD) diskutiert (vgl. Heinen/Kerres/Schiefner-Rohs 2013). Im Zuge dessen wurden verschiedene Apps und Web 2.0-Anwendungen diskursiv reflektiert und auf ihre Anwendbarkeit im Schulalltag überprüft und getestet. Im Sinne des maßgeblich durch das TPCK-Modell von Mishra und Koehler (2006) geprägten Diskurses zu digitalen Kompetenzen von Lehrenden wurden in diesem Seminar alle relevanten Facetten, also die Verknüpfung des pädagogischen Wissens, Inhaltswissens und technischen Wissens gefördert.

In der *zweiten Interventionsgruppe* (IG_2, Perspektive von Lernenden), welche sich im bildungswissenschaftlichen Seminar mit sozialen Beziehungen in Schule und Unterricht beschäftigten, wurden digitale Medien als didaktisches Mittel implementiert. Das Seminar war so aufgebaut, dass digitale Medien gezielt für die Vermittlung von Wissen und das damit verbundene Erreichen von Lernzielen genutzt wurden. Beispielsweise wurde Onlinerecherchen durchgeführt, Onlinestrukturen zur Speicherung von Daten genutzt, eine kollaborative Sammlung von Wissen via HackMD und Padlets genutzt oder Socrative zur Wiederholung von Wissen verwendet. Unter Berücksichtigung des theoretischen Hintergrundes, dem TPCK-Modell von Mishra und Koehler (2006), wurden vor allem die Dimensionen Technological Knowledge und Pedagogical Knowledge *(*TPK-Knowledge*)* im Seminar gefördert. Zu beachten ist, dass in

beiden Seminaren der Medieneinsatz reflektiert wurde, indem verschiedene Medien und ihre Einsatzmöglichkeiten für den eigenen Unterricht kritisch diskutiert wurden (Brendel 2017).

Zur Vergleichbarkeit der Wirksamkeit der Lehrformate dient ein drittes Seminar als Kontrollgruppe. Hier wurden keine digitalen Medien genutzt und thematisch die Entwicklung und Erprobung von eigenen Unterrichtskonzepten umgesetzt. Vor dem Hintergrund des TPCK-Modell von Mishra und Koehler (2006) wurden hier das pädagogische und inhaltliche Wissen der Studierenden gefördert.

Stichprobe

Zum ersten Messzeitpunkt zu Beginn des Semesters wurden insgesamt 62 Lehramtsstudierende (IG_1: n = 15; IG_2: n = 23; KG: n = 24) der Universität Potsdam befragt. Knapp über die Hälfte (57 %) der Studierenden waren weiblich. 64 % der Teilnehmenden studierten den lehramtsbezogenen Bachelor (M_{Alter} = 26,5 Jahre, SD = 4,9). Die Teilnehmenden studieren am häufigsten die Fächer Deutsch, Englisch, Sport und Geschichte. Studierende der Interventionsgruppe I studieren alle das Fach Geografie.

Am Ende des Semesters zum zweiten Messzeitpunkt (T2) nahmen von den Studierenden, die zum ersten Zeitpunkt teilgenommen hatten, 40 Lehramtsstudierende erneut teil (IG_1: n = 11; IG_2: n = 17; KG: n = 12). Die Ausfallquote von 35 % vom ersten zum zweiten Messzeitpunkt entstand vor allem durch Abbruch bzw. Wechsel des Seminars. Um einen systematischen Ausfall in Hinblick auf die jeweils überprüften digitalen Kompetenzbereiche auszuschließen, wurde der MCAR-Test durchgeführt, welcher ein unsystematisches Fehlen aufzeigte ($\chi^2 = 44.741$; $df = 64$; $p > 0.05$).

Instrumente

In Anlehnung an die von der Kultusministerkonferenz (2016) beschriebenen Bereiche digitaler Kompetenzen wurden in allen Seminaren die digitalen Kompetenzeinschätzungen mit sieben Subskalen der Lehramtsstudierenden zu Beginn und zum Ende des Semesters erfasst (Rubach/Lazarides, under review). Das Antwortformat war für alle Items fünfstufig (1 = stimme gar nicht zu, 5 = stimme voll und ganz zu).

Die erste Subskala ‚*Suchen und Verarbeiten*' erfasst selbsteingeschätzte Fähigkeiten, sich im digitalen Raum zu organisieren, Informationen zu suchen und zu verarbeiten. Die Subskala enthielt zwei Items und zeigte eine akzeptable (T1: $\alpha = .68$) bis gute Reliabilität (T2: $\alpha = .81$). Ein Beispielitem ist „Ich kann auf Grundlage meiner Suchinteressen relevante Quellen in digitalen Umgebungen identifizieren und nutzen."

‚*Kommunikation und Kooperieren*‘ wird durch Items abgebildet, die die selbsteingeschätzte Fähigkeit erfassen, im digitalen Raum durch Prozesse des Teilens und Bearbeitens von Dateien zusammenzuarbeiten und geeignete Medien zur Kommunikation zu wählen. Die Subskala wurde mit vier Items erfasst und wies eine akzeptable Reliabilität auf (T1: $\alpha = .72$; T2: $\alpha = .79$). Ein Beispielitem ist „Ich wähle digitale Medien für gezielte Kommunikationsmöglichkeiten situationsgerecht aus.“

‚*Produzieren und Präsentieren*‘ bildet die Fähigkeit ab, Programme und Apps bedarfsgerecht zu nutzen und diese zur Weiterverarbeitung von Dateien zu nutzen. Die Subskala wurde mit zwei Items erfasst und hatte eine akzeptable (T1: $\alpha = .64$) bis gute Reliabilität (T2: $\alpha = .83$). Ein Beispielitem ist „Ich kenne mehrere Apps und Programme und kann diese bedarfsgerecht anwenden.“

‚*Schützen und sicher Agieren*‘ erfasst die Fähigkeit, reflektiert im digitalen Raum zu handeln und dabei die Relevanz der Gesundheit und Privatsphäre zu beachten. Die Subskala wurde mit drei Items erfasst und wies eine akzeptable Reliabilität auf (T1: $\alpha = .75$; T2: $\alpha = .70$). Ein Beispielitem ist „Ich kenne die Gefahren und Risiken in digitalen Umgebungen, berücksichtige und reflektiere diese.“

‚*Problemlösen und Handeln*‘ bildet die Fähigkeit ab, Programme im digitalen Raum bedarfsgerecht als Werkzeug zu nutzen, Lernprozesse durch die Nutzung digitaler Medien anzuregen und inhaltsbezogen im digitalen Raum zu handeln. Die Subskala wird mit vier Items erfasst und zeigte eine akzeptable (T1: $\alpha = .75$) bis gute Reliabilität (T2: $\alpha = .84$). Ein Beispielitem ist „Ich kann digitale Umgebungen und Werkzeuge zum persönlichen Gebrauch anpassen.“

‚*Analysieren und Reflektieren*‘ bildet die selbsteingeschätzte Fähigkeit ab, Handlungsmöglichkeiten im digitalen Raum zu kennen und analysieren zu können. Die Subskala wurde mit drei Items erfasst und wies ebenfalls eine akzeptable Reliabilität auf (T1: $\alpha = .75$; T2: $\alpha = .75$). Ein Beispielitem ist „Ich erkenne Chancen und Risiken des Mediengebrauchs und kann diese reflektieren.“

‚*Unterrichten und Implementieren*‘ beschreibt die selbsteingeschätzte Fähigkeit, Lernpotenziale digitaler Medien für die Unterrichtsgestaltung zu erkennen. Die Subskala wurde mit drei Items erfasst und wies eine gute Reliabilität auf (T1: $\alpha = .80$; T2: $\alpha = .89$). Ein Beispielitem ist „Ich erkenne die Potenziale der Nutzung digitaler Medien für die Vermittlung von Unterrichtsinhalten.“

Tabelle 1: Deskriptive Statistik, hier Mittelwerte zwischen den Interventionsgruppen

	Interventionsgruppe 1: Lehrkräfteperspektive				*Interventionsgruppe 2: Perspektive von Lernenden*				*Kontrollgruppe*			
	Pretest		*Posttest*		*Pretest*		*Posttest*		*Pretest*		*Posttest*	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
SUCH	4.18	0.72	4.36	0.50	3.94	0.58	4.12	0.74	4.17	0.54	4.25	0.54
KOMMU	4.00	0.64	4.57	0.30	4.06	0.55	4.41	0.49	4.38	0.48	4.54	0.46
PRODU	4.00	0.88	4.18	0.81	4.06	0.90	4.21	0.73	4.41	0.58	4.38	0.64
SCHUTZ	3.50	0.88	3.94	0.71	3.41	0.76	3.88	0.73	4.09	0.83	4.08	0.81
PROBL	3.10	0.96	3.89	0.72	2.97	0.48	3.72	0.58	3.55	0.59	3.80	0.69
ANALY	4.07	0.77	4.15	0.78	3.61	0.68	3.75	0.51	4.12	0.31	4.03	0.76
UNTERR	3.73	0.78	4.39	0.57	3.63	0.73	3.96	0.69	4.21	0.31	4.31	0.61

Anmerkung: SUCH: Suchen, Verarbeiten und Aufbewahren; KOMMU: Kommunizieren und Kooperieren; PRODU: Produzieren und Präsentieren; SCHUTZ: Schützen und sicher Agieren; PROBL: Problemlösen und Handeln; ANALY: Analysieren und Reflektieren; UNTERR: Unterrichten und Implementieren

Statistisches Vorgehen

Mit dem Ziel der Veränderungsmessung wurde in der vorliegenden Studie überprüft, inwieweit sich die digitalen Kompetenzeinschätzungen der Teilnehmer/innen der jeweiligen Gruppen im Vergleich vom ersten zum zweiten Messzeitpunkt verändert haben. Zur Überprüfung signifikanter Veränderungen der selbsteingeschätzten digitalen Kompetenzen von Lehramtsstudierenden von T1 zu T2 wurde der Wilcoxon-Vorzeichen-Rangtest genutzt.

5 Ergebnisse

In Tabelle 1 sind deskriptive Werte der Kompetenzbereiche beider Interventionsgruppen und der Kontrollgruppe verdeutlicht. Es ist auffällig, dass die selbsteingeschätzten digitalen Kompetenzen in den Interventionsgruppen im Posttest jeweils höher ausgeprägt sind als im Pretest. Aus diesem Grund ist zunächst zu vermuten, dass die Seminarteilnahme der Interventionsgruppen einen positiven Effekt auf die digitalen Kompetenzeinschätzungen der Lehramtsstudierenden ausübte. Bei der Betrachtung der Mittelwerte der Kontrollgruppe zeigt sich ein heterogenes Bild, da es zum Posttest sowohl höhere als auch geringere Einschätzungen eigener digitaler Kompetenzen im Vergleich zum Pretest gibt.

In Tabelle 2 sind die Ergebnisse von Unterschiedsanalysen des Wilcoxon-Vorzeichen-Rangtest aufgeführt. Die Ergebnisse der Vergleiche der Ränge über die Zeit in den einzelnen Gruppen zeigen, dass die digitalen Kompetenzeinschätzungen in der ersten Interventionsgruppe (Lehrkräfteperspektive) nach der Teilnahme am Seminar signifikant höher ausfallen als zu Beginn des Semi-

nars. Im Detail sind die Kompetenzeinschätzungen in den Bereichen *Kommunizieren und Kooperieren*, *Problemlösen und Handeln* als auch *Unterrichten und Implementieren* in dieser Gruppe zu T2 höher als zu T1. In der zweiten Interventionsgruppe (Perspektive von Lernenden) zeigten sich nach der Teilnahme am Seminar höhere Einschätzungen eigener digitaler Kompetenzen in den Bereichen *Kommunizieren und Kooperieren*, *Schützen und sicher Agieren* als auch *Problemlösen und Handeln* als zu Beginn des Seminars. In der Kontrollgruppe hingegen werden keine signifikanten Anstiege der selbsteingeschätzten digitalen Kompetenzfacetten deutlich.

Tabelle 2: Ergebnisse des Wilcoxon-Tests für die Interventionsgruppen

	Pretest	Posttest	Signifikanztest	
	Mittlerer Rang	*Mittlerer Rang*	*Z*	p
Interventionsgruppe 1: Lehrkräfteperspektive				
SUCH	5.50	3.10	-1.08[a]	.203
KOMMU	3.50	6.56	-2.33[a]	.010*
PRODU	2.83	4.17	-0.42[a]	.375
SCHUTZ	2.00	4.80	-1.72[a]	.055
PROBL	2.50	5.83	-2.57[a]	.005*
ANALY	3.50	4.67	0.00[a]	.563
UNTERR	6.50	4.81	-1.93[a]	.035*
Interventionsgruppe 2: Perspektive von Lernenden				
SUCH	4.50	7.93	-1.34[a]	.091
KOMMU	3.67	8.00	-2.43[a]	.006*
PRODU	6.40	6.57	-0.56[a]	.311
SCHUTZ	4.30	9.28	-1.96[a]	.026*
PROBL	3.00	8.36	-3.25[a]	.000*
ANALY	5.61	10.90	-0.13[a]	.466
UNTERR	7.60	8.91	-1.57[a]	.060
Kontrollgruppe				
SUCH	5.25	3.50	-0.63[a]	.383
KOMMU	4.00	5.50	-1.31[a]	.143
PRODU	3.75	2.50	0.00	.625
SCHUTZ	5.00	3.25	-0.17[b]	.500
PROBL	3.67	5.67	-1.37[a]	.094
ANALY	4.50	2.00	-0.41[b]	.375
UNTERR	4.50	5.40	-0.54[a]	.326

Anmerkung: Hinweise zu den Abkürzungen siehe Tabelle 1; * $p < 0.05$, [a] basiert auf negativen Rängen, [b] basiert auf positiven Rängen

6 Diskussion

Zusammenfassend zeigen die Ergebnisse, dass durch Seminare mit unterschiedlichen didaktischen Ansätzen eine Förderung digitaler Kompetenzeinschätzungen bei Lehramtsstudierenden stattfinden kann. Schlussfolgernd wird angenommen, dass die Lehrkräftebildung zur Entwicklung digitaler Kompetenzeinschätzungen von Lehramtsstudierenden beitragen kann, wenn differenzierte Förderangebote unter Berücksichtigung konkreter digitaler Kompetenzbereiche umgesetzt werden.

Interessant ist, dass ausschließlich Teilnehmende der Interventionsgruppe 1 (Lehrkräfteperspektive) zum zweiten Messzeitpunkt höhere Einschätzungen im Kompetenzbereich ‚Unterrichten und Implementieren' berichten. Hier könnte in Anlehnung an Bastian und Aufenanger (2015) sowie Bauer (2011) angenommen werden, dass die Kompetenz, Medien didaktisch sinnvoll im eigenen Fach-Unterricht einzusetzen, dann gefördert wird, wenn Studierende im Rahmen ihres Studium die „Lehrkräfteperspektive" einnehmen – hier zählt die aktive Auseinandersetzung mit Möglichkeiten des Medieneinsatzes sowie die Planung des Einsatzes. Ein weiterer Erklärungsansatz ist die Verknüpfung von pädagogischem, inhaltlichem und technischem Wissen. Im Seminar mit dem Lehrformat „Lehrkräfteperspektive" entwickelten Lehramtsstudierende fachbezogenen Unterricht unter Nutzung digitaler Medien. Hier können also Annahmen des TPCK-Modells bestätigt werden – erst die Verknüpfung aller drei Wissensfacetten wirkt auf Kompetenzen zum Einsatz digitaler Medien (Mishra/Koehler 2006). Eine mögliche Implikation dieses Befundes ist, dass Studierende angeleitet werden sollten, fachbezogene Unterrichtskonzepte unter Berücksichtigung digitaler Medien zu entwickeln, wenn Lehrkräftebildung darauf abzielt, Lehramtsstudierende im Einsatz digitaler Medien im Unterricht kompetent zu schulen.

In beiden Seminarkonzepten ist der didaktisch sinnvolle Einsatz von Medien positiv für die Förderung der Kompetenzbereiche ‚Kommunizieren und Kollaborieren' und ‚Problemlösen und Handeln'. Das schließt an internationale Studien an, welche hervorheben, dass sogenannte 21st Century Learning Skills wie eben Kooperation und Problemlösen im digitalen Raum durch eine reflektierte Verwendung digitaler Technologien gefördert werden (Niess/Gillow-Wiles 2017). Angenommen werden könnte, dass für diese Kompetenzfacetten besonders Reflexionsprozesse relevant sind, die in beiden Seminargruppen stattfanden.

Eine Limitation der vorliegenden Studie ist, dass nur Studierende einer Universität untersucht wurden. Zukünftig sollten Interventionen jedoch in größeren Stichproben und an mehreren Standorten implementiert und evaluiert werden, um die Generalisierbarkeit ihrer Wirksamkeit zu prüfen. Des Weiteren basiert die Studie ausschließlich auf Selbsteinschätzungen eigener

Kompetenzen. Für weitere Interventionen wäre auch der Bezug zu Fachinhalten und pädagogischen Inhalten stärker zu fokussieren. Trotz dieser Limitationen leistet die vorliegende Studie bedeutsame Hinweise zur Konzeption und Entwicklung von Angeboten in der Lehrkräftebildung, welche als Fördermöglichkeiten digitaler Kompetenzeinschätzungen angesehen werden können.

Literatur

Bastian, Jasmin/Aufenanger, Stefan (2015): Medienbezogene Vorstellungen von (angehenden) Lehrpersonen. In: Schiefner-Rohs, Mandy/Gómez Tutor, Claudia/Menzer, Christine (Hrsg.): Lehrer.Bildung.Medien. Herausforderungen für die Entwicklung und Gestaltung von Schule (Grundlagen der Berufs- und Erwachsenenbildung). Baltmannsweiler: Schneider-Verlag, S. 19–34.

Bauer, Petra (2011): Vermittlung von Medienkompetenz und medienpädagogischer Kompetenz in der Lehrerausbildung. In Köhler, Thomas/Neumann, Jörg (Hrsg.): Wissensgemeinschaften. Digitale Medien – Öffnung und Offenheit in Forschung und Lehre. Münster: Waxmann, S. 294–303.

Bertelsmann Stiftung (2017): Studierende sind keine digitalen Enthusiasten. Digitalisierung ist in den deutschen Hochschulen angekommen. Technisch sind diese zwar gut ausgestattet. Didaktische Potentiale aber bleiben trotz guter Infrastruktur oft ungenutzt. https://www.bertelsmann-stiftung.de/de/presse/pressemitteilungen/pressemitteilung/pid/studierende-sind-keine-digitalen-enthusiasten/ (Abfrage: 16.05.2019)

Brendel, Nina (2017): Reflexives Denken im Geographieunterricht. Eine empirische Studie zur Bestimmung von Schülerreflexion mithilfe von Weblogs im Kontext Globalen Lernens. Münster: Waxmann.

Eickelmann, Birgit (2010): Digitale Medien in Schule und Unterricht erfolgreich implementieren. Eine empirische Analyse aus Sicht der Schulentwicklungsforschung (Empirische Erziehungswissenschaft, Bd. 19). Zugl.: Dortmund, Techn. Univ., Diss., 2009. Münster: Waxmann.

Eickelmann, Birgit/Lorenz, Ramona/Endberg, Manuela (2016): Die Relevanz der Phasen der Lehrerausbildung hinsichtlich der Vermittlung didaktischer und methodischer Kompetenzen für den schulischen Einsatz digitaler Medien in Deutschland und im Bundesländervergleich. In: Bos, Wilfried/Lorenz, Ramona/Endberg, Manuela/Eickelmann, Birgit/Kammerl, Rudolf/Welling, Stefan (Hrsg.): Schule digital – der Länderindikator 2016. Kompetenzen von Lehrpersonen der Sekundarstufe I im Umgang mit digitalen Medien im Bundesländervergleich. Münster: Waxmann, S. 148–179.

Farjon, Dann/Smits, Anneke/Voogt, Joke (2019): Technology integration of pre-service teachers explained by attitudes and beliefs, competency, access, and experience. In: Computers & Education 130, S. 81–93.

Heinen, Richard/Kerres, Michael/Schiefner-Rohs, Mandy (2013): Bring your own device. Private, mobile Endgeräte und offene Lerninfrastrukturen an Schulen. In: Karpa, Dietrich/Eickelmann, Birgit/Grafe, Silke (Hrsg.): Digitale Medien und Schule. Zur Rolle digitaler Medien in Schulpädagogik und Lehrerbildung. Immenhausen bei Kassel: Prolog-Verlag, S. 129–145.

Herzig, Bardo (2007): Medienpädagogik als Element professioneller Lehrerausbildung. In: Sesink, Werner/Kerres, Michael/Moser, Heinz (Hrsg.): Jahrbuch Medienpädagogik 6. Medienpädagogik – Standortbestimmung einer erziehungswissenschaftlichen Disziplin (Jahrbuch Medien-Pädagogik). Wiesbaden: VS Verlag für Sozialwissenschaften, S. 283–297.

Kaufman, Kristina (2014): Information Communication Technology. Challenges & Some Prospects from Preservice Education to the Classroom. In: Mid-Atlantic Education Review, 2, H. 1, S. 1–11.

Klassen, Robert M./Tze, Virginia M. C. (2014): Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. In: Educational Research Review 12, S. 59–76.

Koh, Joyce H. L./Divaharan, Shanti (2013): Towards a TPACK-fostering ICT instructional process for teachers. Lessons from the implementation of interactive whiteboard instruction. In: Australasian Journal of Educational Technology 29, H. 2, S. 233–247.

Krauskopf, Karsten/Zahn, Carmen/Hesse, Friedrich. W. (2015): Cognitive Processes Underlying TPCK: Mental Models, Cognitive Transformation, and Meta-conceptual Awareness. In: Angeli, Charoula/Valanides, Nicos (Hrsg.): Technological pedagogical content knowledge. Exploring, developing, and assessing TPCK. New York: Springer, S. 41–61.

Kultusministerkonferenz (2012): Medienbildung in der Schule. Beschluss der Kultusministerkonferenz vom 8. März 2012. http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2012/2012_03_08_Medienbildung.pdf (Abfrage: 16.05.2019)

Kultusministerkonferenz (2016): Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2016/Bildung_digitale_Welt_Webversion.pdf (Abfrage: 16.05.2019)

Mishra, Punya/Koehler, Matthew J. (2006): Technological Pedagogical Content Knowledge. A Framework for Teacher Knowledge. In: Teachers College Record 108, H. 6, S. 1017–1054.

Rubach, Charlott/Lazarides, Rebecca (under review). Entwicklung und Validierung einer Skala zur Selbsteinschätzungen digitaler Kompetenzen im Umgang mit Medien bei Lehramtsstudierenden.

Schiefner-Rohs, Mandy (2015): Lehrerinnen- und Lehrerbildung und digitale Medien – Herausforderungen entlang der Lehrerbildungskette. In Schiefner-Rohs, Mandy/Gómez Tutor, Claudia/Menzer, Christine (Hrsg.): Lehrer.Bildung.Medien. Herausforderungen für die Entwicklung und Gestaltung von Schule. Baltmannsweiler: Schneider-Verlag, S. 119–128.

Tondeur, Jo/Roblin, Natalie Pareja/van Braak, Johan/Fisser, Petra/Voogt, Joke (2013): Technological pedagogical content knowledge in teacher education: in search of a new curriculum. In: Educational Studies 39, H. 2, S. 239–243.

Tondeur, Jo/Scherer, Ronny/Baran, Evrim/Siddiq, Fazilat/Valtonen, Teema/Sointu, Erkko (2019): Teacher educators as gatekeepers: Preparing the next generation of teachers for technology integration in education. In: British Journal of Educational Technology 50, H. 3, S. 1189–1209.

Voogt, Joke/McKenney, Susan (2017): TPACK in teacher education: are we preparing teachers to use technology for early literacy? In: Technology, Pedagogy and Education 26, H. 1, S. 69–83.

Zusammenhänge zwischen handlungsorientierten und thematischen Lerngelegenheiten und der DaZ-Kompetenz angehender Lehrkräfte

Tobias Schroedler und Ilse Stangen

Abstract

Kompetenzen für den Umgang mit Mehrsprachigkeit und Schülerinnen und Schülern, die Deutsch als Zweitsprache sprechen, sind für (zukünftige) Lehrerkräfte an deutschen Schulen von elementarer Wichtigkeit. Im Projekt DaZKom wurde ein theoretisch begründetes Modell zur Messung von Kompetenzen angehender Lehrkräfte im Umgang mit Deutsch als Zweitsprache entworfen und anschließend der DaZKom-Test entwickelt. An der Universität Hamburg wurde der Test im Prä-Post-Design auf das Seminar „Grundlagen der Sprachbildung für den Fachunterricht" mit insgesamt N = 296 Studierenden angewendet, um weiterführende Reformen der Lehrerkräftebildung für den Umgang mit Sprache im Fachunterricht an der Universität zu überprüfen. Dem theoretischen Ansatz der Angebots-Nutzungs-Modelle folgend hängen Kompetenzzuwächse Lernender neben den individuellen Eingangsvoraussetzungen und institutionellen Rahmenbedingungen auch mit der Nutzung entsprechender Lerngelegenheiten zusammen. Die Ergebnisse der ersten Teilkohorte (N = 143) zeigten, dass universitäre, thematische Lerngelegenheiten zum Umgang mit Deutsch als Zweitsprache und Mehrsprachigkeit signifikant positiv mit der gemessenen Kompetenz zusammenhängen. Mit der Gesamtkohorte (N = 296) konnten diese Ergebnisse repliziert werden, jedoch zeigte sich für die durchgeführten Analysen der handlungsorientierten Lerngelegenheiten ein weniger aussagekräftiges Bild.

1 Problemstellung

Lehrkräfte an deutschen Schulen sind zunehmend konfrontiert mit Schülerinnen und Schülern, die Deutsch als Zweitsprache (DaZ) sprechen, weshalb

Kompetenzen für den Umgang mit DaZ und Mehrsprachigkeit im Fachunterricht von elementarer Wichtigkeit sind. Unter anderem als Folge der ersten PISA-Ergebnisse, welche dem deutschen Schulsystem mangelnde Chancengleichheit von Kindern mit Migrationshintergrund attestierten, reagierten die Bundesländer und lehramtsausbildenden Hochschulen in Deutschland mit einer Umstrukturierung hin zu bzw. einer Etablierung von Angeboten, in welchen Lehramtsstudierende ebensolche Kompetenzen im Bereich DaZ/Mehrsprachigkeit erwerben können (vgl. Baumann/Becker-Mrotzek 2014; vgl. Baumann 2017).

An der Universität Hamburg wird seit 2016 auf ein (für diese Universität) neuartiges Konzept eines integrierten Modells zur Lehrerkräftebildung für den Umgang mit Mehrsprachigkeit im Fachunterricht gesetzt. Durch die Entwicklung dieses Modells wird die Thematik bewusst als Querschnittsthema in der Masterphase des Lehramtsstudiums verankert. Studierende besuchen ein Grundlagenseminar, in welchem sie fachübergreifendes Grundlagenwissen über Mehrsprachigkeit sowie Sprachbildung im Fachunterricht erhalten (vgl. Schroedler/Grommes 2019). Auf dieses Grundlagenwissen kann dann in fachdidaktischen Begleitveranstaltungen fachspezifisch eingegangen werden (vgl. Schroedler/Lengyel 2018). Darüber hinaus können Studierende in Wahlpflichtmodulen weitere Veranstaltungen im Themenfeld Mehrsprachigkeit/sprachlich-kulturelle Heterogenität belegen.

Um die oben beschriebenen Reformbemühungen für die DaZ-/Mehrsprachigkeitsausbildung angehender Lehrkräfte in Hamburg zu überprüfen und um die Kompetenzentwicklung Lehramtsstudierender zu messen, wurde – wie an einer Reihe weiterer lehramtsausbildender Universitäten – auch an der Universität Hamburg seit 2016 der DaZKom-Test eingesetzt. Erste Ergebnisse der Evaluationsstudie mit Daten der ersten Teilkohorte wurden in Stangen/Schroedler/Lengyel (2019, im Druck) publiziert. Eine zentrale Frage, der sowohl in diesem Beitrag, als auch in der zwar zunehmenden, jedoch immer noch überschaubaren Literatur nachgegangen wird, ist die nach möglichen Prädiktoren für Kompetenzzuwächse.

Die hier vorgestellte Studie untersucht den Zusammenhang zwischen handlungsorientierten und thematischen Lerngelegenheiten und der DaZ-Kompetenz angehender Lehrkräfte der mathematisch-naturwissenschaftlichen Fächer. Die publizierten Ergebnisse der Messung der Kompetenzentwicklung angehender Lehrerkräfte mit Hilfe des DaZKom-Tests (vgl. Hammer et al. 2015; vgl. Ehmke et al. 2018), zeigen unterschiedliche Befunde. Stangen/Schroedler/Lengyel (2019, im Druck) berichten beispielsweise, dass DaZ-bezogene thematische Lerngelegenheiten, selbst regressionsanalytisch und unter Kontrolle einer Reihe weiterer Variablen, den zweitstärksten vorhersagenden Beitrag auf die gemessenen Kompetenzzuwächse haben (den stärksten Prädiktor für die Post-Testergebnisse in dieser Studie stellen die Prä-Testergebnisse

dar). In der bislang umfangreichsten veröffentlichten Studie, in welcher Daten von zwölf Universitätsstandorten gemeinsam überprüft wurden, bestehen zwar signifikante bivariate Korrelationen zwischen der DaZ-Kompetenz und ebendiesen thematischen Lerngelegenheiten, regressionsanalytisch sind sie jedoch kein signifikanter Prädiktor für gute Ergebnisse im DaZKom-Test (vgl. Ehmke/Lemmrich 2018). Paetsch et al. (2019) berichten, dass unter Berliner Studierenden, die am DaZKom-Test teilnahmen, zumindest die außeruniversitären Lerngelegenheiten keinen signifikanten Prädiktor darstellten.

2 Theoretischer Hintergrund

Dem theoretischen Ansatz der Angebots-Nutzungs-Modelle (vgl. Helmke/Schrader 2010) folgend, hängen Kompetenzzuwächse Lernender neben den individuellen Eingangsvoraussetzungen und institutionellen Rahmenbedingungen auch mit der Nutzung entsprechender Lerngelegenheiten zusammen. Inwieweit dies für den universitären Kontext und im Besonderen für den Bereich der DaZ-Kompetenz zutrifft, wurde bisher jedoch kaum untersucht. Auch für andere Wissensbereiche von Lehramtsstudierenden wurden Zusammenhänge mit universitären Lerngelegenheiten eher selten untersucht (vgl. Doll et al. 2018) und die Ergebnisse aktueller Studien weisen in unterschiedliche Richtungen.

So fanden Kunina-Habenicht et al. (2013) keine statistisch signifikanten Zusammenhänge zwischen den genutzten Lerngelegenheiten und dem pädagogischen Wissen; in den Studien von Tachtsoglou/König (2018) konnten dagegen mittlere Zusammenhänge gezeigt werden. In einer umfangreichen Studie mit fast 1300 Lehramtsstudierenden an 18 Universitätsstandorten in Deutschland und Österreich, fanden König et al. (2017) auf der Fallebene statistisch signifikante Effekte der Nutzung pädagogischer Lerngelegenheiten für den Umgang mit Heterogenität. Jentsch et al. (eingereicht) untersuchten neben dem Einfluss sprachbezogener Merkmale der Studierenden auch den Zusammenhang zwischen der Nutzung von Lerngelegenheiten und dem deutschdidaktischen (N = 320) und erziehungswissenschaftlichen (N = 1698) Professionswissen. In dieser Studie konnte für beide Gruppen ein positiver Zusammenhang zwischen der Nutzung universitärer Lerngelegenheiten und dem deutschdidaktischen und erziehungswissenschaftlichen Wissen gezeigt werden.

In Bezug auf die Entwicklung relevanter Kompetenzen von Lehramtsstudierenden für den Umgang mit DaZ und Mehrsprachigkeit im Unterricht, kommt es je nach Universitätsstandort, Curriculum und individueller Neigung zu unterschiedlich stark genutzten Lerngelegenheiten. Naturgemäß unterscheiden sich Lehramtsstudiengänge stark zwischen den Bundesländern und zusätzlich auch zwischen den Institutionen innerhalb eines Bundeslandes. Somit unter-

scheiden sich auch die Gelegenheiten für Studierende, sich mit DaZ-/mehrsprachigkeitsrelevanten Inhalten in ihrem Studium zu befassen, hinsichtlich ihrer Struktur, Menge und Qualität (vgl. Ehmke/Lemmrich 2018).

Im Projekt DaZKom wurde zwischen 2012 und 2015 ein theoretisch begründetes Modell zur Messung von Kompetenzen angehender Lehrkräfte im Umgang mit DaZ entworfen (vgl. Koch-Priewe 2018). Dies geschah unter anderem im Kontext der seit den 1990er Jahren zunehmenden Lehrkräfteprofessionsforschung sowie dem Versuch, an Hochschulen erlangte Kompetenzen messbar zu machen. Basierend auf diesem Modell wurde anschließend der DaZKom-Test entwickelt, welcher das theoretische Kompetenzkonstrukt valide misst (vgl. Hammer et al. 2015; vgl. Carlson/Präg 2018).

Theoretisch fundiert der DaZKom-Test auf einem Modellkonstrukt bestehend aus einem Struktur- und einem Entwicklungsmodell (vgl. Ohm 2018). Das Strukturmodell basiert auf Analysen relevanter Curricula für die DaZ-/Mehrsprachigkeitsausbildung deutscher Lehramtsstudierender und besteht aus drei Dimensionen (Fachregister, Mehrsprachigkeit und Didaktik), welche sich in sechs Subdimensionen einteilen lassen, welche wiederum aus 20 Facetten bestehen (vgl. Ohm 2018, S. 74). Im Entwicklungsmodell wird die DaZ-Kompetenz als „kontextabhängige Leistungsdisposition" (Ohm 2018, S. 73) erfasst. Es basiert auf dem *five stage model of skill acquisition* nach Dreyfus und Dreyfus (1986). So wird die gemessene DaZ-Kompetenz mit Hilfe von Aufgaben entlang der oben beschriebenen Dimensionen erfasst und kategorisiert nach dem Stufenmodell des Fähigkeitserwerbs (ausführlichere Beschreibungen finden sich in Ohm 2018 sowie in weiteren Kapiteln in Ehmke et al. 2018).

Wie oben berichtet, zeigen Studien der Lehrkräfteprofessionsforschung (in anderen Bereichen) signifikante Zusammenhänge zwischen der Nutzung universitärer Lerngelegenheiten und dem Kompetenzerwerb. Für den spezifischen Bereich der universitären DaZ-bezogenen Lerngelegenheiten werden Zusammenhänge mit der DaZ-Kompetenz von Ehmke/Lemmrich (2018) berichtet. Allerdings handelt es sich dabei um Daten aus einem querschnittlichen Design (ibid., S. 213). Hingegen wird in der hier vorliegenden Untersuchung der Aufbau von DaZ-Kompetenz entlang der universitären Lerngelegenheiten im Längsschnitt untersucht.

Wie angedeutet, berichten Stangen/Schroedler/Lengyel (2019, im Druck) in ihrer an der Universität Hamburg durchgeführten Studie sowohl eine Vielzahl signifikanter Zuwächse der wahrgenommenen Lerngelegenheiten in ihrem quasi-Längsschnitt als auch eine beträchtliche Vorhersagekraft der wahrgenommenen Lerngelegenheiten für die erworbene Kompetenz ihrer Studierenden. Die Lerngelegenheiten unterteilen sich in theoretische (16 Items) und handlungsorientierte Lerngelegenheiten (acht Items), welche von Ehmke und anderen entwickelt und erprobt wurden (vgl. Ehmke/Lemmrich 2018). Die thematischen Lerngelegenheiten fragen nach der Häufigkeit des Auftretens der

im Studium behandelten Inhalte mit Relevanz für die DaZ-Kompetenz (beispielsweise Inhalte der theoretischen Linguistik oder Methoden des sprachbewussten Unterrichts wie Scaffolding). Die handlungsorientierten Lerngelegenheiten fragen danach, wie häufig im Studium der am Test teilnehmenden Studierenden relevante Praxisaufgaben, wie beispielsweise das Erstellen von Sprachförderplänen für Schülerinnen und Schüler mit DaZ, behandelt wurden. In einem Regressionsmodell, welches ein korrigiertes R^2 von .45 liefert, haben die Lerngelegenheiten einen Anteil von .21. Die Befunde beschränken sich allerdings lediglich auf die erste untersuchte Teilkohorte (N = 143) und betrachten ausschließlich die 16 Items der Skala der thematischen Lerngelegenheiten. Im Folgenden werden Analysen anhand der Hamburger Gesamtkohorte (N = 296) aufgezeigt, welche sowohl die Entwicklungen der thematischen als auch der handlungsorientierten Lerngelegenheiten beleuchten.

3 Fragestellungen

In Stangen/Schroedler/Lengyel (2019, im Druck) deutete sich für die erste Teilkohorte (N = 143) an, dass sich eine Vielzahl (elf von 16) der wahrgenommenen thematischen Lerngelegenheiten zwischen den Testzeitpunkten signifikant erhöht. Ferner zeigte sich, dass im Prä-Test lediglich eine von 16 Lerngelegenheiten signifikant mit dem Testergebnis des DaZKom-Tests korreliert. Im Post-Test waren es dann zwölf von 16 thematischen Lerngelegenheiten. Hier schließt sich die erste Forschungsfrage dieses Beitrags an:

(i) Können die Ergebnisse der Teilkohorte (N = 143) aus Stangen/Schroedler/Lengyel (2019, im Druck) im Bereich der thematischen DaZ-bezogenen Lerngelegenheiten für die Gesamtkohorte (N = 296) repliziert werden?

Da sich in explorativen Analysen zunächst kaum Zusammenhänge für die handlungsorientierten Lerngelegenheiten zeigten, wurde in den Arbeiten zur ersten Teilkohorte (N = 143) in Stangen/Schroedler/Lengyel (2019, im Druck) auf belastbare Analysen ebendieser Skala verzichtet. Für ein vollumfänglicheres Bild der Daten der nun zur Verfügung stehenden Gesamtkohorte (N = 296), werden ebendiese acht Items ebenfalls analysiert und vorgestellt. Ferner soll im finalen Teil dieses Beitrags diskutiert werden, in welchem Zusammenhang die handlungsorientierten Lerngelegenheiten mit der DaZ-Kompetenz stehen (im Vergleich zu den thematischen Lerngelegenheiten).

(ii) Zeigen sich die Zusammenhänge zwischen den wahrgenommenen Lerngelegenheiten und der DaZ-Kompetenz auch für die handlungsorientierten DaZ-bezogenen Lerngelegenheiten?

4 Methodisches Vorgehen

Seit Beginn des Jahres 2016 wird im Rahmen des Projekts ProfaLe[1] die Hamburger Lehrkräftebildung in verschiedenen Teilfeldern reformiert. Wie oben beschrieben, beinhalten diese Reformbemühungen die Etablierung eines integrierten Modells für Sprachbildung im Fachunterricht, in welchem der Themenbereich DaZ/Mehrsprachigkeit bewusst als Querschnittsthema in mehreren Veranstaltungen verankert sein soll. Aus dieser Struktur, in welcher bewusst auf ein explizites additives DaZ-Modul verzichtet wird, ergibt sich ein ganz besonderes Interesse an einer Messung von DaZ- bzw. mehrsprachigkeitsrelevanten Lerngelegenheiten.

Zwischen dem Sommersemester 2016 und dem Wintersemester 2017/18 haben rund 300 Lehramtsstudierende der Universität Hamburg an dem oben genannten Seminar „Grundlagen der Sprachbildung für den Fachunterricht" teilgenommen (vgl. Schroedler/Grommes 2019). Im Rahmen der Seminarveranstaltung wurde der DaZKom-Test (vgl. Hammer et al. 2015) im Prä-Post-Design mit allen Studierenden am Semesterbeginn und nochmal am Ende der darauffolgenden vorlesungsfreien Zeit, anschließend an ein vier- bis sechswöchiges Schulpraktikum, circa sechs Monate später durchgeführt. Ziel war es, den Kompetenzerwerb über einen Zeitraum abzubilden, in dem Studierende sowohl das beschriebene Grundlagenseminar als auch verknüpfte fachdidaktische Begleitseminare sowie möglicherweise weitere Wahlpflichtseminare mit Sprachbezug besuchten. Hierbei wurden, neben der DaZ-Kompetenzmessung, auch die Nutzung von handlungsorientierten und thematischen Lerngelegenheiten im Bereich DaZ/Mehrsprachigkeit durch 24 Items (16 thematische DaZ-bezogene und acht handlungsorientierte DaZ-bezogene Lerngelegenheiten) anhand einer fünfstufigen Skala (von 1 = gar nicht bis 5 = in mehreren Lehrveranstaltungen) erhoben. Dies wurde durch vier Skalen im DaZKom-Test operationalisiert, welche unterschiedlich viele Items zu universitären und außeruniversitären Berührungspunkten mit DaZ beinhalten (eine präzise Schilderung der Skalen und Items findet sich in Ehmke und Lemmrich 2018, S. 204–206). Die so erhobenen Lerngelegenheiten wurden auf bivariate Zusammenhänge mit der DaZ-Kompetenz, welche methodisch auf einer IRT-Skalierung der Textitems basiert, zu beiden Messzeitpunkten analysiert. Ferner wurde die Entwicklung der Lerngelegenheiten zwischen den Testzeitpunkten beobachtet und mittels T-Tests auf ihre Signifikanz überprüft.

---

1 Das Projekt *Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen* (ProfaLe) der Universität Hamburg wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung" von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung gefördert.

5 Ergebnisse

In Anlehnung an die in Ehmke und Lemmrich (2018) berichteten Skalen- und Itemkennwerte für die im DaZKom-Test gemessenen Lerngelegenheiten, zeigt Tabelle 1 die Kennzahlen für die Hamburger Gesamtkohorte (N = 296). Die Skalen zu den Lerngelegenheiten wurden von Ehmke und anderen entwickelt (vgl. ibid.). Ein Vergleich mit der Normierungsstudie (N = 496) zeigt, dass die Kennwerte der Skalen in einem ähnlichen Bereich liegen. Der Mittelwert der thematischen Lerngelegenheiten ist um 0.1 höher, der der Handlungen um 0.31. Die Standardabweichungen sind etwas niedriger, Cronbachs α nahezu identisch. Eine Überprüfung auf Signifikanz der Unterschiede lässt die Datenverfügbarkeit derzeit nicht zu.

Tabelle 1: Skalenkennwerte zu DaZ-bezogenen Lerngelegenheiten

Skalen/Items	Min	Max	MW	SD	α
DaZ-Themen	1.38	5.00	2.78	0.68	0.89
DaZ-Handlungen	1.00	4.88	2.10	0.61	0.80

Ebenfalls in Anlehnung an Ehmke und Lemmrich (2018), zeigen die Abbildungen 1 und 2 die prozentualen Häufigkeiten der Nutzung von DaZ-bezogenen Lerngelegenheiten (thematische Lerngelegenheiten in Abbildung 1 und handlungsorientierte Lerngelegenheiten in Abbildung 2). Die Abbildungen nennen die jeweiligen Lerngelegenheiten in der absteigenden Reihenfolge ihrer Mittelwerte. Hierbei zeigen sich einige interessante Unterschiede zwischen den Hamburger Daten und der Normierungsstichprobe. So sind bei den thematischen Lerngelegenheiten in der Normierungsstichprobe Inhalte zu „Linguistik“ die am meisten wahrgenommenen Lerngelegenheiten. In Hamburg ist der Mittelwert ebendieses Items lediglich an Position 7. Konträr dazu sind Themen wie Bildungssprache oder Scaffolding in der Normierungsstudie relativ weit unten zu finden, in Hamburg jedoch augenscheinlich deutlich prominenter.

Abbildung 1: Prozentuale Häufigkeit der Nutzung von thematischen Lerngelegenheiten (Posttest)

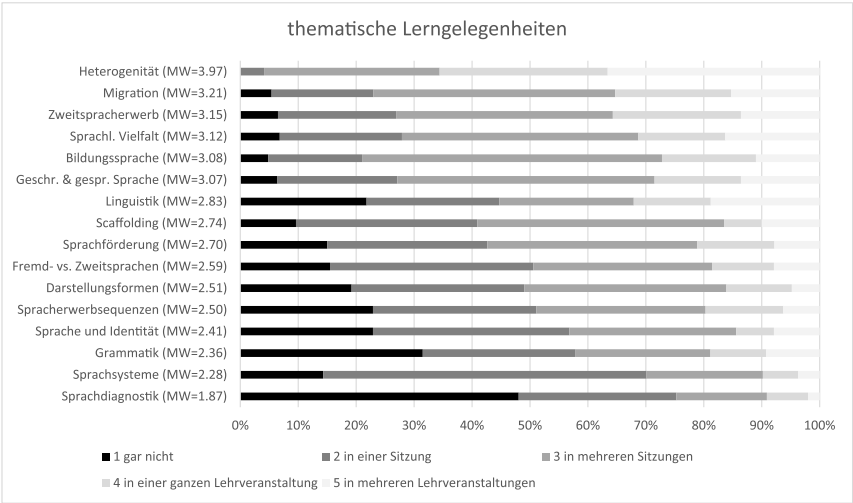
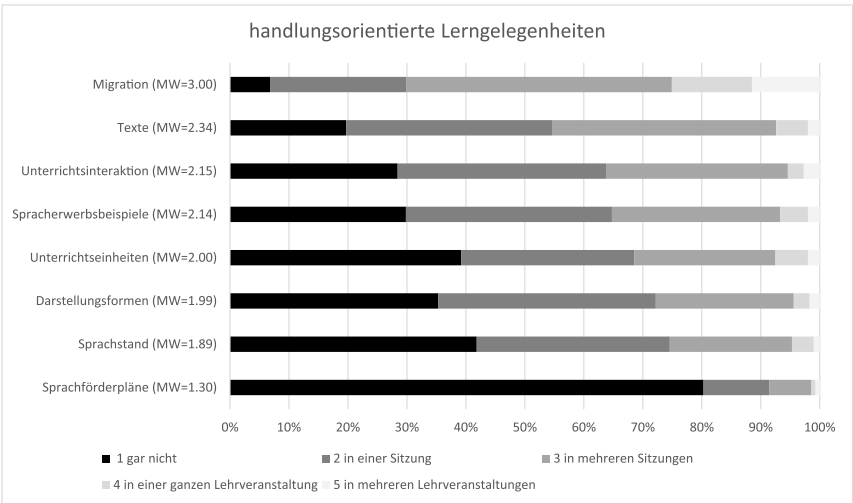


Abbildung 2: Prozentuale Häufigkeit der Nutzung von handlungsorientierten Lerngelegenheiten (Posttest)



Wie in den Forschungsfragen ausgeführt, soll in diesem Beitrag überprüft werden, inwiefern die Lerngelegenheiten mit der gemessenen DaZ-Kompetenz zusammenhängen. In Analysen für die erste Teilkohorte zeigte sich, dass sich elf der thematischen Lerngelegenheiten zwischen Prä- und Posttest signifikant erhöhten und zwölf ebendieser Lerngelegenheiten signifikant mit dem Testergebnis (im Posttest) korrelierten (nach Spearman). In Bezug auf die wahrge-

nommenen thematischen Lerngelegenheiten und deren Zusammenhang mit der gemessenen DaZ-Kompetenz zeigen sich in der Gesamtkohorte ähnliche Ergebnisse.

Tabelle 2: Entwicklung der thematischen Lerngelegenheiten und Zusammenhang mit erreichter DaZ-Kompetenz, absteigend nach den Mittelwerten zu Messzeitpunkt 2 (MWt_2)

Inhalt	**MWt_1**	**SDt_1**	**MWt_2**	**SDt_2**	**ΔMW**	**r t_1**	**r t_2**
Umgang mit Heterogenität	3.92	1.008	3.97	.915	.04	n.s.	n.s.
Migration und Mehrsprachigkeit	2.88	1.195	3.21	1.076	.33**	n.s.	.171**
Phänomene des Zweitspracherwerbs	2.46	1.241	3.15	1.098	.69**	n.s.	.167**
Sprachliche Vielfalt in der Schule	2.91	1.192	3.12	1.128	.21**	n.s.	.137*
Erwerb von Bildungssprache	2.66	1.126	3.08	.973	.42**	n.s.	.131*
Unterschiede zwischen mündlich und schriftlich geprägter Sprache	2.76	1.297	3.07	1.074	.31**	n.s.	.196**
(Teil-)Gebiete der Linguistik (z. B.: Syntax, Semantik, Morphologie)	2.82	1.528	2.83	1.403	.01	.238**	.229**
Unterstützung des sprachlichen Lernprozesses durch Scaffolding	1.83	1.097	2.74	1.052	.91**	n.s.	.176**
Sprachförderung	2.36	1.154	2.70	1.116	.34**	n.s.	n.s.
Unterschiede zwischen Fremd- und Zweitspracherwerb	2.22	1.133	2.59	1.133	.37**	n.s.	.124°
Sprachliche Anforderungen verschiedener Darstellungsformen	2.30	1.132	2.51	1.073	.21**	n.s.	n.s.
Erwerbsequenzen sprachlicher Entwicklung	2.18	1.194	2.50	1.166	.32**	n.s.	.203**
Sprache und Identität	2.25	1.166	2.41	1.145	.16*	n.s.	.123*
Grammatik des Deutschen	2.08	1.322	2.36	1.273	.28**	.173**	.152**
Sprachsysteme von Zuwanderungssprachen	1.51	.852	2.28	.918	.77**	n.s.	.118*
Sprachstandsdiagnose	1.81	1.111	1.87	1.046	.05	n.s.	.121*
Gesamt	2.42	0.689	2.76	0.690	.34**	n.s.	.222**

Anmerkung: Die markierten Δ*MW*-Werte sind mit °p< .10, *p < .05, **p < .01, signifikant unterschiedlich, die markierten r Werte korrelieren mit °p< .10, *p < .05, **p < .01 mit dem DaZKom-Testergebnis

Tabelle 2 zeigt, dass 13 der 16 Lerngelegenheiten zwischen den Messzeitpunkten signifikant zunehmen; während zu Messzeitpunkt 1 lediglich zwei Lerngelegenheiten signifikant mit dem Testergebnis korrelieren, sind es zum zweiten Messzeitpunkt zwölf. So zeigen die Analysen zur Gesamtkohorte sogar ein ähnliches Bild, wie die zur ersten Teilkohorte. Es ist anzumerken, dass bei allei-

nigen Analysen der zweiten Teilkohorte weniger Lerngelegenheiten signifikant mit dem Testergebnis korrelieren.

Tabelle 3: Entwicklung der handlungsorientierten Lerngelegenheiten und Zusammenhang mit erreichter DaZ-Kompetenz, absteigend nach MWt_2

Handlung	**MWt_1**	**SDt_1**	**MWt_2**	**SDt_2**	**ΔMW**	**r t_1**	**r t_2**
Sich zum Thema *Migration* und Mehrsprachigkeit im schulischen Kontext auseinandergesetzt.	2.83	1.133	3.00	1.050	.17*	n.s.	.138*
Für das eigene Studienfach typische *Texte* hinsichtlich ihrer sprachlichen Besonderheiten für SchülerInnen mit DaZ analysiert	1.76	1.013	2.34	.925	.58**	n.s.	.174**
Die Verwendung von Sprache in authentischen *Unterrichtsinteraktionen* analysiert	1.93	1.081	2.15	.962	.22**	n.s.	n.s.
Konkrete *Spracherwerbsbeispiele* von Lernenden mit DaZ analysiert?	1.60	.885	2.14	.969	.54**	n.s.	n.s.
Sprachsensible *Unterrichtsstunde(n)* entworfen	1.61	.961	2.00	1.019	.39**	n.s.	n.s.
Typische *Darstellungsformen* und ihre Versprachlichung und damit verbundene Anforderungen an SchülerInnen mit DaZ analysiert	1.59	.920	1.99	.922	.40**	n.s.	n.s.
Den *Sprachstand* von SchülerInnen mit DaZ anhand von authentischen Schülertexten oder -äußerungen diagnostiziert	1.69	.949	1.89	.927	.20**	n.s.	n.s.
Individuelle *Sprachförderpläne* für SchülerInnen mit DaZ für das eigene Unterrichtsfach aufgestellt	1.22	.655	1.30	.691	.08	-.12*	-.134*
Gesamt	1.78	0.633	2.10	0.612	.32**	n.s.	n.s.

Anmerkung: Die markierten Δ*MW*-Werte sind mit *p < .05, **p < .01, signifikant unterschiedlich, die markierten r Werte korrelieren mit *p < .05, **p < .01 mit dem DaZKom-Testergebnis

Im Bereich der handlungsorientierten Lerngelegenheiten in Tabelle 3 zeigt sich eine andere Sachlage. Zwar nehmen sieben der acht Lerngelegenheiten zwischen den Messzeitpunkten signifikant zu, jedoch korrelieren lediglich drei der Konstrukte signifikant mit dem Posttestergebnis (hiervon eines negativ). Tabelle 3 zeigt ferner, dass die handlungsorientierten Lerngelegenheiten auf der fünfstufigen Skala deutlich seltener wahrgenommen werden als die thematischen Lerngelegenheiten. Ein Mittelwert von 2 bedeutet beispielsweise, dass Studierende lediglich in einer einzelnen Lehrveranstaltungssitzung mit dem jeweiligen Inhalt in Kontakt gekommen sind. Einzige Ausnahme ist hierbei die

Lerngelegenheit „Migration“, welche als handlungsorientierte Lerngelegenheit allerdings der thematischen Lerngelegenheit „Migration und Mehrsprachigkeit“ sehr nahekommt (beide Items korrelieren mit einem Koeffizienten von 0.61, $p = .000$, miteinander).

Eine Auffälligkeit in Tabelle 3 stellt der signifikant negative Zusammenhang zwischen der handlungsorientierten Lerngelegenheit „Sprachförderpläne“ dar. Die sehr niedrigen Mittelwerte lassen bereits vermuten, dass es Probleme mit der Verteilung der Antwort geben könnte. Tests zu Schiefe ($G_1 = 2.58$) und Kurtosis ($w = 7.09$) bestätigen, dass die Verteilung derartig schief ist, dass sinnvolle Analysen, wie in diesem Fall die Überprüfung des Zusammenhangs der Lerngelegenheit mit dem Gesamttestergebnis, nicht möglich sind.

Zusammenfassend lassen sich die oben skizzierten Forschungsfragen wie folgt beantworten. Die Ergebnisse zu den thematischen Lerngelegenheiten (Forschungsfrage 1) aus Stangen/Schroedler/Lengyel (2019, im Druck) können durch die Analysen anhand der Daten der Gesamtkohorte verlässlich repliziert werden. Vor allem bei den thematischen Lerngelegenheiten zeigt sich eine gewisse Prominenz bestimmter Inhalte. So finden sich die Inhalte ‚Umgang mit Heterogenität‘, ‚Bildungssprache‘ und ‚Scaffolding‘ beispielsweise deutlich höher in der sortierten Liste der 16 Inhalte als in der Normierungsstudie von Ehmke und Lemmrich (2018). Dieses Ergebnis ist insofern interessant, als dass auf diese drei Inhalte in der curricularen Reform der Hamburger Lehrerkräftebildung für den Umgang mit DaZ besonderer Wert gelegt wurde. Die Zusammenhänge zwischen den handlungsorientierten Lerngelegenheiten und der DaZ-Kompetenz (Forschungsfrage 2) zeigen ein komplexeres Bild. Einerseits erhöhen sich fast alle Lerngelegenheiten zwischen den Messzeitpunkten signifikant, was für eine gewisse Effektivität des neu gestalteten Lehrangebots während der Semester spricht, in welchen die Prä-Post-Tests durchgeführt wurden. Andererseits zeigen sich kaum Zusammenhänge zwischen den wahrgenommenen Lerngelegenheiten und der DaZ-Kompetenz.

6 Diskussion

Regressionsanalytisch konnte in Stangen/Schroedler/Lengyel (2019, im Druck) identifiziert werden, dass die Menge der wahrgenommenen Lerngelegenheiten – neben den Prätestergebnissen – der einflussreichste Prädiktor für die DaZ-Kompetenz im Posttest ist. Die oben skizzierten Ergebnisse zeichnen deutlich vor, dass im ProfaLe-Modellversuch im Bereich der thematischen Lerngelegenheiten den Lehramtsstudierenden ein solides Angebot gemacht wird, welches für die DaZ-Kompetenzentwicklung gewinnbringend genutzt wird. So zeigen die Analysen zu den thematischen Lerngelegenheiten anhand der Daten der Hamburger Gesamtkohorte, dass nahezu alle diese Lerngelegenheiten in einem

deutlichen positiven Zusammenhang mit der erreichten DaZ-Kompetenz stehen. Im Kontext einer angestrebten Verbesserung der Lehrkräftebildung in Bezug auf eine Professionalisierung angehender Lehrkräfte für den Umgang mit DaZ und Mehrsprachigkeit im Schulalltag, lässt sich hieraus schließen, dass diesen Lerngelegenheiten curricular Beachtung geschenkt werden sollte. Im Hamburger Modellversuch scheint dies – naturgemäß mit Einschränkungen – relativ gut funktioniert zu haben. Die insgesamt schwachen Ergebnisse der DaZ-Kompetenz im Posttest (u. a. berichtet in Schroedler/Grommes 2019; Stangen/Schroedler/Lengyel 2019, im Druck) zeigen jedoch auch auf, dass die Lerngelegenheiten noch weiter ausgebaut werden müssten, um zufriedenstellendere Ergebnisse in Bezug auf die DaZ-Kompetenz zu erzielen.

Der Umstand, dass die handlungsorientierten Lerngelegenheiten nur in sehr geringem Umfang mit der DaZ-Kompetenz in Verbindung zu stehen scheinen, lässt sich zum jetzigen Zeitpunkt nur schwer erklären. Stangen/Schroedler/Lengyel (2019, im Druck) zeigen mit den Daten der ersten Hamburger Teilkohorte, dass die thematischen Lerngelegenheiten in einem Regressionsmodell eine starke Vorhersagekraft besitzen, während die handlungsorientierten Lerngelegenheiten keinen signifikanten Beitrag leisten. Möglicherweise sind die handlungsorientierten Lerngelegenheiten in den Curricula der Hamburger Lehrkräftebildung insgesamt deutlich zu geringfügig angesiedelt, um in den oben beschriebenen Analysen eine größere Rolle zu spielen.

Insgesamt zeigen die Befunde zum einen, dass sich die Nutzung der thematischen und handlungsorientierten DaZ-bezogenen Lerngelegenheiten im Verlauf des Lehramtsstudiums an der Universität deutlich erhöht und ein größerer Umfang genutzter thematischer Lerngelegenheiten positiv mit den Ergebnissen im DaZKom-Test zusammenhängt. Diese Ergebnisse stützen die einleitend beschriebenen positiven Zusammenhänge zwischen Lerngelegenheiten und Professionswissen aus dem Forschungsbereich der Lehrkräftebildung in anderen Feldern. Zum anderen zeigen die fehlenden Zusammenhänge zwischen einer gestiegenen Anzahl genutzter handlungsorientierter Lerngelegenheiten und den Ergebnissen im DaZKom-Test, dass hier sowohl Forschungs- als eventuell auch der Bedarf einer curricularen Anpassung besteht. Die Unterschiede in der Gesamthöhe zwischen den gemessenen Mittelwerten der thematischen und handlungsorientierten Lerngelegenheiten zeigen, dass handlungsorientierte Lerngelegenheiten in der universitären Lehrerkräftebildung generell einen geringeren Umfang einnehmen.

Literatur

Baumann, Barbara (2017): Sprachförderung und Deutsch als Zweitsprache in der Lehrerbildung – ein deutschlandweiter Überblick. In: Becker-Mrotzek, Michael/Rosenberg, Peter/Schroeder, Christoph/Witte, Annika (Hrsg.): DaZ in der Lehrerbildung – Modelle und Handlungsfelder. Münster: Waxmann, S. 9–26.

Baumann, Barbara/Becker-Mrotzek, Michael (2014): Sprachförderung und Deutsch als Zweitsprache an deutschen Schulen: Was leistet die Lehrerbildung? Überblick, Analysen und Handlungsempfehlungen. Köln: Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache.

Becker-Mrotzek, Michael/Rosenberg, Peter/Schroeder, Christoph/Witte, Annika (Hrsg.) (2017): DaZ in der Lehrerbildung – Modelle und Handlungsfelder. Münster: Waxmann.

Carlson, Sonja A./Präg, Désirée (2018): Der Prozess der Aufgabenentwicklung im DaZKom-Projekt: von der Rahmenkonzeption bis zur Pilotierung des Testinstruments. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 93–108.

Doll, Jörg/Buchholtz, Nils/Kaiser, Gabriele/König, Johannes/Bremerich-Vos, Albert (2018): Nutzungsverläufe für fachdidaktische Studieninhalte der Fächer Deutsch, Englisch und Mathematik im Lehramtsstudium. Die Bedeutung der Lehrämter und der Zusammenhang mit Lehrinnovationen. In: Zeitschrift für Pädagogik 64, 4, S. 511–532.

Dreyfus, Hubert L./Dreyfus, Stuart E. (1986): Mind over machine. The Power of Human Intuition and Expertise in the Era of The Computer. New York: Free Press, S. 16–51.

Ehmke, Timo/Lemmrich, Svenja (2018): Bedeutung von Lerngelegenheiten für den Erwerb von DaZ-Kompetenz. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 201–220.

Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.) (2018): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann.

Hammer, Svenja/Carlson, Sonja A./Ehmke, Timo/Koch-Priewe, Barbara/Köker, Anne/Ohm, Udo/Rosenbrock, Sonja/Schulze, Nina (2015): Kompetenz von Lehramtsstudierenden in Deutsch als Zweitsprache. Validierung des GSL-Testinstruments. In: Blömeke, Sigrid/Zlatkin-Troitschanskaia (Hrsg.): Kompetenzen von Studierenden. Weinheim u. a.: Beltz Juventa 2015 (Zeitschrift für Pädagogik 61), S. 32–54.

Helmke, Andreas/Schrader, Friedrich-Wilhelm (2010): Determinanten der Schulleistung. In: Rost, Detlef H. (Hrsg.): Handwörterbuch Pädagogische Psychologie. 4. Auflage. Weinheim: Beltz Psychologische Verlags Union, S. 90–102.

Jentsch, Armin/Doll, Jörg/Stangen, Ilse/Meyer, Dennis/Kaiser, Gabriele (eingereicht): Die Bedeutung von Lerngelegenheiten und sprachbezogenen Merkmalen für den Erwerb von deutschdidaktischem und erziehungswissenschaftlichem Wissen bei Lehramtsstudierenden. In: Zeitschrift für Erziehungswissenschaft.

Koch-Priewe, Barbara (2018): Das DaZKom-Projekt – ein Überblick. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 7–38.

König, Johannes/Ligtvoet, Rudy/Klemenz, Stefan/Rothland, Martin (2017): Effects of opportunities to learn in teacher preparation on future teachers' general pedagogical knowledge: Analyzing program characteristics and outcomes. In: Studies in Educational Evaluation 53, S. 122–133.

Kunina-Habenicht, Olga/Schulze-Stocker, Franziska/Kunter, Mareike/Baumert, Jürgen/Leutner, Detlev/Förster, Doris/Lohse-Bossenz, Hendrik/Terhart, Ewald (2013): Die Bedeutung der Lerngelegenheiten im Lehramstsstudium und deren individuelle Nutzung für den Aufbau des bildungswissenschaftlichen Wissens. In: Zeitschrift für Pädagogik 59, S. 1–23.

Ohm, Udo (2018): Das DaZKom-Modell von DaZ-Kompetenz bei angehenden Lehrkräften. In Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 73–91.

Paetsch, Jennifer/Darsow, Annkathrin/Wagner, Fränze Sophie/Hammer, Svenja/Ehmke, Timo (2019): Prädiktoren des Kompetenzzuwachses im Bereich Deutsch als Zweitsprache bei Lehramtsstudierenden. In: Unterrichtswissenschaft. Zeitschrift für Lernforschung 47, H. 1, S. 51–77.

Schroedler, Tobias/Lengyel, Drorit (2018): Umgang mit sprachlich-kultureller Heterogenität im Fachunterricht – Was kann die erste Phase der Lehrerbildung leisten? In: SEMINAR 4/2018 Interkulturelles Lernen in Schule und Seminar, S. 6–20.

Schroedler, Tobias/Grommes, Patrick (2019): Learning about Language: Preparing Pre-Service Subject Teachers for Multilingual Classroom Realities. In: Journal of Language Learning in Higher Education 9, H. 1, S. 223–240.

Stangen, Ilse/Schroedler, Tobias/Lengyel, Drorit (2019, im Druck): Kompetenzentwicklung für den Umgang mit Deutsch als Zweitsprache und Mehrsprachigkeit im Fachunterricht: Universitäre Lerngelegenheiten und Kompetenzmessung in der Lehrer(innen)bildung. In: Zeitschrift für Erziehungswissenschaft.

Tachtsoglou, Sarantis/König, Johannes (2018): Der Einfluss von Lerngelegenheiten in der Lehrerausbildung auf das pädagogische Wissen angehender Englischlehrkräfte. In: Journal for Education Research Online 10, H. 2, S. 3–33.

Performanznahe und videobasierte Messung von DaZ-Kompetenz bei Lehrkräften

Skalierung und dimensionale Struktur des Testinstruments

Svenja Lemmrich, Sarah-Larissa Hecker, Stephanie Klein, Timo Ehmke, Anne Köker, Barbara Koch-Priewe und Udo Ohm

Abstract

Die vorliegende Studie verfolgt im Rahmen des DaZKom-Video-Projekts das Ziel, Deutsch-als-Zweitsprache(DaZ)-Kompetenz bei (angehenden) Lehrkräften performanznah mit einem videobasierten Testinstrument zu erfassen. Dieser Beitrag prüft die psychometrische Qualität des DaZKom-Video-Testinstrumentes und untersucht die Dimensionalität des Konstruktes *DaZ-Kompetenz* sowie Zusammenhänge zwischen der DaZ-Kompetenz mit soziodemographischen Angaben und DaZ-bezogenen Lerngelegenheiten. Dabei handelt es sich um eine Erprobungsstudie im Rahmen eines Testentwicklungsprozesses. Getestet wurden Lehramtsstudierende sowie Lehrkräfte unterschiedlicher Klassenstufen. Die Ergebnisse der Item- und Skalenanalyse zeigen eine befriedigende bis sehr gute psychometrische Qualität. Die Items haben eine sehr gute Passung zum Rasch-Modell und erreichen eine zufriedenstellende Trennschärfe. Die Ergebnisse zur dimensionalen Struktur des Konstrukts weisen darauf hin, dass es sich eher um ein eindimensionales Konstrukt handelt und nicht um ein zwei- oder dreidimensionales Konstrukt. In Bezug auf soziodemographische Angaben und DaZ-bezogene Lerngelegenheiten zeigt sich, dass die Erstsprache Deutsch sowie insbesondere Zusatzqualifikationen wie das DaZ-Zertifikat oder auch Forschungserfahrung im Bereich DaZ, statistisch signifikant mit der DaZ-Kompetenz von (angehenden) Lehrkräften zusammenhängen. Der Umfang der DaZ-bezogenen Lerngelegenheiten korreliert niedrig mit dem Testergebnis. Die Befunde dieser Erprobungsstudie werden genutzt, um das Testinstrument weiterzuentwickeln.

1 Einleitung

Sprache spielt im Fachunterricht eine zentrale Rolle: Sie ist der Schlüssel zu fachunterrichtlichen Inhalten, ohne den Schülerinnen und Schüler keinen Zugang zu diesen erhalten. Sprache im Fachunterricht hat besondere Eigenschaften (bspw. schriftsprachlich geprägt, spezifische Operatoren, vgl. dazu Schmölzer-Eibinger 2013) und ist deshalb für viele Schülerinnen und Schüler, insbesondere auch für solche mit Deutsch als Zweitsprache, ein Hindernis für ihren Lernerfolg (Köker 2018; Schmölzer-Eibinger 2013). Lehrkräfte aller Unterrichtsfächer müssen die fachsprachlichen Anforderungen an ihre Schülerinnen und Schüler kennen und Kompetenzen besitzen, sie nicht nur fachlich, sondern ebenso in sprachlicher Hinsicht dabei zu unterstützen, Zugang zu Unterrichtsinhalten zu erhalten (Thürmann/Vollmer 2013). Allerdings liegen hierzu weder Ausbildungs- noch Performanzstandards vor, wie Lehrkräfte im Bereich DaZ ausgebildet werden müssen (Koch-Priewe 2018) und es ist unklar, welche Lerngelegenheiten zur Kompetenzentwicklung der Lehrkräfte im Bereich Deutsch als Zweitsprache (DaZ) beitragen (Baumann 2017). Das *DaZKom-Video*-Projekt[1] setzt hier an mit dem Ziel, universitäre und außeruniversitäre Lerngelegenheiten im Bereich DaZ zu evaluieren und zu verbessern. Der vorliegende Beitrag greift auf Vorarbeiten aus dem Vorgängerprojekt *DaZKom*[2] zurück, die bereits ein DaZ-Kompetenzmodell sowie einen Paper-Pencil-Test beinhalten (Köker et al. 2015). Für eine Erfassung professioneller Kompetenzen von Lehrkräften, die nicht allein Wissen fokussiert, sondern möglichst handlungsnah operiert, muss über Papier- und Bleistift-Formate hinausgegangen werden (Aufschnaiter/Blömeke 2010). Dies setzt das Projekt mithilfe von Video-Vignetten und einem mündlichen Antwortformat um. Dadurch soll nicht nur Wissen, sondern auch Können erfasst werden. Beides sind Teil des Professionalitätsverständnisses nach Baumert et al. (2011). Der vorliegende Beitrag untersucht, inwieweit das videobasierte Testinstrument geeignet ist, das Konstrukt DaZ-Kompetenz möglichst performanznah zu erfassen, prüft verschiedene Skalierungsmodelle und untersucht als Indikatoren für die Konstruktvalidität die Zusammenhänge zwischen der gemessenen DaZ-Kompetenz mit sozio-

1 Projekt *DaZKom-Video* (2017–2019): Performanznahe Messung von Deutsch-als-Zweitsprache-Kompetenz bei (angehenden) Lehrkräften; Projektleitung: Prof. Dr. B. Koch-Priewe, Dr. A. Köker, Prof. Dr. U. Ohm, Prof. Dr. T. Ehmke, gefördert vom BMBF (Teil des Clusters KoKoHs II)

2 Projekt DaZKom (2015–2017): Professionelle Kompetenzen angehender LehrerInnen (Sek I) im Bereich Deutsch als Zweitsprache; Projektleitung: Prof. Dr. B. Koch-Priewe, Dr. A. Köker, Prof. Dr. U. Ohm, Prof. Dr. T. Ehmke, gefördert vom BMBF (Teil des Clusters KoKoHs I)

demographischen Merkmalen und DaZ-bezogenen Lerngelegenheiten der befragten (angehenden) Lehrkräfte.

2 Forschungslage: Deutsch als Zweitsprache – Kompetenz empirisch erfassen

In den folgenden beiden Abschnitten wird die Forschungslage dargestellt mit zweierlei Fokussierung: Zum einen wird auf die dimensionale Struktur von DaZ-Kompetenz eingegangen. Zum anderen wird die Bedeutung von Wahrnehmen und Handeln für performanznahe DaZ-Kompetenz näher beschrieben.

Theoretische Fundierung der Dimensionen von DaZ-Kompetenz

Im Rahmen des Vorgängerprojektes *DaZKom* wurde bereits ein Kompetenzmodell entwickelt, welches die inhaltliche Struktur von DaZ-Kompetenz darstellt (Köker et al. 2015). Als drei Hauptdimensionen wurden hier *Fachregister*, *Mehrsprachigkeit* und *Didaktik* identifiziert, die wiederum in Subdimensionen untergliedert sind, welche sich dann in einzelne Facetten ausdifferenzieren lassen (ebd.). In Bezug auf unterrichtliches Handeln von Lehrkräften wirken diese drei Dimensionen meist gemeinsam, wobei jedoch davon ausgegangen wird, dass je nach Situation meist eine Dimension fokussiert wird (Ohm 2018). Die Dimension *Fachregister* beschreibt das Medium Sprache, welches als Basis für jegliche Lehr-Lern-Prozesse dient, die in (schulischen) Institutionen stattfinden und dessen Umfang und Funktion sich Fachlehrkräfte bewusst sein müssen. Die Dimension *Mehrsprachigkeit* fokussiert den Lernprozess und betont besondere Anforderungen an Schülerinnen und Schüler mit DaZ, die von Lehrkräften berücksichtigt werden sollten. Die Dimension *Didaktik* ergänzt die beiden ersten Dimension um didaktische Kompetenzen, die notwendig sind, um einen Fachunterricht zu gestalten, der zweitsprachliche Entwicklungsprozesse berücksichtigt und fokussiert damit den Lehrprozess (Köker et al. 2015).

Wahrnehmung und Handlung als Indikatoren für Performanz

Im Folgenden wird auf die situationsspezifischen Fähigkeiten *Wahrnehmung* und *Handlung* eingegangen, welche in dieser Studie als Indikatoren für Performanz in DaZ-relevanten Unterrichtssituationen dienen (Blömeke/Gustafsson/Shavelson 2015). Die Definition von *Wahrnehmung* wird in der Forschung zu Lehrkräfteprofessionalisierung meist auf das Konzept der *professional vision* nach Goodwin (1994) bezogen (Lindmeier 2013). Goodwin beschreibt in seiner Arbeit, dass Expertinnen und Experten bedeutungsreiche Situationen für ihren Tätigkeitsbereich gezielter wahrnehmen als Novizinnen und Novizen, da ihre

Wahrnehmung dahingehend geschult ist (ebd. 1994). Van Es und Sherin (2002) greifen das Konzept auf und definieren zwei Komponenten der *professionellen Wahrnehmung: noticing* und *knowledge-based reasoning,* wobei ersteres eben diese gefilterte Wahrnehmung beschreibt und letzteres die Fähigkeit, auf Basis des professionellen Wissens Schlussfolgerungen aus der wahrgenommenen Situation zu ziehen (van Es/Sherin 2002). Expertinnen und Experten können demnach auf professionelles Wissen zurückgreifen, um das Wahrgenommene in Konzepte und Theorien einzuordnen und deshalb präziser darauf reagieren. Dies unterscheidet sie von Novizinnen bzw. Novizen, welche meist eine reine Beschreibung dessen liefern, was sie beobachten (Seidel/Blomberg/Stürmer 2010).

Professionelle *Handlungskompetenz* besteht aus Professionswissen und individuellen Fähigkeiten sowie motivationalen, volitionalen und sozialen Aspekten und ist abhängig von domänenspezifischen Handlungsanforderungen (Weinert 2001; Baumert/Kunter 2006). Lindmeier (2013) erfasst in ihrer Studie Handlungskompetenz als *aktionsbezogene Kompetenz* und meint damit die Fähigkeit der Lehrkräfte, ihr „professionelles Wissen spontan und unmittelbar für Situationen an[zu]wenden" (ebd., S. 51).

Das DaZKom-Video-Testinstrument verwendet Testitems, die diese theoretische Struktur aufgreifen. Authentische Unterrichtssituationen werden durch Videostimuli präsentiert und mit gezielten Wahrnehmungsaufgaben und Handlungsaufgaben verbunden, um so die DaZ-Kompetenz der (angehenden) Lehrkräfte möglichst performanznah zu messen.

3 Forschungsfragen

Vor dem Hintergrund der Forschungslage sollen folgende Forschungsfragen in dieser Studie beantwortet werden:

1. Inwieweit kann eine ausreichende psychometrische Qualität der Items des DaZKom-Video-Testinstrumentes festgestellt werden?

Bezugnehmend auf andere Studien (u. a. Lindmeier 2013), die Performanztests nutzen, kann davon ausgegangen werden, dass eine ausreichende psychometrische Qualität erreicht werden kann. Jedoch ist das Format dieser Studie mit der Kombination aus Video-Vignetten und mündlichen Antworten sehr innovativ. Insbesondere zum mündlichen Antwortformat liegen bislang nur wenig bis keine Erfahrungswerte vor. Des Weiteren hat die Erprobungsstudie das Ziel, aus dem Itempool eine Itemauswahl zu treffen. Dabei sollen zufriedenstellende psychometrische Kennwerte für den Gesamttest erreicht werden und zugleich sollen die Items (a) alle drei Dimensionen *Fachregister, Mehrsprachigkeit, Di-*

daktik sowie (b) die beiden Aufgabentypen *Wahrnehmung* und *Handlung* möglichst gleichverteilt abdecken.

2. Welche dimensionale Struktur weist das Testinstrument auf?

Aufgrund von Befunden des nicht-videobasierten DaZKom-Tests kann vermutet werden, dass bei einer videobasierten Messung das Konstrukt ebenfalls nicht eindimensional strukturiert ist. Orientiert an der Definition von Performanz könnte es zum einen eine zweidimensionale Struktur aufweisen (Wahrnehmungs- und Handlungsdimension). Orientiert an der Definition von DaZ-Kompetenz könnte es auch eine dreidimensionale Struktur aufweisen (Fachregister, Mehrsprachigkeit, Didaktik), siehe dazu Abschnitt 2.

3. Welche Zusammenhänge bestehen zwischen Personenmerkmalen, der individuellen DaZ-Erfahrung, der Nutzung von DaZ-bezogenen Lerngelegenheiten und der DaZ-Kompetenz von Testpersonen?

Diese Zusammenhänge zu erfassen dient einerseits der Konstruktvalidierung, andererseits sollen sie Aufschluss über die Beschaffenheit von DaZ-kompetenzförderlichen Lerngelegenheiten geben. Dabei wird angenommen, dass Probandinnen und Probanden, die bereits DaZ-bezogene Lerngelegenheiten hatten, im Kompetenztest bessere Ergebnisse erzielen. Außerdem wird davon ausgegangen, dass DaZ-erfahrene Lehrkräfte einen performanzorientierten Test besser bewältigen können als unerfahrene Studierende.

4 Methodisches Vorgehen

Das Testinstrument DaZKom-Video

Der auf Tablets dargebotene Test enthält 16 Video-Vignetten, die DaZ-relevante Unterrichtssituationen zeigen, mit jeweils zwei dazugehörigen Items. Ein Item aus dem Bereich Wahrnehmung (*Was nehmen Sie wahr?)* und eines aus dem Bereich Handlung (*Wie würden Sie anstelle der Lehrkraft in dieser Situation handeln?* oder *Sie sind Lehrkraft in dieser Situation, wie reagieren Sie wörtlich?*). Die Teilnehmerinnen und Teilnehmer sahen jeweils eine zufällige Auswahl von zehn der 20 Videos und antworteten jeweils mündlich auf die Fragen. Die Audioantworten der Probanden wurden mithilfe eines Headsets und von Tablet-PCs aufgezeichnet. Die Videos sind den Dimensionen des DaZ-Kompetenzmodells zugeordnet (*Fachregister, Mehrsprachigkeit und Didaktik*). Basierend auf Expertinnen- und Expertenratings wird davon ausgegangen, dass sich zwar häufig alle Dimensionen des Modells in jedem Video wiederfinden,

jedoch jede gezeigte Situation einen Fokus hat, welcher einer konkreten Dimension zugeordnet werden kann. Für die Ratings wurden drei Expertinnen und Experten das Modell (sowie die theoretische Fundierung des Modells), die Videos des Testinstruments und die dazugehörigen Kodierleitfäden vorgelegt, mit der Bitte, die Items den (Sub-)Dimensionen des Modells zuzuordnen. Für die Auswertung der Probandinnen- und Probandenantworten wurde mithilfe von Expertinnen- und Experteninterviews zu jedem Item ein Kodierleitfaden entwickelt, der mit Ankerbeispielen aus Erprobungsstudien unterfüttert ist (vgl. dazu auch Hecker et al., eingereicht). Alle Items werden durch einen dreistufigen Antwortcode kodiert: Code 2 (die DaZ-relevanten Aspekte wurden situationsspezifisch gänzlich erfasst und in einen weiteren Kontext eingeordnet), Code 1 (die DaZ-relevanten Aspekte wurden teilweise situationsspezifisch erfasst) und Code 0 (die DaZ-relevanten Aspekte wurden nicht erfasst). Für diese Teilstudie wurden Code 2 und Code 1 zusammengefasst und dementsprechend nur Code 1 (*richtig*) und 0 (*falsch*) vergeben. Die Kodiererübereinstimmung bei der Doppelkodierung für die dreistufige Kodierung lag zwischen $0.6 < PÜ < 0.8$.

Soziodemographische Angaben und DaZ-bezogene Lerngelegenheiten

Neben der DaZ-Kompetenz wurden Personenmerkmale wie Geschlecht und Studien- bzw. Unterrichtsfächer (6 Items) erfasst. Zudem wurden DaZ-relevante Lerngelegenheiten durch eine Skala mit 8 Items ($\alpha = 0.83$) (Ehmke/Lemmrich 2018) und zwei Einzelitems erfasst. Für ersteres wurde eine fünfstufige Likert-Skala verwendet: (1) *gar nicht*, (2) *in einer Sitzung*, (3) *in mehreren Sitzungen*, (4) *in einer ganzen Lehrveranstaltung*, (5) *in mehreren Lehrveranstaltungen*. Dabei wurden die Items wie folgt abgefragt: „In Veranstaltungen in Ihrem Lehramtsstudium sowie in Fort- und Weiterbildungen haben Sie bereits für das eigene Studienfach typische Texte hinsichtlich ihrer sprachlichen Besonderheiten für Schülerinnen und Schüler mit Deutsch als Zweitsprache analysiert (z. B. mathematische Textaufgaben)?“

Stichprobe und Untersuchungsablauf

Für die Durchführung der Studie wurden N = 137 Personen befragt, davon sind 83.2 % weiblich und 16.8 % männlich bzw. 77.4 % Lehramtsstudierende deutscher Universitäten und 22.6 % Lehrkräfte. Es waren fast alle Studien- bzw. Unterrichtsfächer (Mathematik, Deutsch, Englisch, Physik, Biologie, Chemie, Geographie, Sachunterricht, Geschichte, Musik, Kunst, Religion, Politik, Sport) vertreten, wobei 61.9 % das (Studien-)Fach Deutsch angaben. 13.4 % gaben an, Zusatzqualifikationen im Bereich DaZ zu besitzen. 88.3 % gaben an, Deutsch als Erstsprache zu haben, wobei 8.8 % eine andere Erstsprache haben, darunter vor allem Russisch, Polnisch und Kroatisch (2.9 % haben keine Angabe gemacht).

Für die technische Umsetzung des Testinstrumentes auf Tablet-PCs wurde das Programm *LimeSurvey* verwendet, welches die Umsetzung des Tests mithilfe einer App ermöglicht. Die Teilnehmenden sahen auf den Tablets zunächst eine Einführung in das Testinstrument. Diese besteht zum einen aus einer technischen Einführung, zum anderen ist eine Beispiel-Vignette mit dazugehörigen Beispielitems und -antworten enthalten. Nach der Einführung bearbeiteten die Probandinnen und Probanden zunächst den Videotest, bevor sie dann, ebenfalls auf den Tablets, die Fragen zu soziodemographischen Daten und den Lerngelegenheiten beantworteten. Insgesamt dauerte die Erhebung 60 Minuten.

Statistisches Vorgehen

Die Dateneingabe und Auswertung der Häufigkeiten (Codes des DaZKom-Video-Tests, soziodemographische Daten und Lerngelegenheiten) erfolgten in SPSS 25 (IBM 2018). Die Item- und Skalenanalysen wurden auf Basis des Rasch-Modells (Rost 2004) mithilfe des Programms *ConQuest* (Adams/Wu/Wilson 2015) durchgeführt. Die DaZ-Kompetenz der Probandinnen und Probanden wurde als Personenmesswert mithilfe von Personenfähigkeitsschätzern, Weighted Likelihood Estimates (WLEs), bestimmt. *Personenfähigkeitsschätzer* impliziert, dass die Beantwortung eines Items einerseits davon abhängt, welche Fähigkeiten (in dieser Studie DaZ-Kompetenz) Probandinnen und Probanden mitbringen und andererseits auch berücksichtigt, wie hoch die Itemschwierigkeit ist (Wilson 2005). Es werden also hiermit nicht die tatsächlichen Antworten abgebildet, sondern die Wahrscheinlichkeit, die sich aus der Personenfähigkeit und Itemschwierigkeit ergibt, eine richtige Antwort zu erzielen (ebd.). Um die Dimensionalität des Konstruktes zu untersuchen, wurden mit *ConQuest* ein-, zwei- und dreidimensionale Modelle geschätzt.

5 Ergebnisse

Psychometrische Qualität

Das erste Kriterium für die psychometrische Qualität des entwickelten Instruments ist die Passung zum Rasch-Modell. In Tabelle 1 sind hierzu die Ergebnisse der IRT-Skalierung dargestellt. Es zeigt sich, dass die Items im Mittel eine sehr gute Passung zum Rasch-Modell aufweisen ($1.0 < MNSQ$[3] < 1.25; OECD 2005). Einige Items haben eine weniger gute Passung ($Min = 0.85$; $Max = 1.14$), jedoch befindet sich zumindest der *Max*-Wert noch in der Spanne des zufrie-

3 *Weighted Mean Square Fit Statistics (*MNSQ) zeigt an, ob es bei einem Item eine gute Übereinstimmung zwischen der beobachteten und der erwartete Lösungswahrscheinlichkeit gibt.

denstellenden Item-Fits. Das zweite Kriterium ist die Trennschärfe der Items, da diese der Indikator dafür ist, wie verlässlich das Instrument zwischen Personen mit hoher und Personen mit niedriger Merkmalsausprägung unterscheiden kann. Items mit einer Trennschärfe von unter 0.30 wurden aus testtheoretischen Gründen aus den Berechnungen ausgeschlossen, mit der Ausnahme von vier Items, die eine Trennschärfe von 0.21 bis 0.25 (*Min* = 0.21) aufweisen. Diese wurden in den Berechnungen belassen, da es sich bei zweien davon jeweils um eines der Partneritems der Dimension *Wahrnehmung* zur Dimension *Handlung* handelt und das jeweils andere Item der Dimension *Handlung* eine zufriedenstellende Trennschärfe aufweist. Ein weiteres Itempaar wurde aus inhaltlichen Gründen im Test belassen, obwohl beide Items eine Trennschärfe von kleiner 0.3 aufweisen. Insgesamt wurden acht Items (also vier Vignetten) aus der finalen Skalierung ausgeschlossen. Die Trennschärfe der Items ist danach zufriedenstellend ($M = 0.44$; $SD = 0.12$). Die mittlere Itemschwierigkeit ($M = 1.56$; $SD = 0.88$) und die Personenfähigkeit ($M = 0.08$; $SD = 1.10$) überlappen allerdings zu wenig. Dies legt nahe, dass die Items zu schwierig sind bzw. die Personenfähigkeit zu gering. Die EAP-Reliabilität beträgt eindimensional skaliert $\alpha = 0.62$.

Tabelle 1: Item- und Skalenkennwerte

	M	SD	Min	Max
Weighted fit MNSQ	1.01	0.07	0.85	1.14
Trennschärfe	0.44	0.12	0.21	0.74
Itemschwierigkeit	1.56	0.88	-0.71	3.05
WLE	0.08	1.10	-2.73	2.77

Dimensionalitätsprüfung

Die Dimensionalität des eingesetzten Testinstrumentes wurde mithilfe von ein-, zwei- und dreidimensionalen IRT-Modellen geprüft. Die Fit-Indices des Modellvergleichs (vgl. Abb. 1) sind in Tabelle 2 dargestellt. Die Devianz ist die Abweichung vom Idealwert und sollte dementsprechend möglichst niedrig sein. Sie ist, genau wie das AIC (Akaike's Information Criterion), „ein Maß für Anpassungsgüte des geschätzten Modells an die vorliegenden empirischen Daten (Stichprobe) unter Berücksichtigung des Modells" (Moosbrugger 2007, S. 390). Aus dem AIC sind die Informationskriterien BIC und CAIC hervorgegangen, wobei BIC viele unnötige Modellparameter (Missachtung des Sparsamkeitsprinzips) bestraft und CAIC im Gegensatz zu AIC den Stichprobenumfang mit berücksichtigt. All diese Maße sollten möglichst niedrig sein (ebd.). Die Informationskriterien für das eindimensionale Modell sind für die Devianz, den BIC und CAIC niedriger als für das zweidimensionale Modell (Skalierung nach den

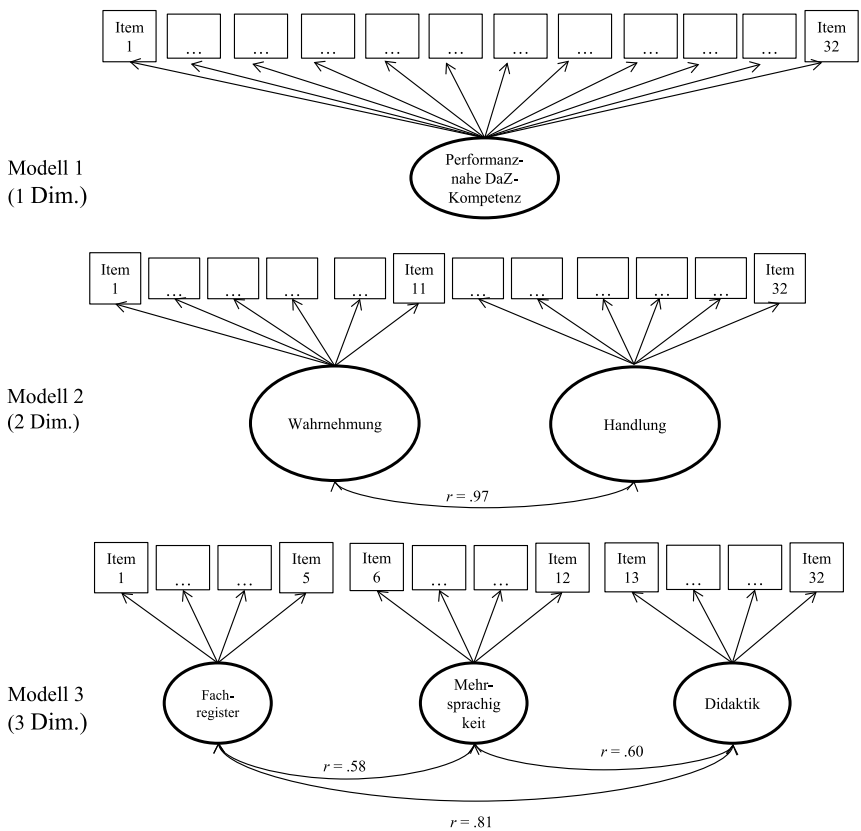
Dimensionen *Wahrnehmung* und *Handlung*) und für das dreidimensionale Modell (Skalierung nach den Dimensionen *Fachregister*, *Mehrsprachigkeit* und *Didaktik*). Dieses Ergebnis zeigt, dass es sich – entgegen der aufgestellten Hypothese – offenbar eher um ein eindimensionales Konstrukt handelt.

Tabelle 2: Dimensionalitätsvergleich

	N	Parameter	Devianz	AIC	BIC	CAIC
1-dimensional	137	31	1757	1819	1910	1942
2-dimensional	137	33	1789	1855	1951	1984
3-dimensional	137	36	1778	1850	1955	1991

In Abbildung 1 zum Modellvergleich wird in Modell 2 deutlich, dass die zwei Dimensionen *Wahrnehmung* und *Handlung* stark korrelieren (r = 0.97), was darauf hinweist, dass das Konstrukt der *performanznahen DaZ-Kompetenz* hier keine zwei trennbaren Dimensionen aufweist.

Abbildung 1: Modellvergleich



In Modell 3 wird deutlich, dass Dimension 2 mit Dimension 1 (r = 0.58) und mit Dimension 3 (r = 0.60) eine mittlere Korrelation aufweist, Dimension 1 und 3 (r = 0.81) dagegen hoch korrelieren. Die Abgrenzung nach Dimensionen ist für die Dimensionen *Fachregister* (1) und *Didaktik* (3) offenbar schwierig, wohingegen die Dimension *Mehrsprachigkeit* (2) sich eher als separates Teilkonstrukt von den beiden anderen abgrenzt.

Zusammenhänge von DaZ-Kompetenz mit weiteren Variablen wie soziodemographischen Merkmalen und Lerngelegenheiten

In Tabelle 3 sind die Korrelationen ausgewählter soziodemographischer Merkmale und DaZ-Lerngelegenheiten mit dem Personenfähigkeitsschätzer (WLE) für die DaZ-Kompetenz (eindimensionale Skalierung) dargestellt. Weibliche und männliche Testpersonen unterscheiden sich nicht in ihrer DaZ-Kompetenz. Testpersonen, die Deutsch als Erstsprache haben, erreichen im Test höhere Werte als Personen, die eine andere Sprache als Erstsprache sprechen. Deutsch als Studien- bzw. Unterrichtsfach steht in keinem signifikanten Zusammenhang mit der DaZ-Kompetenz. Testpersonen mit Zusatzqualifikationen im Bereich DaZ (DaZ-Zertifikat) hingegen erreichen im Test erwartungsgemäß statistisch hochsignifikant höhere Testwerte als Personen ohne DaZ-Zertifikat. Gleiches gilt für Personen, die Forschung im Bereich DaZ betreiben. Lehrerfahrung allein zeigt dagegen keinen statistisch signifikanten Zusammenhang mit der Fähigkeit, die Items des Tests zu bearbeiten, ebenso wenig wie die Summe der DaZ-bezogenen Lerngelegenheiten. Letzteres ist nicht erwartungskonform, weshalb die Rolle der Lerngelegenheiten im Anschluss auf Einzelitemebene detaillierter untersucht wurde. Von acht Items der DaZ-Lerngelegenheiten-Skala (LG) kann eines identifiziert werden (LG_8), welches statistisch signifikant mit dem WLE korreliert: Die Probandinnen und Probanden, die angeben, in Lehrveranstaltungen bereits sprachsensible Unterrichtsstunden entworfen zu haben, erzielen im Test höhere Werte als Personen, die dies noch nicht gelernt haben (r = 0.17*).

Tabelle 3: Korrelationen der Ergebnisse aus dem DaZKom-Video-Test mit Personenmerkmalen

	WLE
Geschlecht (m = 0; w = 1)	0.01
Erstsprache (Deutsch = 1; andere = 0)	0.18*
Fach Deutsch (nein = 0, ja = 1)	0.03
DaZ-Zertifikat (nein = 0, ja = 1)	0.22**
Lehrerfahrung DaZ (nein = 0; ja = 1)	0.13
Forschung DaZ (nein = 0; ja = 1)	0.24**
Skala DaZ-Lerngelegenheiten (DaZ-LG)	0.14

* Die Korrelation ist auf dem Niveau von .05 (2-seitig) signifikant
** Die Korrelation ist auf dem Niveau von .01 (2-seitig) signifikant

6 Zusammenfassung, Diskussion und Ausblick

Ziel der vorliegenden Studie war es, die psychometrische Qualität und die dimensionale Struktur des Testinstruments zur performanznahen Messung von DaZ-Kompetenz zu prüfen und Zusammenhänge zwischen der gemessenen DaZ-Kompetenz mit soziodemographischen Merkmalen und DaZ-Lerngelegenheiten zu untersuchen. Die zentralen Erkenntnisse der Untersuchung sind folgende:

1. Das untersuchte performanznahe Testinstrument zeigt – nach Ausschluss von 8 Items – eine gute Passung zum Rasch-Modell und die Items weisen eine zufriedenstellende Trennschärfe auf. Die Skalenreliabilität ist ausreichend. Noch nicht optimal ist die Passung von Itemschwierigkeit zu den Fähigkeiten der getesteten Probandinnen und Probanden. Diese Testitems sind etwas zu schwer für die aktuelle Stichprobe.
2. Die angenommene Mehrdimensionalität konnte nicht bestätigt werden. Das eindimensionale Modell wies die beste Passung auf.
3. Bezüglich der Zusammenhänge der Testergebnisse mit soziodemographischen Merkmalen und Lerngelegenheiten konnte festgestellt werden, dass Deutsch als Erstsprache als Merkmal sowie insbesondere Lerngelegenheiten in Form von Zusatzqualifikationen und Forschung im Bereich DaZ einen statistisch signifikanten Zusammenhang mit dem Testergebnis aufweisen. Die Summe der Skala der DaZ-Lerngelegenheiten zeigt wider Erwarten keinen statistisch signifikanten Zusammenhang mit dem Testergebnis.

Die Auswertung des Testinstruments auf Grundlage der Item-Response-Theorie nach ersten Erprobungen bietet den Vorteil, dass das Konstrukt theoretisch

definiert wird und Modelle zunächst geprüft werden, um dann aus den psychometrischen Kennwerten und dem Modellvergleich Rückschlüsse für die weitere Testentwicklung zu ziehen. Im Hinblick auf die psychometrischen Kennwerte kann festgehalten werden, dass die Items aktuell überwiegend noch zu schwer für die Testpersonen sind. Dies wird in Zukunft insofern bearbeitet, als dass die Video-Stimuli mit weiteren Kontextinformationen versehen werden, um eine noch realere Einbettung der Situation zu erreichen und damit die Situationsspezifität zu erhöhen. Gleichzeitig wird die Stichprobe der erfahrenen Lehrkräfte in nachfolgenden Studien vergrößert, sodass mehr Probandinnen und Probanden der Stichprobe bereits einen höheren Grad an Expertise erreicht haben, die benötigt wird, um performanznahe Testsituationen bewältigen zu können (vgl. Kapitel 2).

Im Hinblick auf die Dimensionalität des Konstruktes können unterschiedliche Perspektiven herangezogen werden, um die Ergebnisse zu reflektieren. Die Korrelation zwischen den zwei Dimensionen *Wahrnehmung* und *Handlung* in der zweidimensionalen Skalierung sprechen dafür, dass diese beiden Dimensionen nicht getrennt voneinander für das Konstrukt festgelegt werden können. Hinweise dazu finden sich auch in anderen Arbeiten zur professionellen Wahrnehmung oder zur Handlungskompetenz bei Lehrkräften. Niesen (2018) fasst in ihrer Studie zusammen: „Zunehmend häufen sich Hinweise für die Annahme, dass ein Zusammenhang zwischen professioneller Wahrnehmungs- und Handlungskompetenz von Lehrkräften besteht, dergestalt, dass sich erstere positiv auf letztere auswirkt." (ebd., S. 122). Die kognitionspsychologische Forschung geht davon aus, dass beide Dimensionen in ihrer kognitiven Repräsentation überlappen (Hommel et al. 2001). Die dreidimensionale Struktur des Konstruktes DaZ-Kompetenz bestätigt sich für die Kompetenzstruktur im Bereich des Wissens (Ehmke/Hammer 2018). Die vorliegende Studie gibt jedoch Hinweise darauf, dass im Bereich der performanznahen Messung keine klare Trennung der Dimensionen mehr möglich ist, sondern dass die drei Dimensionen *Fachregister*, *Mehrsprachigkeit* und *Didaktik* einheitlicher auftreten. Einschränkend muss allerdings darauf hingewiesen werden: nicht alle Probandinnen und Probanden haben auch alle Items bearbeitet. Es sind in der Summe also nur wenige Items pro Dimension, die von einer Testperson bearbeitet wurden. In zukünftigen Studien soll die dreidimensionale Struktur des Konstrukts daher noch einmal, unter Berücksichtigung einer größeren Stichprobenzahl, untersucht werden. In der DaZKom-Normierungsstudie sollen rund 300 Probandinnen und Probanden befragt werden, von denen mindestens die Hälfte Lehrkräfte sind. Alle Probandinnen und Probanden werden dann jeweils auch alle Items bearbeiten. Außerdem sollen in zukünftigen Analysen Testlet-Effekte kontrolliert werden, da der Test zu einem Stimulus immer zwei Items vorgibt und dadurch ggf. Kontexteffekte entstehen.

Dass sich Dimension 2 im Modellvergleich klarer von Dimension 1 und 3 abgrenzen lässt, könnte auf die inhaltliche Fokussierung der Items zurückzuführen sein. Items dieser Dimension fokussieren entweder die Subdimension *Zweitspracherwerb* oder *Migration*. Besonders in letzterer spielen häufig Differenzkonstruktionen im Unterricht eine zentrale Rolle, welche für Lehrkräfte offenbar eine besondere Herausforderung darstellen. Der Fortbildungsbedarf im reflexiven Umgang mit ethno-kulturellen Differenzordnungen im Unterricht wurde bereits in einer anderen Studie festgestellt (Klein/Hecker/Lemmrich, eingereicht). Ähnliche Korrelationen der drei Dimensionen zeigen sich im Übrigen auch im Modellvergleich der Skalierung des DaZKom-Testinstrumentes (Ehmke/Hammer 2018).

Im Hinblick auf Zusammenhänge der Testergebnisse mit weiteren Merkmalen sticht insbesondere heraus, dass die Skala für die DaZ-bezogenen Lerngelegenheiten nicht statistisch signifikant mit dem Testergebnis korreliert. Gleichzeitig gibt es zwischen dem Erwerb von Zusatzqualifikationen und dem Testergebnis sehr wohl einen statistisch signifikanten Zusammenhang. Dies ist ein Indikator dafür, dass DaZ-spezifische Lerngelegenheiten offenbar zum Erfolg führen können und ggf. die genutzte Skala zur Erfassung der Lerngelegenheiten a) nicht umfassend genug ist und b) die Skala in Bezug auf handlungsnahe Kompetenzerfassung irrelevante Lerngelegenheiten abfragt. Die Entwicklung der Fragenbogenskalen ist Teil eines Testentwicklungsprozesses. In künftigen Studien wird demnach der Fokus auf die DaZ-bezogenen Lerngelegenheiten verstärkt und eine umfassende Überarbeitung und Erweiterung der Skala vorgenommen, die auch Lerngelegenheiten in Fort- und Weiterbildungen erfasst.

Diese Studie hatte das Ziel, das Testinstrument zu erproben und psychometrische Kennwerte zu gewinnen. Dabei wurden jedoch keine deskriptiven Ergebnisse der erhobenen Daten untersucht. Dies wird in zukünftigen Datenerhebungen mit einem Testinstrument von guter psychometrischer Qualität das Ziel sein, um schließlich Veranstaltungskonzeptionen und lernförderliche DaZ-bezogene Lerngelegenheiten zu identifizieren, mit dem langfristigen Ziel, angehende Lehrkräfte optimal auf den Unterricht vorzubereiten.

Literatur

Adams, Raymond J./Wu, Margaret L./Wilson, Mark R. (2015): ACER ConQuest: Generalised Item Response Modelling Software [Computer software]. Version 4. Camberwell, Victoria: Australian Council for Educational Research.

Aufschnaiter, Claudia v./Blömeke, Sigrid (2010): Professionelle Kompetenz von (angehenden) Lehrkräften erfassen – Desiderata. In: Zeitschrift für Didaktik der Naturwissenschaften 16, S. 361–367.

Baumann, Barbara (2017): Sprachförderung und Deutsch als Zweitsprache in der Lehrerbildung. In: Becker-Mrotzek, Michael/Rosenberg, Peter/Schroeder, Christoph/Witte, Annika (Hrsg.): Deutsch als Zweitsprache in der Lehrerbildung. Münster: Waxmann, S. 9–26.

Baumert, Jürgen/Kunter, Mareike/Blum, Werner/Klusmann, Uta/Krauss, Stefan/Neubrand, Michael (2011): Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schülerinnen und Schülern (COACTIV) – Ein Forschungsprogramm. In: Kunter, Mareike/Baumert, Jürgen/Blum, Werner/Klusmann, Uta/Krauss, Stefan/Neubrand, Michael (Hrsg.): Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann, S. 7–26.

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520.

Blömeke, Sigrid/Gustafsson, Jan-Eric/Shavelson, Richard J. (2015): Beyond dichotomies: Competence viewed as a continuum. In: Zeitschrift für Psychologie 223, H. 1, S. 3–13.

Ehmke, Timo/Hammer, Svenja (2018): Skalierung und dimensionale Struktur des DaZKom-Testinstruments. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 129–148.

Ehmke, Timo/Lemmrich, Svenja (2018): Bedeutung von Lerngelegenheiten für den Erwerb von DaZ-Kompetenz. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 201–219.

Goodwin, Charles (1994): Professional vision. In: American Anthropologist 96, H. 3, S. 606–633.

Hecker, Sarah-Larissa/Klein, Stephanie/Lemmrich, Svenja/Ehmke, Timo (eingereicht): Ist DaZ-kompetentes Wahrnehmen von Unterrichtssituationen notwendig, um DaZ-kompetente Handlungsoptionen zu nennen?

Hommel, Bernhard/Müsseler, Jochen/Aschersleben, Gisa/Prinz, Wolfgang (2001): The theory of event coding (TEC): A framework for perception and action planning. In: Behavioral and Brain Sciences 24, H. 5, S. 849–878.

Klein, Stephanie/Hecker, Sarah-Larissa/Lemmrich, Svenja (eingereicht): Lehrkräfteprofessionalisierung im Kontext von Heterogenität. DaZ-kompetenter Umgang mit Differenzordnungen im Unterricht.

Koch-Priewe, Barbara (2018): Das DaZKom-Projekt – ein Überblick. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 7–37.

Köker, Anne (2018): Zur Relevanz der Ausbildung von Kompetenzen von Lehrkräften im Bereich DaZ. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 57–71.

Köker, Anne/Rosenbrock-Agyei, Sonja/Ohm, Udo/Carlson, Sonja A./Ehmke, Timo/Hammer, Svenja/Koch-Priewe, Barbara/Schulze, Nina (2015): DaZKom – Ein Modell von Lehrerkompetenz im Bereich Deutsch als Zweitsprache. In: Koch-Priewe, Barbara/Köker, Anne/Seifried, Jürgen/Wuttke, Eveline (Hrsg.): Kompetenzerwerb an Hochschulen: Modellierung und Messung. Zur Professionalisierung angehender Lehrerinnen und Lehrer sowie früh-pädagogischer Fachkräfte. Bad Heilbrunn: Klinkhardt, S. 177–206.

Lindmeier, Anke/Heinze, Aiso/Reiss, Kristina (2012): Eine Machbarkeitsstudie zur Operationalisierung aktionsbezogener Kompetenz von Mathematiklehrkräften mit videobasierten Maßen. In: Journal für Mathematik-Didaktik 34, H. 1, S. 99–119.

Lindmeier, Anke (2013): Video-vignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In: Riegel, Ulrich/Macha, Klaas (Hrsg.): Videobasierte Kompetenzforschung in den Fachdidaktiken. Fachdidaktische Forschungen. Band 4. Münster: Waxmann, S. 45–62.

Moosburger, Helfried/Kelava, Augustin (2007): Testtheorie und Fragebogenkonstruktion. Heidelberg: Springer Medizin Verlag.

Niesen, Heike (2018): Förderung mehrsprachigkeitssensibler professioneller Handlungskompetenz angehender Englischlehrkräfte. In: Zeitschrift für Interkulturellen Fremdsprachenunterricht: Didaktik und Methodik im Bereich Deutsch als Fremdsprache 23, H. 1, S. 121–134.

OECD (2005): PISA 2003. Technical report. Paris: OECD.

Ohm, Udo (2018): Das Modell von DaZ-Kompetenz bei angehenden Lehrkräften. In: Ehmke, Timo/Hammer, Svenja/Köker, Anne/Ohm, Udo/Koch-Priewe, Barbara (Hrsg.): Professionelle Kompetenzen angehender Lehrkräfte im Bereich Deutsch als Zweitsprache. Münster: Waxmann, S. 73–91.

Rost, Jürgen (2004): Testtheorie Testkonstruktion. 2. Auflage. Bern: Huber.

Schmölzer-Eibinger, Sabine (2013): Sprache als Medium des Lernens im Fach. In: Becker-Mrotzek, Michael/Schramm, Karen/Thürmann, Eike/Vollmer, Helmut J. (Hrsg.): Sprache im Fach. Sprachlichkeit und fachliches Lernen. Münster: Waxmann, S. 25–40.

Seidel, Tina/Blomberg, Geraldine/Stürmer, Kathleen (2010): „Observer" – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE. In: Bayrhuber, Marianne/Leuders, Timo/Bruder, Regina/Wirtz, Markus (Hrsg.): Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes. Weinheim und Basel: Beltz, S. 296–306.

Thürmann, Eike/Vollmer, Helmut J. (2013): Schulsprache und Sprachsensibler Fachunterricht: Eine Checkliste mit Erläuterungen. In: Röhner, Charlotte/Hövelbrinks, Britta (Hrsg.): Fachbezogene Sprachförderung in Deutsch als Zweitsprache. Theoretische Konzepte und empirische Befunde zum Erwerb bildungssprachlicher Kompetenzen. Weinheim und Basel: Beltz Juventa, S. 121–233.

van Es, Elizabeth A./Sherin, Miriam G. (2002): Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. In: Journal of Technology and Teacher Education 10, H. 4, S. 571–596.

Weinert, Franz E. (2001): Concept of Competence: A Conceptual Clarification. In: Rychen, Dominique S./Salganik, Laura H. (Hrsg.): Defining and Selecting Key Competencies. Göttingen: Hogrefe, S. 45–66.

Wilson, Mark (2005): Constructing Measures. An Item Response Modeling Approach. New York und London: Psychology Press, Taylor & Francis Group.

III Pädagogische Interventionen in der Lehrerbildung

Tobias Feldhoff

Thematische Einleitung

Zu Beginn des Abschnitts in diesem Band zum Thema pädagogischer Interventionen in der Lehrerbildung erfolgt zunächst eine Einführung in das Thema. In dieser wird ein Rahmen gespannt, in dem die nachfolgenden Beiträge verortet werden. Dies soll den Lesern und Leserinnen einen Überblick verschaffen, was sie in den fünf Beiträgen erwartet. Um den Autorinnen und Autoren der Beiträge gerecht zu werden, ist vorab anzumerken, dass die Autorinnen und Autoren nicht den expliziten Auftrag hatten, ihre Maßnahmen und Projekte im Bereich der Lehrerbildung in der Logik von Interventionen darzustellen. Die Beiträge wurden durch die Herausgebenden des Bandes den Interventionen zugeordnet.

Definition und Ziele von Interventionen

Der Begriff „Intervention" leitet sich aus dem lateinischen Verb „intervenio" ab, was übersetzt *dazwischenkommen, einschreiten, einwirken auf* oder *unterbrechen,* bedeutet (Pons, Online Übersetzung). In unserem heutigen Sprachgebrauch werden unter einer Intervention allgemein eine oder mehrere Maßnahmen verstanden, mit deren Hilfe in „einen laufenden Prozess eingegriffen" (Leutner 2010, S. 63) wird, mit dem Ziel etwas entweder im Vorhinein zu verhindern (präventiv) oder im Nachhinein zu verbessern (reaktiv) (vgl. Mittag/Bieg 2010, S. 10).

Pädagogische Interventionen verfolgen zumeist das Ziel, mittels spezifischer Maßnahmen die Handlungskompetenz von Personen und über diese vermittelt, deren Handeln zu verbessern (vgl. Hurrelmann/Holler 1988, S. 81 ff.). Die Personen können sowohl pädagogische Professionelle (z. B. Lehrpersonen, Dozierende) als auch deren Klientinnen und Klienten bzw. Adressaten (zum Beispiel Schülerinnen und Schüler, Studierende) sein.

Interventionsmaßnahmen können direkt auf die Verbesserung der Handlungskompetenz einer Person zielen. Im Feld der Lehrerbildung sind dies zumeist Lehrpersonen oder angehende Lehrpersonen (Lehramtsstudierende oder

Referendarinnen und Referendare). Ein Beispiel für eine solche Intervention ist ein Trainingsprogramm zur Differenzierung im Fachunterricht Mathematik. Interventionen können aber auch indirekt auf Faktoren, zum Beispiel die Motivation oder die Einstellung einer Person zielen, über die vermittelt deren Handlungskompetenz beeinflusst werden. Darüber hinaus können Interventionen jedoch auch darauf abzielen, gleichzeitig direkt und indirekt die Handlungskompetenz zu verbessern.

Je nach Ziel ist es jedoch nicht immer möglich mit einer Maßnahme direkt bei den Adressaten anzusetzen. In solchen Fällen zielen, Interventionen auf die Veränderung der Umwelt der Adressaten (i. d. R. andere Personen, Strukturen oder Organisationen), um über diese die Handlungskompetenz positiv zu beeinflussen (vgl. Hurrelmann/Holler 1988, S. 81 ff.). In der Lehrerbildung sind solche Interventionen zumeist auf der Mesoebene (zum Beispiel Schule, Hochschulseminare, Universität, Studienseminare), teilweise aber auch auf der Makroebene (Bundesland) verortet. Solche Interventionen verfolgen somit verschiedene Ziele (direkte und indirekte) auf unterschiedlichen Ebenen. So wird beispielsweise eine Interventionsmaßnahme zur Verbesserung der Beratungskompetenz von angehenden Lehrpersonen auf der Ebene der Dozierenden an der Universität und den Seminarleiterinnen und Seminarleitern der Studienseminaren angesiedelt.

Programm und Evaluation von Interventionen

Die Maßnahmen und Ziele einer Intervention werden in einem sogenannten Programm festgehalten (vgl. Mittag/Bieg 2010). Es enthält eine Beschreibung der konkreten Aufgaben und Handlungen, die im Rahmen der einzelnen Maßnahmen ausgeführt werden sollen. Darüber hinaus sollte ein Interventionsprogramm mit den darin enthaltenen kausalen Wirkungserwartungen aus empirischen Studien abgeleitet und „wissenschaftlich und theoretisch fundiert“ (Mittag/Bieg 2010, S. 33) sein. Die tatsächliche Wirkung und damit die Zielerreichung der Intervention sollte durch eine begleitende Evaluation überprüft werden (ebd.).

Aus methodischer Sicht ist die Überprüfung von Wirkung im Sinne von Kausalität an vier zentrale Bedingungen geknüpft (Reinecke 2005, S. 12): 1. Hypothesen über den Wirkungszusammenhang, die belastbar und aus Theorien abgleitet sind; 2. das Vorhandensein von korrelativen Zusammenhängen in den Daten, anhand derer der Wirkungszusammenhang überprüft werden kann; 3. der Ausschluss bzw. die Kontrolle von möglichen weiteren Faktoren, die einen Einfluss auf den Wirkungszusammenhang haben könnten; 4. eine zeitlich logische Differenz zwischen unabhängiger und abhängiger Variable. Die ersten beiden Bedingungen können in gleicher Weise durch die Analyse von Quer- und Längsschnittdaten eingelöst werden. Bei den beiden letzten Bedingungen

bieten Längsschnittstudien deutliche Vorteile: „With a longitudinal design, the time ordering of events is often relatively easy to establish, while in cross-sectional designs this is typically impossible“ (Gustafsson 2010, S. 79). Die mögliche Verzerrung durch nicht berücksichtigte weitere Einflussfaktoren stellt die Forschung in den Sozialwissenschaften vor das Problem, dass weder theoretisch noch empirisch die möglicherweise große Anzahl potentieller alternativer Einflussfaktoren eingeschränkt werden kann. Aus methodischer Sicht bieten randomisierte Experimentalstudien mit Kontrollgruppendesign die besten Möglichkeiten der Kontrolle von Drittvariablen (für eine detaillierte Beschreibung siehe Slavin 2010). Die dahinterliegende Annahme ist, dass durch die zufällige Zuteilung zur Treatment- und Kontrollgruppe sich die nicht berücksichtigten Einflussfaktoren auf beide Gruppen gleichmäßig verteilen und somit die Ergebnisse nicht verzerrt werden. Sie sind vor allem in der Psychologie und in der anglo-amerikanischen Schulforschung zum Beispiel im Rahmen der „Comprehensive School Reforms“ in den USA (Rowan et al. 2009) sehr verbreitet. Die Umsetzung randomisierter Experimentalstudien ist oftmals mit einigen Herausforderungen verbunden, wie etwa ethische Gründe bei der Vorenthaltung eines Treatments in der Kontrollgruppe; Probleme bei der Randomisierung, vor allem wenn die Adressaten Teil eines Mehrebenensystems sind oder neben direkten auch indirekte Ziele verfolgt werden; die Sicherstellung, dass das Treatment (Maßnahmen im Rahmen der Intervention) in der Treatment-Gruppe ausreichend genau umgesetzt wird (treatment fidelity) und in der Kontrollgruppe nicht, etc. (für eine detaillierte Betrachtung im schulischen Kontext siehe Maag Merki 2014). Am Beispiel von Evaluationen in Form randomisierter Experimentalstudien wird besonders deutlich, dass Interventionen und deren Evaluation generell im Vorfeld genau aufeinander abgestimmt werden müssen, um deren Wirkungen auch präzise und verlässlich empirisch überprüfen zu können.

Leutner (vgl. 2010) weist daraufhin, dass Interventionen im pädagogischen Bereich oftmals eng mit Diagnostik verbunden sind. Diagnostik hat im Kontext von Intervention zunächst einmal die Funktion den Verbesserungsbedarf möglichst valide und reliabel zu identifizieren. Dies gilt sowohl für den Bereich präventiver Maßnahmen, bei denen ein potentielles Risiko identifiziert wird (in der Lehrerbildung zum Beispiel ein potentiell bevorstehender Burn-Out von Lehrpersonen), als auch für den Bereich von reaktiven Maßnahmen, bei denen ein Defizit identifiziert wird, dass behoben werden soll (zum Beispiel mangelnde diagnostische Fähigkeit von Lehrpersonen). Die Diagnostik kann sich spezifisch auf die betroffene Person oder Personengruppe richten. Eine weitere Möglichkeit besteht darin, auf eine spezifische Diagnostik zu verzichten und stattdessen die Adressaten „als Repräsentanten einer Gruppe von Personen zu betrachten, bezüglich derer man aus der bisherigen Forschung weiß, dass sie bestimmte Eigenschaften aufweisen“ (ebd., S. 18).

Beiträge zu Interventionen in der Lehrerbildung in diesem Band

Die „Qualitätsoffensive Lehrerbildung“ des Bundesministeriums für Bildung und Forschung (BMBF) kann als eine großangelegte Intervention in der Lehrerbildung auf Makroebene interpretiert werden, auch wenn sie in vielen Aspekten von der skizzierten Beschreibung von pädagogischen Interventionen abweicht. Sie hat mit ihrer Intervention Einfluss auf die Förderung von Lehre und Forschung an Universitäten und Pädagogischen Hochschulen in Deutschland genommen. Über diese hat sie vermittelt ihrerseits verschiedene Interventionen an den beteiligten Universitäten und Pädagogischen Hochschulen initiiert. Einige diese Interventionen in der Lehrerbildung werden in den folgenden Beiträgen vorgestellt. Die Vorstellung erfolgt anhand zentraler o. g. Merkmale von Interventionen: Ziele und Maßnahmen, Zielgruppe und Ort der Intervention, Evaluation, Forschungsdesign.

Bevor auf die Beiträge im Einzelnen eingegangen wird, werden die Artikel übergreifen diskutiert. Zunächst fällt auf, dass in der Mehrzahl Studierende zumeist des Lehramts, Adressaten und Universitäten oder pädagogischer Hochschulen die Orte der Interventionsmaßnahmen sind (vgl. Altmann et al. 2019; Klewin/Köker/Störtländer 2019). Weitere Adressaten und Orte sind Lehrpersonen in Gemeinschaftsschulen im Rahmen von Fortbildungen im Beitrag von Schmaltz und Leuders (2019) und Referendarinnen und Referendare in Schule und Studienseminar im Beitrag von Feierabend und Bartels (2019). Alle Maßnahmen haben die Professionalisierung der Adressaten zumeist in Form einer Verbesserung der Handlungskompetenz als (angehende) Lehrperson (zum Beispiel im Bereich Kernkompetenzen, Umgang mit Heterogenität) zum Ziel. Diese sollen über Faktoren, wie Selbstwirksamkeit, Überzeugungen, handlungsleitende Kognitionen oder eines forschenden Habitus vermittelt, beeinflusst werden. So unterschiedlich die Maßnahmen auch sind, alle setzen auf Reflexion als einen zentralen Mechanismus zum Aufbau der Handlungskompetenz. Dabei werden unterschiedliche Strategien eingesetzt, um den Reflexionsprozess zu fördern (Videovignetten, Forschendes Lernen). Zur Evaluation wird ein breites Spektrum verschiedener Designs eingesetzt; Kontrollgruppendesigns, längsschnittliches Mixed-Method-Design, qualitative Längsschnittstudie und querschnittliche Befragung mittels Fragebögen.

Auffällig ist, dass bei keinen der Maßnahmen eine spezifische Diagnostik der Ausgangspunkt darstellt, weder auf der Ebene der Adressaten noch auf der Ebene der Orte der Interventionen. Stattdessen wird zumeist auf Basis des Forschungsstandes von einem Defizit bzw. einer aufzubauenden Kompetenz hingewiesen. Implizit wird davon ausgegangen, dass durch die Maßnahmen ein höherer Kompetenzerwerb erfolgt, als durch die bisherige Praxis. Die wird jedoch nicht erwähnt.

Die einzelnen Beiträge im Überblick

Der Beitrag von Altmann, Weber, Prilop, Kleinknecht und Nückles (1) beschreibt eine Intervention, die an der Universität Freiburg in Seminaren der Bildungswissenschaften angesiedelt ist und sich an Studierende des Lehramts für Gymnasien richtet. Die Intervention hat zum Ziel die Handlungskompetenz der Lehramtsstudierenden im Bereich der Kernkompetenzen unterrichtlichen Handelns (vgl. Grossman/Hammerness/McDonald 2009) sowie ihre Selbstwirksamkeit zu fördern. Die Kernkompetenzen werden, wie es der Name schon impliziert, als zentrale Kompetenzen des unterrichtlichen Handelns angesehen, die angehende Lehrpersonen in ihrer Ausbildung aufbauen müssen. Um dies zu erreichen, orientiert sich die Intervention u. a. an dem ALACT-Cycle (vgl. Korthage/Vasalos 2005) mit den Phasen Aktion, Reflektion, Entwicklung alternativer Handlungsoptionen, erneute Aktion. Das Kreislaufmodell wird in videobasierten Peerfeedback-Seminaren umgesetzt. Nach der Vermittlung von Inhalten erproben Studierende mittels Microteaching in der Rolle als Lehrpersonen die Kernkompetenzen in kurzen simulierten Unterrichtssituationen. Die Situationen werden per Video aufgezeichnet und die Studierenden geben sich gegenseitig ein Peerfeedback. Anschließend wird die Simulation erneut durchgeführt. Die Intervention im engeren Sinne beschränkt sich auf ein Peerfeedback-Training, das eine der beiden Seminargruppen zusätzlich erhält. Durch gezieltes Training soll das Peer-Feedback differenzierter und wirksamer für den Aufbau der Handlungskompetenz und der Selbstwirksamkeit sein, als das Feedback ohne Training. Die Evaluation besteht aus einem quasi-experimentellen Kontrollgruppendesign mit einer Prä-Post-Befragung zur Selbstwirksamkeit und Handlungskompetenz anhand der Selbsteinschätzung der Studierenden.

Der Beitrag von Schmaltz und Leuders (2) unterscheidet sich in Bezug auf die Zielgruppe und dem Ort der Intervention von den anderen Beiträgen. Die Intervention in Form einer Fortbildung zum Umgang mit Heterogenität im Unterricht richtet sich an Lehrpersonen an Gemeinschaftsschulen in Baden-Württemberg und hat die Verbesserung des Umgangs mit Heterogenität durch die Lehrpersonen und die Förderung ihrer Selbstwirksamkeitserwartungen zum Ziel. Der Umgang mit Heterogenität wird als eine zentrale Kompetenz von Lehrpersonen insbesondere an einer Gemeinschaftsschule angesehen (vgl. Bohl/Budde/Rieger-Ladich 2017), die sich viele Lehrpersonen selbst nicht zutrauen (vgl. Solzbacher 2008), weshalb die Fortbildung auch die Selbstwirksamkeitserwartungen der Lehrpersonen fördern will. Die Fortbildung besteht aus der Vermittlung allgemein- und fachdidaktischer Kompetenzen, der Planung und einer anschließenden gemeinsamen Reflexion einer Unterrichtsstunde für heterogene Lerngruppen durch die teilnehmenden Lehrpersonen. Die Evalua-

tion besteht aus einem Mixed-Methods-Design mit einer standardisierten Befragung mittels eines Fragbogens und Interviews vor und nach der Fortbildung. Die Evaluation fokussiert auf die Untersuchung der Veränderung der Selbstwirksamkeitserwartungen der Teilnehmerinnen und Teilnehmer und der Überzeugungen in Bezug auf heterogene Lerngruppe.

Schöb (3) beschäftigt sich in ihrem Beitrag mit einer Intervention zur Förderung der Lehrkompetenz über den Aufbau handlungsleitender Kognitionen. Die Zielgruppen sind angehende Lehrpersonen, die als Ehrenamtliche im öffentlich-kirchlichen Kontext arbeiten und eine Ausbildertätigkeit anstreben. Die Interventionsmaßnahme besteht aus einer Videofallarbeit, bei dem sich die Studierenden in einem Seminar mehrfach mit videographierten Lehr-Lernsituationen unter Anleitung auseinandersetzen. Diese Auseinandersetzung soll zum Aufbau kognitiver Denkstrukturen zur Diagnose der Lehr-Lernsituationen führen. Über diese soll vermittelt die Qualität und Situationsangemessenheit des eigenen Handelns verbessert werden. In einer Kontrollgruppe sollten sich Studierende die gleichen Inhalte über Vorträge und in Gruppenarbeit ohne Fälle aneignen. In beiden Seminaren fand zudem ein Training von Handlungssituationen statt, bei dem Studierende ein Video einer Lehr-Lernsituation präsentiert bekamen, dass gestoppt wurde und sie die Handlung in der Rolle als Lehrperson fortführen sollten. Die Evaluation besteht aus einem quasi-experimentellen Kontrollgruppendesign, bei dem die Qualität und Angemessenheit des Handelns der Studierenden videographiert und durch Experteninnen und Experten mit Hilfe eines Kategoriensystems bewertet wurden. Zusätzliche wurden bei den Studierenden der Treatmentgruppe die Diagnosekompetenz im Prä-Post-Vergleich untersucht.

Klewin, Köker und Störtländer (4) berichten über eine Intervention zur Professionalisierung von Lehramtsstudierenden im Praxissemester mittels Forschendem Lernen. Die Interventionsmaßnahme zum Forschenden Lernen besteht aus drei Seminaren in den Bildungswissenschaften und zwei Studienprojekten während des Praxissemesters. Im Vorbereitungsseminar vor dem Praxissemester werden den Studierenden Grundlagen zum Forschenden Lernen und zu Forschungsmethoden sowie zur Didaktik und Unterrichtsplanung vermittelt. Im Reflexionsseminar parallel zum Praxissemester reflektieren die Studierenden eine besondere Praxiserfahrung. Im anschließenden Begleitseminar werden die Studierenden bei der Auswertung der Studienprojekte unterstützt. Professionalisierung wird von den Autorinnen und Autoren in dem Beitrag bewusst breit verstanden. Durch Forschendes Lernen sollen Studierende einen forschenden Habitus entwickeln und es soll ihnen helfen, eine reflexive Distanz zur eigenen Praxiserfahrungen herzustellen. Bei dem Evaluationsdesign handelt es sich um eine qualitative Längsschnittstudie mittels Interviews zu drei Zeit-

punkten (direkt nach dem Praxissemester, während des Vorbereitungsdiensts, nachdem Berufseinstieg) zu den Lerngelegenheiten im Praxissemester und zur Bedeutung des Forschenden Lernens. Im Beitrag werden Ergebnisse der ersten Befragung präsentiert.

Beim letzten Beitrag von Feierabend und Bartels (5) handelt es sich nicht um eine Intervention, sondern um eine Untersuchung zum Zusammenhang von Anforderungen im niedersächsischen Vorbereitungsdienst und dem Belastungsempfinden angehender Lehrkräfte. Ausgangspunkt sind die Herausforderungen denen sich Lehrpersonen, aber auch insbesondere angehende Lehrpersonen im Vorbereitungsdienst gegenübersehen und die zu einem hohen Belastungsempfinden führen können. Als theoretischer Hintergrund wird ein Belastungs-Beanspruchungs-Modell genutzt, das zwischen objektiven Anforderungen und subjektiv empfundenen Belastungen unterscheidet. Referendarinnen und Referendare sind gleichzeitig schon Lehrende, aber noch Lernende. Sie müssen einerseits als Novizen den Anforderungen der Schule und des Unterrichts und anderseits der Anforderung der Studienseminare (insb. Unterrichtsentwürfe und -besuche) gerecht werden. In dem Beitrag wird mit Hilfe einer einmaligen standardisierten Befragung untersucht, inwieweit die Anforderungen an 16 niedersächsischen Studienseminaren variieren, ob sich Referendarinnen und Referendare stärker durch die Schule oder das Studienseminar belastet fühlen und inwieweit strukturelle Anforderungen mit dem Belastungsempfinden zusammenhängen.

Literatur

Bohl, Thorsten/Budde, Jürgen/Rieger-Ladich, Markus (Hrsg.) (2017): Umgang mit Heterogenität in Schule und Unterricht: Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen. Bad Heilbrunn: Julius Klinkhardt.

Grossman, Pam/Hammerness, Karen/McDonald, Morva (2009): Redefining teaching, reimagining teacher education. In: Teachers and Teaching: theory and practice 15, 2, S. 273–289.

Gustafsson, Jan-Eric (2010): Longitudinal designs. In: Creemers, Bert P. M./Kyriakides, Leonidas/Sammons, Pamela (Hrsg.): Methodological Advances in Educational Effectiveness Research. Quantitative Methodology Series. Milton Park: Routledge, S. 77–101.

Hurrelmann, Klaus/Holler, Birgit (1988): Pädagogische Intervention. In: Hörmann, Georg/Nestmann, Frank (Hrsg.): Handbuch der psychosozialen Intervention Wiesbaden: VS Verlag für Sozialwissenschaften, S. 81–92.

Korthagen, Fred/Vasalos, Angelo (2005): Levels in reflection: Core reflection as a means to enhance professional growth. In: Teachers and Teaching: theory and practice 11, 1, S. 47–71.

Leutner, Detlev (2010): Perspektiven pädagogischer Interventionsforschung. In: Hascher, Tina/Schmitz, Bernhard (Hrsg.): Pädagogische Interventionsforschung. Theoretische Grundlagen und empirisches Handlungswissen. Weinheim und München: Juventa, S. 63–72.

Maag Merki, Katharina (2014). Conducting intervention studies on school improvement: An analysis of possibilities and constraints based on an intervention study of teacher cooperation. Journal of Educational Administration, 52, 5, S. 590–616.

Mittag, Waldemar/Bieg, Sonja (2010): Die Bedeutung und Funktion pädagogischer Interventionsforschung und deren grundlegende Qualitätskriterien. In: Hascher, Tina/Schmitz, Bernhard (Hrsg.): Pädagogische Interventionsforschung. Theoretische Grundlagen und empirisches Handlungswissen. Weinheim und München: Juventa, S. 31–72.

Reinecke, Jost (2005): Strukturgleichungsmodelle in den Sozialwissenschaften. München: Oldenburg.

Rowan, Brian/Correnti, Richard/Miller, Robert J./Camburn, Eric M. (2009): School Improvement by Design. Lessons from a Study of Comprehensive School Reform Programs. Consortium for Policy Research in Education, S. 1–53.

Slavin, Robert E. (2010): Experimental studies in education. In: Creemers, Bert P. M./Kyriakides, Leonidas/Sammons, Pamela (Hrsg.): Methodological Advances in Educational Effectiveness Research. Quantitative Methodology Series. Milton Park: Routledge, S. 102–114.

Solzbacher, Claudia (2008): Was denken Lehrerinnen und Lehrer über individuelle Förderung? In: Pädagogik 60, 3, S. 38–42.

Förderung von Kernkompetenzen in der Lehramtsausbildung durch videobasiertes Microteaching und Peerfeedback

Anna F. Altmann, Kira Elena Weber, Christopher Neil Prilop, Marc Kleinknecht und Matthias Nückles

Abstract

Das Fokussieren von *Core Practices* ist ein aussichtsreicher Ansatz, um Kernkompetenzen unterrichtlichen Handelns aufzubauen. Solche Kernkompetenzen können im geschützten Raum von Microteachings trainiert werden, die anschließend durch videogestütztes Peerfeedback reflektiert werden. Wir boten zwei videobasierte Seminare an ($N = 37$, mit und ohne Peerfeedback-Training) und untersuchten die Selbstwirksamkeit sowie die Akzeptanz des Konzepts. Die *Core Practices* „Erklären", „Fragen stellen" und „Unterrichtsstörungen vorbeugen" wurden in jeweils vier Sitzungen bearbeitet. Bei zwei von drei *Core Practices* nahm die Selbstwirksamkeit durch die Microteachings zu, wobei das Peerfeedback-Training keinen Unterschied machte. Die Bewertungen des Peerfeedbacks zeigten, dass es den Studierenden offenbar gelang, ein hohes Maß an Empathie zu realisieren, ihre Verbesserungsvorschläge jedoch als weniger produktiv wahrgenommen wurden. Die Evaluation ergab eine sehr hohe Akzeptanz der Seminare. Microteachings und videogestütztes Peerfeedback werden von Lehramtsstudierenden offenkundig sehr geschätzt und könnten daher in Master of Education Studiengängen verstärkt eingesetzt werden. Zukünftige Studien sollten die Frage klären, inwiefern das skizzierte Trainingssetting zu einem objektiven Zuwachs an Kompetenzen in der Ausführung von Kerntätigkeiten des Unterrichtens führt.

1 Problemstellung

Die Planung und Durchführung von Unterricht steht im Mittelpunkt des beruflichen Handelns von Lehrkräften. Um diese Anforderungen erfolgreich zu bewältigen, benötigen Lehrkräfte neben professionellem Wissen entsprechende Handlungskompetenzen. In der Lehrkraftforschung wird dabei sogenannten

Core Practices eine zentrale Rolle zugeschrieben. Nach Grossman, Hammerness und McDonald (2009) handelt es sich um Tätigkeiten von Lehrkräften, die (1) beim Unterrichten häufig vorkommen, (2) auch von Lehramtsstudierenden bereits ausgeführt und mit zunehmender Übung beherrscht werden können, (3) auf empirischer Forschung basieren sowie (4) das Potential haben, den Lernerfolg von Schülerinnen und Schülern zu steigern. Um Studierende auf die schulpraktischen Ausbildungsanteile vorzubereiten, sollten *Core Practices* bereits in universitären Lehrveranstaltungen trainiert werden. Daher stellt sich die Frage, wie universitäre Lehrveranstaltungen aufgebaut werden sollten, um solche Kerntätigkeiten zu trainieren, damit angehende Lehrkräfte Kompetenzen in den entsprechenden Tätigkeitsbereichen aufbauen können.

2 Theoretischer Hintergrund

Grossman et al. (2009) verstehen *Core Practices* als elementare unterrichtliche Tätigkeiten, die nichtsdestotrotz den komplexen und ganzheitlichen Charakter unterrichtlichen Handelns erkennen lassen. So können das Geben von Erklärungen, das Anleiten von Gruppenarbeiten, das Geben von Feedback, das Stellen denkanregender Fragen, das Leiten von Diskussionen oder das Vorbeugen von Unterrichtsstörungen als *Core Practices* gelten. Um Kernkompetenzen in Bezug auf diese Tätigkeiten zu trainieren, können der ALACT-Cycle (Korthagen/Vasalos 2005) und der Learning Cycle (McDonald et al. 2013) herangezogen werden. Der ALACT-Zyklus beschreibt, wie eine Tätigkeit einzuüben ist, indem diese zunächst ausgeführt wird (*Action*), dann betrachtet (*Looking back at the action*) und reflektiert wird (*Awareness of essential aspects*), danach Handlungsalternativen generiert werden (*Creating alternative methods of action*) und zuletzt die Tätigkeit erneut durchgeführt wird (*Trial*). Der Schlüssel liegt also darin, eine Tätigkeit durchzuführen und diese nach einem Reflexions- und Überarbeitungsprozess erneut zu erproben. Der Learning-Cycle von McDonald et al. (2013) berücksichtigt zusätzlich das deklarative Wissen. Er beschreibt, wie Lernende vom deklarativen Wissen zur praktischen Tätigkeit gelangen können. Es wird also zunächst eine Tätigkeit vorgestellt und etwas über die Tätigkeit gelernt (z. B. durch Modellieren). Dann folgen zwei Phasen, in denen die Tätigkeit geplant, geübt und durchgeführt wird. In der letzten Phase wird die Tätigkeit schließlich analysiert.

Microteaching

Im Learning-Cycle von McDonald et al. (2013) ist Microteaching ein zentrales Element, um *Core Practices* einzuüben. Microteaching ist eine Methode, bei der angehende Lehrkräfte vor einer kleinen Gruppe von Lernenden oder Mitstudie-

renden Unterrichtssituationen realisieren, um bestimmte unterrichtliche Handlungen einzuüben. So können Kerntätigkeiten des Unterrichtens unter vereinfachten Bedingungen trainiert werden. Die Länge einer Microteaching-Einheit beträgt ca. zehn bis zwanzig Minuten und die zu übende Fertigkeit wird anhand von Videoaufzeichnungen analysiert. Die Videoaufzeichnungen dienen dem Geben von Feedback, das als notwendiges Element erachtet wird, um Kompetenzen in der Ausführung von Kerntätigkeiten auszubilden (Grossman et al. 2009).

Feedback

In der Lehrkraftausbildung regt Feedback Reflexionsprozesse und eine kritische Analyse des eigenen Unterrichtshandelns an (Hammerness et al. 2005). Mehrere Studien wiesen nach, dass Feedback die professionelle Kompetenz von Lehrkräften fördert und somit auch positiven Einfluss auf die Leistungen von Schülerinnen und Schülern hat (Weber/Gold/Prilop/Kleinknecht 2018).

Feedback kann durch Peers (z. B. Mitstudierende) oder durch Personen mit einem höheren Expertisegrad (z. B. Dozierende) gegeben werden. Peers wird im Unterschied zu Experten keine klare „knowledge authority“ zugeschrieben (Gielen/Peeters/Dochy/Onghena/Stuyven 2010, S. 305), was eine aktive Diskussion des Feedbacks erleichtert und somit zu einem tiefergehendem Verständnis führen kann. Expertenfeedback hingegen kann eine passive Haltung des Feedbackempfängers begünstigen und dadurch zu einer oberflächlichen Verarbeitung des Feedbacks beitragen. In Peerfeedback-Situationen werden aufgrund eines – verglichen mit asymmetrischen Interaktionen – leichter herzustellenden gegenseitigen Vertrauens zudem Fehler eher offen angesprochen, was wiederum eine tiefergehende Reflexion verstärkt (Topping 2005). Allerdings zeigte sich, dass Anfänger beim Geben von Feedback weniger theoriegeleitete Kriterien nutzten oder seltener konkrete Situationen beschrieben als Experten (Prins/Sluijsmans/Kirschner 2006). Daher sollten angehende Lehrkräfte beim Geben von Feedback unterstützt werden.

Selbstwirksamkeit

Angehende Lehrkräfte sollten durch die Ausbildung Vertrauen in das eigene unterrichtliche Handeln entwickeln können. Solche Selbstwirksamkeitserwartungen werden in der Psychologie als Überzeugungen einer Person „neue oder schwierige Anforderungssituationen auf Grund eigener Kompetenz bewältigen zu können“ definiert (Schwarzer/Jerusalem 2002). Bezogen auf Lehrkräfte zeigte sich, dass diejenigen mit hoher Selbstwirksamkeit stärker motiviert sind, erfolgreich zu unterrichten und insgesamt einen herausfordernderen Unterricht gestalten. Lehrkräfte mit niedriger Selbstwirksamkeit hingegen trauen sich

nur wenig zu, was dazu führen kann, dass sie weniger komplexe Unterrichtsaktivitäten planen und sich weniger um schwache Schülerinnen und Schüler kümmern. Neben diesen Effekten auf das unterrichtliche Handeln stellen Selbstwirksamkeitserwartungen auch einen protektiven Faktor dar, da sie mit einer höheren Berufszufriedenheit, einer geringeren Anfälligkeit für Stress sowie höherem beruflichen Engagement zusammenhängen (Chesnut/Burley 2015).

Bei Lehramtsstudierenden kann insbesondere das erfolgreiche eigene Unterrichtshandeln kann zu einem Anstieg der Selbstwirksamkeit führen, aber auch stellvertretende Erfahrungen in Form des Beobachtens von Kommilitonen und Kommilitoninnen können die Selbstwirksamkeit positiv beeinflussen. Als weitere Quelle zum Aufbau von Selbstwirksamkeitserfahrungen zählt Feedback von Peers oder Dozierenden (Morris/Usher/Chen 2017). Aus den genannten Gründen kann Microteaching eine effektive Quelle zum Aufbau von Selbstwirksamkeitserwartungen darstellen, da es direkte und indirekte Erfahrungen sowie Gelegenheiten zu Feedback ermöglicht.

3 Fragestellungen und Hypothesen

Auf Grundlage dieser theoretischen Überlegungen haben wir ein Seminarkonzept für Lehramtsstudierende entwickelt und durchgeführt, in welchem wir *Core Practices* des Unterrichtens mittels Microteaching und videogestütztem Peerfeedback trainiert haben. Als unabhängige Variable variierten wir, ob das Geben von Peerfeedback durch ein vorgeschaltetes kurzes Feedbacktraining unterstützt wurde (Experimentalgruppe) oder nicht (Kontrollgruppe). Unsere Forschungsfragen lauteten:

1. Zeigen sich höhere Selbstwirksamkeitswartungen und ein höherer subjektiver Kompetenzzuwachs bei Studierenden, die vorab ein Peerfeedback-Training erhalten haben, gegenüber jenen, die kein Training zum Peerfeedback bekommen haben?
2. Wie werden das gegebene und das erhaltene Feedback bewertet? Gibt es Unterschiede in der Bewertung zwischen den Gruppen und zwischen verschiedenen Dimensionen des Feedbacks?
3. Wie wurde das Seminar insgesamt evaluiert?

4 Methode

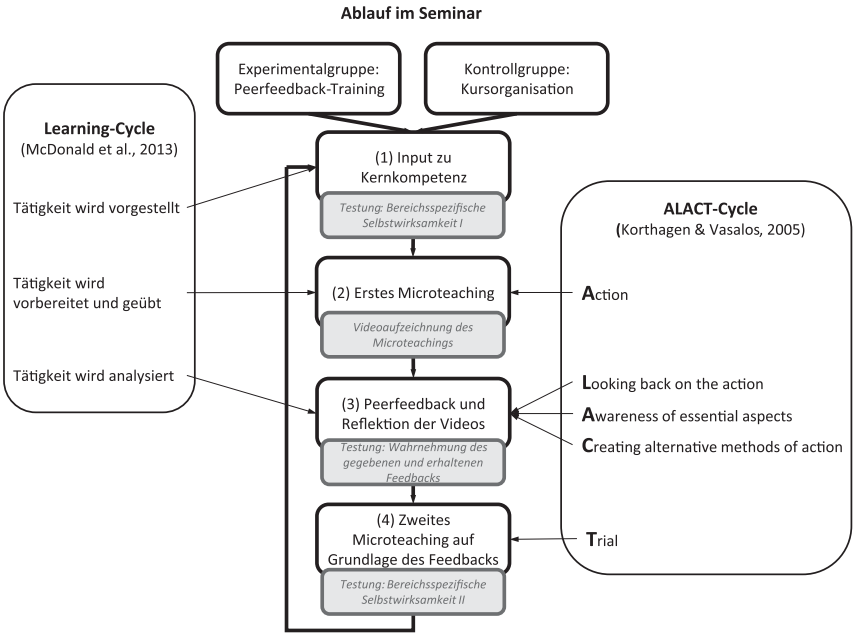
Teilnehmende und Versuchsplan

Wir boten zwei Parallelseminare im Rahmen des bildungswissenschaftlichen Begleitstudiums in der gymnasialen Lehramtsausbildung an. Es nahmen 34 Lehramtsstudierende (17 Studierende pro Seminar, $M_{\text{Alter}} = 24.94$, $SD_{\text{Alter}} = 3.04$, $M_{\text{Semester}} = 8.40$, $SD_{\text{Semester}} = 2.18$) teil. Die Studierendengruppen unterschieden sich weder hinsichtlich ihres Alters, $F(1, 27) = 1.41$, *ns,* noch ihres Studiensemesters signifikant voneinander, $F(1, 27) = 0.01$, *ns*. Es handelte sich um eine quasi-experimentelle Studie, bei der in einem Seminar in der ersten Sitzung ein kurzes Peerfeedback-Training durchgeführt wurde (Experimentalgruppe), während das andere Seminar kein solches Training erhielt (Kontrollgruppe).

Ablauf der Seminare

Über das Semester wurden die drei *Core Practices Erklären*, *Fragenstellen* und *Unterrichtsstörungen* trainiert. Eine Kerntätigkeit wurde entsprechend des Learning-Cycle von McDonald et al. (2013) in vier aufeinanderfolgenden Sitzungen bearbeitet (siehe Abb. 1). (1) In der ersten Sitzung wurde zu der Kerntätigkeit anhand eines Lehrvortrags das relevante theoretische sowie empirische Wissen vermittelt. (2) In der zweiten Sitzung führten die Studierenden reihum in Kleingruppen eine kurze Unterrichtssequenz durch (Microteaching). Die Mitstudierenden nahmen die Rolle der Schülerinnen und Schüler ein. Die Unterrichtssequenz zur jeweiligen Kerntätigkeit bereiteten die Studierenden anhand eines Leitfadens zu Hause vor. Alle Microteaching-Aktivitäten wurden auf Video für die Feedback-Sitzung (Sitzung 3) aufgezeichnet. Zusätzlich verwendeten die Studierenden Beobachtungsbögen, auf denen Notizen zu Aspekten der Kerntätigkeit gemacht werden konnten. (3) In der dritten Sitzung gaben sich die Studierenden in ihren Peergroups gegenseitig Feedback. Für das Feedback zogen die Studierenden die Beobachtungsbögen und die Videos heran. (4) In der vierten und letzten Sitzung zu einer Kerntätigkeit wurde die Unterrichtseinheit erneut durchgeführt. Die Studierenden überarbeiteten hierfür zu Hause auf Grundlage des erhaltenen Feedbacks ihren Unterrichtsplan.

Abbildung 1: Ablauf des Seminars, Darstellung der Messzeitpunkte sowie Darstellung der beiden Lernzyklen (Korthagen & Vasalos, 2005; McDonald et al., 2013) bezogen auf das Design des Seminars



5 Material

Peerfeedback-Training

In der Experimentalgruppe wurde in der ersten Seminarsitzung ein Peerfeedback-Training durchgeführt, das 60 Minuten dauerte. Zunächst wurden der Nutzen und die Funktionen von Feedback besprochen. Danach wurden die Feedback-Kriterien *Konkretheit*, *Empathie* und *Aktivierung* von Prilop, Weber und Kleinknecht (2018) vorgestellt. Es wurde mit den Studierenden diskutiert, wie man beim Feedback-Geben diese drei Kriterien berücksichtigen kann und wie Feedback formuliert werden sollte. Anhand eines Beispielvideos wurden das Beobachten von Unterrichtssituationen und das Feedbackgeben unter Berücksichtigung der Kriterien geübt. Im der Kontrollgruppe wurde kein Feedbacktraining durchgeführt, sondern Zeit für die Vorbesprechung des Seminarablaufs investiert.

Selbstwirksamkeit und subjektiver Kompetenzzuwachs

Vor und nach der Microteaching-Sitzung zu einer Kerntätigkeit wurde die Selbstwirksamkeit in Bezug auf die im Fokus stehende Kerntätigkeit erhoben (siehe Abb. 1). Dazu wurde ein Fragebogen angelehnt an Schwarzer und Jerusalem (2002) entwickelt, der 12 Items zur wahrgenommenen Selbstwirksamkeit auf einer sechsstufigen Skala („*lehne voll ab*" bis „*stimme voll zu*") enthielt (z. B. „*Ich fühle mich in der Lage, eine Erklärung nach den gelernten Prinzipien zu geben*"). Drei weitere Items erfragten den erlebten Kompetenzzuwachs (z. B. „*Ich habe viele neue Inhalte zum Thema „Erklären" gelernt*"). Die Reliabilität beider Skalen erwies sich für die verschiedenen Kerntätigkeiten und Messzeitpunkte durchgängig als gut bis sehr gut (die Cronbachs Alphas lagen zwischen $\alpha = .79$ bis $\alpha = .85$).

Feedback-Fragebogen

Nach der Seminarsitzung, in welcher sich die Microteaching-Gruppen gegenseitig Feedback gaben, wurde ein Fragebogen ausgefüllt, in dem zum einen das gegebene Feedback wie auch das erhaltene Feedback auf einer sechsstufigen Skala („*lehne voll ab*" bis „*stimme voll zu*") bewertet wurde. Der Fragebogen bestand jeweils aus zehn Items, die sich an den Kriterien konkret, empathisch und aktivierend orientierten (z. B. „*Ich habe zu konkreten Situationen Feedback gegeben*" (gegebenes Feedback zur Facette *Konkretheit*) oder „*Das Feedback hat mich motiviert*" (erhaltenes Feedback zur Facette *Empathie*). Die Reliabilität der Skalen erwies sich für die drei Kerntätigkeiten in Hinblick auf das gegebene und das erhaltene Feedback als sehr gut (Cronbachs $\alpha = .87$ bis $\alpha = .92$).

6 Ergebnisse

Zur Beantwortung der Frage, inwiefern das Feedbacktraining die wahrgenommene Selbstwirksamkeit und den subjektiven Kompetenzzuwachs beeinflusste, wurden drei multivariate zweifaktorielle Varianzanalysen mit Messwiederholung für die Tätigkeiten *Erklären*, *Fragen stellen* und *Unterrichtsstörungen* berechnet. Die Selbsteinschätzungen vor dem ersten sowie nach dem zweiten Microteaching (vgl. Abb. 1) bildeten hierbei den Messwiederholungsfaktor und der Vergleich zwischen Experimental- und Kontrollgruppe den Zwischengruppenfaktor. Abhängige Variablen waren die eingeschätzte Selbstwirksamkeit und der eingeschätzte Kompetenzzuwachs. Deskriptive Statistiken für die Skalen sind in Tab. 1 aufgeführt.

Tabelle 1: Deskriptive Statistik für die Skalen Selbstwirksamkeit und Kompetenzzuwachs

		Kontrollgruppe *M (SD)*		Experimentalgruppe *M (SD)*	
		prä	post	prä	post
Erklären	Selbstwirksamkeit	4.88 (.42)	5.17 (.51)	4.61 (.47)	5.16 (.55)
	wahrgenommene Kompetenz	4.14 (.41)	4.33 (.47)	3.91 (.47)	4.27 (.63)
Fragen stellen	Selbstwirksamkeit	4.91 (.48)	5.35 (.44)	4.64 (.75)	5.06 (.63)
	wahrgenommene Kompetenz	4.23 (.66)	4.54 (.66)	4.42 (.73)	4.35 (.81)
Unterrichtsstörungen	Selbstwirksamkeit	4.58 (.48)	5.02 (.51)	4.54 (.37)	4.41 (.57)
	wahrgenommene Kompetenz	4.84 (.29)	4.96 (.50)	4.77 (.43)	4.79 (.32)

Anmerkungen: Werte in beiden Skalen von 1 („lehne voll ab") bis 6 („stimme voll zu")

Die multivariate Varianzanalyse zur Kerntätigkeit *Erklären* ergab einen signifikanten Messwiederholungseffekt. Sowohl die Selbstwirksamkeitseinschätzungen, $F(1, 21) = 21.80$, $p < .001$, $\eta^2 = .51$, als auch die Kompetenzeinschätzungen, $F(1, 23) = 4.87$, $p < .05$, $\eta^2 = .18$, fielen nach den Microteachings höher aus als davor. Auch bei der Tätigkeit *Fragenstellen* zeigten sich höhere Selbstwirksamkeitseinschätzungen nach den Microteachings, $F(1, 21) = 14.02$, $p < .001$, $\eta^2 = .40$, allerdings kein höherer subjektiver Kompetenzzuwachs, $F(1, 21) = .64$, *ns*. Entgegen unserer Erwartung (siehe Forschungsfrage 1) konnten für beide Tätigkeiten keine signifikanten Gruppenunterschiede gefunden werden (Interaktion zwischen Gruppe und Messwiederholung: $Fs < 1.7$, *ns*). Bei der Kernkompetenz *Umgang mit Unterrichtsstörungen* zeigte sich entgegen der erwarteten Richtung, dass die Gruppe mit Feedbacktraining nach den Microteachings ihre Selbstwirksamkeit niedriger einschätzte als vorher, während die Einschätzungen in der Kontrollgruppe hingegen höher ausfielen (Interaktionseffekt: $F(1, 24) = 5{,}24$, $p = .05$, $\eta^2 = .18$).

Zur Beantwortung der zweiten Forschungsfrage (Unterschiede in Hinblick auf das gegebene und erhaltene Feedback) aggregierten wir die Bewertungen des gegebenen und erhaltenen Feedbacks über die Kerntätigkeiten hinweg (vgl. Tab. 2).

Tabelle 2: Deskriptive Statistik für die Skalen zur Bewertung des gegebenen und des erhaltenen Peerfeedbacks sowie für die Abschlussevaluation

		Kontrollgruppe *M (SD)*	Experimentalgruppe *M (SD)*
Bewertung des gegebenen Feedbacks insgesamt	Konkretheit	5.22 (.63)	5.19 (.65)
	Empathie	5.33 (.47)	5.08 (.74)
	Aktivierung	4.69 (.74)	4.28 (.89)
Bewertung des erhaltenen Feedbacks insgesamt	Konkretheit	5.35 (.59)	5.23 (.70)
	Empathie	5.44 (.59)	5.35 (.64)
	Aktivierung	5.00 (.73)	4.57 (.81)
Abschlussevaluation		1.37 (.28)	1.36 (.38)

Anmerkungen: Werte in beiden Skalen von 1 („lehne voll ab") bis 6 („stimme voll zu"). Bewertung des Seminars in der Abschlussevaluation in Schulnoten (1 = „sehr gut" bis 6 = „ungenügend")

Wir berechneten für jede der beiden abhängigen Variablen (Bewertung des gegebenen und des erhaltenen Feedbacks) jeweils eine zweifaktorielle Varianzanalyse mit den Feedback-Dimensionen Konkretheit, Empathie und Aktivierung als Messwiederholungsfaktor sowie der Kontroll- und Experimentalgruppe als Zwischengruppenfaktor. Ähnlich wie bei Forschungsfrage 1 zeigten sich auch bei Forschungsfrage 2 keine signifikanten Unterschiede zwischen den beiden Seminargruppen (für beide Haupteffekte: $Fs < .84$).

Der signifikante Haupteffekt Feedback-Dimensionen zeigte hingegen, dass die verschiedenen Dimensionen des gegebenen Feedbacks sehr unterschiedlich bewertet wurden, $F(2, 19) = 37.02$, $p < .001$, $\eta^2 = .80$. So wurde die Dimension *Aktivierung* in beiden Gruppen (mit und ohne Feedbacktraining) als geringer ausgeprägt bewertet als die Dimensionen *Konkretheit* und *Empathie* (*Aktivierung* versus *Empathie:* $t(21) = 8.48$, $p < .001$, $d = .95$, *Aktivierung* versus *Konkretheit:* $t(21) = 6.76$, $p < .001$, $d = .95$). Die Dimensionen *Konkretheit* und *Empathie* unterschieden sich nicht signifikant, $t(22) = 0.03$, *ns*. Beim erhaltenen Feedback resultierte ebenfalls ein Haupteffekt Feedbackdimensionen, $F(1,30) = 27.51$, $p < .001$, $\eta^2 = .71$. Konkret wurde die Dimension *Aktivierung* sowohl niedriger als die Dimension *Empathie*, $t(24) = 6.96$, $p < .001$, $d = .88$, als auch niedriger als die Dimension *Konkretheit*, $t(24) = 6.53$, $p < .001$, $d = .71$, bewertet. Darüber hinaus wurde auch die Dimension *Konkretheit* niedriger bewertet als die Dimension *Empathie,* $t(24) = 2.09$, $p < .05$, $d = .18$.

Die Abschlussevaluation wies insgesamt auf eine hohe Akzeptanz beider Seminare hin. Die im Mittel gegebene Schulnote lag bei 1.36 ($SD = .33$) und die Gruppen unterschieden sich hinsichtlich ihrer Evaluation nicht signifikant, $t(27) = 0.08$, *ns*.

7 Diskussion

Wir stellten ein Trainingsseminar zu Kerntätigkeiten des Unterrichtens vor, welches wir in einem quasi-experimentellen Kontrollgruppendesign evaluierten. Das Training basiert auf theoretischen Überlegungen zum Erwerb von Lehrkompetenzen („Lernzyklen"), wie sie von Korthagen und Vasalos (2005) sowie McDonald et al. (2013) formuliert wurden. Darauf aufbauend realisierten wir in unseren Seminaren Microteaching sowie videogestütztes Peer-Feedback als zentrale instruktionale Elemente. In Bezug auf die unabhängige Variable Peerfeedback-Training zeigten sich keine signifikanten Unterschiede zwischen den beiden Seminargruppen. Allerdings zeigten sich über beide Gruppen hinweg differenzielle Muster in Bezug auf Veränderungen der wahrgenommenen Selbstwirksamkeit und Kompetenz. Während bei den Tätigkeiten *Erklären* und *Fragen* stellen beide Seminargruppen im Verlauf des Seminars eine positive Entwicklung wahrnahmen, zeigte sich beim Thema *Unterrichtsstörungen* in den beiden Seminargruppen ein gegenläufiger Trend: Die Werte stiegen bei der Kontrollgruppe, während sie bei der Experimentalgruppe im Seminarverlauf abnahmen. Es ist möglich, dass die Studierenden mit Feedbacktraining suboptimale Verhaltensweisen im Umgang mit Unterrichtsstörungen bei ihren Kommilitonen genauer und kritischer beschrieben (Prins et al. 2006) und daher die Selbstwirksamkeit der Studierenden abnahm.

In Bezug auf die Bewertung des Peerfeedbacks zeigte sich, dass die Dimension Aktivierung am schwächsten bewertet wurde, während die Dimension Empathie (zumindest in Bezug auf das erhaltene Feedback) die höchsten Werte erhielt. Die Studierenden nahmen also ihre Kommilitonen als besonders empathisch war, vermissten jedoch tendenziell im erhaltenen Feedback konkrete Hinweise, was sie zukünftig bei der Ausübung der unterrichtlichen Tätigkeiten besser machen könnten. Dieses Ergebnis könnte ein Hinweis auf eine mögliche Einschränkung von Peerfeedback gegenüber Expertenfeedback sein.

Insgesamt macht die Evaluation des Seminars deutlich, dass Microteachings und videobasiertes Peerfeedback sehr akzeptierte und motivierende Maßnahmen zum Training von Kerntätigkeiten des Unterrichtens sind. Neben der sehr guten Gesamtnote betonten die Studierenden außerdem, dass sie die Seminare als hilfreich erachteten, da sie das Unterrichten im geschützten Rahmen üben konnten. Für die meisten unserer fortgeschrittenen Studierenden (8. Semester) war es das erste Mal, dass sie beim Unterrichten auf Video aufgezeichnet wurden. Diesbezüglich meldeten sie zurück, dass sie eine Hemmschwelle überwinden mussten und es ein ungewohntes Gefühl war, sich selbst auf Video zu betrachten, sie gleichzeitig aber in dieser Methode einen „riesigen Nutzen" sahen. In diesem Zusammenhang kann das Peerfeedback vor allem durch das hohe Maß an Empathie hier besonders dazu beigetragen haben, dass die Studierenden sich in einem geschützten Raum erlebten.

Zukünftige Studien sollten den hier vorgestellten Trainingsansatz systematischer und unter kontrollierteren Bedingungen untersuchen, als es im Rahmen der beiden Seminare im Feld möglich war. Aus diesem Grund planen wir derzeitige eine Serie experimenteller Studien mit dem Ziel, die optimale Kombination und Abfolge der skizzierten instruktionalen Elemente zu identifizieren sowie die Vor- und Nachteile von Peerfeedback gegenüber Feedback durch Experten genauer zu erforschen. Obwohl Peerfeedback für Lehrende eine effiziente Möglichkeit ist, in Seminaren mit Feedback zu arbeiten, konnten wir beobachten, dass die Rückmeldungen häufig oberflächlich waren (geringe Aktivierung) und sich nicht immer auf theoretisch bedeutsame Merkmale der Kerntätigkeiten bezogen. Dessen ungeachtet hat die vorliegende Feldstudie deutlich gezeigt, dass das Seminarkonzept, welches auf theoretisch etablierten Modellen des Fertigkeitserwerbs in der Lehrkraftbildung basiert (Korthagen/Vasalos 2005; McDonald et al. 2013) und handlungsorientierte Elemente wie Microteaching und videogestütztes Peerfeedback realisiert, von den Lehramtsstudierenden in hohem Maße als gewinnbringend erlebt wurde.

Literatur

Chesnut, S. R. & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. Educational Research Review, 15, 1–16.

Gielen, S., Peeters, E., Dochy, F., Onghena, P. & Stuyven, K. 2010. Improving the effectiveness of peer feedback for learning. Learning and instruction, 20, 304–315.

Grossman, P., Hammerness, K. & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. Teachers and Teaching: theory and practice, 15(2), 273–289.

Korthagen, F. & Vasalos, A. (2005). Levels in reflection: Core reflection as a means to enhance professional growth. Teachers and teaching, 11(1), 47–71.

Morris, D. B., Usher, E. L. & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. Educational Psychology Review, 29(4), 795–833.

Prilop, C. N., Weber, K. E. & Kleinknecht, M. (2018, August 29–31). Assessing pre-service teachers' adaptive peer feedback expertise: Effects of a field experience. Paper presented at the EARLI SIG 1 Conference, Helsinki, Finland.

Prins, F. J., Sluijsmans, D. M. & Kirschner, P. A. (2006). Feedback for general practitioners in training: Quality, styles, and preferences. Advances in Health Sciences Education, 11(3), 289.

Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. Zeitschrift für Pädagogik, 44. Beiheft: Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, 28–53.

Topping, K. J. (2005). Trends in peer learning. Educational psychology, 25(6), 631–645.

Weber, K. E., Gold, B., Prilop, C. N. & Kleinknecht, M. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online-and video-based self-reflection and feedback intervention. Teaching and Teacher Education, 76, 39–49.

Veränderung der Selbstwirksamkeitserwartungen von Lehrkräften in einer Fortbildung zum Umgang mit Heterogenität im Unterricht

Claudia Schmaltz und Timo Leuders

Abstract

In diesem Artikel gehen wir der Frage nach, über welche Selbstwirksamkeitserwartungen Lehrkräfte, die an einer Fortbildung zum Umgang mit Heterogenität im Unterricht teilnehmen, verfügen und wie sich die Selbstwirksamkeitserwartungen dieser Lehrkräfte durch die Teilnahme an der Fortbildung verändern. Selbstwirksamkeitserwartungen fassen wir dabei als breites Konstrukt. Im Fokus unserer Analyse ist folglich nicht eine quantifizierbare Erhöhung, sondern die Frage, welchen Einfluss die Fortbildung auf die Selbstwirksamkeitserwartungen der teilnehmenden Lehrkräfte hat. Hierzu wurden in einem Mixed-Methods-Design die Selbstwirksamkeitserwartungen der TeilnehmerInnen erfasst. In der hier dargestellten Fallstudie zeigte sich, dass der Fokus der Lehrkräfte auf der Frage liegt, wie sie der „Herausforderung Heterogenität" begegnen können und dass sich die Ausprägungen und Bezüge ihrer Selbstwirksamkeitserwartungen verändern. Nach der Fortbildung äußern sich die befragten Lehrkräfte unabhängig von Ausprägung und Bezug ihrer Selbstwirksamkeitserwartungen zuversichtlicher, diese Herausforderung im Alltag meistern zu können.

1 Problemstellung

Der Umgang mit Heterogenität durch Berücksichtigung der individuellen Voraussetzungen der Lernenden wird auch als „Schlüsselproblem der beruflichen Tätigkeit von Lehrkräften" (Trautmann/Wischer 2011, S. 105) bezeichnet. Im Rahmen einer Fortbildung, die Differenzierungsstrategien vermittelt, können sich die Selbstwirksamkeitserwartungen von Teilnehmenden nicht nur erhöhen, sondern aufgrund stärkerer Sensibilisierung für die Herausforderungen

auch verändern und sogar zunächst einmal verringern (Lipowsky 2014). Die Studie will dieses Phänomen als bedeutsame Komponente von Lernwegen herausarbeiten und empirisch belegen.

2 Theoretischer Hintergrund

Unbestritten scheint, „dass der Umgang mit Differenz und Heterogenität eine der zentralen schulischen Herausforderungen darstellt" (Bohl/Budde/Rieger-Ladich 2017, S. 7). Auch wenn diese Herausforderung bei Weitem nicht neu ist (vgl. Klafki/Stöcker 1976), wird die Frage, wie eine Lehrkraft „eine Gruppe von Schülern zeitgleich unterrichten kann, ohne die Bedürfnisse der Einzelnen zu vernachlässigen" (Trautmann/Wischer 2011, S. 7) in den letzten Jahren wieder stärker diskutiert und hat für die Professionalisierung von Lehrkräften eine große Bedeutung gewonnen (Beck et al. 2010; Trautmann/Wischer 2013).

Vielerorts wird an die Einstellung der Lehrkräfte appelliert: „Heterogenität soll nicht als Problem oder Belastung betrachtet, sondern als Chance gesehen und für gegenseitige Lernarrangements genutzt werden" (Wischer 2007, S. 423). Was genau diese Chance jedoch ausmacht, wird meist nicht näher erläutert (Budde 2012). Wischer (2008) spricht in diesem Zusammenhang auch von einer „Diskrepanz zwischen den zahlreich eingebrachten Forderungen und der unterrichtlichen Realentwicklung" (ebd. S. 715). Viele Lehrkräfte zweifeln laut einer Befragung von Kunze und Solzbacher (2008) an ihrem Können: Ein Großteil der befragten 180 Lehrkräfte sieht in der individuellen Förderung aller Schülerinnen und Schüler ein wichtiges Ziel der Entwicklung von Schule und Unterricht. 90 % der Lehrkräfte gaben aber an, dass es nicht möglich sei, jedem Kind gerecht zu werden. „Viele Lehrkräfte verlieren an Selbstbewusstsein; sie fühlen sich inkompetent und glauben nicht, dass sie ihre Defizite aufholen können" (Solzbacher 2008, S. 41). Hier deutet sich an, dass in Fortbildungen zum Umgang mit Heterogenität im Unterricht neben dem Professionswissen und den Einstellungen auch die Selbstwirksamkeitserwartungen von Lehrkräften adressiert werden müssen (Tschannen-Moran/McMaster 2009).

Schwarzer und Werner (2001) definieren Selbstwirksamkeitserwartung „als die subjektive Gewissheit, neue oder schwierige Anforderungssituationen aufgrund eigener Kompetenz bewältigen zu können" (ebd. S. 496). Betont wird dabei, dass es sich bei den Anforderungen nicht um einfache, durch routiniertes Handeln zu lösende Herausforderungen handelt, sondern „um solche, mit einem Schwierigkeitsgrad, der Anstrengung und Ausdauer für die Bewältigung erforderlich macht" (Schwarzer/Werner 2001, S. 496).

Bandura (1997) unterscheidet vier Einflussquellen auf die Selbstwirksamkeitserwartungen, von denen eigene Erfahrungen den größten Einfluss haben. Selbstwirksamkeitserwartungen können sich auf die Bewältigung von Anforde-

rungen ganz allgemein oder auf bestimmte Bereiche beziehen. Bei Lehrkräften beziehen sie sich darauf, die vielfältigen beruflichen Herausforderungen auch unter suboptimalen Bedingungen zu meistern. „Lehrer mit hoher Selbstwirksamkeit gestalten einen insgesamt herausfordernderen Unterricht, sie unterstützen Schüler bei der Erzielung von Lernfortschritten und haben mehr Geduld sowie Zuwendung für lernschwache Schüler, weil sie sich selbst mehr zutrauen, stärker motiviert sind und eine hohe Verantwortung für einen erfolgreichen und verständlichen Unterricht empfinden" (Schwarzer/Jerusalem 2002, S. 40). Edelmann (2002) betont, dass der Lehrer-Selbstwirksamkeit besonders in schulischen Reformprozessen eine wichtige Rolle zukommt: „[…] die Überzeugung eigener Handlungsmächtigkeit begünstigt die Bereitschaft der Individuen zu einer Veränderung, indem sie das Vertrauen in die eigene Handlungskompetenz stärkt" (ebd. S. 18). Es gilt aber auch hier, dass alleine das Zutrauen noch nicht zur erfolgreichen Bewältigung der Anforderungen führt: „Wenn Vorwissen oder Fähigkeiten nicht gegeben sind, nutzt die Selbstwirksamkeit relativ wenig, da sie eine Überschätzung eigener Möglichkeiten bedeuten kann" (Warner/Schwarzer 2009, S. 637).

In der hier berichteten Fortbildung zum Umgang mit Heterogenität im Unterricht wurde daher der Fokus zum einen auf die Vermittlung allgemein- und fachdidaktischer Kompetenzen gelegt, zum anderen war die Fortbildung so angelegt, dass alle teilnehmenden Lehrkräfte zu mehreren Zeitpunkten der Fortbildung aufgefordert waren, selbst Unterricht für heterogene Lerngruppen zu planen und ihre Überlegungen gemeinsam mit den anderen TeilnehmerInnen zu reflektieren (Leuders/Schmaltz/Erens 2018; Schmaltz 2019). Die Ausschreibung der Fortbildung richtete sich explizit an Lehrkräfte an Gemeinschaftsschulen, einer Schulform, die seit dem Schuljahr 2012/2013 in Baden-Württemberg existiert und das gemeinsame Lernen aller Schülerinnen und Schüler durch „an individuellem und kooperativem Lernen orientierte Unterrichtsformen" (Schulgesetz BW 2015) vorsieht. Trautmann und Wischer (2013) sehen gerade für diese Lehrkräfte einen großen Bedarf, sich im Umgang mit Heterogenität im Unterricht fortzubilden. „Da die meisten Lehrkräfte an Gemeinschaftsschulen ihre Ausbildung bereits hinter sich haben und Erfahrungen vor allem im herkömmlichen Regelschulsystem vorweisen können, ist es aus unserer Sicht notwendig […], ihnen Gelegenheit zu eröffnen, sich weiter zu professionalisieren" (Trautmann/Wischer 2013, S. 58). Aktuellen Studien zum Professionalisierungsbedarf von Lehrkräften an Gemeinschaftsschulen bestätigen dies (Bohl/Wacker 2016; Leuders/Vent-Schmidt 2017).

3 Fragestellung

In diesem Beitrag behandeln wir die Frage, über welche Selbstwirksamkeitserwartungen Lehrkräfte verfügen, die an einer Fortbildung zum Umgang mit Heterogenität teilnehmen und welchen Einfluss die Fortbildung auf die Selbstwirksamkeitserwartungen hat. Da „Umgang mit Heterogenität" als Gegenstand der Selbstwirksamkeitserwartungen breit gefächert ist und vielfältige allgemeindidaktische und fachdidaktische Aspekte umfasst müssen die hier adressierten Selbstwirksamkeitserwartungen als ein breites Konstrukt angesehen werden. Daher liegt das Hauptaugenmerk nicht auf einer pauschalen Erhöhung oder Verringerung, sondern auf der Analyse der Veränderung insbesondere auch hinsichtlich einer Verschiebung des subjektiven Fokus der Selbstwirksamkeitserwartungen.

4 Methodisches Vorgehen

Neben den von den Lehrkräften angefertigten Unterrichtsplanungen wurden in einem Mixed-Methods-Design zum einen die Selbstwirksamkeitserwartungen (Schwarzer/Schmitz 1999; 10 Items; Beispielitem: „Ich bin mir sicher, dass ich mich in Zukunft auf individuelle Probleme der Schüler noch besser einstellen kann") und die Überzeugungen bezogen auf heterogene Lerngruppen (Lang et al. 2010; 14 Items; Beispielitem: „Durch die Verschiedenheit der Kinder in meiner Klasse ergibt sich die Möglichkeit, dass Kinder voneinander lernen") der Lehrkräfte (n = 50) quantitativ erfasst. Zum anderen wurden mit ausgewählten Lehrkräften (n = 8) vor und nach der Fortbildung leitfadengestützte Interviews zu ihrer Unterrichtsplanung, ihrem Verständnis von Differenzierung und ihren Erwartungen an die Fortbildung geführt und mit der qualitativen Inhaltsanalyse (Mayring 2015) ausgewertet.

5 Ergebnisse

Zu Fortbildungsbeginn zeigt sich ein Zusammenhang der Selbstwirksamkeitserwartungen der Lehrkräfte mit einer positiven Sichtweise auf heterogene Lerngruppen (κ= .474**, p <.01**). Insgesamt schätzen die befragten Lehrkräfte ihre Selbstwirksamkeit auf einer Likertskala von 1–4 mit einem Mittelwert M von 2.89 (SD 0.36; Cronbachs α = .70) relativ hoch ein. Ein Zusammenhang der Selbstwirksamkeitserwartungen mit den von den Lehrkräften angefertigten Unterrichtsplanungen zeigt sich in der quantitativen Befragung nicht.

Bei der Analyse der Interviews lassen sich Veränderungen der Selbstwirksamkeitserwartungen der teilnehmenden Lehrkräfte identifizieren. Hierbei

zeigt sich, dass sich der Fokus der Selbstwirksamkeitserwartungen im Laufe der Fortbildung verändert.

Exemplarisch werden im Folgenden zwei Fälle dargestellt. Die Unterrichtsplanungen der beiden Lehrkräfte wiesen bereits zu Beginn der Fortbildung mehrheitlich differenzierende Elemente auf. Es ist daher von einer hohen fachlichen Kompetenz der beiden Lehrkräfte auszugehen.

Frau Müller erläutert zu Beginn der Fortbildung, dass sie im nächsten Schuljahr an einer Gemeinschaftsschule unterrichten werde. Sie sei „noch sehr motiviert" (FM Prä 16:49) und freue sich darauf. Sie hat bereits Kontakt zu Lehrkräften, die an einer Gemeinschaftsschule unterrichten, aufgenommen. „Um einfach in, ja in diese Arbeit reinzukommen. Das ist mein Job" (FM Prä 16:49). Sie gibt an, dass sie „dieses Arbeiten auch gut" (FM Prä 17:10) finde. „Ich finde es gut, den Schülern einfach diese unterschiedlichen Niveaustufen anzubieten, weil sich viele wirklich damit leichter tun. Und viele auch, ja, ein einfacheres Arbeiten haben werden" (FM Prä 17:15). Ebenso berichtet sie aber auch, dass die Lehrkräfte, mit denen sie sich ausgetauscht hat, die Differenzierung in ihrem Unterricht inzwischen wieder „zurückgefahren haben. Weil das einfach nicht möglich war" (FM Prä 19:08). Sie erwähnt in diesem Zusammenhang die zeitliche Mehrbelastung und gibt an, dass sie diese derzeit nicht bewältigen könnte. „Ich denke, da muss man auch ganz dringend noch daran arbeiten. Weil dieser Mehraufwand und diese Zeit, ganz ehrlich, ich könnte es nicht leisten, im Moment" (FM Prä 19:08). Im weiteren Verlauf des Gesprächs gibt Frau Müller zu bedenken, dass die Umstellung auf eine Gemeinschaftsschule und die damit verbundene innere Differenzierung „einfach ein großer Schritt" (FM Prä 22:50) sei, nicht nur für Eltern und Schüler, sondern eben auch für die Lehrkräfte. „Und als Lehrer muss man da irgendwie auch mitkommen. Mit diesen Schritten. Es sind nicht nur die Schüler und die Eltern die mitkommen müssen, sondern auch wir Lehrer und da wird man manchmal so ein bisschen überfahren oder übergangen und das ja, wirkt sich dann aus. Ist ganz klar, auf unsere Arbeit und ist eigentlich schade dann" (FM Prä 23:04). Im Anschluss an die Fortbildung äußert Frau Müller, dass sie festgestellt habe, „dass es gar nicht so kompliziert ist zu differenzieren" (FM Post 01:53) und sie mit Blick auf das nächste Schuljahr „wesentlich entspannter" (FM Post 07:16) sei. „Ich habe einfach auch im Gespräch mit den anderen mitgekriegt, es kocht jeder nur mit Wasser. Also es ist alles, ja, alles gut" (FM Post 05:49). Sie erläutert, dass der Austausch und die Reflexion in der Fortbildung für sie „auch schon ganz arg wertvoll" (FM Post 07:16) war und ihr „sehr gut getan" hat (FM Post 07:16). Sie hofft, dass es auch an ihrer Schule im Kollegium gelingen wird, sich gegenseitig zu entlasten. „Also das haben wir schon gesagt. Was dann nachher kommt, das muss man sehen. In der Planung sieht alles immer schön aus. Wie es nachher in der Realität ist, werden wir sehen, aber ich denke, wir

sind auf einem guten Weg und können da auch auf Sachen zurückgreifen“ (FM Post 07:16).

Während sich Frau Müller vorab insbesondere um die zeitliche Mehrbelastung sorgt, scheint sie nach der Fortbildung zuversichtlich, die auf sie zukommenden Herausforderungen gemeinsam mit ihren Kolleginnen und Kollegen bewältigen zu können.

Frau Schmidt unterrichtet bereits an einer Gemeinschaftsschule. Sie gibt an, sich für die Fortbildung angemeldet zu haben, weil sie „teilweise ein bisschen überordert damit“ (FS Prä 09:23) sei. Bezugnehmend auf die großen Leistungsunterschiede in ihrer Klasse macht sie sich insbesondere um die schwächeren Schülerinnen und Schüler Sorgen: „Ich weiß nicht wie ich das noch auffangen soll. Ich weiß es nicht und ich habe auch irgendwie als Lehrperson ein bisschen Angst, ehrlich gesagt, davor, dass ich irgendwann an den Punkt komme, wo ich es nicht mehr packe, die irgendwie noch zu halten. […] Wo ich an den Punkt komme, dass ich quasi fast schon aufgebe und sage, ich weiß gar nicht mehr, wie ich es machen soll alleine“ (FS Prä 09:45). Im weiteren Verlauf des Gesprächs äußert Frau Schmidt, dass die Arbeit für sie im Alltag derzeit „ein bisschen unbefriedigend“ (FS Prä 10:35) und Differenzierung „auf jeden Fall eine Herausforderung“ (FS Prä 12:10) sei. Der Austausch mit anderen Lehrkräften helfe ihr zu sehen, „dass alle eigentlich wirklich tatsächlich im gleichen Boot sitzen und sagen: Theorien schön und gut, aber so ist es halt einfach im Alltag und ich kann nicht mehr stemmen“ (FS Prä 12:48). […] „also helfen tut es wirklich. Hört sich ein bisschen komisch an, aber allein die Tatsache den anderen geht es nicht anders“ (FS Prä 13:54). Frau Schmidt gibt an, sich im Alltag bereits gefragt zu haben, ob es an ihr liege, dass sie die Herausforderungen an einer Gemeinschaftsschule als belastend empfindet: „Bin das nur ich? Kriege ich es nicht gebacken?“ (FS Prä 14:18). Nach der Fortbildung konzentrieren sich die Aussagen von Frau Schmidt auf die Frage, wie das Gelernte nun in ihrem Alltag umgesetzt werden kann. Konkrete Zweifel, an ihrer eigenen Fähigkeit, dies zu tun, äußert sie nicht mehr. Sie betont nun, dass es nicht alleine an ihr liege, ob das zukünftig gelänge: „Weil letztendlich ist es ja nicht eine Sache nur von den einzelnen Kollegen die hier sitzen“ (FS Post 05:06). Hierfür seien vielmehr Zeit und Austausch im Kollegium notwendig. „Okay, jetzt fehlt halt so, oder jetzt sollte noch so eine Zeit kommen wo man wirklich danach Zeit hat, sich hinzusetzen, das auf seine Schule anzupassen, mit Kollegen austauschen – und zwar wirklich austauschen“ (FS Post 05:02). Sie führt aus, dass das nur „Sinn macht […] wenn man wirklich Stunden dafür bereitgestellt bekommt“ (FS Post 06:27), weil solche Veränderungen der Unterrichtskultur ja „das komplette Kollegium stemmen“ (FS Post 06:35) muss. Für das Gelingen einer solchen Veränderung seien die Motivation und die Bereitschaft der Lehrkräfte, bestehende Strukturen zu hinterfragen, von großer Bedeutung: „[…] und halt die Motivation natürlich […] sodass man auch bereit ist, irgendwas:

okay, wir probieren es aus, wir werfen mal etwas irgendwie nochmal um und so“ (FS Post 07:00).

Frau Schmidt äußert vor der Fortbildung Zweifel an ihrer Kompetenz. Am Ende der Fortbildung betont sie, dass nicht einzelne Lehrkräfte, sondern die Zusammenarbeit im Kollegium, die Bereitstellung (zeitlicher) Ressourcen sowie die Motivation der Beteiligten ihrer Meinung nach entscheidend dafür sind, ob sich die Unterrichtskultur an einer Schule verändert.

6 Diskussion

Die dargestellten Fälle zeigen auf, dass sich nicht nur die Ausprägungen der Selbstwirksamkeitserwartungen dieser an einer Fortbildung zum Umgang mit Heterogenität im Unterricht teilnehmenden Lehrkräfte, sondern auch die Bezüge ihrer Selbstwirksamkeitserwartungen im Laufe der Fortbildung verändert haben.

In den Interviews liegt der Fokus der Befragten weniger auf dem erworbenen Professionswissen als vielmehr auf der Frage, ob und wie es ihnen zukünftig gelingen kann, der „Herausforderung Heterogenität“ für sich zufriedenstellend zu begegnen.

Die gemeinsame Arbeit und Reflexion an konkretem Material (hier: Unterrichtsplanungen) und deren Würdigung in der Fortbildung ermöglichten den teilnehmenden Lehrkräften, sich selbst als wirksam zu erleben. „Dieses Erleben eigener Wirksamkeit ist […] eine wichtige Schlüsselvariable für das Lernen von Lehrkräften. […] Deshalb ist es auch Aufgabe von Fortbildern, Lehrpersonen darin zu unterstützen, die Wirksamkeit ihres Handelns zu erleben“ (Lipowsky 2014, S. 9)

Nach der Fortbildung verändert sich die Perspektive der Lehrkräfte dahingehend, dass sie ihren Weg zum Umgang mit Heterogenität im Unterricht verstärkt als Prozess wahrnehmen. In den im Rahmen dieser Studie durchgeführten Interviews gaben alle Lehrkräfte am Ende der Fortbildung an, dass sie sich für den Umgang mit Heterogenität im Unterricht gestärkt fühlen und durch den Austausch und die Beschäftigung mit konkreten Lösungen in der Fortbildung wahrgenommen haben, dass andere Lehrkräfte auf die gleichen Schwierigkeiten treffen wie sie.

Die Analysen zeigen, dass Fortbildungen auf verschiedenen Ebenen die Selbstwirksamkeitserwartungen von Lehrkräften verändern, insbesondere dass eine Verschiebung des Fokus stattfinden kann: Von einer umfassenden vorgegebenen methodischen Differenzierungsstrategie hin zu lokaleren Strategien auf Aufgabenebene oder vom Maßstab hoher Anforderungen zu einem entlastenden Vergleich mit anderen Lehrkräften an anderen Schulen. Interpretiert man diese Befunde im Sinne der Selbstbestimmungstheorie von Deci & Ryan

(2008), so kann man feststellen, dass die in den Blick genommenen Wirkungen von Fortbildungen auf die Selbstwirksamkeitserwartung nicht allein darin bestehen müssen, diese über eine Kompetenzsteigerung und ein Kompetenzerleben zu verändern (was erst nach längerer Zeit der Umsetzung im Unterricht einsetzen kann), sondern auch durch soziale Eingebundenheit in eine Gruppe Gleichgesinnter und durch Autonomie in kleineren Handlungsfeldern.

Die Übertragbarkeit der hier vorgestellten Ergebnisse muss nicht nur aufgrund der Tatsache, dass es sich um eine Fallstudie handelt, eingeschränkt werden. Durch ihre Tätigkeit an Gemeinschaftsschulen ist bei den befragten Lehrkräften der Umgang mit Heterogenität in besonderem Maße expliziter Gegenstand ihrer täglichen Arbeit (Trautmann/Wischer 2013). Da alle befragten Lehrkräfte an einer Fortbildung zum Umgang mit Heterogenität im Unterricht teilgenommen haben, ist zudem davon auszugehen, dass ihre Motivation zur eigenen Professionalisierung besonders hoch ist und die TeilnehmerInnen keine repräsentative Stichprobe darstellen (Lipowsky 2010).

Literatur

Bandura, Albert (1997): Self-efficacy: the exercise of control. New York: Freeman

Bohl, Thorsten/Budde, Jürgen/Rieger-Ladich, Markus (Hrsg.) (2017): Umgang mit Heterogenität in Schule und Unterricht: grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen. Stuttgart: UTB GmbH; Bad Heilbrunn: Julius Klinkhardt

Bohl, Thorsten/Wacker, Albrecht (Hrsg.). (2016): Die Einführung der Gemeinschaftsschule in Baden-Württemberg: Abschlussbericht der wissenschaftlichen Begleitforschung (WissGem). Münster: Waxmann.

Budde, Jürgen (2012): Problematisierende Perspektiven auf Heterogenität als ambivalentes Thema der Schul- und Unterrichtsforschung. In: Zeitschrift für Pädagogik 58,4, S. 522–540

Deci, Edward L./Ryan, Richard M. (2008): Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. In: Canadian Psychology 49, S. 182–185

Edelmann, M. (2002): Gesundheitsressourcen im Beruf: Selbstwirksamkeit und Kontrolle als Faktoren der multiplen Stresspufferung. Weinheim: Beltz.

Klafki, Wolfgang/Stöcker, Hermann (1976): Innere Differenzierung des Unterrichts. In: Zeitschrift für Pädagogik, 22, 4, S. 497–523.

Kunze, Ingrid/Solzbacher, Claudia (2008): Individuelle Förderung in der Sekundarstufe I und II. Baltmannsweiler: Schneider-Verlag Hohengehren.

Lang, Eva/Grittner, Frauke/Rehle, Cornelia/Hartinger, Andreas (2010): Das Heterogenitätsverständnis von Lehrkräften im jahrgangsgemischten Unterricht der Grundschule. In: Hagedorn, Jörg/Schurt, Verena/Steber, Corinna/Waburg, Wiebke (Hrsg.): Ethnizität, Geschlecht, Familie und Schule: Heterogenität als erziehungswissenschaftliche Herausforderung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 315–331

Leuders, Timo/Vent-Schmidt, Kirsten (2017): Fachgruppenarbeit am Beispiel Mathematik. Den Blick auf Differenzierung mit einem fachlichen Fokus verbinden. In: Pädagogik 2017, 6, S. 42–46

Leuders, Timo/Schmaltz, Claudia/Erens, Ralf (2018): Entwicklung einer Fortbildung zu allgemeindidaktischen und fachdidaktischen Aspekten des Differenzierens. In: Biehler, Rolf/Lange, Thomas/Leuders, Timo/Rösken-Winter, Bettina/Scherer, Petra/Selter, Christoph (Hrsg.): Mathematikfortbildungen professionalisieren. Wiesbaden: Springer, S. 281–297

Lipowsky, Frank (2010): Lernen im Beruf. In: Müller, Florian/Eichenberger, Astrid/Lüders, Manfred/Mayr, Johannes: Lehrerinnen und Lehrer lernen – Konzepte und Befunde zur Lehrerfortbildung. Münster: Waxmann, S. 51–72

Lipowsky, Frank (2014): Sich selbst als wirksam erfahren – ein Schlüssel für erfolgreiches Lernen von Lehrerinnen und Lehrern: Interview mit Prof. Dr. Frank Lipowsky. In: Bertelsmann Stiftung (Hrsg.), Podium Schule 2013/2014, S. 8–9

Mayring, Philipp (2015): Qualitative Inhaltsanalyse: Grundlagen und Techniken (12., überarb. Aufl.). Weinheim/Basel: Beltz

Schmaltz, Claudia (2019): Heterogenität als Herausforderung für die Professionalisierung von Lehrkräften: Entwicklung der Unterrichtsplanungskompetenz im Rahmen einer Fortbildung. Springer VS: Wiesbaden

Schwarzer, Ralf/Schmitz, Gerdamarie S. (1999): Skala Lehrer-Selbstwirksamkeit. http://www.redibw.de/db/ebsco.php/search.ebscohost.com/login.aspx%3fdirect%3dtrue%26db%3dpdx%26AN%3dPT9004398%26site%3dehost-live

Schwarzer, Ralf/Jerusalem, Matthias (2002): Das Konzept der Selbstwirksamkeit. In: Zeitschrift für Pädagogik, Beiheft 44, S. 28–53

Solzbacher, Claudia (2008): Was denken Lehrerinnen und Lehrer über individuelle Förderung? In: Pädagogik, 60, 3, S. 38–42

Trautmann, Matthias/Wischer, Beate (2011): Heterogenität in der Schule Eine kritische Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften

Trautmann, Matthias/Wischer, Beate (2013): Professionalisierung von Lehrkräften für die Arbeit an Gemeinschaftsschulen. In: Bohl, Thorsten/Meissner, Sibylle (Hrsg.): Expertise Gemeinschaftsschule: Forschungsergebnisse und Handlungsempfehlungen für Baden-Württemberg. Weinheim/Basel: Beltz, S. 47–60

Tschannen-Moran, Megan/McMaster, Peggy (2009): Sources of Self-Efficacy: Four Professional Development Formats and Their Relationship to Self-Efficacy and Implementation of a New Teaching Strategy. In: Elementary School Journal, S. 228–245.

Warner, Lisa M./Schwarzer, Ralf (2009): Selbstwirksamkeit bei Lehrkräften. In: O. Zlatkin-Troitschanskaia et al. (Hrsg.): Lehrerprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung. Weinheim: Beltz, S. 629–640

Wischer, Beate (2007): Wie sollen LehrerInnen mit Heterogenität umgehen? Über „programmatische Fallen“ im aktuellen Reformdiskurs. In: Die deutsche Schule: DDS; Zeitschrift für Erziehungswissenschaft, Bildungspolitik und pädagogische Praxis, 99, 4, S. 422–433

Wischer, Beate (2008): „Binnendifferenzierung ist ein Wort für das schlechte Gewissen des Lehrers. In: Erziehung & Unterricht, 158, 9–10, S. 714–722

Transfer vom Wissen zum Können

Potenzial videofallbasierten Lernens zum Aufbau handlungsleitender Kognitionen

Sabine Schöb

1 Ausgangslage

Die Lehrer- und Erwachsenenbildung(-sforschung) sieht sich in der Diskussion um die Professionalität Lehrender mit dem Verhältnis von Theorie und Praxis bzw. konkreter dem Zusammenhang berufsrelevanten Wissens und praktischen Könnens konfrontiert (z. B. Nittel 2000; Syring et al. 2015, S. 668). Es bedarf – in der Wirksamkeit geprüfter – Konzepte, die den Aufbau handlungsleitender Kognitionen zur Diagnose von Lehr-Lernsituationen sowie zur Auswahl situationsangemessener Vorgehensweisen und deren erfolgreicher Umsetzung im Lehrhandeln fördern (z. B. Seidel et al. 2015, S. 85; Shavelson 2012, S. 30).

Hier setzt die präsentierte Untersuchung an, die sich der Übertragbarkeit der durch Videofallarbeit geförderten kognitiven Kompetenz auf die eigene unterrichtliche Praxis von Lehrenden widmet und dabei die Verbindung von berufsrelevantem Wissen und praktischem Können als „die beiden Quellen von Professionalität" (Nittel 2000, S. 71) in den Blick nimmt.

Aufbauend auf die Annahmen der Adaptive Control Thought Theory (Anderson 1982; 2014) wird ein Trainingskonzept entworfen, das in die Praxis der Lehrerausbildung integriert realisiert wurde. Darauf folgend werden die über einen Pre-Posttest-Vergleich gewonnenen Befunde zur Förderung von Lehrkompetenz über einen Aufbau handlungsleitender Kognitionen durch Videofallarbeit vorgestellt und kritisch diskutiert.

2 Konzeptionelle Grundlagen und Fragestellung

Gemäß den Diskursen zur Professionalisierung bzw. Expertise von Lehrenden ist kompetentes Lehrhandeln durch ein hohes Maß an Situativität gekennzeichnet (z. B. Baumert/Kunter 2006, S. 489; Nittel 2000, S. 85). So bildet die für die inhaltliche, zeitliche und soziale Strukturierung von Lehr-Lernprozessen notwendige (fach-)didaktische Kompetenz Lehrender nicht nur für die professio-

nelle Vorbereitung von Lehr-Lernsituationen eine unerlässliche Grundlage (Bromme 2008, S. 162 f.). Vielmehr stellt sie eine je Situation spezifisch hervorzubringende, nicht überdauernde Leistung dar (vgl. Goeze 2016, S. 24; Nittel 2000, S. 85). Konkreter kann sie als Fähigkeit zu 1) einer differenzierten Erfassung von Lehr-Lernsituationen und Reaktion auf deren zentralen Gegebenheiten unter 2) einer Berücksichtigung der Perspektive der Lernenden und 3) der Heranziehung didaktisch-methodischen Wissens (z. B. Berliner 2001, S. 469 f.; Goeze 2016, S. 34; Tietgens 1988, S. 37 f.) gefasst werden.

Die urteilskräftige Nutzung theoretischen Wissens in der konkreten Lehr-Lernsituation und damit die Bewältigung der Differenz von Wissen und Können (vgl. Kade 1990, S. 50; Tietgens 1988, S. 37) sind an berufsförmiges Deuten gebunden, durch das „zwischen singulären Beobachtungen von Interaktionssituationen einerseits und abstrakten pädagogischen Erkenntnissen andererseits Zusammenhänge" hergestellt werden, „Wissen auf konkrete Fälle" (Nittel 1998, S. 4) übertragen und dieses als handlungsleitende Kognitionen für die Gestaltung von Lehr-Lernsituationen verfügbar wird. Können setzt eine Verinnerlichung von Wissen als Handlungsbasis voraus, die im Zuge der Kontextualisierung des Wissens in der Anwendungssituation, durch eine sensible Einlassung auf situative und wechselnde Gegebenheiten erfolgt (vgl. Bromme 1992/2014, S. 123; König 2010, S. 61).

Die Arbeit mit videographierten Fällen, die Dokumentationen regelmäßiger Bestandteile (fach-)didaktischer Schlüsselsituationen darstellen, stellt eine Möglichkeit dar, sich mit realer Praxis auseinanderzusetzen und sowohl vermitteltes theoretisches Wissen anzuwenden als auch Lehrverhalten unter Rekonstruktion von Erfahrungswissen zu hinterfragen und darüber Strategien zur Bewältigung praktischer Probleme zu generieren bzw. zu differenzieren (vgl. z. B. Goeze/Hartz 2008, S. 68 f.; Krammer/Reusser 2005, S. 40 f.; Seidel et al. 2011, S. 260; Tripp/Rich 2012, S. 728 f.). Zahlreiche Studien zum videofallbasierten Lernen belegen dessen Potenzial zur Förderung professioneller Lehrkompetenz, fokussieren dabei jedoch auf die Förderung berufsrelevanten Wissens oder kognitiver Kompetenzen und blenden Fragen des Transfers der erworbenen Kognitionen in unterrichtliches Handeln sowie der Veränderung der Qualität der Gestaltung von Lehr-Lernprozessen weitgehend aus (vgl. z. B. Gaudin/Chaliès 2015, S. 54; Goeze 2016, S. 132 f., Krammer et al. 2016, S. 369; Sun/van Es 2015, S. 2). Dies ist mitunter dem Fehlen geeigneter Testverfahren und videobasierter Tools zur Erfassung von Lehrkompetenz insgesamt und der Erfassung des Übergangs von Wissen in Können im Speziellen geschuldet (vgl. Blomberg et al. 2013, S. 103 f.; Seidel et al. 2015, S. 85; Seidel/Thiel 2017, S. 12 f.).

Gemäß der von dem Kognitionspsychologen Anderson (1982, 1983) entwickelten Adaptive Control of Thought Theory gestaltet sich der Erwerb flexibel einsetzbarer, handlungsleitender Kognitionen als dreistufiger Prozess, bei dem

durch wiederholtes Handeln in Form von Übung zuvor separat vorliegende Wissenselemente zusammengeführt, zu Prozeduren gefestigt und zu handlungsleitenden Regeln generalisiert werden (vgl. Anderson 2014, S. 196; Bromme 1992/2014, S. 128; Lipowsky 2015, S. 88). Wendet man die Annahmen der Theorie auf das hier verfolgte Untersuchungsinteresse des Aufbaus der Kompetenz zur Diagnose von Lehr-Lernsituationen und deren Transfer in Handeln an, so lassen sich die drei Lernstufen wie folgt darstellen:

- *1. Stufe:* In der kognitiven Phase (deklaratives Wissen) erfolgt der Erwerb von Faktenwissen, das aus der Umwelt aufgenommen und in Form von Propositionen gespeichert wird. Dabei werden durch eine deklarative Enkodierung die Regeln für das Diagnostizieren gelernt, wobei die einzelnen, dafür relevanten kognitiven Operationen der Wahrnehmungsdifferenzierung, Perspektivenübernahme und Theorieverwendung für sich als Sinneinheiten abgelegt werden und unverbunden bleiben (vgl. Anderson 1982, S. 369 f.; 2014, S. 196).
- *2. Stufe:* In der assoziativen Phase (prozedurales Wissen) werden durch die wiederholte Auseinandersetzung mit videographierten authentischen, durch perspektivisches und theoretisches Wissen zur Deutung angereicherten Lehr-Lernsituationen die kognitiven Operationen zur Diagnose von Lehr-Lernsituationen zusammengefasst, zu einem Schemata verdichtet und darüber in eine prozedurale Form überführt. Es wird die Methode zur Durchführung der Prozedur in Form von Wenn-Dann-Verknüpfungen ausgearbeitet (vgl. Anderson 1982, S. 370; 2014, S. 196).
- *3. Stufe:* In der autonomen Phase (Automatisierung) wird schließlich die Ausbildung der Prozedur für die Fähigkeit zur Diagnose von Lehr-Lernsituationen weiter verfeinert und als handlungsleitende Regel verinnerlicht. Dabei wird das Vorgehen durch die wiederholte Auseinandersetzung über einzelne Erfahrungen hinweg zunehmend routinisiert und abstrahiert, wobei die Authentizität der videographierten Fallbeispiele eine Übertragbarkeit der Prozedur auf eine durch vergleichbare Situationsmerkmale und Verhaltensanforderungen (Wenn-Dann-Verknüpfungen) gekennzeichnete eigene unterrichtliche Praxis der Lehrenden wahrscheinlich macht (vgl. z. B. Anderson 2014, S. 200; Bromme 1992/2014, S. 121 f.; König 2010, S. 59 f.).

Inwiefern der Transfer der Prozedur gelingt steht im Zentrum der hier präsentierten Untersuchung, die sich mit der Frage beschäftigt, ob anhand der instruktional unterstützten, wiederholten Auseinandersetzung mit videographierten Lehr-Lernsituationen eine kognitive Denkstruktur zur Diagnose von Lehr-Lernsituationen so aufgebaut und eingeübt werden kann, dass diese als Handlungsgrundlage verinnerlicht wird und sich auf die Qualität sowie Situationsangemessenheit von Lehrhandeln auswirkt?

3 Methodisches Vorgehen

Aufbauend auf dem Prinzip der nutzeninspirierten Grundlagenforschung (vgl. Schrader/Goeze 2011), die praktischen Nutzen mit wissenschaftlichen Erkenntnissen verbinden möchte, setzt die Interventionsstudie direkt in der Ausbildungspraxis von angehenden Lehrkräften in Form eines Train-the-Trainer-Kurses an, um zugleich eine möglichst hohe interne und externe Validität sicherzustellen.

Design und Stichprobe

Die im Pre-Posttest-Design realisierte quasi-experimentelle Interventionsstudie wurde mit 82 angehenden Lehrkräften (30 % weiblich, 70 % männlich, Durchschnittsalter 26,8 Jahre) durchgeführt, die im Blick auf die institutionellen Voraussetzungen und Themenbereiche unter vergleichbaren Bedingungen als ehrenamtliche Kräfte im öffentlich-kirchlichen Kontext einer Wohlfahrts- und Hilfeorganisation arbeiten und eine Ausbildertätigkeit anstreben. Es wurden sechs standardisierte Erhebungen umgesetzt. Der Kurs umfasste 32 Stunden und war in die drei Phasen der Einführung, des Trainings und des Abschlusses unterteilt (s. Tab. 1).

Tabelle 1: Ablauf des Train-the-Trainer-Kurses

<table>
<tr><th>Pretest-Phase B: Transfer</th><th>Pretest-Phase A: Lernerfolg</th><th colspan="3">Interventionsphase</th><th>Posttest-Phase A: Lernerfolg</th><th>Posttest-Phase B: Transfer</th></tr>
<tr><td rowspan="2">Pretest: Lehrhandeln</td><td rowspan="2">Kontrollvariablen
Pretest: Fallanalyse
Einführende Grundlagen
Wissenstest</td><td colspan="3">Kontrollbedingung: Vermittlung theoretischer Konzepte ohne Fälle</td><td rowspan="2">Posttest: Fallanalyse
Kontrollvariablen und Akzeptanz</td><td rowspan="2">Posttest: Lehrhandeln</td></tr>
<tr><td colspan="3">Experimentalbedingung:</td></tr>
<tr><td></td><td></td><td>Didaktisierter Fall 1</td><td>Didaktisierter Fall 2</td><td>Didaktisierter Fall 3</td><td></td><td></td></tr>
<tr><td>konstant</td><td>konstant</td><td colspan="3">Experimental- vs. Kontrollbedingung</td><td>konstant</td><td>konstant</td></tr>
<tr><td>vor Beginn</td><td>Einführung: 8 Stunden</td><td colspan="3">Training: 16 Stunden</td><td>Abschluss: 8 Stunden</td><td>nach Abschluss</td></tr>
</table>

In der Interventionsphase wurde die Art der Aneignung des inhaltlich identischen theoretischen Grundlagenwissens in der Trainingsphase des Kurses als unabhängige Variable variiert (Experimentalbedingung (N = 38): Bearbeitung didaktisierter Videofälle in Dreiergruppen; Kontrollbedingung (N = 44): Vorträge und Gruppenarbeit ohne Fälle). Die Kurse wurden von drei für den her-

kömmlichen Lehrgang zuständigen und in dessen Umsetzung langjährig erfahrenen Praktikern geleitet, die auf Basis eines einheitlichen Skripts je eine Kontroll- und eine Experimentalbedingung gestalteten.

Datenerhebung und -auswertung

Als Pretest und Posttest wurde ein Lehrtraining standardisiert durchgeführt, bei dem die teilnehmenden Lehrenden eine zu einem herausfordernden Zeitpunkt „eingefrorene" Videofallsituation fortführten, wobei ihr Lehrhandeln videographiert wurde. Die Kompetenz zum situativen Handeln wurde auf der Grundlage der Videoaufzeichnungen der Gestaltung der Lehr-Lernsituation durch die Teilnehmenden erfasst. Auf der Basis der aus der Lehr-Lernforschung bekannten Qualitätskriterien des Unterrichts (vgl. Helmke 2012) wurden passend zu den kognitiven Operationen der Theorieverwendung und Perspektivenübernahme 1) die Strukturierung der Stoffvermittlung, 2) die Passung des Methodeneinsatzes, 3) die Teilnehmerorientierung und 4) die Teilnehmeraktivierung als Beobachtungskategorien definiert und je Kategorie drei Merkmale zur näheren Beschreibung von Ausmaß und Art der Berücksichtigung der Qualitätskriterien festgelegt[1]. Die Beurteilung bezog sich auf die gesamte Sequenz von durchschnittlich 10 Minuten Dauer als Analyseeinheit und erfolgte anhand eines Ratingbogens (vierstufige Skala von 1 = trifft zu bis 4 = trifft), der von zwei Codierenden unabhängig voneinander gekreuzt wurde. Dabei wurden mit .68 bis .75 nach Krippendorffs Alpha akzeptable bis gute Übereinstimmungswerte für die einzelnen Qualitätsmerkmale des Unterrichts erreicht (vgl. Milne/Adler 1999).

Die Auswertung wurde um eine Einschätzung der Angemessenheit der Weiterführung in Bezug auf den Stoppmoment der Videofallsituation ergänzt. Zur Erfassung der Situationsangemessenheit wurden jeweils die kognitiven Operationen der Diagnosekompetenz auf die Handlungssituation adaptiert. Erfasst wurde 1) das Aufgreifen von zentralen, in der Fallsituation gegebenen Situationsmerkmalen, 2) das Aufgreifen von vorab gezeigtem Kursleiter-Teilnehmer-Interaktionsverhalten in der Fallsituation sowie 3) das Aufgreifen von in der Fallsituation angelegten didaktisch-methodischen Vorgehensweisen (vierstufige Skala von 1 = trifft zu bis 4 = trifft nicht zu). In der Einschätzung wurden vergleichbar gute Übereinstimmungswerte erreicht.

---

1 In der Kategorie der Teilnehmerorientierung wurden beispielsweise folgende Merkmale gewählt: 1) Es werden Verknüpfungen der Inhalte mit Erfahrungen aus der Lebenswelt der Teilnehmenden hergestellt; 2) die didaktisch-methodische Umsetzung des Lehr-Lernprozesses erfolgt unter Fokus auf die Lernenden; 3) die Lehrperson berücksichtigt die individuellen Wissensstände der Teilnehmenden im Unterrichtsaufbau/Lernprozess.

In der Experimentalbedingung wurde zudem über die Reflexion der eigenen Vorgehensweise der Lehrenden bei der Fallanalyse zum Zeitpunkt vor und nach der Trainingsphase untersucht, ob die Fallarbeit zu einer Prozeduralisierung von Wissen führt. Es wurden leitfadengestützte Interviews mit den teilnehmenden Lehrenden geführt, in denen sie gebeten wurden 1) ihr eigenes Vorgehen zu erläutern und 2) zu begründen, warum sie das Vorgehen gewählt haben. Bei der strukturierenden inhaltsanalytischen Codierung der transkribierten Interviews (Mayring 2010) wurde aufbauend auf die theoretischen Annahmen zur Prozeduralisierung von Wissen (Anderson 1982, S. 369 f.; 2014, S. 196) 1) die Art der Darstellung des eigenen Vorgehens (offenes bzw. auf einzelne Aspekte fokussiertes Vorgehen vs. entschiedenes bzw. umfassend angelegtes Vorgehen) und 2) die Art der Begründung (keine/kaum Begründung vorhanden vs. viel/ausführliche Begründung) zeilenweise ausgewertet, um zu prüfen, ob sich bei den Lehrenden zunehmend ein „Wissen, wie etwas zu tun ist" (Hasler Roumois 2007, S. 46) zeigt: die Lehrenden also wissen, was sie tun und begründen können, warum sie es so und nicht anders machen.

4 Ergebnisse

Für die hier fokussierte Forschungsfrage ist die Veränderung der Qualität und der Situationsangemessenheit des Lehrhandelns von Interesse. Diese Erkenntnisse werden anschließend an theoretische Annahmen zum professionellen Lehrhandeln rückgebunden und es werden Implikationen für Forschung und Praxis herausgearbeitet.

Darstellung der Ergebnisse

Die Auswertung der videographierten Handlungssituationen in ihrer Qualität erfolgte über eine Bildung von Mittelwerten der jeweils drei Itemwerte je Beobachtungskategorie. Je kleiner der Mittelwert, desto besser stellt sich die Unterrichtsqualität in den einzelnen Kategorien dar, da die Teilnehmenden in ihrem Lehrhandeln die zugrunde liegenden Kriterien überwiegend bis durchgängig erfüllt haben (Skala von 1 = trifft zu bis 4 = trifft nicht zu).

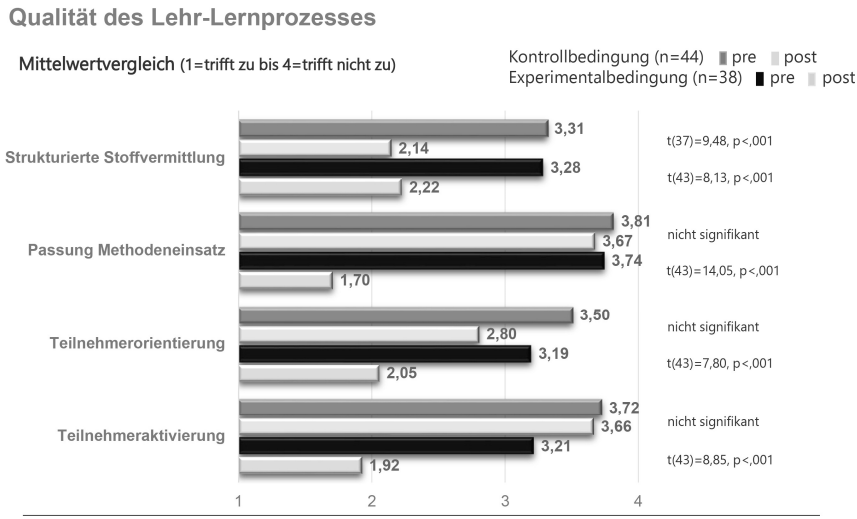
In den Pretestergebnissen zeigen sich kaum Unterschiede je Kursbedingung. Alle Teilnehmenden schneiden in ihrem Lehrhandeln weniger gut ab. Eine didaktisch strukturierte Stoffvermittlung ist eher nicht gegeben, das Vorgehenskonzept erscheint im Aufbau eher nicht stimmig (Strukturiertheit der Stoffvermittlung M = 3,29; SD = 0,42). In der Vorgehensweise konzentrieren sich die Lehrenden überwiegend auf den Einsatz nur einer Methode, die eine eher geringe Kohärenz zum thematisierten Inhalt aufweist (Passung des Methodeneinsatzes M = 3,45; SD = 0,63). Gleichzeitig erfolgt die Umsetzung der

Lehr-Lernsituation unter hoher Konzentration auf den Stoff und die eigene Lehrperson und wird eher nicht an den Erfahrungen und Wünschen der beteiligten Lernenden ausgerichtet (Teilnehmerorientierung M = 3,29; SD = 0,43). Die Lehrenden legen wenig Wert auf die Aufmerksamkeit und aktive Beteiligung der Lernenden am Lehr-Lernprozess und ermöglichen eher keine Arbeitsphasen, in denen die Lernenden sich eigenaktiv mit Inhalten auseinandersetzen (Teilnehmeraktivierung M = 3,21; SD = 0,51).

Betrachtet man die zu Beginn und zum Ende des Kurses realisierten Handlungssituationen im Vergleich, lässt sich für alle Teilnehmenden eine Veränderung in der Gesamtbeurteilung der Qualität feststellen, wobei sich die Veränderung bei den Kontrollgruppenmitgliedern auf wenige Kategorien beschränken und nur die Experimentalgruppenmitglieder in allen Kategorien gleichbleibende oder bessere Beurteilungen erzielen können. Während die Teilnehmenden der Kontrollgruppe nur im Bereich der Strukturierung der Stoffvermittlung einen signifikanten Zuwachs verzeichnen können[2], zeichnen sich die Vorgehensweisen der Experimentalgruppenmitglieder darüber hinaus auch durch einen vielfältigeren und zu den Inhalten passenderen Einsatz von Methoden aus (von M = 3,74; SD = 0,41 auf M = 1,70; SD = 0,62; t = 14,05; p = 0,000; d = 0,82). Zudem konzentriert sich ihr Vorgehen im Posttest gegenüber dem Pretest weniger auf sie als Lehrende und die Vermittlung des Stoffs, sondern orientiert sich stärker am Vorwissen und den Vorerfahrungen ihrer Lernenden (Teilnehmerorientierung von M = 3,19; SD = 0,53 zu M = 2,05; SD = 0,49; t = 7,80; p = 0,000; d = 0,71). Auch legen sie in der Ausrichtung ihres Vorgehens höheren Wert auf die kognitive Aktivierung ihrer Teilnehmenden (von M = 3,21; SD = 0,54 zu M = 1,92; SD = 0,61; t = 8,85; p = 0,000, d = 0,80) (vgl. Abb. 1).

2 Der positive Effekt des herkömmlichen Kurses in Bezug auf eine strukturierte Stoffvermittlung könnte auf die Art der dort gestellten Aufgaben und Übungen zurückzuführen sein, die mitunter didaktische Modelle bereithielten, die für Planungs- und Umsetzungsentwürfe für Kurse genutzt werden sollten.

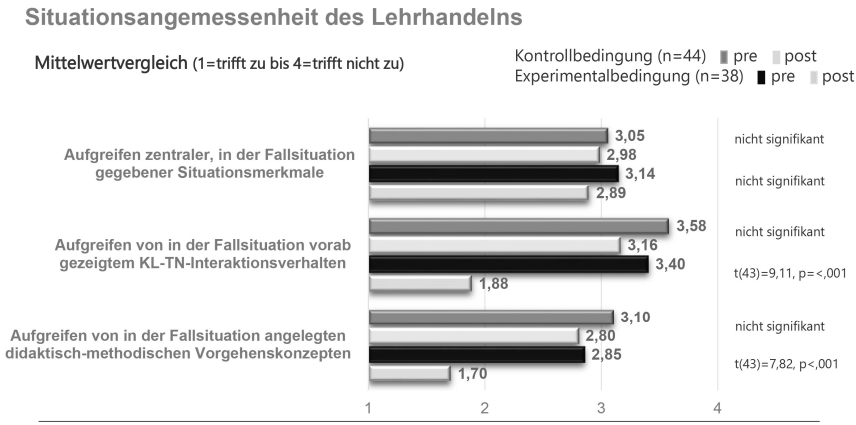
Abbildung 1: Veränderung der Qualität des Lehr-Lernprozesses



Die Angemessenheit des im Lehrtraining gezeigten Lehrhandelns in Bezug auf das Lehr-Lerngeschehen im videographierten Freeze-Fall kann im Pretest für alle Teilnehmenden in allen drei Dimensionen als eher weniger passend beurteilt werden (Situationsmerkmale M = 3,08; SD = 0,76; Interaktionsverhalten M = 3,49; SD = 0,48; didaktisch-methodisches Vorgehen M = 2,98; SD = 0,83). Viele der Teilnehmenden neigen unabhängig von ihrer Gruppenzugehörigkeit dazu, nicht an den gezeigten Fall anzuknüpfen und entweder die Situation bei Null beginnend und nach einem eigenen Konzept umzusetzen oder sie fokussieren auf einige wenige Aspekte oder Teilnehmerreaktionen. Diese behalten sie gelegentlich im Blick, allerdings, auf Kosten der Wahrnehmung der Gesamtgruppe der Lernenden.

Im Posttest zeigt sich bei den Kontrollgruppenmitgliedern diesbezüglich keine nennenswerte Veränderung, wohingegen die Teilnehmenden der Experimentalbedingung sich zwar nicht in der Dimension des Aufgreifens von Situationsmerkmalen der Freeze-Situation verbessern (M = 2,89; SD = 0,76), jedoch das Interaktionsverhalten des Kursleitenden und der Lernenden in der Fallsituation deutlich umfassender berücksichtigen (von M = 3,40; SD = 0,42 auf M = 1,88; SD = 0,61; t = 9,11; p = 0,000; d = 0,80). Zudem versuchen sie, passender an das darin dargelegte Vorgehenskonzept anzuknüpfen (von M = 2,85; SD = 0,51 auf M = 1,70; SD = 0,57; t = 7,80; p = 0,000; d = 0,82) (vgl. Abb. 2).

Abbildung 2: Veränderung der Situationsangemessenheit des Lehrhandelns



Der bei den Teilnehmenden der Experimentalbedingung vor und nach der Trainingsphase erhobene Grad der Prozeduralisierung der kognitiven Operationen zur Diagnose von Fällen ist vor dem Beginn der Auseinandersetzung mit den didaktisierten Videofällen gering ausgeprägt. So sind die Vorgehensbeschreibungen der befragten Teilnehmenden durch Offenheit und Unentschiedenheit geprägt. Die Teilnehmenden handeln interessensgeleitet und haben dabei oft nur einzelne Aspekte im Fokus. Eine Begründung des Vorgehens erfolgt in der Regel nicht. Drei Teilnehmeräußerungen werden im Folgenden exemplarisch dargestellt:

> „Ich schaue mir die Fälle einfach mal an und sehe dann, was ich dazu sagen kann.“ [TN3][3]

> „So richtig weiß ich noch nicht wie ich am besten rangehen werde, ich mache das vom Inhalt des Falls abhängig. Je nachdem worum es geht, finden sich sicher Ansatzpunkte für mich.“ [TN14]

> „Ich bin noch unentschieden, wie ich an die Analyse rangehen werde, am meisten interessiert mich die Sicht der Teilnehmenden, da werde ich in jedem Fall schwerpunktmäßig ansetzen.“ [TN27]

Nach der Bearbeitung von drei didaktisierten Fällen in der Trainingsphase zeigt sich ein anderes Bild. Die Teilnehmenden wirken entschiedener und haben ein

3 Die Teilnehmenden der Experimentalbedingung wurden mit laufenden Nummern von 1–38 gekennzeichnet.

klares und umfassendes Vorgehenskonzept, das sie anteilig begründen. So zeigt sich bei den drei exemplarischen Teilnehmenden folgende Veränderung:

> „Inzwischen habe ich ein klares Vorgehenskonzept, das gut funktioniert und mit dem ich zu einer umfassenden Betrachtung von solchen Fällen komme.“ [TN3]

> „Eigentlich stecken in den Situationen oft vergleichbare Fragen drin, weswegen sie sich ähnlich erschließen lassen. Erstmal ist es gut sich einen Überblick zu verschaffen, dann versuche ich zu verstehen, was in den Lehrenden und Lernenden vorgeht und wie man die Situation gut gestalten könnte.“ [TN14]

> „Wichtig scheint es mir, das Kernproblem des Falls herauszuarbeiten und daraus dann Vorgehensideen abzuleiten. Schließlich sollten diese zu den Zielen, Inhalten und Teilnehmenden passen.“ [TN27]

Betrachtet man die Äußerungen vor und nach der Trainingsphase im Vergleich zueinander, lässt sich eine deutliche Veränderung der Qualität des Wissens über das Vorgehen bei der Fallanalyse feststellen, die durch eine Wissenskompilierung und die Ausbildung einer Handlungsfolge für die Diagnose von Lehr-Lernsituationen gekennzeichnet ist.

Diskussion der Ergebnisse

Die Ergebnisse zeigen, dass das in der Studie eingesetzt videofallbasierte Angebot die Kompetenz zum situativen Handeln signifikant besser fördern kann als das traditionell umgesetzte Angebot. Die anhand der Fallanalysen trainierten kognitiven Operationen der Situationsbeschreibung, Perspektivenübernahme und Theorieverwendung werden als Deutungs- bzw. Reaktionsmuster für das unterrichtliche Handeln der angehenden Lehrenden bedeutsam, was gemäß Anderson (2014, S. 199 f.) anhand einer Ausbildung von Prozeduren erklärt werden kann. Für eine Prozeduralisierung spricht zudem die sich durch das Training signifikant verändernde Art der Vorgehensreflexion der Lehrenden, die zunehmend durch ein „Wissen, wie etwas zu tun ist“ (Hasler Roumois 2007, S. 46) geprägt ist. Somit lässt sich mit Blick auf die Forschungsfrage zum Transferpotenzial der instruktional unterstützten Auseinandersetzung mit Videofällen insgesamt festhalten, dass diese zu einer Veränderung von Lehrhandeln beitragen kann und sich insbesondere die über die eingesetzten didaktisierten Videofälle gezielt geförderten kognitiven Operationen der Perspektivenübernahme und Theorieverwendung als relevante Handlungsgrundlage erweisen bzw. sich auf die Qualität und Situationsangemessenheit des Lehrhandelns auswirken.

Die Befunde lassen sich in den wissenschaftlichen Diskurs zur möglichen Überwindung trägen Wissens (Renkl 1996, S. 79) einordnen und bieten einen Ansatz zur Untersuchung des bis dato wenig untersuchten Übergangs zwischen Wissen und Können. Zudem verdeutlichen sie zentrale Einflussfaktoren für die Gestaltung von Aus- und Fortbildungsangeboten zur Förderung des professionellen Wissens und Könnens Lehrender.

Der zentrale Nutzen des dargelegten Vorhabens besteht jedoch in einer verbesserten Praxis der Erwachsenbildung/Weiterbildung. Mit der anwendungsorientierte Entwicklungs- und Forschungsstrategie wurde Wert auf eine feldnahe und bedarfsgerechte Konzeption des Lernangebots sowie eine hohe interne und externe Validität der Ergebnisse durch das quasi-experimentelle Design der Untersuchung gelegt, um eine Implementation des in der Wirkung geprüften Konzepts und dessen Bausteine in die Aus- und Weiterbildungspraxis zu begünstigen (vgl. Goeze 2016, S. 129; Gräsel 2010, S. 8; Schrader 2010, S. 90 f.).

Grenzen der Studie bestehen neben der kleinen Stichprobengröße und einer Durchführung mit natürlichen Gruppen, in der Heranziehung von Lehrtrainings statt der Untersuchung des realen Lehrhandelns. Offen bleibt, ob und wie sich ein Transfereffekt des videofallbasierten Lernens im Handlungsalltag der Lehrenden zeigt. So bedarf es weiterer, in Form von Design-Based-(Implementation-)Research (z. B. Penuel/Fishman 2012) in die Praxis integriert umgesetzter Untersuchungen zu Fragen der Transferwirkung von Fallarbeit (Seidel/Thiel 2017), die auch andere Bildungskontexte sowie erfahrene Lehrkräfte und deren alltägliches Handeln in einem Vorher-Nachher-Vergleich einschließen. Auch könnten längsschnittliche Designs, die über wiederholte Messzeitpunkte die Veränderung von Kompetenzen nachzeichnen, weiteren Aufschluss zur Nachhaltigkeit der Effekte bieten (vgl. Krammer et al. 2016, S. 369; vgl. auch Blomberg et al. 2013, S. 106 f.).

5 Fazit

In der vorliegenden Arbeit wurde ein Konzept instruktional unterstützten videofallbasierten Lernens zur Professionalisierung von Lehrenden in der Transferwirkung betrachtet. Die erzielten Befunde geben Auskunft über den möglichen Aufbau kognitiver Kompetenzen, die das Handeln Lehrender leiten und zu dessen Optimierung beitragen können. Sie unterstreichen die Notwendigkeit, sich dem Forschungsdesiderat der transferwirksamen Gestaltung videofallbasierter Aus- und Weiterbildungskonzepte für Lehrende weiter zu widmen. Im Zuge nutzeninspiriert angelegter Vorhaben können Lernmaterialien und Lernumgebungen entstehen, die nicht nur eine valide Prüfung der Wirksamkeit

ermöglichen sondern aufgrund ihrer anwendungsorientierten Konzeption zugleich attraktiv für eine Nachnutzung durch die Bildungspraxis werden.

Literatur

Anderson, John R. (1982): Acquisition of cognitive skill. In: Psychological Review 89, H. 4, S. 369–406.

Anderson, John R. (2014): Kognitive Psychologie. 7., erw., überarb. und neu gestalt. Aufl. Berlin, Heidelberg: Springer.

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520.

Berliner, David C. (2001): Learning about and learning from expert teachers. In: International Journal of Educational Research 35, H. 5, S. 463–482.

Blomberg, Geraldine/Renkl, Andreas/Sherin, Miriam G./Borko, Hilda/Seidel, Tina (2013): Five research-based heuristics for using video in pre-service teacher education. In: Journal for Educational research Online 5, H. 1, S. 90–114.

Bromme, Rainer (1992/2014): Der Lehrer als Experte: Zur Psychologie des professionellen Wissens. Reprint der Originalausgabe von 1992. Münster: Waxmann.

Bromme, Rainer (2008): Lehrerexpertise. In: Schneider, Wolfgang/Hasselhorn, Marcus (Hrsg.): Handbuch der Pädagogischen Psychologie. Göttingen: Hogrefe, S. 159–167.

Gaudin, Cyrill/Chali⊠s, Sébastian (2015): Video viewing in teacher education and professional development: A literature review. In: Educational Research Review 16, S. 41–67.

Goeze, Annika (2016): Professionalitätsentwicklung von Lehrkräften durch videofallbasiertes Lernen – Voraussetzungen, Prozesse, Wirkungen. Bielefeld: Bertelsmann.

Goeze, Annika/Hartz, Stefanie (2008): Die Arbeit an Fällen als Medium der Professionalisierung von Lehrenden. In: Report – Zeitschrift für Weiterbildungsforschung 31, H. 3, S. 68–78.

Gräsel, Cornelia (2010): Stichwort: Transfer und Transferforschung im Bildungsbereich. In: Zeitschrift für Erziehungswissenschaft 13, H. 1, 7–20.

Hasler Roumois, Ursula (2007): Studienbuch Wissensmanagement. Grundlagen der Wissensarbeit in Wirtschafts-, Non-Profit- und Public-Organisationen. Zürich: Orell Füssli.

Helmke, Andreas (2012): Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. 4., überarb. Aufl. Seelze: Klett-Kallmeyer.

Kade, Silvia (1990): Handlungshermeneutik. Qualifizierung durch Fallarbeit. Bad Heilbrunn: Klinkhardt.

König, Johannes (2010): Lehrerprofessionalität – Konzepte und Ergebnisse der internationalen und deutschen Forschung am Beispiel fachübergreifender, pädagogischer Kompetenzen. In: König, Johannes/Hofmann, Bernhard (Hrsg.): Professionalität von Lehrkräften – Was sollen Lehrkräfte im Lese- und Schreibunterricht wissen und können? Berlin: DGLS, S. 40–105.

Krammer, Kathrin/Hugener, Isabelle/Biaggi, Sandro/Frommelt, Manuela/Fürrer Auf der Maur, Gabriela/Stürmer, Kathleen (2016): Videos in der Ausbildung von Lehrkräften: Förderung der professionellen Unterrichtswahrnehmung durch die Analyse von eigenen bzw. fremden Videos. In: Unterrichtswissenschaft 44, H. 4, S. 357–372.

Krammer, Kathrin/Reusser, Kurt (2005): Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. In: Beiträge zur Lehrerbildung 23, H. 1, S. 35–50.

Lipowsky, Frank (2015): Unterricht. In: Wild, Elke/Möller, Jens (Hrsg.): Pädagogische Psychologie. Berlin: Springer, S. 69–105.

Mayring, Philipp (2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 11., aktual. u. überarb. Auflage. Weinheim: Beltz.

Milne, Marcus J./Adler, Ralph W. (1999): Exploring the Reliability of Social and Environmental Disclosures Content Analysis. In: Accounting, Auditing & Accountability Journal 12, H. 2, S. 237–256.

Nittel, Dieter (1998): Das Projekt „Interpretationswerkstätten". Zur Qualitätssicherung didaktischen Handelns. In: Grundlagen der Weiterbildung – Praxishilfen 9.20.30.9, S. 1–16.

Nittel, Dieter (2000): Von der Mission zur Profession. Stand und Perspektiven der Verberuflichung der Erwachsenenbildung. Bielefeld: Bertelsmann.

Penuel, William R./Fishman, Barry J. (2012): Large-scale science education intervention research we can use. In: Journal of Research in Science Teaching 49, H. 3, S. 281–304.

Renkl, Andreas (1996): Träges Wissen: Wenn Erlerntes nicht genutzt wird. In: Psychologische Rundschau 47, H. 2, S. 78–92.

Schrader, Josef (2010): Mediengestützte Fallarbeit. Grundlagen und Zielsetzungen eines Forschungs- und Entwicklungsprojekts zur Kompetenzentwicklung von Lehrenden in der Erwachsenenbildung. In: Schrader, Josef/Hohmann, Reinhard/Hartz, Stefanie (Hrsg.): Mediengestützte Fallarbeit. Konzepte, Erfahrungen und Befunde zur Kompetenzentwicklung von Erwachsenenbildnern. Bielefeld: Bertelsmann, S. 71–124.

Schrader, Josef/Goeze, Annika (2011). Wie Forschung nützlich werden kann. In: Report – Zeitschrift für Weiterbildungsforschung 34, H. 2, S. 67–76.

Seidel, Tina/Stürmer, Kathleen/Blomberg, Geraldine/Kobarg, Mareike/Schwindt, Katharina (2011): Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? In: Teaching and Teacher Education 27, H. 2, S. 259–267.

Seidel, Tina/Stürmer, Kathleen/Schäfer, Stefanie/Jahn, Gloria (2015): How preservice teachers perform in teaching events regarding generic teaching and learning components. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 47, H. 2, S. 84–96.

Seidel, Tina/Thiel, Felicitas (2017): Standards und Trends der videobasierten Lehr-Lernforschung. In: Zeitschrift für Erziehungswissenschaft 20, H. S1, S. 1–21.

Sun, Jennifer/van Es, Elisabeth A. (2015): An exploratory study on the influence that analyzing teaching has on preservice teachers' classroom practice. In: Journal of Teacher Education 66, S. 201–214.

Syring, Marcus/Bohl, Torsten/Kleinknecht, Marc/Kuntze, Sebastian/Rehm, Markus/Schneider, Jürgen (2015): Videos oder Texte in der Lehrerbildung? Effekte unterschiedlicher Medien auf kognitive Belastung und die motivational-emotionalen Prozesse beim Lernen mit Fällen. In: Zeitschrift für Erziehungswissenschaft 18, H. 4, S. 667–685.

Shavelson, Richard J. (2012): An approach to testing and modeling competencies. In: Blömeke, Sigrid/Zlatkin-Troitschanskaia, Olga/Kuhn, Christiane/Fege, Judith (Eds.): Modeling and Measuring competencies in Higher Education: Tasks and Challenges. Boston: Sense, S. 29–43.

Tietgens, Hans (1988): Professionalität für die Erwachsenenbildung. In: Gieseke, Wiltrud (Hrsg.): Professionalität und Professionalisierung. Bad Heilbrunn: Klinkhardt, S. 28–75.

Tripp, Tonya R./Rich, Peter J. (2012): The influence of video analysis on the process of teacher change. In: Teaching and Teacher Education 28, H. 5, S. 728–739.

Lerngelegenheiten im Praxissemester aus Absolvierendensicht

Ergebnisse aus der VFL-Praxis-Studie

Gabriele Klewin, Anne Köker und Jan Christoph Störtländer

Abstract

Welche Bedeutung hat das Forschende Lernen für Studierende im Praxissemester und welche Lerngelegenheiten nutzen sie für sich? In dieser qualitativen Studie berichten 16 Studierende in qualitativen Interviews von ihren Erfahrungen. Mittels qualitativer Inhaltsanalyse konnte eine Reihe von „wahrgenommenen" und „verhinderten" Lerngelegenheiten identifiziert werden. Diese werden interpretiert, in den Diskurs um das Praxissemester eingeordnet, und die Befunde werden mit Blick auf hochschuldidaktische Aspekte sowie neu gewonnene Erkenntnisinteressen diskutiert.

1 Problemstellung

Das mit einem Alter von ca. 50 Jahren nicht mehr ganz junge Konzept des Forschenden Lernens (Bundesassistentenkonferenz 1970; Fichten 2010; Huber 2014) hat durch die Einführung von Praxissemestern in der Mehrheit der Bundesländer einen bis dahin nicht erlangten Stellenwert in der Lehrerbildung erfahren (Weyland/Wittmann 2017; Fichten/Weyland 2018). So auch in Nordrhein-Westfalen, wo angehende Lehrkräfte im Masterstudium eine rund fünfmonatige Praxisphase an einer Schule absolvieren, in der sie Unterrichtsvorhaben verwirklichen und parallel Forschungsprojekte durchführen und in Studienberichten dokumentieren sollen.

Während seitens der Universitäten und der Bildungsadministration mit dem Forschenden Lernen hohe akademische und bildungspolitische Erwartungen (Fichten 2010) verbunden sind, sehen andere beteiligte Akteure das Konzept weitaus kritischer. So stehen für die Studierenden – ersten Forschungsbefunden zum Praxissemester in NRW zufolge (Fichten/Weyland 2018; van Ackeren/Herzig 2016) – die eigenen Unterrichtserfahrungen im Vordergrund, und sie empfinden die Studienprojekte zuweilen als belastend. Diese Erkennt-

nisse spiegeln sich auch in ersten Evaluationsergebnissen der Universität Bielefeld (Streblow/Brandhorst 2016) wider. Den zentralen Fragen, welche Lerngelegenheiten Studierende und junge Lehrpersonen im Praxissemester als für sie bedeutsam wahrnehmen und welche Rolle das Forschende Lernen dabei spielt, geht die Studie VFL-Praxis (Vielfalt Forschenden Lernens im Praxissemester) nach, in der je drei Interviews mit (angehenden) Lehrerkräften in unterschiedlichen Phasen ihrer Laufbahn geführt und qualitativ inhaltsanalytisch ausgewertet werden. Ergebnisse, die aus den ersten 16 Interviews gewonnen wurden, liegen diesem Beitrag zugrunde.

2 Forschungslage

Theoretische Begründung für das Forschende Lernen

Obwohl hochschuldidaktisch der Begriff Forschendes Lernen in aller Munde ist, fehlt noch immer eine einheitliche Vorstellung und Definition des Konzepts (Koch-Priewe/Thiele 2009; Huber 2014). Eine mögliche ist die von Huber, der sich bereits im Zuge der Bundesassistentenkonferenz mit dem Forschenden Lernen befasst hat:

> „Forschendes Lernen zeichnet sich vor anderen Lernformen dadurch aus, dass die Lernenden den Prozess eines Forschungsvorhabens, das auf die Gewinnung von auch für Dritte interessanten Erkenntnissen gerichtet ist, in seinen wesentlichen Phasen – von der Entwicklung der Fragen und Hypothesen über die Wahl und Ausführung der Methoden bis zur Prüfung und Darstellung der Ergebnisse in selbstständiger Arbeit oder in aktiver Mitarbeit in einem übergreifenden Projekt – (mit)gestalten, erfahren und reflektieren.“ (Huber 2009, S. 11)

In Übereinstimmung mit Fichten (2017, S. 31) ist die Selbstständigkeit der Studierenden ein wichtiges Kriterium für Forschendes Lernen als hochschuldidaktische Konzept. Darüber hinaus ist die Teilhabe an Forschung eine entscheidende Zielsetzung.

Mit der Einbindung in schulische Praxisphasen haben die Ziele und Erwartungen einen anderen Akzent erhalten und scheinen dadurch gleichermaßen schlicht wie anspruchsvoll: So soll das Forschende Lernen einen Beitrag zur Professionalisierung der Studierenden leisten und die angehenden Lehrerkräfte darin unterstützen, eine reflexive Distanz zum eigenen Lehrerhandeln einzunehmen. Die Hoffnung, solche Effekte zu befördern, gilt vor allem für die Implementierung von Elementen Forschenden Lernens in Praxisphasen, die die Studierenden davor bewahren sollen, Vorbilder aus der schulischen Praxis unhinterfragt zu übernehmen. Eben jene unerwünschten ‚Meisterlehre-Effekte‘ konnten Studien zu Praxisphasen, die ohne Forschungsanteile durchgeführt

wurden, nachweisen (Arnold et al. 2011; Arnold/Gröschner/Hascher 2014; Rothland/Boecker 2014).

Kunter et al. (2011), die in Anlehnung an Eraut (2004) berufsbezogene Lernsituationen in formale, nonformale und informelle Lerngelegenheiten unterscheiden, sehen in der Erstausbildung mit ihrer universitären Phase und anschließenden Praxisphase (dem Referendariat) das wichtigste formale Lernangebot für Lehrkräfte. Jedoch lassen gerade die oben genannten Studien vermuten, dass das formale Lernangebot der praktischen Ausbildung doch eine Vielzahl an informellen, d. h. reaktiven (zwar bewussten, aber nicht reflektierten) und impliziten (unbewussten) Lernsituationen (Eraut 2004) beinhaltet.

Das Forschende Lernen kann aus dieser Perspektive als Versuch definiert werden, die Praxisphase um absichtsvolle Lernsituationen zu erweitern, um informelle, also reaktive und implizite Lernsituationen aufzudecken und zu hinterfragen.

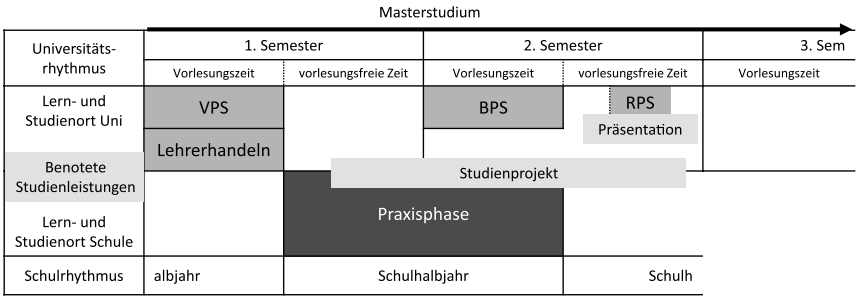
Das Forschende Lernen in der Lehrerausbildung an der Universität Bielefeld

Die Erwartung, die Entwicklung einer forschenden Haltung zu fördern, liegt seit dem Wintersemester 2015/16 auch der Ausgestaltung des Praxissemesters in Nordrhein-Westfalen zugrunde, folglich auch der in Bielefeld, die in diesem Beitrag betrachtet werden soll. Beim Forschenden Lernen, wie es dem „Leitkonzept zur standortspezifischen Ausgestaltung des Bielefelder Praxissemesters“ und auch unserer Studie definitorisch zugrunde liegt, „[…] handelt es sich also um die theoriegeleitete und selbstreflexive Auseinandersetzung mit dem Handlungsfeld Schule unter einer klar formulierten Fragestellung und mittels eines auf Forschungsmethoden gestützten […] und/oder kriteriengeleiteten Vorgehens“ (Bielefelder Leitkonzept 2011, S. 8).

Das Forschende Lernen konkretisiert sich hier in zwei studentischen Forschungsprojekten, die in Studienberichten dokumentiert werden. Die Studienberichte stellen eine wesentliche Grundlage der Leistungsbewertung im Praxissemester dar (vgl. Bielefelder Leitkonzept 2011 und Handreichungen zum Praxissemester 2014), daneben werden in allen Studienfächern weitere mündliche und/oder schriftliche Prüfungen abgenommen.

In der untenstehenden Grafik (vgl. Abb. 1) wird der Ablauf des gesamten berufsfeldspezifischen Studienjahrs zur Komplexitätsreduktion nur bezüglich der bildungswissenschaftlichen Studienanteile dargestellt.

Abbildung 1: Stilisierte Darstellung des Ablaufs des berufsfeldbezogenen Praxisjahrs in Bezug auf das Studienfach Bildungswissenschaften



Die Studierenden der Universität Bielefeld gehen in der Regel im zweiten Mastersemester in die schulische Praxisphase, das gesamte berufsfeldbezogene Studienjahr beginnt jedoch bereits im ersten Mastersemester mit vorbereitenden Seminaren in allen relevanten Studienfächern.

Die von der Fakultät für Erziehungswissenschaft verantworteten Module (s. Abb. 1), die die bildungswissenschaftliche Vorbereitung, Begleitung und Reflexion rahmen und regeln, bestehen aus den vorbereitenden Elementen *Fallforschung/Forschendes Lernen* (*VPS*) und Lehrerhandeln, in denen einerseits „Methodenkenntnisse(n) im Bereich qualitativer und quantitativen Forschung sowie deren Anwendungsmöglichkeiten im schulischen Kontext" vermittelt werden, andererseits „Modelle und Konzepte der Allgemeinen Didaktik und Grundlagen der Unterrichtsplanung und -analyse" (Modulbeschreibung[1]) thematisiert und mögliche weitere Bereiche aus dem breiten Spektrum schulpädagogischer Forschungsfelder angeschnitten werden. Diese Seminare „sollen die Studierenden dazu befähigen, während des Praxissemesters vor dem Hintergrund bildungswissenschaftlicher Theorieansätze die Erfahrungen im Handlungsfeld Schule zu analysieren und forschend zu reflektieren" (ebd.).

Dieselben Lehrenden, die die VPS-Seminare geleitet haben, begleiten im darauffolgenden Semester die Studierenden im Rahmen der *bildungswissenschaftlichen Begleitseminare* (*BPS*), betreuen und bewerten die bildungswissenschaftlich ausgerichteten Studienprojekte und benoten die Präsentationen im *Reflexionsseminar (RPS)*. Als Leistungsanforderung wird im vorbereitenden Semester in den VPS-Seminaren und im Seminar Lehrerhandeln jeweils eine Studienleistung von den Studierenden verlangt, die nicht benotet wird. Benotet werden die beiden Studienberichte, die über die studentischen Forschungsprojekte geschrieben werden. Einer der Berichte kann in den Bildungswissen-

1 Die Modulbeschreibung bezieht sich auf die Studiengänge Haupt-, Real- und Gesamtschule sowie Gymnasium, Gesamtschule https://ekvv.uni-bielefeld.de/sinfo/publ/modul/27461154.

schaften angesiedelt sein. Nach Beendigung der schulischen Praxisphase haben die Studierenden drei Monate Zeit diese Arbeiten zu verfassen. In diese Zeit fallen in der Regel auch die benoteten Reflexionskolloquien, die in den Bildungswissenschaften und in den meisten Fächern ebenfalls Teil der Leistungsanforderungen sind.

Forschungsstand zum Forschenden Lernen in Praxisphasen

Nach der Euphorie der ersten Jahre werden auch im akademischen Diskurs kritische Stimmen lauter, die die Wirksamkeit der Verbindung „Forschendes Lernen in der Praxisphase" hinterfragen und Nachweise für die Leistungsfähigkeit des Konzeptes einfordern. In der Tat ist die theoretische Grundlegung des Konzeptes eher programmatisch (Katenbrink/Wischer 2015) und über die Standorte wenig konsistent (Feindt/Wischer 2017) ausgerichtet. Und auch die empirische Befundlage hierzu ist bislang noch sehr dünn (Fichten/Weyland 2018), was sich jedoch angesichts vieler Projekte in der Qualitätsoffensive Lehrerbildung, die zu diesem Bereich forschen, zunehmend ändert (z. B. König/Rothland/Schaper 2018; Artmann et al. 2018).

Abgesehen davon, dass die Frage nach der Wirksamkeit hochschulischer Lehr-Lern-Prozesse generell sehr jung ist, die Forschung hierzu noch in den Kinderschuhen steckt (Koch-Priewe et al. 2015) und die Bearbeitung dieser Frage – wie jeglicher Nachweis der Wirksamkeit von Lehr-Lern-Prozessen – alles andere als trivial ist (Terhart 2012), ergeben sich für die Überprüfung, ob die in das Konzept gesteckten Hoffnungen berechtigt sind, besondere Herausforderungen. So ist die generelle Komplexität der Wirkketten von Lehr-Lern-Prozessen durch die spezifischen Akteurskonstellationen im Praxissemester (Heinrich/Klewin 2018) noch potenziert. Zwar liegt die Verantwortung für die Anbahnung der forschenden Haltung bei den Universitäten, doch spielen bei der Umsetzung – neben den Studierenden selbst – mehrere Hochschullehrende und Lehrende der Zentren für schulpraktische Lehrerausbildung aus jeweils mindestens drei Domänen (je nach Schulform zwei bis drei Fächer und Bildungswissenschaften), die Mentorinnen und Mentoren, Schulleitungen und Kollegien an den betreuenden Schulen und die Kommilitoninnen und Kommilitonen jeweils eine relevante Rolle. Es ist zu vermuten, dass jeder Akteur dieser Praxisphase andere Schwerpunkte setzt, andere Zielvorstellungen bezüglich des Praxissemesters hat und vor allem die Relevanz Forschenden Lernens und dessen Umsetzung und Umsetzbarkeit unterschiedlich hoch einschätzt.

Hinzu kommt, dass das Praxissemester mit Anteilen Forschenden Lernens in der Lehrerausbildung noch eine junge Einrichtung ist. In Nordrhein-Westfalen wurde das Praxissemester erstmals 2015 landesweit an allen lehrerbildenden Hochschulen etabliert, was zugleich auch eine Vielzahl von begleitenden Forschungsprojekten evoziert hat. Hier liegen die Ergebnisse sachgemäß noch

nicht in der Breite vor. Forschungsbefunde stammen hier zumeist aus Teilstudien umfassenderer Forschungsprojekte oder qualitativer Untersuchungen mit kleiner Stichprobe, aber richtungweisenden Erkenntnissen. So ermittelten Camman, Darge, Kasper und König (2018) durch eine standardisierte Befragung von Studierenden (und Expertinnen und Experten) zu den Anforderungen und Schwierigkeiten von und mit Studienprojekten eine Diskrepanz der eher als gering eingeschätzten forschungsbezogenen Kompetenz der Studierenden vor dem Praxissemester mit den nach der Praxisphase konstatierten moderaten Schwierigkeiten mit der Bewältigung der Anforderungen des Forschenden Lernens. Homt und van Ophuysen (2018) befragten Studierende, wie sie in den vorbereitenden Seminaren auf die Studienprojekte vorbereitet wurden und ermittelten durch die inhaltsanalytische Auswertung der Interviews, die vor der Praxisphase durchgeführt wurden, ein sehr heterogenes Bild. Herzmann und Liegmann konnten durch die inhaltsanalytische Auswertung von Studienprojektberichten eine „Dominanz wirkungsorientierter Zugänge" (2018, S. 61) ausmachen. Die hier vorgestellte Studie lässt sich mit einer ähnlichen übergeordneten Fragestellung in diese Reihe einordnen, erweitert mit ihrer Ausrichtung dennoch das Spektrum.

3 Fragestellungen der VFL-Praxisstudie

Die VFL-Praxisstudie wird unter der übergeordneten Fragestellung durchgeführt, welche Bedeutung das Forschende Lernen im gesamten berufsfeldbezogenen Studienjahr aus der Perspektive der Studierenden für ihren Professionalisierungsprozess hat. Wir interessieren uns hier in einem bildungstheoretischen Verständnis (Klafki 2007; Koch-Priewe/Köker/Störtländer 2016; Köker/Störtländer 2017) für die Bedeutung: Dient das Forschende Lernen den Studierenden insofern für ihre Professionalisierung, als dass sie Elemente ihres angestrebten Berufes für sich erschließen können, und sind sie für dieses Berufsfeld durch das Forschende Lernen erschlossen? Wir legen unserer Betrachtung an dieser Stelle bewusst noch keine der gängigen Professionalisierungsansätze (Herzmann/König 2017) zugrunde, sondern nähern uns offen unserem Erkenntnisinteresse.

Eine zweite Frage – der hier jedoch nicht nachgegangen wird – bezieht sich darauf, ob sich diese Bedeutung im Laufe der Zeit bzw. der voranschreitenden beruflichen Laufbahn ändert. Wir betrachten Professionalisierung als einen zu Berufsbeginn noch nicht abgeschlossenen und womöglich insgesamt nicht abschließbaren Prozess, zieht man zum einen die sich stetig wandelnden Anforderungen des Bildungssystems in Betracht, wie sie aus bildungshistorischer Perspektive zu bestehen scheinen, und bedenkt man zum anderen, dass auch

die individuelle Entwicklung von Menschen über ihre Lebensspanne hinweg verschiedene Phasen durchläuft.

Hieraus wird deutlich, dass wir ein breites Verständnis von Professionalisierung haben, das eine zum Teil selbstständig verantwortete Entwicklungsnotwendigkeit der Person in berufsbezogenen Bereichen umfasst.

4 Methodisches Vorgehen

An der Studie haben bislang 16 Studierende teilgenommen, die zu dem ersten Durchgang des Praxissemesters in Nordrhein-Westfalen gehörten. Es handelt sich bei ihnen sozusagen um „Pioniere", weswegen ein gewisser Kohorteneffekt nicht auszuschließen ist. Weder Studierende noch Lehrende des Praxissemesters konnten in der Vorbereitung und Durchführung des Studienjahres auf gesättigte Erfahrungswerte zurückgreifen, viele Routinen bestanden noch nicht. Insofern befanden sich alle Beteiligten, auch die Akteure der Schulen und der Studienseminare, mehr oder weniger gemeinsam auf Erkundung in unbekannten Gewässern.

Die VFL-Praxis-Studie wurde als qualitativer Längsschnitt mit drei geplanten Erhebungszeitpunkten angelegt. Der erste Erhebungszeitpunkt, aus dem auch die vorliegenden Daten stammen, fand gegen Ende der Masterphase statt. Hier lag das Praxissemester schon so weit zurück, dass alle auf das Praxissemester bezogenen Leistungen bereits erbracht worden waren. Die Erhebung fand bewusst außerhalb jeglicher Bewertungskontexte oder Abhängigkeiten statt, um einen, nicht durch Phänomene wie soziale Erwünschtheit, defensive Reflexion oder andere Dominanzmanöver beeinträchtigten, reflexiven Raum eröffnen zu können (Klewin/Köker/Störtländer 2019 i. V.; Häcker 2017; Reichenbach 2007). Die 16 Teilnehmenden wiesen eine hohe Heterogenität bezüglich der späteren Schulformen und Fächer auf.

Die zweite Erhebung sollte gegen Ende der zweiten Ausbildungsphase, des Vorbereitungsdienstes, durchgeführt werden. Der dritte Erhebungszeitpunkt sollte im Frühsommer des Jahres 2020 liegen, in dem die Studierenden, wenn sie ihr Berufsziel weiter verfolgt haben, bereits mehrere Jahre im Schuldienst verbracht haben werden. Aufgrund der besonderen Situation des ersten Durchgangs und Problemen bei der zweiten Befragung, wird der erste Durchgang nicht wie geplant weiter befragt, sondern es wird ein zweiter Durchlauf im Wintersemester 2019/2020 starten, mit dem der Längsschnitt realisiert werden soll. Die Daten des ersten Erhebungszeitpunkts sind dennoch ertragreich zur Analyse der Anfänge des Praxissemesters und wurden deshalb nicht verworfen; sie bilden die Grundlage der nachfolgenden Ergebnisdarstellung.

Die Erhebung wurde mit halbstandardisierten Interviews durchgeführt. Der Impuls für den episodischen bzw. narrativen Einstieg (Flick 1996) lautete: „Lass doch bitte mal das Vorbereitungssemester und das Praxissemester selbst möglichst detailliert Revue passieren.

Nach diesem Einstieg erfolgte eine Fokussierung auf Lerngelegenheiten und mögliche Zusammenhänge mit universitärer Rahmung (Vorbereitung, Begleitung, Studienprojekte, Reflexionsprüfung). Hier haben wir uns an insgesamt acht Leitfragen orientiert, von denen zwei in dem vorliegenden Beitrag im Mittelpunkt stehen. Sie lauten:

- Wenn du noch mal zusammenfassend guckst, welches waren die Gelegenheiten, in denen du am meisten gelernt hast?
- Wie ordnest du die Studienprojekte in diesen Kontext ein?

Die Interviews, die eine Länge von acht bis 40 Minuten hatten, wurden induktiv inhaltsanalytisch ausgewertet (Kuckartz 2014). Die Analyse umfasste mehrere Materialdurchgänge: In einem ersten Materialdurchgang wurde das Datenmaterial offen inhaltlich unter der übergeordneten Fragestellung kodiert, welche Lerngelegenheiten die befragten Studierenden thematisieren. In den nachfolgenden Durchgängen wurden die Codes geordnet und weiter ausdifferenziert, sodass wir am Material insgesamt sechs Kategorien 1. Ordnung mit insgesamt 47 Unterkategorien entwickeln konnten.

5 Ergebnisse

Unsere bisherige Analyse ergibt, dass die Studierenden die in den Interviews benannten Lerngelegenheiten sehr unterschiedlich einschätzen. Die Darstellung der Ergebnisse richtet sich nach den Kategorien 1. Ordnung „wahrgenommene Lerngelegenheiten“ und „verhinderte Lerngelegenheiten“, wobei im Folgenden auf die Kategorien 2. Ordnung „Studienprojekte als wahrgenommene Lerngelegenheiten“ und „Studienprojekte als verhinderte Lerngelegenheiten eingegangen wird“.

Bei wahrgenommenen Lerngelegenheiten handelt es sich um solche Studienprojekte, die entweder tatsächlich auf Grundlage eines eigenen Erkenntnisinteresses durchgeführt worden sind und erkenntnisreich verlaufen sind, wo also Studierende die Gelegenheit wahrgenommen haben bzw. wahrnehmen konnten, ihrem Erkenntnisinteresse systematisch zu folgen und so forschend zu lernen, oder solche, bei denen sich erst im Verlauf in einem phänomenologischen bzw. epistemischen Sinn die Wahrnehmung eingestellt hat, dass hier etwas Bedeutsames durch das Forschende Lernen erschlossen wird. Der Begriff der wahrgenommenen Lerngelegenheiten ist in diesem Doppelsinn zu verste-

hen. Bei verhinderten Lerngelegenheiten handelt es sich u. a. um solche Studienprojekte, die entweder nicht durchgeführt werden konnten, obwohl sie im Vorfeld als erkenntnisversprechend eingeschätzt worden sind, oder um Situationen, die sich in der Retrospektive der Studierenden als vielversprechend herausgestellt haben, bei denen aber zu der Zeit, als das interessierende Phänomen „akut" aufgetreten ist, Hinderungsgründe bestanden haben, so dass es nicht zu diesem Studienprojekt gekommen ist.

Zu der Kategorie der als Lerngelegenheiten wahrgenommenen Studienprojekte lässt sich eine Reihe von Ausdifferenzierungen und Begründungen aus den Interviews mit den Studierenden gewinnen, dargestellt am Kategorienbaum in Tabelle 1.

Tabelle 1: Ausschnitt des Kategoriensystems – „Studienprojekte als wahrgenommene Lerngelegenheiten"

Wahrgenommene Lerngelegenheiten			
	Studienprojekte als wahrgenommene Lerngelegenheit		
		Eigenes Erkenntnisinteresse Beeindruckt von Ergebnissen Auswertungsprozess erkenntnisreich Theorie-Praxis-Anwendung Im Nachhinein ertragreich Studienprojekte sinnvoll für	
			Perspektiverweiterung Distanznahme Verbindung mit Unterricht Außenperspektive auf eigenen Unterricht Wissenschaftlichen Habitus

Die Studienprojekte werden dann als bedeutsam beschrieben, wenn es den Studierenden möglich war, dem eigenen Erkenntnisinteresse nachzugehen (das gilt umgekehrt auch für das Gegenteil der verhinderten Lerngelegenheiten, siehe unten). Weiterhin stellen sie sich als positiv dar, wenn ein Bezug zwischen Theorie und Praxis gelungen ist, worüber einige Studierende explizit erstaunt sind. Gesprochen wird dann häufig über die „Anwendung" von Theorie auf Praxis. Ebenso beeindruckend sind für manche Studierende die Befunde ihrer Studienprojekte, auch wird mitunter der systematische Auswertungsprozess als erkenntnisreich beschrieben. Zahlreiche Fundstellen weisen darauf hin, dass die Studienprojekte eher in der Retrospektive als ertragreich wahrgenommen werden – ein Indiz, dem wir im qualitativen Längsschnitt weiter nachgehen.

Wir finden im Datenmaterial eine Verdichtung bei den Stellen, bei denen die Studierenden beschreiben, wofür die Studienprojekte sinnvoll gewesen sind. In Bezug auf den eigenen Unterricht stellt es sich als sinnvoll heraus, diesen

überhaupt zum Gegenstand gemacht zu haben, auch weil dies eine Außenperspektive auf das eigene Handeln erlaubt. In einer weiteren Perspektive gilt das auch, wenn durch die Studienprojekte eine kritisch-reflexive Distanznahme zur Institution Schule (und der eigenen Rolle in ihr) ermöglicht wurde und insgesamt eine Perspektiverweiterung gelungen ist. Der wissenschaftliche Habitus, zu dem die Studienprojekte „zwingen", wird als sinnvoll und bewahrenswert angesehen.

Einer erfolgreichen Durchführung der Studienprojekte steht eine Reihe von Hinderungsgründen entgegen (s. Tab. 2).

Tabelle 2: Ausschnitt des Kategoriensystems – „Studienprojekte als verhinderte Lerngelegenheiten"

Verhinderte Lerngelegenheiten			
	Studienprojekte als verhinderte Lerngelegenheit		
		Doppelrolle Unterricht + Forschung Zeitmangel Keine echte Forschung Relevanz BiWi gering Vorgegebene Themen Studienskizzen im Vorhinein Sinn erst im Nachhinein erkannt Feldbedingungen Studienprojekt als Leistungserbringung Belastung	
			durch zwei Studienprojekte durch zu viele Prüfungsleistungen durch zeitgleiche Erwerbstätigkeit für die Schule

Die oben als positiv beschriebene Doppelrolle von Forschung und Unterricht wird mitunter als aversiv wahrgenommen, da für diese Studierenden während der schulischen Praxisphase das Unterrichten im Vordergrund stand. Insgesamt waren Zeitmangel und häufig auch die Bedingungen im Feld hinderlich, vor allem dann, wenn sich zuvor geplante Projekte nicht durchführen ließen und dann schnell umgeschwenkt werden musste. Wenn die Studierenden nicht ihrem eigenen Erkenntnisinteresse folgen konnten, sondern gezwungen waren, „Auftragsforschung" zu betreiben, weil Themen und sogar Fragestellungen bis hin zu Instrumenten der Erhebung und Auswertung vorgegeben wurden, wird dies als negativ eingeschätzt. Studentische Forschung wird manchmal auch als „Forschung light" wahrgenommen, die nicht mit der „echten" Forschung mithalten kann, wobei hier manchmal paradigmatische Hindernisse zum Tragen kommen, etwa dann, wenn unter Forschung nur quantifizierende oder hypo-

thesenprüfende Untersuchungen oder komplexe Wirkungsforschung mit Interventionsstudien verstanden werden.

Aus organisatorischen Gründen war es (und ist es bis dato) notwendig, die Studienprojektskizzen, also erste Überlegungen über die angestrebten Forschungsvorhaben, sehr früh zu Beginn der Vorbereitungsphase auf das Praxissemester einzureichen, was von den befragten Studierenden als hinderlich wahrgenommen wird. Weiterhin wird die Kopplung der Studienprojekte mit der benoteten Leistungserbringung in einem recht hohen Maß (berechnet man das Verhältnis zwischen ECTS-Punkten und Noten, die im Master zur Verfügung stehen) als ungünstig beschrieben, dies ist jedoch explizit so gewollt, da der unterrichtliche Teil, anders als im Vorbereitungsdienst, unbenotet bleibt, wenngleich er auch eine hohe Relevanz für die Studierenden birgt. Zur Relevanz fällt auf, dass die Studierenden je nach Schulform einen hohen Wert auf Fachlichkeit und fachdidaktische Fragestellungen legen und die Relevanz des bildungswissenschaftlichen Studienanteils z. T. als insgesamt gering erachten.

Ähnlich wie bei den wahrgenommenen Lerngelegenheiten wird auch bei den verhinderten eine Verdichtung deutlich, hier unter dem Aspekt der Belastung. Das Praxissemester ist von den Studierenden unentgeltlich zu verrichten, aber ein verpflichtender Studienanteil. Zahlreiche der befragten Studierenden waren auf eine Erwerbstätigkeit angewiesen, die zur Zeit der Praxisphase nicht selten ruhen musste, was zu prekären Situationen führte. Zwei Studienprojekte durchführen zu müssen, wird von mehreren Befragten als Belastung deklariert. Einige führen die insgesamt hohe Prüfungsdichte im Praxissemester an. Auch für die Einzelschulen scheinen aus Sicht mancher Studierender die Studienprojekte eine Belastung darzustellen. Hier kann aber auch ein Zusammenhang damit vermutet werden, ob an einer Einzelschule ein eher forschungsaffines Klima herrscht oder ob Dynamiken der Abwehr die forschend lernenden Studierenden einhegen.

6 Diskussion

Insgesamt zeigt sich ein ambivalentes Bild bei der Frage danach, ob und wenn ja, welche Lerngelegenheiten mit den Studienprojekten verbunden sind. Die Ergebnisse zeigen auch, wie differenziert sich die Studierenden in der Reflexion ihrer Tätigkeiten im Rahmen der Interviews auseinandersetzen.

Einige Studierende beschreiben einen – den Erwartungen an das Forschende Lernen (siehe 2.1. und 2.2) entsprechenden – Erkenntnisgewinn hinsichtlich Perspektiverweiterung, Distanznahme zur Praxis, Verknüpfung von Theorie und Praxis sowie des wissenschaftlichen Habitus. Häufig erfolgt dies allerdings erst in der nachträglichen Beurteilung. Einschränkend muss allerdings gesagt werden, dass unklar ist, inwieweit die Äußerungen auch auf soziale

Erwünschtheit zurückzuführen sind und, sollte dies nicht der Fall sein, ob die beschriebenen Erkenntnisgewinne von Dauer sind. Die Frage, welche Faktoren bedingen, ob die studentische Forschung auf individueller Ebene als Lerngelegenheit wahrgenommen wird oder nicht, kann anhand der vorliegenden Auswertung nicht beantwortet werden. Es können allerdings Hinweise für die Hochschuldidaktik aus den Ergebnissen abgeleitet werden, sowohl allgemein über die Rahmenbedingungen des Praxissemesters als auch über die jeweilige curriculare Gestaltung.

So lässt sich sowohl aus den positiven als auch aus den negativen Aussagen zu Lerngelegenheiten herausarbeiten, dass die eigenständige Themenwahl und damit die Ermöglichung, das eigene Erkenntnisinteresse zu verfolgen ein wichtiger curricularer Baustein ist. Dies entspricht dem Kennzeichen „Selbstständigkeit", wie es von Huber (2009) und Fichten (2017) für Forschendes Lernen benannt wird (siehe oben). Um dieser Frage weiter nachgehen zu können, wären weitere Untersuchungen notwendig, die explizit auch die Seminarkonzeptionen bzw. die durchgeführte Begleitung von Studienberichten mit einbeziehen.

Widersprüchlich stellen sich die Ergebnisse zur Parallelität von Unterrichten und Forschen dar. Während auf der einen Seite gerade die Verknüpfung des Forschungsprojekts mit dem eigenen Unterricht als gewinnbringend angesehen wird, wird auf der anderen Seite diese Doppelrolle problematisiert. Gerade weil im Praxissemester für Studierende das Unterrichten im Vordergrund steht (Weyland/Wittmann 2017; Leonhard/Herzog 2018), könnte die stärkere Ausrichtung der Forschungsprojekte auf den eigenen Unterricht die Relevanz des Forschenden Lernens erhöhen (Fichten/Weyland 2018). Die Herausforderungen, die sich aus der Doppelrolle als Unterrichtende und Forschende ergeben, müssten dann in der Begleitung thematisiert werden.

Bei den allgemeinen Rahmenbedingungen stechen insbesondere die immer wieder genannten zu hohen Belastungen hervor. Betrachten Studierende die Studienprojekte vorwiegend als Belastung, die sie von den als zentral bewerteten Unterrichtserfahrungen abhält, dann werden die Anforderungen „abgearbeitet", weil sie leistungsrelevant sind. Diese Ergebnisse decken sich mit anderen Evaluationsergebnissen (Fichten/Weyland 2018). Inwieweit hier Veränderungen in den Studiengängen nötig und möglich sind, muss auf Basis weiterer Evaluationen beurteilt werden. Für Studierende, die neben dem Praxissemester noch erwerbstätig sein müssen oder Familienarbeit leisten, müssen jedoch auch jetzt schon individuelle Vereinbarkeiten noch stärker geregelt werden.

Auch wenn die genannten Ansätze zur Weiterentwicklung der universitären Vorbereitung und Begleitung des Praxissemesters nur kleine Beiträge sind, so können sie aus unserer Sicht helfen, das Praxissemester so zu gestalten, dass es auch und im oben erläuterten Sinne professionalisierungsförderlich ist.

Literatur

Ackeren, I. van/Herzig, S. (2016): Hochschulbeiträge zum Praxissemester. Die Bedeutung von Studienprojekten (S. 4–6). In: Ministerium für Schule und Weiterbildung (Hrsg.), Das Praxissemester auf dem Prüfstand. Zur Evaluation des Praxissemesters in Nordrhein-Westfalen. SchuleNRW, Beilage November 2016.

Arnold, K./Hascher, T./Messner, R./Niggli, A./Patry, J./Rahm, S. (Hrsg.) (2011): Empowerment durch Schulpraktika. Bad Heilbrunn: Klinkhardt.

Arnold, K./Gröschner, A./Hascher, T. (Hrsg.) (2014): Schulpraktika in der Lehrerbildung. Münster u. a.: Waxmann.

Artmann, M./Berendonck, M/Herzmann, P./Liegmann, A. B. (Hrsg.) (2018): Professionalisierung in der Lehrerbildung. Qualitative Forschung aus Bildungswissenschaft und Fachdidaktik. Bad Heilbrunn: Klinkhardt, S. 74–92.

Bielefelder Leitkonzept: Universität Bielefeld/ZfsL Bielefeld/ZfsL Minden/ZfsL Paderborn (2011): Leitkonzept zur Standortspezifischen Ausgestaltung des Bielefelder Praxissemesters. Erprobungsfassung vom 12.10.2011. Bielefeld.

Bundesassistentenkonferenz (1970): Forschendes Lernen – Wissenschaftliches Prüfen. Schriften der Bundesassistentenkonferenz 5, Bonn.

Cammann, F./Darge, K./Kaspar, K./König, J. (2018). Anforderungen des Forschenden Lernens im Praxissemester: Entwicklung eines Modells und erste empirische Befunde zur Validität. In: Herausforderung Lehrer_innenbildung – Zeitschrift zur Konzeption, Gestaltung und Diskussion 1, H. 1, S. 17–34.

Eraut, M. (2004): Informal learning in the workplace. In: Studies in Continuing Education 26, No. 2, S. 247–273.

Feindt, A./Wischer, B. (2017): Begründungen, Ziele und Formen Forschenden Lernens – ein Reflexionsangebot für den Einstieg. In: Schüssler, R./Schöning A./Schwier, V./Schicht, S./Gold, J./Weyland, U. (Hrsg.): Forschendes Lernen im Praxissemester. Zugänge, Konzepte, Erfahrungen. Bad Heilbrunn: Klinkhardt, S. 139–146.

Fichten, W. (2010): Forschendes Lernen in der Lehrer*innenausbildung. In: Eberhardt, U. (Hrsg.): Neue Impulse in der Hochschuldidaktik. Wiesbaden: Springer VS, S. 127–182.

Fichten, W. (2017): Forschendes Lernen in der Lehrerbildung. In. Schüssler, R./Schöning A./Schwier, V./Schicht, S./Gold, J./Weyland, U. (Hrsg.): Forschendes Lernen im Praxissemester. Zugänge, Konzepte, Erfahrungen. Bad Heilbrunn: Klinkhardt, S. 30–38.

Fichten, W./Weyland, U. (2018): Empirie zu Forschendem Lernen: Analyse und Perspektiven unter Berücksichtigung der Evaluation von Praxissemestern. In: Neuber, N./Paravicini, W./Stein, M. (Hrsg.): Forschendes Lernen – the wider view. Eine Tagung des Zentrums für Lehrerbildung der Westfälischen Wilhelms-Universität Münster. Münster: WTM, S. 47–58.

Flick, U. (1996): Psychologie des technisierten Alltags. Soziale Konstruktion und Repräsentation technischen Wandels in verschiedenen kulturellen Kontexten. Wiesbaden: Springer VS.

Häcker, T. (2017): Grundlagen und Implikationen der Forderung nach Förderung von Reflexivität in der Lehrerinnen- und Lehrerbildung. In: Berndt, C./Häcker, T./Leonhard, T. (Hrsg.). Reflexive Lehrerbildung revisited. Traditionen – Zugänge – Perspektiven. Bad Heilbrunn: Klinkhardt, S. 21–45.

Heinrich, M./Klewin, G. (2018): Forschendes Lernen im Praxissemester – Governanceanalysen eines bildungspolitischen Programms. In: Ukley, N./Gröben, B. (Hrsg.): Forschendes Lernen im Praxissemester. Begründungen, Befunde und Beispiele aus dem Fach Sport. Wiesbaden: Springer VS, S. 3–25.

Herzmann, P./König, J. (2015): Lehrerberuf und Lehrerbildung. Stuttgart: UTB.

Herzmann, P./Liegmann, A. B. (2018): Studienprojekte im Praxissemester. Wirkungsforschung im Kontext Forschenden Lernens. In: Lehrerbildung auf dem Prüfstand 11, H. 1, S. 46–65.

Homt, M./van Ophuysen, S. (2018): Die Studienprojekte im Praxissemester. Wie nehmen Studierende die universitäre Vorbereitung in den Projektseminaren wahr? In: Herausforderung Lehrer_innenbildung – Zeitschrift zur Konzeption, Gestaltung und Diskussion 1, H. 1, S. 77–89.

Huber, L. (2014): Forschungsbasiertes, Forschungsorientiertes, Forschendes Lernen: Alles dasselbe? Ein Plädoyer für eine Verständigung über Begriffe und Unterscheidungen im Feld forschungsnahen Lehrens und Lernens. In: Das Hochschulwesen 62., H. 1+2, S. 22–29.

Huber, L. (2009): Warum Forschendes Lernen nötig und möglich ist. In Huber, L./Hellmer J./Schneider, F. (Hrsg.): Forschendes Lernen im Studium. Bielefeld: Universitätsverlag Webler, S. 9–35.

Katenbrink, N./Wischer, B. (2015): „Zum Glück hatte mein Projekt ja was mit Geschichte zu tun“ – Reflexivität als Herausforderung forschenden Lernens. In: Katenbrink N./Kunze I./Solzbacher, C. (Hrsg.): Brücken bauen – Praxisforschung zu Übergängen im Bildungssystem. Münster: MV-Wissenschaft, S. 221–241.

Klafki, W. (2007): Neue Studien zu Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik. Weinheim: Beltz.

Klewin, G./Köker, A./Störtländer, J. (2019 i. V.): Veranlasste und benotete Reflexion: ein unmögliches Prüfungsformat? Eine empirische Untersuchung zur Wahrnehmung und Bewertung von Reflexionsprüfungen im Praxissemester durch Studierende.

Koch-Priewe, B./Köker, A./Seifried, J./Wuttke, E. (Hrsg.) (2015): Kompetenzerwerb an Hochschulen: Modellierung und Messung. Zur Professionalisierung angehender Lehrerinnen und Lehrer sowie frühpädagogischer Fachkräfte. Bad Heilbrunn: Klinkhardt.

Koch-Priewe, B./Köker, A./Störtländer, J. C. (2016): Die bildungstheoretische Didaktik und die kritisch-konstruktive Didaktik. In: Porsch, R. (Hrsg.): Einführung in die Allgemeine Didaktik. Ein Lehr- und Arbeitsbuch für Lehramtsstudierende. Stuttgart: utb, S. 101–132.

Koch-Priewe, B./Thiele, J. (2009): Versuch einer Systematisierung der hochschuldidaktischen Konzepte zum Forschenden Lernen. In: Roters, B./Schneider, R./Koch-Priewe, B./Thiele, J./Wildt, J. (Hrsg.): Forschendes Lernen im Lehramtsstudium. Hochschuldidaktik – Professionalisierung – Kompetenzentwicklung. Bad Heilbrunn, S. 271–292.

Köker, A./Störtländer, J. C. (2017): Studentische Forschung im Praxissemester – Empirische Zugänge zur eigenen Praxis durch Forschendes Lernen. In: Burger, T./Miceli, N. (Hrsg.): Empirische Forschung im Kontext Schule. Einführung in theoretische Aspekte und methodische Zugänge. Wiesbaden: Springer VS, S. 359–371.

König, J./Rothland, M./Schaper, N. (Hrsg.) (2018): Learning to Practice, Learning to Reflect? Wiesbaden: Springer VS.

Kuckartz, U. (2014): Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung. Weinheim u. a.: Beltz Juventa.

Kunter, M./Kleickmann, T./Klusmann, U./Richter, D. (2011): Die Entwicklung professioneller Kompetenz von Lehrkräften. In: Kunter, M./Baumert, J./Blum, W./Klusmann, U./, Krauss, S./Neubrand, M. (Hrsg.): Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster et al.: Waxmann, S. 55–68.

Leonhard, T./Herzog, S. (2018): Was Langzeitpraktika leisten (können) – empirische und konzeptionelle Erkundungen. In: Rothland. M./Schaper, N. (Hrsg.): Forschung zum Praxissemester in der Lehrerbildung, Lehrerbildung auf dem Prüfstand 11, H. 1 (Themenheft) S. 5–23.

Reichenbach, R. (2007): Kaschierte Dominanz – leichte Unterwerfung. Bemerkungen zur Subtilisierung der pädagogischen Autorität. In: Zeitschrift für Pädagogik 53, H. 5, S. 651–659.

Rothland, M./Boecker, S. (2014): Wider das Imitationslernen in verlängerten Praxisphasen. Potenzial und Bedingungen Forschenden Lernens im Praxissemester. In: Die Deutsche Schule 106, H. 4, S. 386–397.

Streblow, L./Brandhorst, A. (2016): Ergebnisse der BiSEd-Befragungen der Studierenden zur Entwicklung von Interesse und epistemologischen Überzeugungen im Studienverlauf sowie ausgewählte Befunde der Befragungen weiterer Akteursgruppen. Präsentation im Rahmen des Symposiums der Bielefeld School of Education „Empirische Befunde zum Praxissemester“ am 05. April 2016.

Terhart, E. (2012): Wie wirkt Lehrerbildung? Forschungsprobleme und Gestaltungsfragen. In: Zeitschrift für Bildungsforschung 2, H. 1, S. 3–21.

Weyland, U./Wittmann, E. (2017): Praxissemester en vogue. In: Schüssler, R./Schöning, A./Schwier, V./Schicht, S./Gold, J./Weyland, U. (Hrsg.), Forschendes Lernen im Praxissemester. Zugänge, Konzepte, Erfahrungen. Bad Heilbrunn: Klinkhardt, S. 17–29.

Unterrichtsbesuche am laufenden Band

Anforderungen und Belastungsempfinden im niedersächsischen Vorbereitungsdienst

Marike Feierabend und Frederike Bartels

Abstract

In der vorliegenden Studie werden, in Anlehnung an das transaktionale Stressmodell von Lazarus/Folkman (1984), Zusammenhänge zwischen strukturellen Anforderungen an niedersächsischen Studienseminaren und dem psychosozialen Belastungsempfinden (Rohmert/Rutenfranz 1975) von Lehramtsanwärtern und -anwärterinnen untersucht.

Die Ergebnisse stammen aus der quantitativen Querschnittstudie „Belastungen im Referendariat" (Bel^{Air}), an der (N = 366) Lehramtsanwärter und -anwärterinnen des Grund-, Haupt- und Realschullehramts an 16 niedersächsischen Studienseminaren teilnahmen.

Statistische Analysen auf individueller sowie auf Ebene der Studienseminare zeigen, dass sich die strukturellen Anforderungen an niedersächsischen Studienseminaren z. T. erheblich voneinander unterscheiden. Die Daten deuten darauf hin, dass sich Anwärter und Anwärterinnen durch ihre praktische Tätigkeit an der Schule weniger belastet fühlen als durch Anforderungen an den Studienseminaren. Analysen bestätigen z. T. die Annahme, dass sich das Belastungsempfinden von angehenden Lehrkräften in Abhängigkeit struktureller Anforderungen signifikant voneinander unterscheidet.

Die Erkenntnisse aus der hier vorgestellten Studie werden mit Blick auf die Vergleichbarkeit struktureller Rahmenbedingungen diskutiert.

1 Einleitung und Problemaufriss

Lehrkräfte sind im Schulalltag mental stark gefordert. Dies kann zu einem hohen psychosozialen Belastungsempfinden führen (Kliebisch 2009; Klippert 2004). Studien dokumentieren, dass sich die erhöhte mentale Belastung in die-

ser Berufsgruppe bereits während der Ausbildungsphase manifestiert, speziell bei Lehrkräften im Vorbereitungsdienst[1] (kurz: LiVD; Braun 2017; Drüge/Schleider/Rosati 2014). Christ/van Dick/Wagner (2004) beschreiben das Referendariat sogar als einen „Abschnitt in der Lehrerausbildung, der mit den stärksten Belastungen für die angehenden Lehrkräfte einhergeht[.]" (S. 113).

Die Ursachen für das Belastungsempfinden der angehenden Lehrkräfte wurden bisher allerdings nur marginal erforscht (u. a. Braun 2017). Es fehlen u. a. Erkenntnisse darüber, in welchem Zusammenhang das Belastungsempfinden von LiVD mit den strukturellen Anforderungen in der zweiten Ausbildungsphase steht. Diese Frage nach einer potentiellen Belastung ausgelöst durch Ausbildungsbedingungen steht im Fokus der vorliegenden Studie. Am Beispiel Niedersachsen werden vergleichend die Ausbildungsanforderungen an verschiedenen Studienseminaren für das Lehramt an Grund-, Haupt und Realschulen (GHRS) untersucht und es wird aufgeklärt, inwieweit die strukturellen Anforderungen im Zusammenhang zur empfundenen Belastung von LiVD stehen.

2 Theoretischer Hintergrund

Nach dem Belastungs-Beanspruchungs-Modell (vgl. Rohmert/Rutenfranz 1975, S. 8) handelt es sich bei Belastungen um objektive Faktoren, welche von außen auf einen Menschen einwirken. Es zeigt sich, dass die Belastung wächst, je mehr Belastungsfaktoren vorliegen (vgl. Rohmert/Rutenfranz 1975, S. 25). Die Auswirkungen, die durch die Auseinandersetzung mit der Belastung entstehen, werden als Beanspruchungen bezeichnet und interindividuell erlebt (vgl. Drüge/Schleider/Rosati 2014, S. 360). Im Rahmen dieser Arbeit wird das psychosoziale Belastungsempfinden betrachtet. Dies bezeichnet in Anlehnung an das transaktionale Stressmodel das subjektive Empfinden einer Person in Bezug auf Belastungen, welche in Abhängigkeit der eigenen Fähigkeiten und Fertigkeiten sowie situativen Bedingungen individuell wahrgenommen und bewertet und ggf. als stressrelevant eingeschätzt werden (in Anlehnung an Lazarus/Folkman 1984; Rohmert/Rutenfranz 1975).

Mit Blick auf den Lehrberuf verweisen Studienergebnisse auf eine erhöhte Anzahl an Lehrkräften, die eine hohe psychosoziale Belastung aufweisen (u. a. Kliebisch 2009; Klippert 2004). Als Ursachen für die Belastung konnten u. a.

1 Der Vorbereitungsdienst wird allgemein auch als Referendariat bezeichnet. Die hier fokussierte Gruppe des GHRS-Lehramts befindet sich formal jedoch im Vorbereitungsdienst, da die Bezeichnung Referendariat nur für Anwärter und Anwärterinnen des Gymnasiallehramts gilt.

eine (weniger gute) Lehrer-Schüler-Beziehung, Zeitdruck und fehlende Erholungspausen identifiziert werden (vgl. Aldrup et al. 2018; Krause/Dorsemagen 2007, S. 63). Die dabei erlebte Beanspruchung kann unmittelbare Auswirkungen auf den Unterricht haben, indem sie sich auf das Lernen, die Leistung und die Motivation der Schüler und Schülerinnen auswirkt (vgl. Arens/Morin 2016; Klippert 2004; Klusmann et al. 2006; Krause/Dorsemagen 2007). Weitere Folgen der Beanspruchungsreaktionen für die Lehrkräfte sowie auch die Schulorganisation können u. a. Fehlzeiten, Arbeitsunfähigkeitstage oder sogar Berufsunfähigkeit sein (vgl. Drüge/Schleider/Rosati 2014, S. 360; Krause/Dorsemagen 2007, S. 73).

Studien, die sich mit der Ausbildung von angehenden Lehrkräften beschäftigen, zeigen, dass sich bereits bei dieser Gruppe ein erhöhtes Beanspruchungserleben abzeichnet (u. a. Braun 2017; Christ/van Dick/Wagner 2004; Gecks 1990; Schaarschmidt 2005). Klusmann et al. (vgl. 2012, S. 282) finden etwa Belege dafür, dass Anwärter und Anwärterinnen ein vergleichbares Ausmaß an Erschöpfung aufweisen, wie Lehrkräfte, die bereits seit 20 Jahren tätig sind und auch Drüge/Schleider/Rosati (vgl. 2014, S. 358) verweisen auf ähnliche Folgen der Belastung wie bei Lehrkräften (u. a. kognitive Stresssymptome, Gedanken an Berufsaufgabe). Die Ursachen dafür, könnten möglicherweise auch mit den strukturellen Bedingungen der zweiten Ausbildungsphase zusammenhängen.

Belastungsfelder Schule und Studienseminar und rechtliche Rahmenbedingungen

Die angehenden Lehrkräfte befinden sich in der zweiten Phase der Lehrerbildung in einer Doppelrolle (u. a. Lenhard 2004). Sie sind zugleich Lehrende und Lernende.

In der Rolle als Lehrende sind sie u. a. mit den Erwartungen und Ansprüchen der Schule konfrontiert (*Belastungsfeld Schule*). Die Novizen stehen vor der für sie noch neuen Aufgabe den Unterricht regelmäßig vorzubereiten, durchzuführen und nachzubereiten, was auch erfahrene Lehrkräfte herausfordert. Zudem haben LiVD i. d. R. nur wenig Erfahrung im Umgang mit Schülern und Schülerinnen (u. a. Lenhard 2004). Wenn den Anwärtern und Anwärterinnen beim Eintritt in die Berufspraxis eigene Schwächen und Kompetenzdefizite bewusst werden, kann dies zu einer verringerten Selbstwirksamkeitserwartung und einem höheren Stresserleben führen (u. a. Klusmann et al. 2012). Die verschiedenen Belastungsquellen des Referendariats wurden von Kiel/Weiß (vgl. 2015, S. 94) systematisiert. Sie bilden dazu fünf Belastungsfelder (*private Probleme, Kollegen, Unterricht, Schüler/innen, Ausbildung*) und ordnen diesen Feldern potenzielle Belastungsquellen (z. B. fehlende Freizeit) zu.

In der Rolle als Lernende sind LiVD u. a. mit den Erwartungen ihrer Ausbildenden konfrontiert (*Belastungsfeld Studienseminar*). Diese Erwartungen

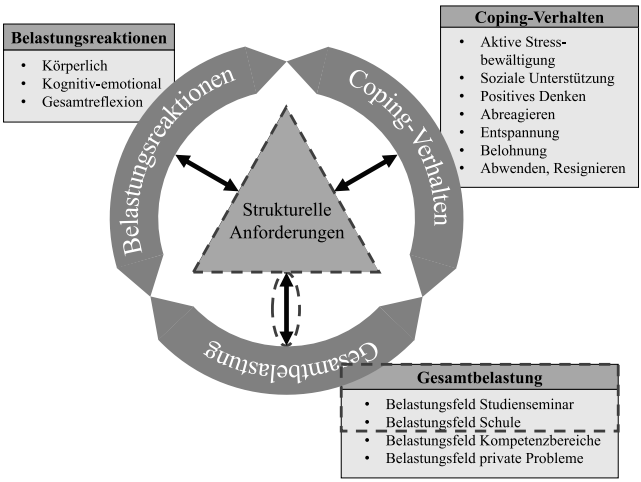
gründen auf den bundeslandspezifischen Ausbildungsverordnungen, welche den rechtlichen Rahmen bilden. In Niedersachsen werden Bestimmungen der zweiten Phase der Lehrerbildung in der „Verordnung über die Ausbildung und Prüfung von Lehrkräften im Vorbereitungsdienst" geregelt. Regelmäßige Unterrichtsbesuche und schriftliche Unterrichtsentwürfe bilden dabei einen wesentlichen Bestandteil der Ausbildungsphase. Diese werden von angehenden Lehrkräften oft als besonders belastend empfunden, auch aufgrund unklarer Bewertungskriterien und mangelnder Transparenz (u. a. Böhner 2010, S. 59) sowie der hohen zeitlichen Belastung, die mit der Verschriftlichung von Unterrichtsentwürfen verbunden ist (vgl. Pille 2013, S. 96). Als besonderer Umstand erweist sich, dass die Ausbildungsverordnung die Anzahl zu absolvierender Unterrichtsbesuche und schriftlicher Unterrichtsentwürfe nicht genau spezifiziert. Dadurch ergeben sich für Studienseminare signifikante Handlungsspielräume. Die Erkenntnislage darüber, wie Studienseminare den rechtlichen Rahmen nutzen bzw. auslegen und inwieweit möglicherweise Zusammenhänge zwischen der standortspezifischen Auslegung und den Belastungen der LiVD bestehen, ist bisher gering.

3 Fragestellungen

Am Beispiel Niedersachsen wird untersucht, inwieweit ein Zusammenhang zwischen den in der Ausbildungsverordnung offen formulierten strukturellen Anforderungen und der Gesamtbelastung in den potenziellen Belastungsfeldern *Studienseminar* und *Schule* besteht. Zur Beantwortung dieser Frage wurde anknüpfend an die Systematisierungen und Überlegungen von Kiel/Weiß (2015), Satow (2012) und Kaluza (2012) ein theoretisches Rahmenmodell entwickelt (Abb. 1). Diesem zufolge wird davon ausgegangen, dass strukturelle Anforderungen, Coping-Verhalten, Gesamt-Belastung und Belastungsreaktionen korrelieren. Im Zentrum der vorliegenden Studie wird ein Teilaspekt der Gesamtuntersuchung dargestellt, der sich mit der Frage nach der Vergleichbarkeit von Ausbildungsvorschriften und dem Belastungsempfinden von LiVD auseinandersetzt. Untersucht wird,

(1) inwieweit bestimmte strukturelle Anforderungen (u. a. die Anzahl an Unterrichtsbesuchen) zwischen Studienseminaren variieren,
(2) in welchem Belastungsfeld (*Schule* oder *Studienseminar*) sich LiVD stärker belastet fühlen und
(3) inwieweit ausgewählte Aspekte struktureller Anforderungen mit dem psychosozialen Belastungsempfinden von LiVD korrelieren.

Abbildung 1: Theoretisches Rahmenmodell der Bel[Air]-Studie



Aufgrund der nicht näher definierten strukturellen Anforderungen bzgl. der Anzahl zu absolvierender Unterrichtsbesuche und anzufertigender schriftlicher Unterrichtsentwürfe (Kurz- und Langversion) wird folgende erste Hypothese aufgestellt:

> H1: Es bestehen signifikante Unterschiede an den verschiedenen Studienseminaren in Niedersachsen in Bezug auf die Anzahl von Unterrichtsbesuchen und schriftlichen -entwürfen.

Wenngleich es Befunde (u. a. Christ/van Dick/Wagner 2004) gibt, die darauf hindeuten, dass sich LiVD stärker in der Rolle als Lernende belastet fühlen, wird in dieser vorliegenden Studie theoriegestützt davon ausgegangen, dass sich die LiVD in ihrer Rolle als Lehrende belasteter fühlen. Die Rolle als Lernende und die dazugehörigen Aufgaben sind den LiVD bereits aus ihrem Lehramtsstudium bekannt. Die Rolle des Lehrenden hingegen enthält neue Elemente für sie, weshalb diese voraussichtlich als belastender empfunden wird. Zusätzlich ist bereits aus Studien zur Lehrerbelastung bekannt, dass die Lehrtätigkeit ein hohes Risiko für Belastungen bergen kann. Demnach wird folgende zweite Hypothese aufgestellt:

> H2: Angehende Lehrkräfte fühlen sich in der Rolle als Lehrende belasteter als in der Rolle als Lernende.

Weiterhin wird angenommen, dass die Anzahl zu absolvierender Unterrichtsbesuche positiv mit dem Belastungsempfinden assoziiert ist. Es wird daher ergänzend zur ersten Hypothese angenommen:

H_Z1: Je mehr Unterrichtsbesuche angehende Lehrkräfte absolvieren müssen, umso belasteter fühlen sie sich im Bereich Studienseminar/Schule.

4 Methodisches Vorgehen

Stichprobe

Grundlage zur Überprüfung der Hypothesen bildet die Querschnittstudie Bel^{Air} (Belastungen im Referendariat), für die im März/April 2018 Teilnehmende von 16 der 21 Studienseminare in Niedersachsen für das GHRS-Lehramt befragt wurden. Das Erhebungsinstrument wurde schriftlich-postalisch ausgegeben. Aufgrund des Anspruchs, eine umfassende summarische Beurteilung zu erhalten, wurden LiVD im dritten Halbjahr des Vorbereitungsdienstes befragt (Abschlussprüfung Sommer 2018). An den 16 Studienseminaren wären laut Auskunft der Studienseminarleitungen N = 566 LiVD erreichbar gewesen. Erhoben werden konnten die Daten von N = 366 LiVD. Die Rücklaufquote kann mit 64,66 % als zufriedenstellend betrachtet werden. Die Teilnehmenden waren im Alter zwischen 23 und 49 Jahren und zu 80 % weiblich. Die Mehrheit der LiVD (N = 202) wurde für das Grundschullehramt ausgebildet.

Erhebungsinstrument

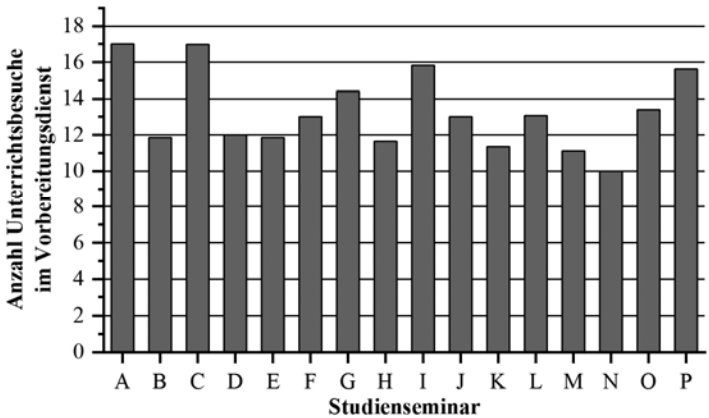
Für die Studie wurde ein Fragebogen entwickelt, der an das Stress- und Coping-Inventar von Satow (2012) angelehnt ist. Die Items wurden unter Einbezug aktueller Forschungsliteratur (u. a. Kiel/Weiß 2015) adaptiert und verändert, so dass sie gezielt auf das Belastungserleben von LiVD fokussieren.

Um die strukturellen Anforderungen an den Studienseminaren zu erfassen, wurden die LiVD im Rahmen eines offenen Antwortformats um Angaben zur Anzahl der Unterrichtsbesuche und zur Anzahl anzufertigender schriftlicher Kurz- und Langentwürfe gebeten. Ein langer Entwurf enthält mehrere Analyseteile (Lernvoraussetzungen, Sachanalyse, Didaktische Analyse, Methodische Analyse), wohingegen ein Kurzentwurf nur einzelne Analyseteile enthält und somit zeitlich weniger beanspruchend ist.

Angaben zu Belastungen („Inwieweit haben Sie sich während des Vorbereitungsdienstes durch folgende/Ereignisse und Probleme überfordert/Unsicherheiten belastet/gefühlt?“) wurden im *Belastungsfeld Studienseminar* über acht Items (z. B. Planung von Unterrichtsbesuchen; α= .91) bzw. im *Belastungsfeld Schule* über neun Items (z. B. Durchführung des regulären Unterrichts; α= .85) erfasst. Die Items und Skalen waren auf einer fünfstufigen Likert-Skala (1 = nicht überfordert/belastet; 5 = sehr stark überfordert/belastet) zu beantworten.

Auswertungen der offenen Antwortformate verdeutlichen, dass ein hohes Maß an Divergenz bezogen auf die erhobenen strukturellen Unterschiede zwischen niedersächsischen Studienseminaren besteht. Laut Selbstauskunft der LiVD liegt das Minimum bei acht zu absolvierenden Unterrichtsbesuchen (N = 1). N = 2 LiVD geben an, 20 Unterrichtsbesuche gehabt zu haben. Die Mehrheit der LiVD wurde laut Selbstauskunft 13 Mal im Unterricht von den Ausbildenden besucht (N = 107). Auch die Auswertungen in Bezug auf die Anzahl der Unterrichtsbesuche in Abhängigkeit des Studienseminars weisen darauf hin, dass eine hohe Variabilität in der Auslegung der Ausbildungsverordnung besteht. Wie in Abbildung 2 verdeutlicht, sind am Studienseminar A im Mittel 17 Unterrichtsbesuche üblich, während am Studienseminar N im Mittel nur 10 Besuche zu erwarten sind. Eine einfaktorielle ANOVA bestätigt, dass es sich um statistisch bedeutsame Unterschiede zwischen den Standorten handelt ($F(15,333) = 83.00$, $p < .001$).

Abbildung 2: Anzahl der Unterrichtsbesuche im Vorbereitungsdienst in Abhängigkeit des Studienseminars (anonymisiert)

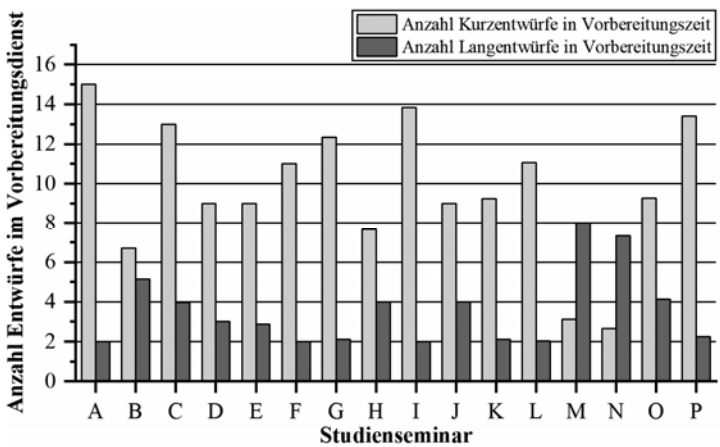


Diese hohe Streuung in den Angaben lässt sich einerseits damit erklären, dass der Handlungsspielraum in Bezug auf die Auslegung rechtlicher Rahmenbedingungen offenbar unterschiedlich genutzt wird. Andererseits könnten die Angaben der LiVD zum Teil auch subjektiv verzerrt sein. Unterschiede lassen sich möglicherweise darüber erklären, dass sogenannte „Mitschauen“ von einigen LiVD zu den Unterrichtsbesuchen gezählt werden, obwohl sich diese Seminargruppenbesuche allgemein von den „klassischen“ Unterrichtsbesuchen unterscheiden. An einigen Studienseminaren finden, nach eigenen Aussagen, keine Mitschauen statt, an anderen sind sie ein freiwilliges Angebot und an

wieder anderen Studienseminaren ist mindestens eine Mitschau verpflichtend. Da in dieser Studie die subjektive Sichtweise der LiVD im Fokus stand, wurden Mitschauen als Unterrichtsbesuche gewertet, wenn diese genannt wurden.

In Bezug auf die schriftlichen Entwürfe zeigt sich ebenso ein heterogenes Bild. Wie in Abbildung 3 zu erkennen, haben die befragten LiVD insgesamt mehr Kurz- als Langentwürfe anzufertigen, wobei auch hier die Einzelangaben deutlich schwanken. Lediglich an zwei Standorten (M, N) verhält es sich umgekehrt. Das Minimum an Kurzentwürfen liegt bei einem (n = 3) und das Maximum liegt bei 18 Kurzentwürfen (N = 1). Auch hier handelt es sich um statistisch bedeutsame Unterschiede zwischen den Studienseminaren (Kurzentwürfe: $F(15,333) = 108.671$, $p < .001$; Langentwürfe: $F(15,333) = 81.603$, $p < .001$).

Abbildung 3: Anzahl von Kurz- und Langentwürfen in der zweiten Phase der Lehrerbildung in Abhängigkeit des Studienseminars (anonymisiert)



Deskriptive Analysen weisen darüber hinaus darauf hin, dass LiVD sich im Allgemeinen belastet fühlen, wobei sie sich insgesamt stärker durch die Anforderungen vom *Studienseminar* belastet fühlen (N = 349; $M = 3.45$, $SD = .82$) als durch solche im *Belastungsfeld Schule* (n = 349; $M = 2.29$; $SD = .65$) (*Hypothese 2*). Mittelwertausprägungen auf Einzelitem-Ebene weisen dabei auf ein besonderes Unsicherheitsempfinden mit Blick auf die Planung von Unterrichtsbesuchen ($M = 3.56$, $SD = 1.03$) und der Anfertigung schriftlicher Entwürfe ($M = 3.67$, $SD = 1.00$) hin.

Anhand der Spearman-Rangkorrelation lässt sich allerdings kein statistisch signifikanter Zusammenhang zwischen der Anzahl an Unterrichtsbesuchen und dem Belastungsempfinden im *Belastungsfeld Studienseminar* absichern ($r = .07$, n.s.) (*Hypothese 3*). Das Belastungsempfinden lässt sich daher nicht unmittelbar in Zusammenhang mit den im Rahmen dieser Studie ausgewählten Aspekten struktureller Bedingungen bringen. Um dennoch weiterführend zu

überprüfen, ob Unterschiede im Belastungsempfinden zwischen LiVDs, die mehr bzw. weniger Unterrichtsbesuche haben, bestehen, wurde anhand des t-Tests auf Mittelwertunterschiede zwischen LiVD mit vielen (> 13) und wenigen (< 13) Unterrichtsbesuchen geprüft. Die Gruppen wurden am Median (Md = 13) dichotomisiert. Es lässt sich beobachten, dass sich LiVD mit wenigen Unterrichtsbesuchen signifikant weniger belastet im *Belastungsfeld Studienseminar* fühlen (n = 227: M = 3.40; SD = .72) als LiVD mit vielen Unterrichtsbesuchen (n = 122: *M* = *3*.63; *SD* = .72, t = -2.58, p < .001, d = .29).

Darüber hinaus lässt sich ein signifikant positiver und starker Zusammenhang zwischen den Belastungen in den Bereichen *Schule* und *Studienseminar* beobachten (*r* = .40, p < .001). Mit Blick auf das *Belastungsfeld Schule*, lässt sich feststellen, dass das Belastungsempfinden signifikant stärker ausgeprägt ist, umso mehr Unterrichtsbesuche zu absolvieren sind (N = 349: *r* = .13, p = .02). Weiterführende differenzielle Analysen anhand des t-Tests zeigen dabei, dass diejenigen, die weniger als 13 Unterrichtsbesuche haben (*Md* = 13.00), sich substantiell weniger belastet fühlen (N = 232: *M* = *2*.21; *SD* = .65) als LiVD, die viele Unterrichtsbesuche haben (n = 117: *M* = *3*.63; *SD* = .72, t = -3.23, p < .001, d = .36).

6 Diskussion

Zusammengefasst lässt sich festhalten, dass die strukturellen Anforderungen – zumindest in Bezug auf die Unterrichtsbesuche und schriftlichen Entwürfe – erwartungsgemäß an den einzelnen Ausbildungsstandorten variieren. Auch das Belastungsempfinden ist zumindest im *Belastungsfeld Studienseminar* überdurchschnittlich stark bei LiVD ausgeprägt. Da die Unsicherheiten und das Überforderungsempfinden allerdings nicht über die Anzahl der Unterrichtsbesuche vorhergesagt werden können, muss nach weiteren Prädiktoren für das psychosoziale Belastungsempfinden gesucht werden.

Es muss darauf hingewiesen werden, dass diese Ergebnisse auf den subjektiven Sichtweisen der LiVD beruhen. Dementsprechend ist es möglich, dass die von den LiVD angegebene Anzahl an Unterrichtsbesuchen nicht den tatsächlichen allgemeinen Vorgaben der verschiedenen Studienseminare entspricht. Auch die Gründe für die gegebenen Antworten besonders in Bezug auf die Belastungsbereiche sind nicht rekonstruierbar. Weiterhin ist darauf hinzuweisen, dass es sich bei dieser Untersuchung um ein Querschnittdesign bzw. eine Momentaufnahme handelt, weshalb die Ergebnisse von bestehenden Umständen zum Zeitpunkt der Datenerhebung beeinflusst sein können. Es können in der Studie keine Aussagen darüber getroffen werden, welche Bedeutung die Unterschiede zwischen den Studienseminaren insgesamt für die Unterschiede individueller beruflicher Beanspruchung haben. Dennoch zeigt sich, dass die

offen formulierten Ausbildungsanforderungen zu Unterschieden im Belastungserleben zwischen den Studienseminaren führen.

Zurückgewiesen werden musste die Hypothese, dass das Belastungsempfinden an der Schule in der Rolle als Lehrende stärker ausgeprägt ist als am Studienseminar. Dieser Befund deckt sich zwar mit Erkenntnissen aus anderen Studien (vgl. u. a. Christ/van Dick/Wagner 2004), dennoch könnte die enge Korrelation zwischen den *Belastungsfeldern Schule* und *Studienseminar* und die signifikante Korrelation zwischen dem *Belastungsfeld Schule* und den Unterrichtsbesuchen ein Indiz dafür sein, dass sich durch die Vermischung von Lernorten und Anforderungen innerhalb der beiden Belastungsfelder auch Rollen nicht klar voneinander trennen lassen. Weiterführende Untersuchungen sollten diese Differenzkonstruktion stärker in den Blick nehmen. Zudem stellt sich die Frage, inwieweit die strukturellen Bedingungen von Bedeutung für die Ausbildungsqualität sind. Das sollte Gegenstand weiterer Studien sein.

Für die Praxis lässt sich, trotz der beschriebenen Einschränkungen, resümieren, dass vor allem fehlende Transparenz an den Studienseminaren von den LiVD zu Unsicherheit führt. Einheitliche und klare Vorgaben innerhalb eines Bundeslandes dürften das Unsicherheitsempfinden deutlich verringern.

Literatur

Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R. & Trautwein, U. (2018): Student misbehavior and teacher well-being: Testing the mediating role of the teacher-student relationship. In: Learning and Instruction 58, S. 126–136.

Arens, A. K. & Morin, A. J. S. (2016): Relations between teachers' emotional exhaustion and students' educational outcomes. In: Journal of Educational Psychology 108, S. 800–813.

Böhner, M. M. (2010): Die zweite Phase der Lehrkräfteausbildung in Deutschland: Befunde und Optimierungsgrundsätze. In: Journal für Lehrerinnen- und Lehrerbildung 2010, H. 2, S. 58–64.

Braun, A. (2017): Erleben Lehrkräfte und Referendare berufsbezogene Belastungen anders? Berufs(phasen)pezifische Präventionsansätze zur Gesundheitsförderung. Bad Heilbrunn: Klinkhardt.

Christ, O.; van Dick, R. & Wagner, U. (2004): Belastung und Beanspruchung bei Lehrern in der Ausbildung. In: Hillert, A. & Schmitz, E. (Hrsg.): Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern. Ursachen, Folgen, Lösungen. Stuttgart: Schattauer, S. 113–119.

Drüge, M.; Schleider, K. & Rosati, A-S. (2014): Psychosoziale Belastungen im Referendariat – Merkmale, Ausprägungen, Folgen. In: Die Deutsche Schule 106, H. 4, S. 358–372.

Gecks, L. C. (1990): Sozialisationsphase Referendariat – objektive Strukturbedingungen und ihr psychischer Preis. Frankfurt a. M.: Peter Lang.

Kaluza, G. (2012): Gelassen und sicher im Stress. Das Stresskompetenz-Buch – Stress erkennen, verstehen, bewältigen. 4., überarb. Auflage. Berlin Heidelberg: Springer.

Kiel, E. & Weiß, S. (2015): Gesund im Referendariat – Die zweite Ausbildungsphase des Lehramts aus salutogener Perspektive. In: Seminar 21, H. 1, S. 93–106.

Kliebisch, U. W. (2009): Stress verstehen und bewältigen. Hilfen für Lehrerinnen und Lehrer. In: Kliebisch, U. W. & Meloefski, R. (Hrsg.): LehrerGesundheit. Anregungen für die Praxis. Baltmannsweiler: Schneider Verlag Hohengehren, S. 158–176.

Klippert, H. (2004): Lehrerbildung. Unterrichtsentwicklung und der Aufbau neuer Routinen. Weinheim und Basel: Beltz.

Klusmann, U.; Kunter, M.; Trautwein, U. & Baumert, J. (2006): Lehrerbelastung und Unterrichtsqualität aus der Perspektive von Lehrenden und Lernenden. In: Zeitschrift für Pädagogische Psychologie 20, S. 161–173.

Klusmann, U., Kunter, M., Voss, T. & Baumert, J. (2012): Berufliche Beanspruchung angehender Lehrkräfte: Die Effekte von Persönlichkeit, pädagogischer Vorerfahrung und professioneller Kompetenz. In: Zeitschrift für Pädagogische Psychologie 26, H. 4, S. 275–290.

Krause, A. & Dorsemagen, C. (2007): Ergebnisse der Lehrerbelastungsforschung: Orientierung im Forschungsdschungel. In: Rothland, M. (Hrsg.): Belastung und Beanspruchung im Lehrerberuf. Modelle, Befunde, Interventionen. Wiesbaden: VS, S. 52–80.

Lazarus, R. S. & Folkman, S. (1984): Stress, appraisal, and coping. New York: Springer.

Lenhard, H. (2004): Zweite Phase an Studienseminaren und Schulen. In: Blömeke, S.; Reinhold, P.; Tulodziecki, G. & Wildt, J. (Hrsg.): Handbuch Lehrerbildung. Bad Heilbrunn: Klinkhardt, S. 275–290.

Pille, T. (2013): Das Referendariat. Eine ethnographische Studie zu den Praktiken der Lehrerbildung. Bielefeld: transcript.

Rohmert, W. & Rutenfranz, J. (1975): Arbeitswissenschaftliche Beurteilung der Belastung und Beanspruchung an unterschiedlichen industriellen Arbeitsplätzen. Bonn: Der Bundesminister für Arbeit und Sozialordnung.

Satow, L. (2012): „Stress- und Coping-Inventar (SCI). Test- und Skalendokumentation." www.drsatow.de (Abfrage: 20.03.2019).

Schaarschmidt, U. (2005): Halbtagsjobber? Psychische Gesundheit im Lehrerberuf – Analyse eines veränderungsbedürftigen Zustandes. 2. Auflage. Weinheim und Basel: Beltz.

Verordnung über die Ausbildung und Prüfung von Lehrkräften im Vorbereitungsdienst (2017). www.schure.de/20411/apvo-lehr.htm (Abfrage: 20.03.2019).

IV Bedingungen effektiven Unterrichts

Marc Kleinknecht

Unterricht lässt sich aus unterschiedlichen Theorie- und Forschungsperspektiven betrachten und mit verschiedenen Forschungsmethoden untersuchen. Im folgenden Einführungsbeitrag wird im ersten Teil die theoretische Perspektive der Unterrichtsqualitätsforschung dargestellt. Im Zentrum steht das Angebots-Nutzungs-Modell, das versucht, Bedingungen und Wirkungen von Unterricht zu erklären, und damit als Heuristik zur Planung und Analyse dienen kann. Im zweiten Teil wird auf die Frage nach der Perspektive auf Unterricht eingegangen. Es werden aktuelle Befunde zur Güte von Urteilen der Lehrkräfte, Lernenden und externen Beobachtenden zur Unterrichtsqualität vorgestellt. Am Ende der beiden Teilkapitel werden die dem Einführungskapitel folgenden Beiträge zu sieben Studien in einem vereinfacht dargestellten Angebots-Nutzungs-Modell verortet (Teil 1) und die empirischen Ansätze der Studien in ihrer jeweiligen Perspektivität auf Unterrichtsqualität dargestellt und reflektiert (Teil 2).

Effektiver Unterricht – das Angebots-Nutzungs-Modell als Forschungsheuristik

Mit der Frage, was effektiven oder guten Unterricht charakterisiert, beschäftigten sich unterschiedliche Bereiche der Bildungs- und Erziehungswissenschaft. Das Spektrum reicht von theoretischen und normativ begründeten Analysen einer bildungstheoretischen Didaktik bis hin zur empirischen Messung effektiven Lehrpersonenverhaltens vor dem Hintergrund behavioristischer oder kognitivistischer Lerntheorien. In jüngster Zeit ist die Frage nach Unterrichtsqualität mit der empirischen Forschung zu Unterrichtsmerkmalen im Rahmen eines Angebots-Nutzungs-Modells verbunden.

Die empirische Perspektive der deutschsprachigen Unterrichtsqualitätsforschung ist dadurch gekennzeichnet, dass die Qualität von Unterricht an seinen Wirkungen bzw. Effekten gemessen wird. Eine für jeden Unterricht gültige abgeschlossene Liste von Merkmalen oder ein erfolgreiches methodisches Vorgehen kann es nach dieser Definition nicht geben. Die Frage nach der Qualität ist etwa davon abhängig, welche Ziele auf der Wirkungsseite definiert werden (z. B. kognitive oder motivationale Ziele), welche Klassenzusammensetzung

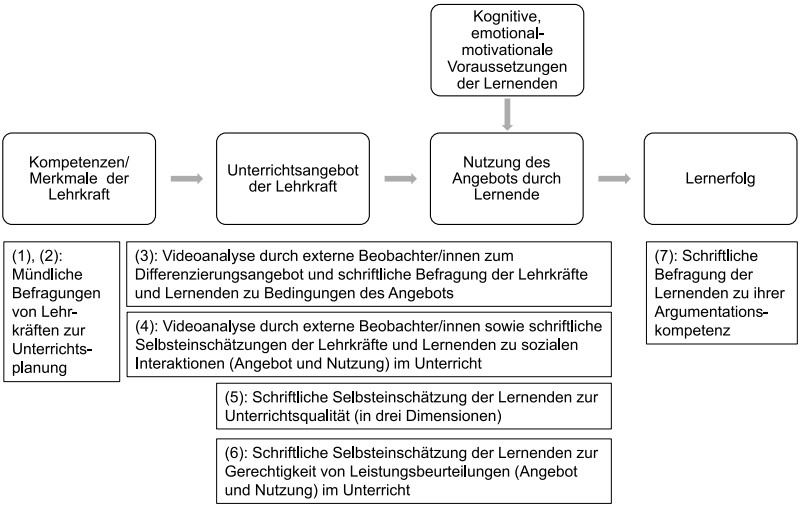
und Schülervoraussetzungen vorzufinden sind und wie die Schülerinnen und Schüler das Unterrichtsangebot nutzen.

Die aktuellen Modellierungen zur Unterrichtsqualität unterscheiden sich von früheren Ansätzen eines Prozess-Produkt-Ansatzes (vgl. Caroll 1963) erstens dadurch, dass weitere Variablen zur Erklärung von Unterrichtserfolg aufgenommen werden, etwa die Nutzung des Unterrichtsangebots durch die Schülerinnen und Schüler oder die fachliche und didaktische Expertise der Lehrkraft. Es wurden insbesondere Annahmen und Befunde der Experten-Novizen-Forschung rezipiert, die auf die hohe Bedeutung von verschiedenen Formen des Wissens auf das professionelle Handeln verweisen. Zweitens orientiert sich die empirische Forschung heute in höherem Maße als früher an theoretischen Vorstellungen zu gelingenden Lehr-Lernprozessen und zur Struktur des Unterrichts. Ausgehend von kognitiven und soziokonstruktivistischen Theorien des Lernens wird Unterricht als Gelegenheitsstruktur für verständnisvolle Lernprozesse verstanden (vgl. u. a. Baumert/Kunter 2006). Diese Struktur wiederum lässt sich in eine Sichtstruktur und eine Tiefenstruktur unterteilen, wobei Sichtstrukturen die Inszenierungsformen (Sozialformen, methodisches Konzept) und die Tiefenstrukturen die Aktionsformen innerhalb dieser Sichtstrukturen (z. B. Frage- und Aufgabenstellungen) umfassen (vgl. Seidel/Thiel 2017). Das Angebots-Nutzungs-Modell wurde von Helmke (2003) in Anlehnung an Fend (1981) aufgegriffen, ohne die system- und handlungstheoretischen Hintergründe zu rezipieren (vgl. Vieluf et al. i.Dr.). Vereinzelt findet sich das Modell auch in englischsprachigen Publikationen (auch deutscher Autorinnen und Autoren), wobei der Diskurs hauptsächlich im deutschsprachigen Raum zu verorten ist.

In diesem deutschsprachigen Kontext ist bereits zur Berichtlegung der TIMSS-Video-Studie aus dem Jahre 1995 der Vorschlag zu finden, die unterschiedlich komplexen Merkmale drei Qualitätsdimensionen zuzuordnen und sie theoretisch an Traditionen in der Pädagogik und Psychologie rückzubinden. Ein Unterricht, der gleichzeitig zu einer hohen kognitiven Verstehensleistung und Motivation führt, zeichnet sich demnach durch eine störungspräventive Klassenführung und klare Strukturierung der Inhalte, ein schülerorientiertes Sozialklima und eine konstruktive Unterstützung sowie ein kognitiv-aktivierendes Unterrichtsangebot aus (vgl. Klieme 2019). Vorwiegend in Studien zum Mathematikunterricht konnte gezeigt werden, dass sich diese drei Dimensionen dazu eignen, die Angebotsseite durch Beobachtung und Befragung von Lehrkräften empirisch zu erfassen. Allerdings liegt nicht für alle Dimensionen und deren Teilmerkmale eine befriedigende Studien- und Datenlage vor, sodass von unterschiedlich hoher Evidenz für die Lernwirkung einzelner Merkmale auszugehen ist. Kognitive Aktivierung etwa gilt als relativ neue Qualitätsdimension, die vor allem im deutschsprachigen Raum und vorwiegend aus fachlicher bzw. fachdidaktischer Perspektive des Mathematik- und Naturwissenschaftsunter-

richts der Sekundarstufe untersucht wurde (für einen Überblick: Praetorius et al. 2018).

Abbildung 1: Verortung der Beiträge im Angebots-Nutzungs-Modell nach Helmke (2003) – sehr vereinfachte Darstellung (ohne die im Text angedeutete Rekursivität, die weitere Verbindungspfeile erforderlich machen würde)



Die folgenden Beiträge lassen sich entlang der Wirkungskette des Angebots-Nutzungs-Modells verorten (vgl. Abb. 1). Diese Orientierung am Angebot-Nutzungs-Modell erlaubt zum einen eine an bisherigen Befunden der Unterrichtsforschung orientierte Wirkungsmodellierung zu möglichen unabhängigen, mediierenden und abhängigen Variablen. Zum anderen dient sie im Rahmen der Professionalisierung von Lehrkräften dazu, das Handeln von Lehrkräften im Unterricht systematisch zu analysieren (vgl. Arnold 2018). Orientiert an den Vorstellungen zur Planung und Gestaltung von Unterricht führt die Planung einer Lehrkraft zu Ziel-, Inhalts-, Methoden- und Medienentscheidungen, die sich in spezifischen Lehrhandlungen im Unterricht zeigen. Die Schülerinnen und Schüler erfahren dieses Angebot als Lerngelegenheit, die sie individuell wahrnehmen und verarbeiten. Gleichzeitig kann ihr individuelles Nutzungsverhalten wieder als Angebot verstanden werden, sodass Angebot und Nutzung in einem rekursiven Verhältnis stehen (vgl. Vieluf et al. i.Dr.). Schließlich zeigt sich ein hoher Lernerfolg der Klasse und der/des einzelnen Lernenden in der erfolgreichen Lösung von Aufgaben auf Performanzebene. Lernerfolg gilt als wichtiger Maßstab für ein wirksames Zusammenspiel von Unterrichtsangebot und -nutzung und ist gleichzeitig wieder Voraussetzung für die Nutzung des Angebots. Zunächst werden die Zielrichtungen der Beiträge mit Hilfe des Angebots-Nutzungs-Modells verdeutlicht, im folgenden Kapitel wird dann stärker auf das methodische Design der Studien eingegangen. Das

methodische Vorgehen im Rahmen der Studien wird bereits in Abbildung 1 benannt und in Abschnitt 2 weiter ausgeführt.

Die Beiträge von Wildgans-Lang et al. (1) und Greiten et al. (2) beschäftigen sich mit dem Planungsverhalten der Lehrkräfte und damit mit dem ersten Kettenglied der Wirkungskette des Angebots-Nutzungs-Modells. Die Studie von Wildgans-Lang et al. (1) untersucht, wie erfahrene Mathematiklehrkräfte bei der Planung ihres Unterrichts Fehler und Fehlvorstellungen der Lernenden antizipieren und ihre unterrichtlichen Maßnahmen diagnosegesteuert planen. Die Autorinnen und Autoren fragen zudem, ob die Lehrkräfte bereits ein Diagnoseprozessmodell verwenden bzw. es als hilfreich für ihre Unterrichtspraxis einschätzen. Die Studie von Greiten et al. (2) vergleicht Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräfte, die im inklusiven Unterricht kooperieren, in ihren Planungsaktivitäten. Es wird untersucht, in welchen Situationen Lehrkräfte ihren Unterricht planen und welche Planungsmittel sie verwenden.

Die Beiträge von Jennek et al. (3), Scherzinger et al. (4) und Nonte et al. (5) fokussieren auf die Qualität von Lehr- und Lernsituationen im Unterricht und damit auf die Frage, wie das Angebot von Lehrkräften gestaltet und von den Schülerinnen und Schülern wahrgenommen bzw. genutzt wird. Das Forschungsprojekt von Jennek et al. (3) untersucht Differenzierungsmaßnahmen im Englischunterricht und deren Abhängigkeit von personalen Merkmalen der Lehrkräfte und Lernenden. Die Studie von Scherzinger et al. (4) geht den Fragen nach, wie pädagogische Interaktionen im Unterricht von Primarschulklassen (mikrogenetisch) verlaufen und sie von Lehrenden und Lernenden (global) eingeschätzt werden. Zudem erörtern die Autorinnen und Autoren, ob die Interaktionsgestaltung davon abhängt, ob eine Klassen- oder Fachlehrperson unterrichtet. Das Projekt von Nonte et al. (5) untersucht, ob sich die von Schülerinnen und Schülern wahrgenommene Unterrichtsqualität in MINT-Klassen und Klassen ohne Profil unterscheidet.

Die Beiträge von Körner et al. (6) und Kempf et al. (7) zielen auf die Nutzungsseite des Angebots-Nutzungs-Modells und untersuchen, wie Schülerinnen und Schüler das unterrichtliche Angebot wahrnehmen bzw. interpretieren und sich durch dieses Angebot ihr Wissen verändert. Körner et al. (6) untersuchen in ihrer Studie, ob Oberstufenschülerinnen und -schüler die Leistungsbewertung der Lehrkräfte in unterschiedlichen Fächern als gerecht empfinden. Zudem wird gefragt, ob die von den Lernenden wahrgenommene Zufriedenheit mit dem Unterricht und die Fachnote des letzten Zeugnisses das Gerechtigkeitsempfinden beeinflusst. Der Beitrag von Kempf et al. (7) geht darauf ein, wie Unterricht das Wissen von Schülerinnen und Schüler verändert. Sie präsentieren einen Kompetenztest zum schriftlichen Argumentieren von Schülerinnen und Schülern im Fach Deutsch der Jahrgangsstufen sieben bis zehn, der vor allem dazu dienen soll, den Lehrkräften ein adaptives, leistungsdiagnosebasiertes Unterrichten zu ermöglichen.

Bedingungen, Qualität und Wirkungen von Unterricht – eine Frage der Perspektive

In der Forschung und in der Schulpraxis sind es Schülerinnen und Schüler, Lehrkräfte oder trainierte Beobachterinnen und Beobachter, die Unterrichtsqualität einschätzen und Auskunft zu Bedingungen und Wirkungen des Unterrichts geben. Auch in den Beiträgen zu diesem Tagungsband werden diese drei Personengruppen herangezogen, um Bedingungen und Wirkungen von Unterricht zu ermitteln und die Qualität der Lehr- und Lernprozesse einzuschätzen. Für die Unterrichtsforschung stellt sich die Frage, inwieweit diese Perspektiven als geeignet angesehen werden können, um valide und reliabel die Qualität von Unterricht vor dem Hintergrund der in Kapitel 1 aufgezeigten Wirkungskette einschätzen zu können.

Mittlerweile zeigen einige Studien, dass die Einschätzungen der drei Personengruppen zu den gleichen Konstrukten der Unterrichtsqualität selten bis überhaupt nicht übereinstimmen (vgl. u. a. Clausen 2002; Baumert/Kunter 2006; Wagner et al. 2016). So konnte etwa Clausen (2002, S. 129) für den Vergleich von Selbsteinschätzungen von Lehrkräften und Schülerinnen und Schülern sowie Videoratings zeigen, dass die durchschnittliche Übereinstimmung von vergleichbaren Skalen bei .16 lag, mit einer Schwankung von -.28 bis .45. In einer Übersicht zu weiteren Studien zeigen Fauth et al. (i.Dr.), dass weitere Studien ähnlich niedrige Übereinstimmungen ermittelten. Zudem verweist diese Übersicht darauf, dass die Einschätzungen bei Skalen der Klassenführung höher liegt als jene zur kognitiven Aktivierung oder Lernunterstützung.

Als Erklärung für diese geringe Übereinstimmung wurde zunächst im deutschen Forschungskontext darauf verwiesen, dass jede Perspektive eine spezifische Information zur Qualität von Unterricht liefert (vgl. Baumert et al. 2004; Baumert/Kunter 2006). Nach diesen Überlegungen zur differenziellen Validität unterschiedlicher Perspektiven auf Unterricht werden Schülerurteilen eine hohe Validität für das subjektive Erleben des Unterrichts zugesprochen. Sie geben Einblick in das individuelle Wohlbefinden und die individuelle Entwicklung der Lernenden. Die geteilte Schülerwahrnehmung in einer Lerngruppe gilt wiederum als guter Indikator für tatsächlich bestehende Unterrichtsroutinen, ohne dass allerdings eine neutrale und differenzierte Beschreibung oder gar eine vertiefte Beurteilung des Unterrichts hinsichtlich vorherrschender didaktischer Konzepte und Prinzipien der Lehrkraft möglich ist. Schüleraussagen eignen sich damit vor allem dann, wenn globale Eindrücke ermittelt werden sollen. Den Einschätzungen von Lehrkräften zum Unterricht kann hingegen eine hohe Aussagekraft zugesprochen werden, wenn es um allgemeine und langfristig wirkende didaktische und methodische Aspekte des Unterrichts geht. Sie können als geeignete Quelle angesehen werden, wenn Planungs- und Gestaltungskonzepte erhoben werden sollen, die über den ein-

zelnen Unterricht hinausgehen. Zudem eignen sich Lehrerauskünfte, um subjektive, handlungsleitende Überzeugungen zu erschließen. Die Validität der Lehrerperspektive erweist sich jedoch als gering, wenn Lehrkräfte ihren eigenen Unterricht anhand bestimmter Merkmale qualitativ bewerten sollen und Vergleiche mit anderen Lehrkräften notwendig werden. Schließlich erweist sich die Beobachterperspektive als am besten geeignet, um Unterrichtsprozesse möglichst objektiv und detailliert zu beschreiben. Sie liefert punktuelle Einblicke in den Unterricht, bei dem theoretische Konzepte herangezogen und unterschiedliche Klassen und Lehrpersonen verglichen werden können. Eine geringe Validität scheint für die Beschreibung und Beurteilung von seltenen Ereignissen zu bestehen. Ebenso kann nur bedingt das individuell adaptive Vorgehen von Lehrpersonen eingeschätzt werden, da keine Informationen zum Vorwissen der Lernenden vorliegen. Beobachtungsdaten liefern zudem keine Informationen dazu, welche handlungsleitenden Kognitionen von Lehrenden und Lernenden dem Unterricht zugrunde liegen. Sie geben auch nur begrenzt Einblick in nicht äußerlich sichtbare kognitive Wahrnehmungs- und Interpretationsprozesse der Schülerinnen und Schüler (vgl. Vieluf i Dr.). Die Validität der Beobachterperspektive ist auch von der Analysestrategie abhängig, wobei zwischen niedrig-, mittel- und hoch-inferenten Auswertungsverfahren und dem Grad der Schlussfolgerungen, der den Entscheidungen zugrunde liegt, unterschieden werden kann (vgl. Seidel/Thiel 2017). Derzeit werden vor allem hoch-inferente Verfahren zur Einschätzung von Unterrichtsqualität herangezogen. Sie ermöglichen eine stärker didaktische Sicht auf Unterricht, weil längere Unterrichtsabschnitte untersucht werden können. Im Vergleich zu niedrig-inferenten Verfahren und in Bezug auf Merkmale der Unterrichtsqualität kann demzufolge eine höhere Validität dieses Verfahrens angenommen werden. Als Nachteil erweisen sich allerdings die geringen Übereinstimmungen zwischen den Raterinnen und Ratern. Je nach fachlicher Kenntnis und subjektiven Vorannahmen können sich stark abweichende Entscheidungen ergeben. Die Vielzahl an relevanten Indikatoren und zu bewertenden Situationen erschweren dabei objektive und reliable Urteile. Hinsichtlich der Vor- und Nachteile von niedrig- bzw. hoch-inferenten Verfahren wird auch von einem „Übereinstimmungs-Genauigkeits-Dilemma“ (Clausen 2002, S. 191) gesprochen.

In einer neueren Publikation zur Erforschung von Unterrichtsqualität wird vorgeschlagen, die verwendeten hochinferenten Items in Befragungs- und Beobachtungsinventaren hinsichtlich deren Referenz auf Verhalten der Lehrkräfte, der Lernenden oder auf Verhalten von Lehrkräften und Lernenden noch systematischer zu reflektieren (vgl. Fauth et al. i.Dr.). Selbsteinschätzungen von Lehrkräften zu ihrer Klassenführung fallen nachweislich anders aus („Ich habe Schwierigkeiten, Sachverhalte klar zu erklären“) als Selbsteinschätzungen dieser Personengruppen, in denen stärker das Schülerverhalten eingeschätzt werden soll („Das Schülerverhalten in dieser Klasse ist ein Problem“). Wagner et al.

(2016) konnten zeigen, dass Schülerinnen- bzw. Schüler- und Lehrereinschätzungen zur Klassenführung dann am meisten übereinstimmten, wenn die Items auf das Schülerverhalten fokussierten. Bei einem Fokus auf das Verhalten von Lehrkräften zeigen sich dagegen in verschiedenen Studien kaum Übereinstimmungen in den Einschätzungen der Lehrkräfte und Lernenden.

Auf Basis dieser Erkenntnisse zu differentiellen Validität der drei Perspektiven auf Unterricht sollen im Folgenden die Forschungsdesigns der bereits dargestellten Beiträge dargestellt und reflektiert werden. Für eine Veranschaulichung der methodischen Vorgehensweise dient die im ersten Teil dieser Übersicht präsentierte Abbildung 1. Sie zeigt im unteren Teil, welche Personengruppen für die Studie mit welchen Methoden zu Bedingungen, Lehr-Lernprozessen oder Wirkungen Auskunft gibt.

Die Beiträge von Wildgans-Lang et al. (1) und Greiten et al. (2) stellen Interviewstudien dar, bei denen Lehrkräfte zu ihren Planungsprozessen befragt werden. Wildgans-Lang et al. (1) nehmen aufgrund von theoretischen Vorstellungen zum adaptiven Unterricht an, dass ein systematisches Nachdenken über Fehlvorstellungen von Schülerinnen und Schülern sich positiv auf den Unterricht auswirkt. Die Interviewdaten interpretieren die Autorinnen und Autoren so, dass die befragten Lehrkräfte neue Diagnosekonzepte für ihre Planungsaktivitäten nutzen würden. Greiten et al. (2) befragen Lehrkräfte aufgrund der Annahme, dass eine ko-konstruktive Unterrichtsplanung von Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräften sich positiv auf die Unterrichtsqualität auswirkt. Die Fragen im Interview fokussieren auf die Planungsaktivitäten, von denen auf Planungskonzepte der Lehrkräfte und implizit auch auf die Unterrichtspraxis geschlossen wird. Für die in beiden Studien untersuchten Lehrkräfte wäre interessant zu erfahren, wie sich der Unterricht durch eine diagnosegestützte Planung verbessert bzw. wie die Lehrkräfte-Teams ihren Unterricht im Alltag gestalten.

Jennek et al. (3) und Scherzinger et al. (4) nutzen Beobachterinnen und Beobachter um die Lehr- und Lernprozesse im Unterricht zu untersuchen. Zudem werden in beiden Studien Lehrende und Lernende befragt. In der Studie von Jennek et al. (3) werden die Differenzierungsmaßnahmen der Lehrkräfte mit einem mittel-inferenten Manual untersucht. Dabei wird angenommen, dass Differenzierungsmaßnahmen die Qualität des Unterrichts erhöhen. Die Nutzung des Angebots und dessen Wirkung wird in dieser Studie allerdings nicht erforscht. Von Interesse ist vielmehr, wie personale Faktoren den Einsatz von Differenzierungsmaßnahmen beeinflussen. Um diese Faktoren zu ermitteln, werden Lehrende und Lernende befragt.

Scherzinger et al. (4) untersuchen in einer Beobachtungsstudie sowohl die Angebots-, als auch die Nutzungsseite, indem sie die Interaktionen zwischen Lehrkräften und Lernenden im Unterricht niedrig-inferent kodieren und mit

der Space Grids Methode die Interaktionsverläufe beschreiben. Zudem befragen sie die Lehrenden und Lernenden zu den Interaktionen und setzen diese Befragungsdaten zu den Beobachtungsdaten in Beziehung. Sie können ermitteln, ob Beobachtungs- und Einschätzungsdaten insgesamt und für die beiden untersuchten Lehrpersonengruppen (Klassen- vs. Fachlehrpersonen) übereinstimmen. In der Studie werden damit Einschätzungen zum Unterrichtsangebot und zur -nutzung aus verschiedenen Perspektiven herangezogen.

Die Studie von Nonte et al. (5) zielt ebenfalls wie die beiden letztgenannten Studien auf die Qualität der Lehr- und Lernprozesse im Unterricht. Im Unterschied zu den beiden Studien werden allerdings ausschließlich Befragungen von Schülerinnen und Schülern zur von ihnen wahrgenommenen Unterrichtsqualität herangezogen. Die Lernenden sollten je fünf bis sieben Aussagen zur kognitiven Aktivierung, Klassenführung und zum unterstützenden Verhalten hochinferent einschätzen. Mit Blick auf die Studien zur Validität der Perspektiven auf Unterrichtsqualität wäre interessant zu wissen, ob Schülerinnen und Schüler die Qualität von Unterricht differenziert einschätzen können bzw. ob eine Einschätzung von Fremdbeobachtenden zu ähnlichem Ergebnissen geführt hätte.

Die Beiträge von Körner et al. (6) und Kempf et al. (7) zielen auf die Lernenden und deren Wahrnehmungen und Leistungen. In der Studie von Körner et al. (6) werden die Schülerinnen und Schüler zu deren Wahrnehmung von Unterrichtsangeboten befragt. Konkret sollen die Schülerinnen und Schüler einschätzen, wie gerecht sie die Notenpraxis ihrer Lehrkräfte empfinden. Als weitere Bedingungen des Unterrichts werden aus Schülerperspektive die Zufriedenheit mit dem Unterricht und die Zeugnisnoten in zwei Fächern erhoben. Es wird theoretisch und auf Basis der Ergebnisse angenommen, dass eine faire Leistungsbewertung aus Sicht von Schülerinnen und Schülern ein zentrales Merkmal von Unterrichtsqualität darstellt. Eventuell könnten Unterrichtsbeobachtung und/oder Lehrpersonenbefragungen weitergehend untersuchen, wie sich eine faire Leistungsbewertung im Unterricht oder in Planungsprozessen der Lehrkräfte zeigt.

Die Studie von Kempf et al. (7) fokussiert zunächst die Lernenden mit ihren Kompetenzen. Mit offenen, geschlossenen und halboffenen Aufgaben wird untersucht, ob Schülerinnen und Schüler Wissen im Bereich des schriftlichen Argumentierens vorweisen können. Angenommen wird, dass Deutschlehrkräfte im Unterrichtsalltag den Test einsetzen und die Ergebnisse zur Planung und Reflexion ihres Unterrichts nutzen können. Weitere Studien sind geplant und könnten eventuell untersuchen, ob die diagnosegestützte Planung der untersuchten Lehrkräfte, deren Deutschunterricht und den Lernerfolg der Schülerinnen und Schüler verbessert.

In den dargestellten Studien untersuchen Forscherinnen und Forscher mit spezifischen Methoden einen Teil oder mehrere Teile der Wirkungskette im Angebot-Nutzungs-Modell. Die Verortung der Studien im Angebots-Nutzungs-Modell und die Einschätzung zu den methodisch gewählten Perspektiven auf Unterrichtsqualität soll es Leserinnen und Lesern ermöglichen, die Studien in ihrem Erkenntnisinteresse und ihrem methodischen Vorgehen zu reflektieren. Die Einführung in den Thementeil soll zudem dazu anregen, über Forschung zum effektiven Unterricht nachzudenken.

Literatur

Arnold, Karl-Heinz (2018): Wirrnis der Unterrichtsmethoden: Konzeptuelle Probleme der unterrichtlichen Lehr-Lernforschung. In: Zeitschrift für Pädagogische Psychologie 32, H. 4, S. 193–204.

Baumert, Jürgen/Blum, Werner/Brunner, Martin/Krauss, Stefan/Kunter, Mareike/Neubrand, Michael (2004): Mathematikunterricht aus Sicht der PISA-Schülerinnen und -schüler und ihrer Lehrkräfte. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs S. 354–414. Münster: Waxmann, S. 354–414.

Baumert, Jürgen/Kunter, Mareike (2006): Stichwort: Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520.

Carroll, John B. (1963): A model of school learning. In: Teacher College Record 64, H. 8, S. 723–733.

Clausen, Marten (2002). Unterrichtsqualität: Eine Frage der Perspektive? Münster: Waxmann.

Fauth, Benjamin/Göllner, Richard/Lenske, Gerlinde/Praetorius, Anna-Katharina/Wagner, Wolfgang (i.Dr.): Who sees what? Conceptual considerations on the measurement of teaching quality from different perspectives. In: Zeitschrift für Pädagogik.

Fend, Helmut (1981): Theorie der Schule. München: Urban & Schwarzenberg.

Helmke, Andreas (2003): Unterrichtsqualität: Erfassen, Bewerten, Verbessern. 1. Auflage. Seelze: Kallmeyersche Verlagsbuchhandlung.

Klieme, Eckhard (2019): Unterrichtsqualität. In: Gläser-Zikuda, Michaela/Harring, Marius/Rohlfs, Carsten (Hrsg.): Handbuch Schulpädagogik. Münster: Waxmann, S. 393–408.

Praetorius, Anna-Katharina/Klieme, Eckhard/Herbert, Benjamin/Pinger, Petra (2018): Generic dimensions of teaching quality: The German framework of the three basic dimensions. In: ZDM Mathematics Education, S. 407–426.

Seidel, Tina/Thiel, Felicitas (2017): Standards und Trends der videobasierten Lehr-Lernforschung. Zeitschrift für Erziehungswissenschaft 20, H. 1, S. 1–21.

Vieluf, Svenja/Praetorius, Anna-Katharina/Rakoczy, Katrin/Kleinknecht, Marc/Pietsch, Marcus (i.Dr.). Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts: eine kritische Auseinandersetzung mit ihrer theoretischen Konzeption. In: Zeitschrift für Pädagogik.

Wagner, Wolfgang/Göllner, Richard/Werth, Sarah/Voss, Thamar/Schmitz, Bernhard/Trautwein, Ulrich (2016): Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. In: Journal of Educational Psychology 108, H. 5, S. 705–721.

Epistemisch-diagnostische Aktivitäten im Diagnoseprozess bei Lehrkräften im Mathematikunterricht

Angelika Wildgans-Lang, Andreas Obersteiner und Kristina Reiss

Abstract

Für die adaptive Gestaltung von Unterricht ist es dienlich, dass Lehrkräfte den Lernstand ihrer Schülerinnen und Schüler kennen, Fehlvorstellungen möglichst zeitnah erkennen und ihnen entgegenwirken. Obwohl seit langem bekannt ist, dass Lehrkräfte Diagnosekompetenzen besitzen sollten, wird der Diagnoseprozess in der ersten Phase der Lehrerbildung oftmals nicht ausführlich besprochen. Auch in der zweiten Phase der Lehrerbildung wird der Diagnoseprozess eher intuitiv und wenig theoretisch fundiert behandelt. Auch wenn Diagnosekompetenzen in den letzten Jahren zunehmend an Bedeutung gewinnen, ist bis heute unklar, wie Lehrkräfte diagnostizieren, die keine explizite Vorbildung in der ersten und zweiten Phase der Lehramtsausbildung erfuhren. Die vorliegende Studie untersucht daher, inwieweit erfahrene Lehrkräfte mit Fehlern und Fehlvorstellungen im Unterricht umgehen, ob sie Fehlvorstellungen im Unterricht diagnostizieren und ob ihr Handeln in einem schematischen Diagnoseprozess abgebildet werden kann.

1 Diagnose von Fehlvorstellungen im Mathematikunterricht

„Teachers are among the most powerful influences on learning." (Hattie 2010, S. 238) Diese plakative These stellt Hattie in Folge seiner Metastudien auf. Folgt man einer Studie von Weinert, Schrader und Helmke (1990), so haben Klassenführung, Diagnosekompetenz, unterrichtsmethodische Kompetenz und Sachwissen als Aspekte der professionellen Kompetenz wesentlichen Einfluss auf die Leistungen der Schülerinnen und Schüler. Insbesondere bei Leistungsheterogenität sind Klassenführung und Diagnosekompetenz dafür besonders wichtige Prädiktoren (Weinert/Schrader/Helmke 1990).

Universitäten sollten fachliche, fachdidaktische und pädagogisch-psychologische Kompetenzen vermitteln. Dennoch wird bemängelt, dass speziell diagnostische Kompetenz an deutschsprachigen Universitäten eher vernachlässigt wird (Oser 1997; Abs 2006; Hesse 2017), auch wenn sie mittlerweile erforscht wird (Südkamp/Praetorius 2017; Leuders/Philipp/Leuders 2018). Diagnostische Kompetenz bzw. Diagnosekompetenz meint hierbei nicht nur die akkurate Diagnose von Lernschwierigkeiten bei Schülerinnen und Schülern, sondern vielmehr die Fähigkeit, im Unterricht Wissen, Kompetenzen und Fehlvorstellungen zu erkennen. Diagnosekompetenz beschreibt folglich die Fähigkeit einer Lehrkraft, entsprechende Informationen zielgerichtet zu sammeln und zu integrieren, um so Unsicherheit zu reduzieren und eine angemessene pädagogische Entscheidung zu treffen (Heitzmann et al. 2019).

Im konstruktivistisch geprägten Mathematikunterricht sind Fehler „die notwendige Voraussetzung für jede erfolgreiche Adaption" (Rollett 1999, S. 71). Fehler bieten somit eine Lerngelegenheit, auch wenn sie definitorisch Abweichungen von der Norm (Oser/Hascher/Spychiger 1999) darstellen.

Der erste Schritt im Diagnoseprozess kann das Erkennen eines Schülerfehlers im Mathematikunterricht sein. Das Wahrnehmen eines Fehlers ist somit zentral für das Diagnostizieren, ob der Fehler aufgrund einer oder mehrerer Fehlvorstellungen entstand ist. Unter Fehlvorstellungen werden vorliegend fehlerhafte Vorstellungen und Konzepte verstanden, die als reproduzierbar gelten (Radatz 1980).

Inwiefern Mathematiklehrkräfte gezielt Diagnoseprozesse initiieren, wie diese ablaufen und ob das Diagnostizieren ihrer Schülerinnen gelingt, ist derzeit Gegenstand von aktuellen Forschungsprojekten. Die von der DFG geförderte interdisziplinäre Forschungsgruppe COSIMA (FOR 2385) untersucht Diagnoseprozesse in mehreren Teilprojekten aus dem Bildungsbereich und der Medizin.

2 Modelle des Diagnoseprozesses

In der Forschungsgruppe COSIMA wird zur Beschreibung des Diagnoseprozesses das Prozessmodell des Scientific Reasoning and Argumentation (Fischer et al. 2014) genutzt. Im Zentrum stehen acht epistemisch-diagnostische Aktivitäten, die im Folgenden im Vergleich zum Arbeitsmodell der NeDiKo-Gruppe beschrieben werden: Problem identifizieren, Fragen stellen, Hypothesen generieren, Artefakte konstruieren, Evidenzen generieren, Evidenzen evaluieren, Schlussfolgerungen ziehen und Prozess bzw. Ergebnis kommunizieren.

Ähnlich wie im parallel entwickelten Modell des Diagnoseprozesses des DFG-Netzwerks Diagnostische Kompetenz (NeDiKo) gilt es im Rahmen einer Diagnose zunächst, die vorliegenden Informationen über Lernende zu verar-

beiten (Herppich et al. 2017), also ihre konkreten Probleme zu identifizieren (Fischer et al. 2014). Die Forschungsgruppe COSIMA geht hierbei von einem Problem bzw. einem Fehler aus, wohingegen dem NeDiKo Netzwerk ein gewisses Diagnoseziel als Ausgangspunkt dient. Eine Diagnose kann direkt gestellt werden, falls genügend Informationen vorliegen (Herppich et al. 2017), was bedeutet, dass Schlussfolgerungen gezogen und der Prozess sowie das Ergebnis kommuniziert werden (Fischer et al. 2014). Alternativ ist eine Hypothesenbildung anhand der bereits gesammelten Evidenzen möglich, sodass weitere Informationen gezielt auf Grundlage dieser Hypothesen eingeholt werden (Herppich et al. 2017). Fischer et al. (2014) bezeichnen dieses mögliche Vorgehen als Evidenzgenerierung. Eine Voraussetzung dafür ist, dass passende Aufgaben zur Verfügung stehen. Ist das nicht der Fall, müssen diese erstellt werden, was Fischer et al. als Artefaktkonstruktion bezeichnen. Nach der Evidenzgenerierung sind die Informationen über den Lernenden zu evaluieren (Fischer et al. 2014) oder zu verarbeiten (Herppich et al. 2017). Der beschriebene parallele Aufbau der beiden Modelle lässt die Annahme zu, dass die epistemisch-diagnostischen Aktivitäten geeignet sind, den Prozess der Diagnose in der Realität abzubilden. Das interdisziplinäre Prozessmodell hat den Vorteil, dass bspw. Erkenntnisse zur Diagnosekompetenzförderung aus anderen Disziplinen in den Lehramtsbereich übertragen werden können.

Für die Betrachtung der Diagnosekompetenz scheint der Mathematikunterricht als Arbeitsfeld besonders geeignet, da er geprägt ist von Aufgaben, die von Lernenden zu lösen sind. Dabei entstehen häufig Fehler, die als Start für den Diagnoseprozess von Lehrkräften dienen. Es liegt an der Lehrkraft, produktiv mit den Fehlern im Klassenzimmer umzugehen (Kuntze/Heinze/Reiss 2008). Bereits in einer frühen systematischen Kategorisierung zu Fehlern im Mathematikunterricht (Radatz 1980) werden die „Möglichkeiten der Fehleranalyse" (Radatz 1980, S. 26) als eines von drei großen Themengebieten der Fachdidaktik der Mathematik angesehen. Damals gab es erste Ansätze, Fehler als wiederholend auftretende Abweichungen von der Norm anzusehen. Wichtig an dieser Betrachtungsweise war, Fehler zu erkennen, die auf Regeln basieren, die Lernende systematisch anwenden (Ginsburg 1977; Radatz 1980; Oser/Hascher/Spychiger 1999), was heutzutage als Fehlvorstellung (Leinhardt/Zaslavsky 1990) bekannt ist. Diagnosekompetenzen wurden zu einem frühen Zeitpunkt unter dem Begriff *diagnostic teaching* (Fowler 1980) eingeführt. Hier wird davon ausgegangen, dass Lernende möglicherweise nicht eine weitere Erklärung benötigen, sondern vielmehr Lehrende die Ursache des Fehlers erkennen und Lernende in diesem Bereich konkret fördern können sollten, um die Fehlvorstellung zu beheben.

Oser beschreibt als einer der ersten einen „ ‚offenen' oder ‚losen' Algorithmus" (2005, S. 125 ff.), der Lehrkräften eine Vorgehensweise an die Hand geben soll. Die ersten beiden der zehn Schritte sind vergleichbar mit der Diagnose,

gefolgt von einer Besprechung des Fehlers und einer anschließenden Förderung der korrekten Vorstellung. Auch Abs (2007) erläutert, dass Lehrkräfte einschätzen müssen, was der Lernende ihnen korrekterweise hätte sagen wollen und welche Fehlvorstellungen eine korrekte Antwort verhinderten. Zusätzlich müssten Lehrende einschätzen, in welcher motivationalen, emotionalen und kognitiven Verfassung sich der Lernende befindet und welche Antwort oder Methode somit die sinnvollste ist (Abs 2007).

Alle diese Modelle gehen jeweils von der expliziten Diagnose von Fehlvorstellungen oder vorhandenen Kompetenzen bei einzelnen Lernenden aus; daher ist es unklar, wie Lehrkräfte diese im Klassenzimmer umsetzen. Neben der Diagnose sind sie gefordert, parallel die Klasse zu führen, fachliche Inhalte zu vermitteln und Methoden auszuwählen (Weinert/Schrader/Helmke 1990). Es erscheint daher sinnvoll, Überlegungen zu Problemen und Schwierigkeiten bereits im Vorfeld des Unterrichts anzustellen. Problemidentifikation, Hypothesenbildung und Artefaktkonstruktion sollten dabei in der Vorbereitung des Mathematikunterrichts durchdacht werden, sodass sie im Unterricht lediglich aufgenommen werden müssen.

Zusammenfassend ist festzustellen, dass das Modell des Scientific Reasoning and Argumentation (Fischer et al. 2014) sich in die bisherige Diagnoseprozessforschung einpassen lässt. Es bleibt allerdings zu überprüfen, ob das eher allgemeine Diagnoseprozessmodell auch die Abläufe der Diagnose im realen Mathematikunterricht beschreiben kann.

3 Fragestellungen zur Anwendung der epistemisch-diagnostischen Aktivitäten

Ziel dieser Studie ist es, den Umgang von Lehrkräften mit Fehlern und Fehlvorstellungen näher zu betrachten. Konkret soll untersucht werden, ob Diagnoseprozesse von Lehrkräften im Unterricht geleistet werden und ob Lehrkräfte sich bereits in der Unterrichtsvorbereitung bzw. während des Unterrichts mit potenziellen Schwierigkeiten und Problemen ihrer Schülerinnen und Schüler befassen. Darüber hinaus ist die Frage zu beantworten, ob entsprechende Überlegungen anhand der epistemisch-diagnostischen Aktivitäten eingeordnet werden können. Es wird außerdem im Rahmen der Studie untersucht, ob Lehrkräfte das Prozessmodell des Scientific Reasoning and Argumentation (Fischer et al. 2014) als hilfreich für die Diagnose empfinden und nutzen würden. Hieraus ergeben sich folgende vier Annahmen:

1. Lehrkräfte identifizieren potentielle Probleme ihrer Schülerinnen und Schüler im Rahmen der Unterrichtsvorbereitung.
2. Lehrkräfte generieren Hypothesen über die Fehlvorstellungen ihrer Schülerinnen und Schüler.
3. Lehrkräfte testen ihre Hypothesen, formulieren Schlussfolgerungen und kommunizieren diese an ihre Schülerinnen und Schüler.
4. Lehrkräfte empfinden es als hilfreich, auf der Grundlage eines Diagnoseprozessmodells, wie das des Scientific Reasoning and Argumentation (Fischer et al. 2014), zu arbeiten.

4 Methodisches Vorgehen

Im Rahmen zweier Masterarbeiten (Steinleitner 2018; Keding 2018) wurden zehn Lehrkräfte einer beruflichen Oberschule (BOS) in Bayern und neun weitere Lehrkräfte an einem Gymnasium (5 weiblich, 14 männlich) interviewt. Zum Zeitpunkt der Studie waren die Lehrkräfte zwischen 26 und 66 Jahren alt ($M = 46{,}26$; $SD = 12{,}56$). Es wurden Lehrkräfte der BOS gewählt, da hier eine sehr heterogene Schülerschaft mit unterschiedlicher Vorbildung aus diversen Schularten unterrichtet wird, die in relativ kurzer Zeit gemeinsam zum Abitur geführt werden soll. Aus diesem Grund sollten in dieser Schulart Diagnoseprozesse vergleichsweise häufig stattfinden. Beide Schularten legen den Schwerpunkt auf die fachliche Ausbildung, weshalb diese Lehrkräfte eher wenig über Diagnosekompetenzen im Studium erfahren haben und vermutlich erst im Referendariat oder der weiteren schulischen Lehrlaufbahn damit konfrontiert wurden. Bei dieser Gruppe sollte die Lehrerfahrung eine zentrale Bedeutung für den Erwerb von Diagnosekompetenzen haben. Von den mitwirkenden Lehrkräften hatten lediglich vier eine Lehrerfahrung von unter fünf Jahren, 13 Lehrkräfte verfügten über eine Lehrerfahrung von mehr als zehn Jahren. Bis auf drei Lehrkräfte hatten alle einen Lehramtsstudiengang für gymnasiales oder berufliches Lehramt absolviert. Die Lehrkräfte ohne abgeschlossenes Lehramtsstudium verfügten zum Zeitpunkt der Befragung über eine mindestens zehnjährige Lehrerfahrung.

Mittels einer leitfragengestützten Beobachtung werden in den Mathematikstunden der teilnehmenden Lehrkräfte Situationen analysiert, die sich zur Diagnose von Fehlvorstellungen von Schülerinnen und Schülern eignen. Hierzu gehören fehlerhafte Antworten von Schülerinnen und Schülern ebenso wie Rückfragen, die wörtlich mitnotiert werden.

Nach der leitfragengestützten Beobachtung der Mathematikstunde werden die unterrichtenden Lehrkräfte zu ihrem Diagnoseprozess am Beispiel der beobachteten Situationen interviewt. Nach Fragen zum Alter, zur Lehrerfahrung und dem Studium schildern die Lehrkräfte ihr Vorgehen bei der Unterrichts-

vorbereitung dieser Stunde. Anschließend wird den Lehrkräften die konkrete Situation, in der eine Schülerin oder ein Schüler eine Verständnisfrage stellte oder eine Schülerin oder ein Schüler eine Frage der Lehrkraft falsch beantwortete, vorgestellt. Hierzu werden sie mittels offener Fragen befragt. Auf dem Leitfaden befinden sich zum Zeitpunkt der Befragung bereits jeweils Antwortkategorien, die eine direkte Einordnung der Antworten ermöglichen. Der letzte Teil des Leitfadeninterviews beinhaltet eine Vorstellung des Modells des Scientific Reasoning and Argumentation (Fischer et al. 2014) und die Einschätzung der Zweckmäßigkeit eines Diagnosemodells.

5 Ergebnisse

Alle Lehrkräfte gaben an, in der Ausbildung kein Diagnosemodell kennengelernt zu haben und kein solches zu nutzen.

> *Erste Annahme:* Lehrkräfte identifizieren potentielle Probleme ihrer Schülerinnen und Schüler im Rahmen der Unterrichtsvorbereitung.

Drei der 19 Lehrkräfte berichteten in ihren Interviews über die Unterrichtsvorbereitung unaufgefordert, dass sie bereits im Vorfeld über mögliche Schülerfehler und Fehlvorstellungen nachdachten – alle drei Lehrkräfte hatten bereits in den Vorstunden Probleme bei den Schülerinnen und Schülern bemerkt und wollten diesen aktiv begegnen. Auf Nachfrage gaben weitere 13 Lehrkräfte an, dass sie im Vorfeld durchaus Vermutungen hatten, in welchen Bereichen Schülerinnen und Schüler Schwierigkeiten haben könnten; diese Vermutungen basierten alle auf der Erfahrung der Lehrkräfte. Drei weitere Lehrkräfte bestätigten auf Nachfrage, dass sie vor der Stunde nicht über Fehlvorstellungen der Schülerinnen und Schüler nachgedacht hatten; alle drei verfügten über zehn Jahre Unterrichtserfahrung und unterrichteten am Gymnasium. Keine der Lehrkräfte gab an, Fehlvorstellungen aus der Literatur, Forschung oder der universitären Ausbildung zu dem unterrichteten Unterrichtsinhalt zu kennen.

> *Zweite Annahme:* Lehrkräfte generieren Hypothesen über die Fehlvorstellungen ihrer Schülerinnen und Schüler.

17 der 19 Lehrkräfte konnten direkt eine oder mehrere Hypothesen über mögliche Fehlvorstellungen benennen, die im direkten Zusammenhang mit dem Thema der beobachteten Unterrichtseinheit standen. Alle elf Lehrkräfte, die zu einer beobachteten Situation in der Unterrichtseinheit – speziell zu einer Rückfrage einer Schülerin oder eines Schülers – befragt wurden, hatten direkt auch eine oder mehrere Hypothesen, welche Schwierigkeiten oder Probleme bei

der Schülerin oder dem Schüler vorlagen. Auch alle Lehrkräfte, die zu einer falschen Antwort ihrer Schülerinnen und Schüler befragt wurden ($n = 10$), hatten direkt eine Hypothese zu den Problemen und Schwierigkeiten der Lernenden.

> *Dritte Annahme:* Lehrkräfte testen ihre Hypothesen, formulieren Schlussfolgerungen und kommunizieren diese an ihre Schülerinnen und Schüler.

Bei der offenen Frage, inwiefern die Lehrkräfte sich bereits bei der Unterrichtsvorbereitung Gedanken zum weiteren Vorgehen bei auftretenden Problemen und Schwierigkeiten gemacht hatten und welche Herangehensweisen vorbereitet waren, gaben sechs der 19 Lehrkräfte an, bereits Aufgaben erstellt zu haben (Artefaktkonstruktion), diese Aufgaben einzusetzen (Evidenzgenerierung) und mit den Schülerinnen und Schülern zu besprechen. Letzteres beinhaltete sowohl eine Evidenzevaluation als auch ein Ziehen von Schlussfolgerungen und eine Kommunikation des Prozesses sowie des Ergebnisses an die Schülerin oder den Schüler. 15 der 19 Lehrkräfte sahen dabei die Kommunikation über das Problem bzw. die Schwierigkeit als essentiell an, auch ohne ihre Hypothese mit Evidenzen vorher validiert zu haben.

> *Vierte Annahme:* Lehrkräfte empfinden es als hilfreich, auf der Grundlage eines Diagnoseprozessmodells, wie das des Scientific Reasoning and Argumentation (Fischer et al. 2014), zu arbeiten.

Alle befragten Lehrkräfte gaben an, intuitiv zu diagnostizieren. Keine Lehrkraft kannte bislang ein Modell, das bei der Diagnose von Schülerleistungen einen Prozess vorgab. Nach der Vorstellung des Modells empfanden jedoch 14 der 19 Lehrkräfte die epistemisch-diagnostischen Aktivitäten sinnvoll. Beispielsweise sahen sie besonders bei schüchternen Schülerinnen und Schülern das Problem, dass diese nicht explizit nachfragen würden. Mit dem Modell wäre es ihrer Ansicht nach möglich, auch diese Schülerinnen und Schüler vollständig zu diagnostizieren. Die Lehrkräfte gaben an, die Schritte des Modells bereits intuitiv zu nutzen, auch wenn die Abstraktheit und Wissenschaftlichkeit des Modells angemerkt wurde. Als eher nebensächliche epistemisch-diagnostische Aktivitäten wurden Artefaktkonstruktion, Evidenzgenerierung und Evidenzevaluierung von fünf der 19 Lehrkräfte – alle an einer BOS tätig – eingeschätzt; zwei der Lehrkräfte (am Gymnasium) sahen die Hypothesengenerierung als eher nebensächlich an.

Obwohl die Lehrkräfte weder das Modell der epistemisch-diagnostischen Aktivitäten noch ein vergleichbares Diagnoseprozessmodell kannten, passten ihre Aussagen und Beschreibungen zu den epistemisch-diagnostischen Aktivitäten des Modells des Scientific Reasoning and Argumentation (Fischer et al. 2014). Eine explizite Vermittlung dieses oder eines vergleichbaren Diagnoseprozessmodells erscheint allerdings sinnvoll, da die Lehrkräfte so zum Diagnostizieren angeregt werden und ihr intuitives Wissen theoretisch fundiert wird.

Die Ergebnisse zeigen unter anderem, dass die befragten Lehrkräfte eher gezielt über Fehler, Fehlvorstellungen und Schwierigkeiten während der Unterrichtsvorbereitung nachdenken, wenn sie in der Vorstunde mit Problemen von Schülerinnen und Schülern konfrontiert wurden. Das heißt im Umkehrschluss, dass die befragten Lehrkräfte über mögliche Fehler und Fehlvorstellungen in Bezug auf den Unterrichtsinhalt nicht allgemein reflektieren, sondern einen konkreten Anlass benötigen. Des Weiteren basieren die Überlegungen zu Fehlern und Fehlvorstellungen wesentlich auf eigener Erfahrung. Keine der Lehrkräfte gab an, Fehlvorstellungen aus der Literatur, Forschung oder der universitären Ausbildung zu kennen. Damit scheinen Theorien und Ergebnisse mathematikdidaktischer Forschung bei den Lehrkräften eher nicht präsent zu sein. Eine verstärkte Beschäftigung mit der Mathematikdidaktik im Studium und eine direkte Anwendung wären daher wünschenswert.

Drei Lehrkräfte an Gymnasien gaben an, keine Probleme in der Unterrichtsvorbereitung zu identifizieren. Aufgrund der kleinen Stichprobe kann dieses Ergebnis nicht verallgemeinert werden. Interessant wäre allerdings die Beantwortung der Frage, ob sich Lehrkräfte an Gymnasien hier prinzipiell von denen an anderen Schulformen unterscheiden. Die starke Betonung der Fachkompetenz könnte eine Ursache dafür sein, dass Überlegungen zu Fehlern und Fehlvorstellungen im Vorfeld einer Unterrichtsstunde eine eher geringe Rolle spielen. Die Fragestellung könnte im Rahmen einer größeren Befragung geprüft werden, um das Wissen über Diagnoseprozesse zu vertiefen.

Die Lehrkräfte konnten schnell auf mögliche Hypothesen zu den Schwierigkeiten und Problemen der Schülerinnen und Schüler zurückgreifen und Hypothesen über die Fehlvorstellungen fragender und falsch antwortender Schülerinnen und Schüler generieren. Somit hat die Stichprobe bereits gezeigt, dass die Hypothesengenerierung als eine wichtige (vgl. Herppich et al. 2017; Fischer et al. 2014; Philipp 2018) – wenn auch selten erwähnte – Aktivität gesehen wird. Dass zwei Lehrkräfte die Hypothesenbildung als eine unnötige Aktivität ansehen, zeigt, dass eine Verwendung des Diagnoseprozessmodels sinnvoll wäre, um den Lehrkräften die Bedeutung von Hypothesen im Zusammenhang mit einer zielgerichteten Evidenzgenerierung und Evidenzevaluierung aufzuzeigen.

Als eine der wichtigsten epistemisch-diagnostischen Aktivitäten sehen die teilnehmenden Lehrkräfte das Kommunizieren ihrer Schlussfolgerungen an. Weniger als die Hälfte der Lehrkräfte hat Hypothesen vorher mit Evidenzen überprüfen wollen. Dies ist möglicherweise im Alltag einem Zeitproblem geschuldet. Es kann in der Praxis ggf. als vertretbar angesehen werden, da Lehrkräfte kontinuierlich diagnostizieren und ihre Diagnose der schnell anpassbaren Förderung dient (vgl. Karst/Klug/Ufer 2017).

Die Notwendigkeit einer Förderung der Vermittlung von Diagnosekompetenzen im Studium wurde von den Lehrkräften klar kommuniziert. Zwar eignen sich Lehrkräfte das Wissen um Diagnoseprozesse im Laufe der Berufstätigkeit an; jedoch sähen sie es als hilfreich an, einen Diagnoseprozess bereits im Studium kennenzulernen und diesen im Unterricht nutzen zu können.

Auch wenn die Teilnehmerzahl an dieser Studie eher gering war, legen die Ergebnisse nahe, dass die epistemisch-diagnostischen Aktivitäten des Modells des Scientific Reasoning and Argumentation (Fischer et al. 2014) den Diagnoseprozess bei Mathematiklehrkräften gut beschreiben und abbilden. Es würde sich daher anbieten, ein solches Modell bereits im Studium zu vermitteln. Auf diese Weise könnte nicht nur die Diagnose von Lernschwierigkeiten, Fehlern und Fehlvorstellungen als wichtiger Aspekt des Professionswissens von Lehrkräften betont, sondern auch gezeigt werden, dass gerade ein theoriebasierter Diagnoseprozess zu effizienten und akkuraten Diagnosen führt und somit die Diagnosekompetenz der Lehrkräfte unterstützt.

Literatur

Abs, Hermann Josef (2006): Zur Bildung diagnostischer Kompetenz in der zweiten Phase der Lehrerbildung. In: Allemann-Ghionda, Cristina/Terhart, Ewald (Hrsg.): Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern. Ausbildung und Beruf. Weinheim und Basel: Beltz. S. 217–234.

Abs, Hermann Josef (2007): Überlegungen zur Modellierung diagnostischer Kompetenz bei Lehrerinnen und Lehrern. In: Lüders, Manfred (Hrsg.): Forschung zur Lehrerbildung. Münster, München u. a.: Waxmann Verlag GmbH. S. 63–84.

Fischer, Frank/Kollar, Ingo/Ufer, Stefan/Sodian, Beate/Hussmann, Heinrich/Pekrun, Reinhard/Neuhaus, Birgit J./Dorner, Birgit/Pankofer, Sabine/Fischer, Martin/Strijbos, Jan-Willem/Heene, Moritz/Eberle, Julia (2014): Scientific Reasoning and Argumentation. Advancing an Interdisciplinary Research Agenda in Education. In: Frontline Learning Research 2, H. 3, S. 28–45 (auch online unter http://journals.sfu.ca/flr/flr/index.php/journal/article/download/96/181).

Fowler, Mary Anne (1980): Diagnostic Teaching for Elementary School Mathematics. In: The Arithmetic Teacher 27, H. 7, S. 34–37 (auch online unter http://www.jstor.org/stable/41189603).

Ginsburg, Herbert (1977): The psychology of arithmetic thinking. In: Journal of Children's Mathematical Behavior 1, H. 4, S. 1–89.

Hattie, John (2010): Visible learning. A synthesis of over 800 meta-analyses relating to achievement. Reprinted. London: Routledge.

Heitzmann, Nicole/Hetmanek, Andreas/Wecker, Christof/Fischer, Martin R./Ufer, Stefan/Schmidmaier, Ralf/Neuhaus, Birgit/Seidel, Tina/Siebeck, Matthias/Stürmer, Kathleen/Obersteiner, Andreas/Reiss, Kristina/Girwidz, Raimund/Opitz, Ansgar/Fischer, Frank (2019): Facilitating Diagnostic Competences in Simulations: A Conceptual Framework and a Research Agenda for Medical and Teacher Education. In: Frontline Learning Research, in review.

Herppich, Stephanie/Praetorius, Anna-Katharina/Hetmanek, Andreas/Glogger-Frey, Inga/Ufer, Stefan/Leutner, Detlev/Behrmann, Lars/Böhmer, Ines/Böhmer, Matthias/Förster, Nathalie/Kaiser, Johanna/Karing, Constance/Karst, Karina/Klug, Julia/Ohle, Annika/Südkamp, Anna (2017): Ein Arbeitsmodell für die empirische Erforschung der diagnostischen Kompetenz von Lehrkräften. In: Südkamp, Anna/Praetorius, Anna-Katharina (Hrsg.): Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen. Münster, New York: Waxmann Verlag GmbH. S. 75–94.

Hesse, Ingrid (2017): Diagnostik für Lehrkräfte. 3. Auflage. Opladen, Toronto: Verlag Barbara Budrich.

Karst, Karina/Klug, Julia/Ufer, Stefan (2017): Strukturierung diagnostischer Situationen im inner- und außerunterrichtlichen Handeln von Lehrkräften. In: Südkamp, Anna/Praetorius, Anna-Katharina (Hrsg.): Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen. Münster, New York: Waxmann Verlag GmbH. S. 102–114.

Keding, Malte (2018): Studie zu den diagnostischen Kompetenzen von Gymnasiallehrkräften im Hinblick auf die epistemisch-diagnostischen Aktivitäten. Master's Thesis (unveröffentlicht). München.

Kuntze, Sebastian/Heinze, Aiso/Reiss, Kristina (2008): Vorstellungen von Mathematiklehrkräften zum Umgang mit Fehlern im Unterrichtsgespräch. In: Journal für Mathematik-Didaktik 29, 3–4, S. 199–222.

Leinhardt, Gaea/Zaslavsky, O. B. (1990): Functions, Graphs, and Graphing: Tasks, Learning, and Teaching (auch online unter https://pdfs.semanticscholar.org/daad/ea88f668bbe754fa2f5d98fe40fced8a0e9a.pdf).

Leuders, Timo/Philipp, Kathleen/Leuders, Juliane (Hrsg.) (2018): Diagnostic Competence of Mathematics Teachers. Unpacking a Complex Construct in Teacher Education and Teacher Practice. Cham: Springer International Publishing.

Oser, Fritz (1997): Standards in der Lehrerbildung. Teil 1. Berufliche Kompetenzen, die hohen Qualitätsmerkmalen entsprechen. In: Beiträge zur Lehrerbildung 15, H. 1, S. 26–37.

Oser, Fritz (2005): Lernen ist schmerzhaft. Weinheim, Basel: Beltz.

Oser, Fritz/Hascher, Tina/Spychiger, Maria (1999): Lernen aus Fehlern. Zur Psychologie des „negativen" Wissens. In: Althof, Wolfgang/Interdisziplinäres Symposium aus Anlaß des 60. Geburtstags von Fritz Oser (Hrsg.): Fehlerwelten. Opladen: Leske Budrich.

Philipp, Kathleen (2018): Diagnostic Competences of Mathematics Teachers with a View to Processes and Knowledge Resources. In: Leuders, Timo/Philipp, Kathleen/Leuders, Juliane (Hrsg.): Diagnostic Competence of Mathematics Teachers. Unpacking a Complex Construct in Teacher Education and Teacher Practice. Cham: Springer International Publishing. S. 109–128.

Radatz, Hendrik (1980): Fehleranalysen im Mathematikunterricht. Braunschweig: Vieweg+Teubner Verlag.

Rollett, Brigitte (1999): Auf dem Weg zu einer Fehlerkultur. Anmerkungen zur Fehlertheorie von Fritz Oser. In: Althof, Wolfgang/Interdisziplinäres Symposium aus Anlaß des 60. Geburtstags von Fritz Oser (Hrsg.): Fehlerwelten. Opladen: Leske Budrich. S. 71–87.

Steinleitner, Samira (2018): Studie zu den diagnostischen Kompetenzen von Lehrkräften an den beruflichen Schulen FOS/BOS im Hinblick auf die epistemisch-diagnostischen Aktivitäten. Master's Thesis (unveröffentlicht). München.

Südkamp, Anna/Praetorius, Anna-Katharina (Hrsg.) (2017): Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen. Münster, New York: Waxmann Verlag GmbH.

Weinert, Franz Emanuel/Schrader, Friedrich-Wilhelm/Helmke, Andreas (1990): Educational expertise. Closing the gap between educational research and classroom practice. In: School Psychology International 11, H. 3, S. 163–180.

Welche Planungssituationen und Planungsmittel beschreiben Sonderpädagoginnen, Sonderpädagogen und Fachlehrkräfte, die im inklusiven Unterricht kooperieren?

Ausgewählte Ergebnisse einer Interviewstudie

Silvia Greiten, Franziska Müller, Matthias Trautmann und Daniel Mays

Abstract

Der Beitrag stellt in Form fokussierender Fallbeschreibungen Ergebnisse aus Interviews mit vier im inklusiven Unterricht kooperierenden Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräften vor. Planungssituationen werden durch die Lehrkräfte innerhalb der schulsystemischen Bedingungen individuell und teilweise ad hoc gestaltet. Sonderpädagoginnen und -pädagogen greifen zur Planung von Unterricht für eine Lerngruppe auf ein breites Spektrum an Planungsmitteln zurück, je nach Förderschwerpunkten bis hin zu Materialien aus mehreren unterschiedlichen Jahrgängen. Regelschullehrkräfte orientieren sich eher jahrgangsbezogen. Die individuellen sonderpädagogischen Förderschwerpunkte werden kaum differenziert beachtet.

1 Problemstellung und theoretischer Hintergrund

Inklusiver Unterricht in der Sekundarstufe I ist bislang kaum untersucht und theoretisch modelliert, obwohl dort durch das Fachlehrerprinzip und eingeschränkte Kooperationsmöglichkeiten zwischen Fachlehrkräften und Sonderpädagoginnen und -pädagogen besondere Probleme zu erwarten sind (Heinrich et al. 2013; Greiten et al. 2016). So ist derzeit unklar, wie sich Kompetenzen für inklusiven Unterricht gestalten müssen, auch da entsprechende Forschung, insbesondere für die Sekundarstufe I, fast vollständig fehlt (vgl. Heinrich et al. 2013; s. auch Übersicht Döbert/Weißhaupt 2013; Moser 2016).

Als eine wesentliche Hürde der Qualifizierung für inklusiven Unterricht erweist sich die Unterrichtsplanung und -gestaltung des Fachunterrichts in der Sekundarstufe I. Obwohl bereits eine Vielzahl didaktischer Modelle zur Verfügung steht, erscheint keines zur Planung inklusiven Unterrichts in der Sekundarstufe geeignet, zum einen, da sie nicht erlauben, zugleich Bildungsstandards als auch den förderbedarfsspezifischen Zielen gerecht zu werden. Zum anderen, da didaktisch-methodische Überlegungen sich vorrangig am Inhalt und dem Wissen über die Lerngruppe als Klasse ausrichten und weniger an Lernvoraussetzungen einzelner Schülerinnen und Schülern. Obwohl im inklusiven Unterricht Formen individualisierten Lernens, verbunden mit Binnendifferenzierung und Förderdiagnostik, erforderlich sind, lassen sich diese mithilfe klassischer didaktischer Modelle kaum umsetzen.

Weiterhin ist zu konstatieren, dass der Prozess der Unterrichtsplanung von Lehrkräften bislang wenig untersucht ist (vgl. Übersicht in Wernke/Zierer 2017), erst recht nicht die Planung durch Sonderpädagoginnen und -pädagogen; dies gilt noch mehr für die kooperative Unterrichtsplanung durch die beiden Professionen. Mit Blick auf die Kooperation von Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräften sind eine Reihe Forschungsdesiderata zu benennen. So wird in der wissenschaftlichen Debatte zwar die Kooperation von Sonderpädagoginnen und -pädagogen und Fachlehrkräften als eine wesentliche Gelingensbedingung für die Entwicklung inklusiven Unterrichts identifiziert. Allerdings ist noch nicht hinreichend geklärt, welche Aufgaben von Lehrkräften (Fachlehrkräften und Sonderpädagoginnen bzw. -pädagogen) überhaupt (zusätzlich) in einem inklusiven Bildungssystem übernommen werden müssen und welche Aufgaben davon arbeitsteilig oder in Kooperation zu bewältigen sind (Melzer et al. 2015). Während außerunterrichtliche Aufgaben teilweise sehr gut auch arbeitsteilig und individuell bearbeitet werden können, erscheinen unterrichtsbezogene Aufgaben, sofern diese auch tatsächlich auf eine gemeinsame Unterrichtung abzielen und nicht das stundenweise „Herausziehen" von Schülerinnen und Schülern mit Förderbedarf aus dem Unterricht praktiziert wird, oftmals nur durch Kooperation – durch gemeinsame Absprachen – bearbeitbar.

Studien zur Kooperation von Lehrkräften beschreiben zumeist Kooperationsformen, Niveaustufen oder Prozesse (vgl. Übersichten in Massenkeil/Rothland 2016 und Trumpa/Franz/Greiten 2016). Zu den bekannteren Modellen zählen die Kooperationsformen von Austausch, arbeitsteiliger Kooperation und Ko-Konstruktion (Gräsel/Fußangel/Pröbstel 2006). Während Studien zu Kooperationsformen von Lehrkräften in ausreichendem Maße vorliegen, trifft dies für Studien zur Kooperation von Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräften mit Bezug zur inklusiven Unterrichtsplanung und -durchführung nicht zu. Sie weisen u. a. auf Unterschiede der Professionen in der Unterrichtsplanung hin (vgl. Weiß/Kollmannsberger/Kiel 2013). Gebhard et al.

(2014, S. 17) merken an, dass Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräfte an einigen Schulen noch keine Teampartner „auf Augenhöhe“ sind. Hildebrandt et al. (2017) betonen, dass Teamteaching, das auf einer gemeinsamen Unterrichtsplanung basiert und Absprachen auch während des Unterrichts erfordert, hohe Potenziale zur Professionalisierung bietet und als eine der effektivsten Kooperationsformen gilt. Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräfte könnten dadurch bereits in der Unterrichtsplanung voneinander profitieren, Rollenklärungen kommunizieren und sich als Team mit der Zuständigkeit für die Klasse entwickeln. Damit aber die sonderpädagogische Perspektive(n) auf Unterricht und individuelle Förderung, sonderpädagogische Theoriekonzepte und sonderpädagogisches Handlungswissen in das Regelschulsystem gelangen kann, bedarf es der Kooperation und explizit einer professionalisierten Kommunikation zwischen Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräften. Die Kommunikation muss sich dabei vor allem auf individuelle Lernvoraussetzungen von Schülerinnen und Schüler und entsprechende Konsequenzen für die Unterrichtsplanung beziehen (Greiten et al. 2016).

In der vom BMBF geförderten IKU-Studie (Interprofessionelle kooperative Unterrichtsreihenplanung in der Sekundarstufe I) (Greiten/Trautmann/Mays 2019) werden u. a. Konzepte zur Planung inklusiven Unterrichts und Bedingungen der Kooperationen untersucht. Dazu werden Tandems von Sonderpädagoginnen und -pädagogen und Fachlehrkräften über die Dauer von drei Jahren begleitet. Ziele sind die Entwicklung eines qualitativ-empirischen Modells der Planungskompetenz für inklusiven Unterricht und eines entsprechenden Qualifizierungskonzepts für Fachlehrkräfte und Sonderpädagoginnen bzw. -pädagogen der Sekundarstufe I.

Bedingungen der kooperativen Planung inklusiven Unterrichts sind u. a. die Planungssituationen und von den Lehrkräften genutzte Planungsmittel. Zu Planungssituationen zeigt eine Teilstudie des IKU-Projektes, dass die Kooperation zur Unterrichtsplanung meist wenig strukturiert und von den Fachlehrkräften ausgehend erfolgt (Publ. i.V.). Das bedeutet, dass die Kooperationsform meist nicht über den Austausch in Form von Absprachen, die dann von den kooperierenden Partnerinnen individuell weiter ausdifferenziert werden, hinausgeht. Zu von den kooperierenden Lehrkräften verwendeten Planungsmitteln gibt es im deutschsprachigen Raum bislang keine Untersuchungen.

Nachfolgend werden Ergebnisse einer Interviewstudie zu Planungssituationen und Planungsmitteln kooperierender Sonderpädagoginnen bzw. -pädagogen und Fachlehrkräfte vorgestellt.

2 Forschungsdesign

Zur Datenerhebung in dieser Teilstudie wurden je fünf Einzelinterviews mit Fachlehrkräften und Sonderpädagoginnen bzw. -pädagogen, die als Tandems kooperieren, geführt. Bei den Interviews handelt es sich um Experteninterviews auf der Basis eines Leitfadens (Gläser/Laudel 2010), der den Interviewten offene Antworten ermöglicht, in der Auswertung die Rekonstruktion individueller Planungskonzepte zulässt und dennoch Vergleichbarkeit gewährleistet. Dies ist für das IKU-Projekt insofern wichtig, als die Konzepte der Unterrichtsplanung sowohl individuell und professionsspezifisch als auch hinsichtlich der Auswirkungen auf die Kooperation untersucht werden. Der Leitfaden umfasst u. a. die Themenfelder Unterrichtsplanung, Kooperation, Fachunterricht und Bedeutung des Sonderpädagogischen Förderbedarfs.

Die für diesen Beitrag ausgewählten Ergebnisse fokussieren die Forschungsfragen:

- Welche Planungssituationen beschreiben im inklusiven Unterricht kooperierende Fachlehrkräfte und Sonderpädagoginnen und -pädagogen?
- Über welche Planungsmittel geben im inklusiven Unterricht kooperierende Fachlehrkräfte und Sonderpädagoginnen und -pädagogen Auskunft?

Zur Untersuchung dieser Forschungsfragen wurden die Transkripte zu den Interviews mit der Qualitativen Inhaltsanalyse unter Anwendung der Technik der inhaltlichen Strukturierung (Kuckartz 2012; Mayring 2015; Rädiker/Kuckartz 2019) ausgewertet. Die thematischen Kategorien *Planungssituationen* und *Planungsmittel* entstammen vorrangig den Leitfragen und sind somit deduktiv entwickelt. Als Mindestbedingung für eine Kodiereinheit wurde ein Satz definiert, als Analyseeinheit galt das gesamte Transkript (Mayring 2015). Über die auf die Leitfragen folgenden Interviewpassagen hinaus wurden im Transkript nach den Kodierregeln auch weitere Inhalte kodiert. In Forschungswerkstätten entwickelten vier Auswerterinnen den Kodierleitfaden auf der Basis von vier Transkripten und validierten diesen zunächst konsensuell (Rädiker/Kuckartz 2019). Dem folgte die Erprobung des Kodierleitfadens mit der Software MAXQDA18 und schließlich die Kodierung des Datensatzes durch je zwei Auswerterinnen, fokussiert auf ausgewählte Kategorien und unterschiedliche Fragestellungen.

Die Übersicht in Tabelle 1 dokumentiert den Auszug aus dem Kodierleitfaden.

Tabelle 1: Auszug aus dem Kodierleitfaden

Kategorie	Definition und Kodierregel
Planungssituation	Situationen, in denen Lehrkräfte Unterricht planen; wird codiert, wenn Interviewte konkrete, individuelle oder kooperative Planungssituationen beschreiben (Bsp. Ort, Zeit, Kommunikationsmedien, Häufigkeiten, Zeitintervalle, zeitlicher Vorlauf zum tatsächlichen Unterricht, Kooperationspartner, Verlauf der kooperativen Planungssituation)
Planungsmittel	Konkrete, materielle Planungsmittel, die während der Planung von Unterricht verwendet werden (Bsp. PC, Software, Schulbücher, Curriculum usw.); wird codiert, wenn Interviewte mit Bezug zu ihrer Unterrichtsplanung auf konkrete Planungsmittel verweisen

Für diesen Beitrag wurde die Analyse der *Planungssituationen* und *Planungsmittel* zu zwei Tandems ausgewählt. Die beiden Tandems unterrichten an einer Hauptschule in Nordrhein-Westfalen und stehen dabei im Kontext derselben schulsystemischen Bedingungen in Bezug auf die Schule, an der sie arbeiten. Die Interviewten werden gekennzeichnet mit FL für Fachlehrkraft und SP für Sonderpädagogin bzw. -pädagoge, ergänzt um Buchstaben zur Personenunterscheidung. Das Kürzel SFB steht für sonderpädagogischen Förderbedarf.

Relevant für die Analyse sind die Begriffe *zieldifferent* und *zielgleich*, da Lehrkräfte damit einhergehende Lern- und Leistungssituationen kontextualisieren: Die Begriffe beziehen sich auf die Abschlussorientierung zum Ende der Sekundarstufe I der jeweiligen Förderbedarfe. Schülerinnen und Schüler mit dem Förderschwerpunkt Lernen und Geistige Entwicklung werden in der Regel zieldifferent unterrichtet, unterliegen damit nicht den an der jeweiligen Schulform geltenden Leistungs- und Qualifikationsanforderungen. Dies bedeutet für die Schulpraxis, dass sich Lehrkräfte sowohl hinsichtlich der Unterrichtsziele, der im Unterricht angebotenen Materialien und Aufgabenstellungen und auch bezüglich der Leistungsanforderungen individuell an den jeweiligen Schülerinnen und Schüler orientieren müssen. Schülerinnen und Schüler mit den Förderschwerpunkten Sozial-Emotionale Entwicklung, Sprache, Körperliche Entwicklung werden, sofern die sonderpädagogische Diagnostik nichts anderes ausweist, zielgleich und damit an den curricularen Vorgaben der Schulform, an der sie beschult werden, unterrichtet. Differenzierungen der Unterrichtsinhalte, -materialien und Leistungsanforderungen sind damit nicht zwingend notwendig.

Im Folgenden werden die Ergebnisse als fokussierte Fallbeschreibung vorgestellt. Zur Dokumentation werden die Belege zu den Transkripten in Klammern eingefügt.

Tandem FLO_SPW

Die Fachlehrerin Frau O. (FLO) orientiert ihre Unterrichtsplanung an den Regelschulschülerinnen und -schüler sowie an den zielgleich zu unterrichtenden Schülerinnen und Schüler mit SFB. Ihre Tandempartnerin Frau H. konzentriert sich als Sonderpädagogin (SPH) auf die zieldifferent zu unterrichtenden Schülerinnen und Schüler.

Die Fachlehrkraft und die Sonderpädagogin nennen auf die Frage zur *Planungssituation* das Lehrerzimmer und die einstündige Mittagspause (FLO 183, 215, 304; SPH 70, 285). Die Fachlehrerin gibt an, gelegentlich auch morgens eine halbe Stunde früher in die Schule zu kommen, oder mit ihrer Tandempartnerin die halbstündige Frühstückspause zu nutzen (FLO 215). Dies geschehe dann meist spontan (FLO 293). Sie planen Unterricht auch während gemeinsamer Unterrichtszeiten in Klassenleiterstunden, Förderbandstunden und Arbeitsphasen von Schüler*innen (FLO 305, 313, 317; SPH 87). Planungssituationen über Mails, Telefonate und weitere soziale Medien nutzen die Tandempartnerinnen selten und diese dienen dann eher für kurzfristige Absprachen oder zur Übermittlung von Informationen im Krankheitsfall (FLO 215, 305; SPH 293).

Die Fachlehrerin bezeichnet die Kommunikation während der Planung als ‚Unterhaltung' über „was wir jetzt wollen" (FLO 181). Sie teilt ihre Vorstellungen über die inhaltliche und organisatorische Gestaltung – ausgerichtet auf die Regelschulschülerinnen und -schüler – mit, und intendiert, dass die Sonderpädagogin „sich dann schon mal auf die Suche macht nach Sachen, die sie mit ihren GLs [Anm. Schülerinnen und Schüler mit SFB] dazu machen kann" (FLO 181).

Das Tandem praktiziert eine Grobplanung mit inhaltlicher Ausrichtung des Unterrichtsvorhabens für die kommenden zwei bis drei Wochen. Die Konkretisierung der Einzelstunden erfolgt dann meist sukzessiv in den genutzten Planungszeiten der jeweiligen Woche, „da unterhalten wir uns eigentlich öfter mal drüber […] fünf Minuten vorher ist es in der Regel nicht" (FLO 181). Es finden auch Absprachen zu getrenntem Unterricht statt „So jetzt wupps. Ich mach das und ich die mit und fertig" (FLO 305), was dazu führt, dass dieser in der Verantwortung der jeweiligen Lehrkraft liegt und damit ohne weitere Kooperation geplant wird. Die Sonderpädagogin betont, dass sie für die Abschlussklasse 10 grundsätzlich alleine plane und unterrichte (SPH 70), d. h. hier werden die Schülerinnen und Schüler mit SFB separiert.

Beide Lehrkräfte äußern sich zu den verwendeten *Planungsmitteln* zur Vorbereitung auf den inklusiven Unterricht. Frau O. verweist zunächst auf den Lehrplan, der die curricularen Anforderungen für das Unterrichtsfach vorgibt

(FLO 81, 265). Hierzu ist anzumerken, dass es sich um den Lehrplan für den Regelschulunterricht handelt. Darüber hinaus spricht sie von „Material, das wir […] in der Schule haben“ (FLO 81), meint damit Unterrichtsmaterialien, die nach Fächern sortiert für alle Kolleg*innen im Schulgebäude zur Verfügung stehen. Für Material zu inklusivem Unterricht verweist sie auf den „GL-Schrank“, in dem „tolle Sachen“ drin seien (FLO 99). GL steht für Gemeinsames Lernen, für den GL-Schrank bedeutet dies aber, dass sich dort Material zu verschiedenen sonderpädagogischen Förderbedarfen befindet. Zudem gibt sie an, Verlagsmaterialien, Internetforen und „privates Material“ zu nutzen (FLO 93, 95). Zu Schulbüchern äußert sie sich kritisch: Sie hätten keine gute Qualität, sie nutze sie als „Aufgabenfundus“ (FLO 383). Frau H. bemängelt, dass in der Schule zu wenig Material für Schüler*innen mit SFB vorhanden sei: „(…) Das haben wir ja gar nicht. Wir haben das Zeug ja gar nicht hier. Wir könnens ja auch gar nicht so einfach bestellen. Aber jetzt wenn man wirklich frei plant, auch für die GLs, wärs schon toll, wenn da auch nen größerer Materialfundus da wäre.“ (FLO 435)

Die Tandempartnerin Frau SPH gibt an, bei der Unterrichtsplanung die Zeugnisse und die Inhalte der Förderpläne zu sichten und dies vor allem für die Kernfächer: „Erst mal schau ich mir die Zeugnisse an. Die Förderpläne. In den Förderplänen ist ja verpflichtend festgelegt, zu schreiben was kann das Kind in Mathe, in Deutsch, in Englisch“ (SPH 108). Für die konkretisierte Unterrichtsplanung nutzt die Sonderpädagogin vorrangig Lehrwerke der Regelschule sowie weitere Lehrwerke (SPH 345), „orientier ich mich an den Büchern“ (SPH 114), „da halte ich mich erstmal an die Bücher“ (SPH 209) und an „Übungshefte“ u. a. für Mathematik (SPH 114). Lehrerhandbücher nennt sie als unterstützend in der Planung (SPH 351). Ebenso wie die Fachlehrerin spricht sie auch von „privatem“ Material (SPH 355).

Tandem FLU_SPL

Die Fachlehrerin Frau U. (FLU) plant den Unterricht vorrangig für die Regelschulschülerinnen und -schüler, die Sonderpädagogin Frau L. (SPL) für die zieldifferent zu unterrichtenden Schülerinnen und Schüler. Sie geben an, gemeinsam Unterricht für die zielgleich zu unterrichtenden Schülerinnen und Schüler zu planen.

Zu *Planungssituationen* äußert sich die Fachlehrerin kaum, nennt dazu Pausen, soziale Medien und E-Mails (FLU 251). Sie beschreibt die Kommunikation während der Planung als kurze Absprachen in der Pause (FLU 255). Die Sonderpädagogin äußert sich dazu: „In der Regel unterhalten wir uns kurz in der Pause darüber, was ansteht und dann macht jeder für sich seine Unterrichtsplanung zuhause […] das sind wirklich oft äh (.) ja fast Tür- und Angelgespräche, wo man mal kurz überlegt: was hast du, was hab ich und da können wir

uns treffen, das kann ich ergänzen." (SPL 77). Zudem gibt sie auch an, in gemeinsamen Klassenleiterstunden und spontan in Unterrichtssituationen Absprachen zu treffen (SPL 87). Ansonsten verdeutlicht sie, dass beide zuhause planen, sie selber abends spät, am Wochenende oder auch eine Woche im Voraus (SPL 19).

Als *Planungsmittel* nennen beide Tandempartnerinnen vorrangig Schulbücher für das jeweilige Fach und den jeweiligen Jahrgang (FLU 57, 331, 333; SPL 19, 31, 109). Sie nutzen auch Schulbücher unterer Jahrgänge (FLU 339; SPS 27), die Fachlehrkraft eher mit der Intention der Wiederholung (FLU 339–341), die Sonderpädagogin für das individuelle Lernen von Schülerinnen und Schüler mit SFB, um Differenzierungen auf deutlich unterschiedlichen Lernniveaus zu ermöglichen (SPL 27–31). Das Schulbuch wird von der Fachlehrkraft als „Grundlage", „Grundgerüst" (FLU 333) verstanden. Darüber hinaus verwendet sie auch Internetseiten (FLU 57), Arbeitsblätter (FLU 27) sowie Materialien aus Fortbildungen (FLU 371). Dem Lehrplan der Regelschule weisen die Tandempartnerin eine orientierende Funktion zu (FLU 69, 333; SPL 51). Die Sonderpädagogin ergänzt in dieser Intention den Förderplan (SPL 51).

Die Beschreibungen der Sonderpädagogin konzentrieren sich auf die Förderschwerpunkte Geistige Entwicklung und Lernen (SPL 19, 26, 31). Im Vergleich zur Fachlehrkraft greift sie auf ein großes Spektrum an Planungsmitteln zurück, was der Transkriptausschnitt veranschaulicht: „Also in ähm Mathe ist das in der einen Lerngruppe so, dass ich da sehr – ja nicht strikt nach Buch vorgehe aber schon relativ nah am Buch arbeite und in dieser Lerngruppe auch in Englisch äh mit dem Buch fast nur arbeite. (…) in der Lerngruppe, wo ich Förderschüler mit der geistigen Entwicklung sitzen habe, da arbeite ich ganz viel äh mit dem Internet und such mir Material zusammen was ich dann halt basteln muss." (SPL 19). Trotz des Unterrichts in der Sekundarstufe I zieht sie zur Planung Schulbücher der Förderschule und Grundschule (SPL 27, 31), vor allem aus den Jahrgängen 1 bis 3, heran. Weniger nutzt sie Materialien, die Verlage zu den Regelschullehrwerken für sog. inklusiven Unterricht anbieten (SPL 117). Internet- und Verlagsmaterialen sucht sie entlang der Attribute basteln, spielerisch, handlungsorientiert (SPL 19, 23, 117). Montessori-Material (SPL 33, 41), das der Anschaulichkeit dient, ergänzt ihre Planungsmittel. Zur Unterrichtplanung recherchiert sie auch in Lehrerhandbüchern (SPL 33).

4 Diskussion und Schlussfolgerungen

Die vorgestellten Ergebnisse basieren auf der Analyse von Interviews mit je zwei Sonderpädagoginnen bzw. -pädagogen und zwei Fachlehrkräften, die als Tandems im inklusiven Unterricht kooperieren. Die vorgestellten Ergebnisse gelten zunächst für die Analyse der Interviews mit den Lehrkräften, beziehen

sich damit auf deren Aussagen und erlauben nur bedingt Schlussfolgerungen für die tatsächliche Planungspraxis, sehr wohl aber Hinweise auf die Planungskonzepte von Lehrkräften. Die Reichweite der Ergebnisse bezieht sich vorrangig auf die Einzelfälle, dennoch lassen sich weiterführende Fragestellungen ableiten sowie Unterschiede und Gemeinsamkeiten diskutieren. Die nachfolgenden Ausführungen setzen dazu Akzente.

Die Ergebnisse beziehen sich auf zwei Tandems derselben Schule und die dort gegebenen schulsystemischen Bedingungen. Die Ergebnisse dokumentieren, dass in dieser Schulpraxis keine durch Vorgaben definierten Planungszeiten zur Verfügung stehen. Die Planungssituationen handeln die Tandems in ihrem Arbeitsalltag aus, gestalten diese individuell und entwickeln für die Kooperation offenkundig Routinen.

Aus den Beschreibungen der Lehrkräfte zu den kooperativen Planungssituationen ergeben sich Fragen zur kommunikativen Gestaltung der Planungssituationen: Es fällt auf, dass beide Tandems ihre Kommunikation als Unterhaltung beschreiben. Dies signalisiert den Modus eines lockeren, sich zufällig strukturierenden Gesprächs, aus dem aber Aufgaben für individuell zu organisierende Planungsprozesse hervorgehen können. Im Unterhaltungsmodus liegt die Dominanz der Entscheidungen zu Unterrichtsinhalten eindeutig bei der Fachlehrkraft. Davon ausgehend plant die Sonderpädagogin vorrangig für Schüler*innen mit SFB, ohne diese Aufteilung in Frage zu stellen.

Die Regelschullehrkräfte aus diesem Sample generalisieren Schülerinnen und Schüler mit SFB als „GLs“ oder „Förderschüler“. Differenzierungen zu den jeweiligen Förderbedarfen (oder zu einzelnen Schülerinnen und Schülern) werden kaum vorgenommen, es dominiert ein gruppen- oder klassenbezogener Blick. Dies wirft die Frage auf, inwiefern Regelschullehrkräfte die jeweiligen Unterstützungsbedarfe der Schülerinnen und Schüler mit SFB diagnostizieren und ihre Planung inklusiven Unterrichts daran ausrichten. Auch wenn die Datengrundlage hier nur vorsichtige Schlussfolgerungen erlaubt, scheint es so zu sein, dass explizite und ausführliche Diagnosen von keiner der beteiligten Personen vorgenommen oder vorgeschlagen werden.

Als Planungsmittel nutzen Fachlehrkräfte vor allem den fachbezogenen Lehrplan. Dieser gilt für die Anforderungen des Regelschulunterrichts und entsprechende Leistungen. Für die sonderpädagogischen Förderschwerpunkte existieren ebenfalls verbindliche Richtlinien und zum Teil eigene Lehrpläne mit entsprechenden Schwerpunktsetzungen. So existieren z. B. für die Förderschwerpunkte Lernen und Geistige Entwicklung eigene Vorgaben, die eine zieldifferente Unterrichtsplanung erfordern (MSW 2019). Bei der Planung ihres Unterrichts gehen jedoch weder die teilnehmenden Sonderpädagoginnen noch die Fachlehrkräfte explizit darauf ein. Zu klären ist, ob sich dies in Förderschulen anders verhält und ob in der Konsequenz sonderpädagogische Curricula im inklusiven Prozess an Bedeutung verlieren.

Beide Sonderpädagoginnen, jedoch nicht die Fachlehrerinnen, nennen zumindest als eine Referenz für ihre Planung die Förderpläne (SPH und SPL) oder auch die Zeugnisse (SPH). Dieses Ergebnis ist bedeutsam, da Zeugnisse für zieldifferent zu unterrichtende Schülerinnen und Schüler Berichtszeugnisse sind und Förderpläne Auskunft über Ergebnisse aus der Diagnostik sowie die individuellen Förderziele und -maßnahmen geben (sollen). Die Bedeutung der Förderpläne und der Zeugnisse sollte für die Unterrichtsplanung durch Sonderpädagoginnen und -pädagogen jedoch nicht überschätzt werden, da die Verweise in den Interviews eher kursorisch ausfallen und so der Eindruck entsteht, Förderpläne dienten überwiegend der Erfüllung administrativer Vorgaben. Ungeklärt ist auch, mit welcher Intensität die Lehrkräfte diese nutzen. Dennoch ist in weiterer Forschung zu untersuchen, inwiefern Regelschullehrkräfte und Sonderpädagoginnen und -pädagogen darauf Bezug nehmen und welche Konsequenzen sich daraus für die kooperative Unterrichtsplanung ergeben.

Zu diskutieren für die Planung inklusiven Unterrichts ist die Problematik des fachfremden Unterrichtens durch Sonderpädagoginnen und -pädagogen. Beide Sonderpädagoginnen, aber nicht die Fachlehrkräfte, nennen Lehrerhandbücher als Planungsmittel. Dies könnte darauf hinweisen, dass sie aus Lehrerhandbüchern aufgrund fehlender fachwissenschaftlicher Ausbildung zu Unterricht, den sie fachfremd planen müssen, Grundlageninformationen zu den fachlichen Inhalten, Lösungsvorschläge zu Aufgaben und didaktische Konzeptionen ableiten. Hieraus ergäbe sich dann ein Auftrag zur Weiterentwicklung didaktischer Literatur.

Für ein vertieftes Verständnis der Planungssituationen und Planungsmittel sind weitere Analysen von Interviews und Planungsgesprächen mit kooperierenden Lehrkräften notwendig. Die Inhalte, die Interaktion der Gesprächspartner sowie deren Rollen und der damit verbundene Aushandlungsprozess während der Unterrichtsplanung sollten untersucht werden, um Konsequenzen für die Lehrerbildung und mögliche Interventionen zur Optimierung dieser Prozesse zu ziehen. Ein Fokus der Forschung sollte sich auch auf das Schülerbild der Lehrkräfte sowie dessen Konsequenz für die Unterrichtsplanung richten.

Literatur

Döbert, Hans/Weishaupt, Horst/Buchmann, Ulrike (Hrsg.) (2013): Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen. Münster: Waxmann.

Gebhard, Simone/Happe, Claudia/Paape, Maike/Riestenpatt, Jule/Vägler, Alexandra/Wollenweber, Kai Uwe/Castello, Armin (2014): Merkmale und Bewertung der Kooperation von Sonderpädagogen und Regelschullehrkräften in inklusiven Unterrichtssettings. In: Empirische Sonderpädagogik 6, H. 1, S. 17–32.

Gläser, Jochen/Laudel, Grit (2010): Experteninterviews und qualitative Inhaltsanalyse. 4. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

Gräsel, Cornelia/Fußangel, Kathrin/Pröbstel, Christian (2006): Lehrkräfte zur Kooperation anregen – Eine Aufgabe für Sisyphos? In: Zeitschrift für Pädagogik 52, H. 2, S. 205–219.

Greiten, Silvia/Franz, Eva-Kristina/Biederbeck, Ina (2016): Wodurch konturiert sich die sonderpädagogische Perspektive und wie gelangt sie in den inklusiven Unterricht an Regelschulen? Befunde aus Gruppendiskussionen zu Erfahrungen aus der Netzwerkarbeit von Sonderpädagoginnen, Sonderpädagogen und Regelschullehrkräften. In: Kreis, Annelies/Wick, Jeannette/Kosorok Labhart, Carmen (Hrsg.): Kooperation im Kontext schulischer Heterogenität. Münster: Waxmann, S. 143–157.

Greiten, Silvia/Trautmann, Matthias/Mays, Daniel (2019): Interprofessionelle kooperative Unterrichtsreihenplanung in der Sekundarstufe I – Skizze des Schulbegleitforschungsprojekts IKU zur Unterrichtsentwicklung im Kontext von Inklusion. In: Stechow, Elisabeth von/Hackstein, Philipp/Müller, Kirsten/Esefeld, Marie/Klocke, Barbara (Hrsg.): Inklusion im Spannungsfeld von Normalität und Diversität. Bad Heilbrunn: Klinkhardt, S. 167–174.

Heinrich, Martin/Urban, Michael/Werning, Rolf (2013): Grundlagen, Handlungsstrategien und Forschungsperspektiven für die Ausbildung und Professionalisierung von Fachkräften für inklusive Schulen. In: Döbert, Hans/Weishaupt, Horst (Hrsg.): Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen. Münster: Waxmann, S. 69–133.

Hildebrandt, Elke/Ruess, Annemarie/Stommel, Sarah/Brühlmann, Olga (2017): Planung im Teamteaching – Potentiale nutzen. In: Schweizerische Zeitschrift für Bildungswissenschaften. 39, H. 3, S. 573–591.

Kuckartz, Udo (2012): Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung. Weinheim u. a.: Beltz Juventa.

Massenkeil, Julius/Rothland, Martin (2016): Kollegiale Kooperation im Lehrerberuf. Überblick und Systematisierung aktueller Forschung. Schulpädagogik Heute 7, H. 13.

Mayring, Peter (2015): Qualitative Inhaltsanalyse: Grundlagen und Techniken. 12., überarb. Auflage. Weinheim u. a.: Beltz.

Melzer, Conny/Hillenbrand, Clemens/Sprenger, David/Hennemann, Thomas (2015): Aufgaben von Lehrkräften in inklusiven Bildungssystemen. Review internationaler Studien. In: Erziehungswissenschaft 26, H. 51, S. 61–80.

Moser, Vera (2016): Professionsforschung. In: Hedderich, Ingeborg/Biewer, Gottfried/Hollenweger, Judith/Markowetz, Rainhard (Hrsg.): Handbuch Inklusion und Sonderpädagogik. Bad Heilbrunn: Julius Klinkhardt, S. 277–287.

MSW (2019): „Richtlinien, Lehrpläne, Rahmenlehrpläne und Bildungspläne". www.schulministerium.nrw.de/docs/Schulsystem/RuL/Richtlinienliste.pdf (Abfrage 28.03.2019).

Rädiker, Stefan/Kuckartz, Udo (2019): Analyse qualitativer Daten mit MAXQDA. Text, Audio und Video. Wiesbaden: Springer Fachmedien Wiesbaden.

Trumpa, Silke/Franz, Eva-Kristina/Greiten, Silvia (2016): Forschungsbefunde zur Kooperation von Lehrkräften. Ein narratives Review. In: Die Deutsche Schule 108, H. 1, S. 80–92.

Weiß, Sabine/Kollmannsberger, Markus/Kiel, Ewald (2013): Sind Förderschullehrkräfte anders? Eine vergleichende Einschätzung von Expertinnen und Experten aus Regel- und Förderschulen. In: Empirische Sonderpädagogik, 5, H. 2, S. 167–186.

Wernke Stephan/Zierer, Klaus (Hrsg.) (2017): Die Unterrichtsplanung: Ein in Vergessenheit geratener Kompetenzbereich?! Status Quo und Perspektiven aus Sicht der empirischen Forschung. Bad Heilbrunn: Klinkhardt.

Wovon hängt ab, wie Lehrkräfte differenzieren?

Eine Re-Analyse von Fragebogen- und Video-Daten der DESI-Studie

Julia Jennek, Anna Gronostaj und Miriam Vock

Abstract

Binnendifferenzierung ist eine Möglichkeit, mit der Heterogenität von Schülerinnen und Schülern mit didaktischen Mitteln umzugehen. In dieser Studie wird untersucht, wie Lehrkraft- und Klassenmerkmale mit dem Einsatz verschiedener Maßnahmen von Binnendifferenzierung zusammenhängen. Dazu wurden Unterrichtsvideos der DESI-Studie unter dem Aspekt Binnendifferenzierung kodiert und die Ergebnisse anschließend mit Fragebogendaten der DESI-Studie in Beziehung gesetzt. Der Mädchenanteil der Klasse und die Leistungsheterogenität sind nur für die Unterstützung durch Mitlernende prädiktiv. Die Lehrkraftmerkmale Berufserfahrung und Fachstudium hängen damit zusammen, ob Lehrkräfte Differenzierung durch unterschiedliche Rollen sowie Differenzierung durch Unterstützung durch Hilfs- und Zusatzmaterial einsetzen.

1 Problemstellung

Binnendifferenzierung gilt als ein Merkmal guten Unterrichts (vgl. Helmke 2015; Meyer 2013) und wird als didaktische Möglichkeit angesehen, angemessen auf die Heterogenität der Lernenden zu reagieren (vgl. Trautmann/Wischer 2008). Lehrkräfte können verschiedene Maßnahmen zur Differenzierung einsetzen, die an die Bedürfnisse und Lernmöglichkeiten der Schülerinnen und Schüler angepasst werden können. Verschiedene Studien zeigen, dass Lehrkräfte bisher wenig differenzieren und wenn, dann durch individuelle Unterstützung (vgl. Jennek/Gronostaj/Vock 2018), zusätzliche Aufgaben (vgl. Klieme et al. 2008) oder indem sie eine Zusammenarbeit mit Mitlernenden anregen (vgl. Hugener/Krammer/Pauli 2008). Noch weniger erforscht ist, wie die ver-

schiedenen Heterogenitätsdimensionen der Klasse und Merkmale der Lehrkraft damit zusammenhängen, wie die Lehrkräfte differenzieren.

2 Theoretischer Hintergrund

Für einen angemessenen Umgang mit der Heterogenität der Schülerinnen und Schüler wurden zahlreiche Konzepte erarbeitet (vgl. Bohl/Batzel/Richey 2012; Dumont 2018). Eines dieser Konzepte ist Binnendifferenzierung, dass seit den 1970er Jahren diskutiert wird (vgl. Trautmann/Wischer 2008). Bisher besteht jedoch nur Einigkeit darüber, dass es sich bei Binnendifferenzierung um didaktische Maßnahmen handelt, die in heterogenen Lerngruppen angewandt werden, um die Lernenden entsprechend ihres Entwicklungsstandes optimal zu fordern und zu fördern (vgl. Trautmann/Wischer 2011). Ungeklärt bleibt, inwiefern die Maßnahmen adaptiv gestaltet sein müssen oder welche Aspekte individueller Förderung als „Binnendifferenzierung" zu betrachten sind (vgl. Dumont 2018).

In der vorliegenden Studie wird Binnendifferenzierung verstanden als verschiedene „didaktische Maßnahmen, die in einer Lerngruppe eingesetzt werden mit dem Ziel, das Lernangebot situationsspezifisch an die unterschiedlichen Voraussetzungen und Lernbedürfnisse einzelner Schülerinnen und Schüler oder an Gruppen von Schülerinnen und Schülern anzupassen. Dazu werden allen möglichst gut passende Lern- und Entwicklungsmöglichkeiten geboten" (vgl. Jennek/Gronostaj/Vock i.Dr.). Binnendifferenzierende Maßnahmen unterteilen wir in die Kategorien

(1) *Differenzierung durch unterschiedliche Aufgaben:* Alle Lernenden arbeiten an einem Thema mit Aufgaben unterschiedlicher Art, Niveaus oder Umfangs.
(2) *Unterschiedliche Art der Unterstützung:* Alle Lernenden erarbeiten gleiche Inhalte mit gleichen Aufgaben und erhalten bei Bedarf Unterstützung durch die Lehrkraft, Mitlernende oder Hilfs-/Zusatzmaterialien.
(3) *Unterschiedliche Rollen:* Alle oder einige Lernende nehmen im Unterrichtsgeschehen oder Gruppenarbeiten unterschiedliche Rollen ein.
(4) *Verschiedene Bearbeitungs- und Lösungswege:* Die Lernenden entscheiden selbstständig über die Bearbeitung von Aufgaben, hierzu gehört das Angebot unterschiedlicher Zugänge zum Thema.
(5) *Differenzierung durch das Produkt:* Produkte entstehen in Abhängigkeit von den Lernvoraussetzungen (z. B. Lernlandkarten) und werden individuell betrachtet.

(6) *Differenzierte Lernziele:* Die Lernenden verfolgen unterschiedliche Lernziele, z. B. als Fundamentum und Additum (für weitere Erläuterungen siehe Jennek/Gronostaj/Vock i.Dr.).

Grundlage für die Entwicklung dieser Systematisierung waren allgemein-didaktische Modelle zur Differenzierung, die, ausgehend von Klafki und Stöcker (1976), die konkrete Umsetzung von Differenzierung im Unterricht abzubilden versuchen (vgl. Dumont 2018). Diese Modelle sind mit verschiedenen Schwierigkeiten behaftet (z. B. starke Fixierung auf differenzierte Aufgaben bei Bönsch 2011), weshalb es uns notwendig erschien, eine neue Systematisierung zu entwickeln (zum Vorgehen siehe Jennek/Gronostaj/Vock i.Dr.).

Die Systematisierung bewegt sich auf Ebene der Sichtstrukturen von Unterricht (vgl. Kunter/Trautwein 2013), da es zunächst darum ging, ein Beobachtungsschema für die Analyse von Unterrichtssituationen zu entwickeln, wenngleich sich die Lernwirksamkeit von Methoden letztlich nur über die Tiefenstrukturen beurteilen lässt. Dennoch sehen wir eine an der Oberfläche beobachtete Binnendifferenzierung als einen ersten Indikator für potentiell lernwirksame Methoden. Ein Beispiel hierfür wäre die Berücksichtigung unterschiedlicher Interessen, etwa durch das Angebot verschiedener Bearbeitungs- und Lösungswege. Es lässt sich auch beobachten, ob die Lehrkraft unterschiedliche Aufgaben zuweist, was zwar an sich noch nicht bedeutet, dass die Passung der Aufgaben zu den einzelnen Schülerinnen und Schülern gelingt, aber es kann als eine erste Voraussetzung für eine bessere Passung gesehen werden.

Studien zur Häufigkeit von Binnendifferenzierung verdeutlichen, dass in der Sekundarstufe I eher wenig differenziert wird: In Befragungen von Lehrkräften zeigt sich, dass im Deutschunterricht differenzierte Aufgaben häufiger eingesetzt werden (vgl. Klieme et al. 2008), im Mathematikunterricht hingegen eher selten (vgl. Schiepe-Tiska et al. 2008). Da Auskünfte von Lehrkräften bzw. Lernenden bestimmten Verzerrungen unterliegen können (vgl. Praetorius 2014), erscheint der Zugang über Videographie geeigneter für eine authentische Abbildung des Unterrichts (vgl. Pauli/Reusser 2006), sowohl auf Ebene der Sicht- als auch der Tiefenstrukturen (vgl. Lotz/Garbiel/Lipowsky 2013). Im Rahmen einer reinen Beobachtung von Unterricht anhand von Videos kann die Häufigkeit differenzierender Maßnahmen quantitativ, mithilfe niedrig-inferenter Beobachtungen gemessen werden. Diese ermöglicht jedoch noch keine Beurteilung der Qualität der Maßnahmen. Hierfür wären andere forschungsmethodische Zugänge notwendig.

In einer Videoanalyse zu Schülerarbeitsphasen im Mathematikunterricht der Sek I zeigten Hugener und Kolleginnen (2008), dass die Lernenden in etwa der Hälfte der Arbeitszeit nach eigener Wahl zusammenarbeiten durften, andere Differenzierungsmethoden wurden selten beobachtet. Ebenfalls mit einer quantitativen Videoanalyse wurden Videos aus der DESI-Studie analysiert (vgl.

Jennek/Gronostaj/Vock 2018). Diese Daten liegen auch der vorliegenden Studie zugrunde. Insgesamt zeigte sich, dass in allen beobachteten Unterrichtsstunden differenziert wurde, hauptsächlich, indem die Lehrkräfte individuell unterstützten, gelegentlich auch die Mitlernenden, selten wurden Hilfs- oder Zusatzmaterialien zur Unterstützung eingesetzt. Differenzierung durch unterschiedliche Rollen konnte häufiger beobachtet werden, differenzierte Aufgaben jedoch kaum. Zusammenhänge zwischen dem Einsatz bestimmter Differenzierungsmaßnahmen und der Klassengröße konnten nicht nachgewiesen werden. Im Zusammenhang mit der Schulform fiel nur ein Befund auf: An Gymnasien konnte der Einsatz differenzierter Aufgaben gar nicht beobachtet werden.

Heterogenität und Klassenkomposition

Ausgangspunkt für binnendifferenzierende Maßnahmen ist die Heterogenität der Lernenden, die verschiedene Dimensionen umfasst, hervorgehoben wird häufig die Leistungsheterogenität. Im Schulalltag wird diese durch Noten dokumentiert, die allerdings keine reinen Leistungsmaße darstellen, sondern in die auch individuelle (z. B. Verhalten, vgl. Decristan/Jude 2017) oder gruppenbezogene Aspekte (z. B. Referenzgruppeneffekte, vgl. Westphal et al. 2016) einfließen. Noten bilden somit die von den Lehrkräften wahrgenommene Leistungsheterogenität der Klasse ab (vgl. Decristan/Jude 2017). Eng mit der Leistung verknüpft ist das Vorwissen der Lernenden. Vertiefte Analysen aus PISA 2003 zeigten, dass Lehrkräfte bei Klassen mit einem (im Mittel) niedriger ausgeprägten Vorwissen den Unterricht enger führten und mehr differenzierende und individualisierende Methoden nutzten (Baumert et al. 2004). Demzufolge ist anzunehmen, dass Lehrkräfte, die eine große Leistungsheterogenität in ihren Klassen wahrnehmen, verstärkt differenzieren.

Eine weitere, von Lehrkräften wahrgenommene Heterogenitätsdimension, ist das Geschlecht der Lernenden. So schreiben Lehrkräfte Mädchen höhere Selbststeuerungskompetenzen zu (vgl. Kuhl/Hannover 2012) und erwarteten von ihnen bessere soziale Fähigkeiten (vgl. Thon 2017), während sie für Jungen häufiger von deviantem Verhalten berichteten (vgl. Lupatsch/Hadjar 2011). Unterschiedliche Differenzierungsmethoden erfordern jedoch kooperatives Verhalten (z. B. Unterstützung von Mitlernenden). Möglicherweise lassen sich Lehrkräfte daher in ihrem Differenzierungsverhalten dadurch beeinflussen, wie stark die Leistungsheterogenität in der Klasse ausgeprägt ist und wie viele Mädchen (denen stereotyp kooperativeres Verhalten unterstellt wird) die Klasse besuchen.

Kompetenz von Lehrkräften zur Differenzierung

Die Umsetzung von Differenzierung stellt hohe Anforderungen an Lehrkräfte (vgl. Lipowsky/Lotz 2015). Ausgehend von deren professioneller Kompetenz (vgl. Baumert/Kunter 2006) erscheint die Facette des fachdidaktischen Wissens insbesondere für solche Differenzierungsmaßnahmen notwendig, welche eine didaktische Reduktion von Inhalten verlangt (vgl. Döhrmann/Kaiser/Blömeke 2008). Es ist davon auszugehen, dass Lehrkräfte, die in ihrem Fach ausgebildet sind (also das Fach, das sie unterrichten, studiert haben), über ein breiteres fachdidaktisches Wissen verfügen und daher mehr differenzieren.

Auch das Ausmaß der Berufserfahrung könnte dafür prädiktiv sein, wie gut es Lehrkräften gelingt, ihren Unterricht flexibel an die Bedürfnisse der Schülerinnen und Schüler anzupassen. Die Berufsbiographieforschung von Lehrkräften verdeutlicht, dass die Professionalisierung bzw. berufliche Sozialisation ein lebenslanger Prozess ist (vgl. Reh/Schelle 2006; Terhart 2011). Insbesondere das implizite Lernen durch Handeln im Beruf ermöglicht es Lehrkräften, Unterrichtssituationen in ihrer Komplexität adäquat wahrzunehmen und mit einem breiten Handlungsrepertoire zu reagieren (vgl. Kunter et al. 2011). Wir vermuten daher, dass Lehrkräfte mit höherer Berufserfahrung mehr differenzieren als unerfahrenere Lehrkräfte.

3 Fragestellungen

In der vorliegenden Studie wird untersucht, inwiefern die Heterogenitätsdimensionen Leistung und Geschlecht der Lernenden und die Lehrkraftmerkmale Berufserfahrung und Fachstudium damit zusammenhängen, welche Differenzierungsmaßnahmen eingesetzt werden.

4 Methodisches Vorgehen

Stichprobe

Die Datengrundlage bildeten $N = 112$ videografierte Unterrichtsstunden der DESI-Studie (Deutsch-Englisch-Schülerleistungen-International) aus $N = 83$ Klassen mit insgesamt $N = 81$ Lehrkräften (vgl. Helmke et al. 2008; DESI-Videos: Helmke 2014[1]). Die Videos zeigen Englischunterricht der neunten Klassen aus allen Schulformen der Sekundarstufe I mit Ausnahme von Förderschu-

1 Wir danken dem Forschungsdatenzentrum (FDZ) am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) für die Bereitstellung der Videodaten.

len (45 Gymnasien, 18 Realschulen, sechs Sekundar-/Gesamtschulen, zehn Hauptschulen, zwei ohne Angabe, Stichprobenverzerrung bedingt durch DESI-Stichprobe). Für diese Studie wurden nur Stunden zu den Themen interkulturelle Bildung, Landeskunde und Lebenswelt einbezogen, die nach Angaben der Lehrkräfte mehrheitlich als typisch für den alltäglichen Unterricht beschrieben wurden (vgl. Marquardt 2011).

Instrumente und Datenanalyse

Auf Basis der genannten Systematisierung von binnendifferenzierenden Maßnahmen (Jennek/Gronostaj/Vock i.Dr.) wurde ein mittel-inferentes Kodiermanual entwickelt, das nach Sichtung der ersten Videos erweitert wurde (deduktiv-induktives Vorgehen). Deduktiv gesetzt wurden die oben beschriebenen Kategorien der Binnendifferenzierung sowie die erste Ebene der Subkategorien. Induktiv wurden diese weiter unterteilt, um abzubilden, ob die Schülerinnen und Schüler die Differenzierung zugewiesen bekamen oder selbstständig nutzten (vgl. Jennek/Gronostaj/Vock 2018). Kodiert wurden alle auftretenden Differenzierungsmaßnahmen eventbasiert von jeweils zwei unabhängig arbeitenden Kodiererinnen, die Analyseeinheit umfasste die gesamte gefilmte Stunde. Die Beobachterübereinstimmung war mit $\kappa = .54$ bzw. $\kappa = .53$ (prozentuale Übereinstimmung: 68 % bzw. 76 %) ausreichend, nur in Einzelfällen setzten die Kodiererinnen unterschiedliche (Sub-)Codes (vgl. Jennek/Gronostaj/Vock 2018).

Für die Analysen zu Lehrkraftmerkmalen und der Heterogenität der Klasse wurde auf die Fragebogendaten aus der DESI-Studie zurückgegriffen (Lehrkräfte: Berufserfahrung, Fachstudium Englisch; Lernende: Leistungsheterogenität, Mädchenanteil der Klasse, Datensatz: Klieme 2012[2]). Die Leistungsheterogenität der Klassen wurde über die Standardabweichung der Noten zu t2 (Halbjahreszeugnis Kl. 9) abgebildet, da um diesen Zeitpunkt herum die Videoaufnahmen angefertigt wurden. Somit spiegeln die Halbjahresnoten die Einschätzung der Leistungsheterogenität zum Zeitpunkt der Datenerhebung wider. Zur Beantwortung der Forschungsfrage wurden multiple Regressionen berechnet.

2 Wir danken dem FDZ am Institut zur Qualitätsentwicklung im Bildungswesen (IQB) für die Bereitstellung der Fragebogendaten.

Tabelle 1: Multiple Regressionsanalysen zur Vorhersage von verschiedenen Differenzierungsformen

Abhängige Variable	*Gesamtanzahl von Differenzierungsmaßnahmen pro Stunde*			*Differenzierung durch unterschiedliche Rollen*			*Differenzierung durch Unterstützung durch Mitschüler/innen*			*Differenzierung durch Unterstützung durch Hilfs- und Zusatzmaterialien*		
	b	ß	*p*	*b*	ß	*p*	*b*	ß	*p*	*b*	ß	*p*
Konstante	**24.05**		**.006**	-0.32		.564	-1.98		.052	**0.86**		**.017**
Lehrkraft hat Englisch studiert[1]	2.17	.05	.705	0.58	.19	.118	0.51	.09	.460	**-0.78**	**-.37**	**.002**
Berufserfahrung der Lehrkraft[2]	0.06	.05	.693	**0.02**	**.27**	**.022**	0.00	.02	.871	0.00	.09	.440
Gymnasium[3]	-0.96	-.04	.775	-0.36	-.19	.104	0.00	.00	.995	-0.21	-.17	.144
Leistungsheterogenität[4]	-3.98	-.08	.546	-0.37	-.10	.383	**1.87**	**.27**	**.019**	0.10	.04	.725
Mädchenanteil in %	0.05	.07	.602	0.01	.13	.281	**0.03**	**.27**	**.022**	0.00	.02	.887
R^2		.015			.148			.157			.207	

Anmerkungen: $N = 76$; b = unstandardisierter Regressionskoeffizient; ß = standardisierter Regressionskoeffizient; [1]0 = nein, 1 = ja; [2]in Jahren; [3]Gymnasium = 1, Nicht-Gymnasium = 0; [4]Standardabweichung der letzten Zeugnisnoten der Klasse in Englisch (Halbjahr neunte Klasse)

5 Ergebnisse

Die Lehrkräfte waren im Mittel seit 16 Jahren im Schuldienst ($SD = 11.6$, $min = 1$, $max = 38$), 70 hatten Englisch studiert, sieben nicht (fehlend = 4). Der Mädchenanteil der Klassen betrug im Mittel 54 % ($SD = 18.6$, $min = 24$ %, $max = 100$ %, bei $N = 81$ Klassen). Die Halbjahresnoten bewegten sich um die Note „befriedigend" ($M = 3.04$, $SD = 0.94$, $N = 1782$).

In den Regressionsanalysen in Tabelle 1 zeigt sich kein Zusammenhang zwischen den fünf Prädiktoren Berufserfahrung, Studium Englisch, Schulform, Leistungsheterogenität sowie dem prozentualen Mädchenanteil der Klasse und der Gesamtzahl der eingesetzten Differenzierungsmaßnahmen. Ein positiver Zusammenhang besteht zwischen dem Einsatz von Differenzierung durch Unterstützung durch Mitlernende und der Leistungsheterogenität der Klasse (standardisierter Regressionskoeffizient $ß = .27$) sowie dem prozentualen Mädchenanteil ($ß = .27$).

Auf Seiten der Lehrkraft hängt das Fachstudium Englisch negativ mit dem Einsatz von Differenzierung durch Hilfs- und Zusatzmaterialien zusammen ($ß = -.37$). Die Berufserfahrung zeigt einen positiven Zusammenhang mit dem Einsatz von Differenzierung durch unterschiedliche Rollen ($ß = .27$). Für andere Differenzierungsmethoden zeigen sich keine systematischen Zusammenhänge. Insgesamt zeigen die Analysen, dass Heterogenitätsmerkmale der Klasse

sowie Merkmale der Lehrkraft mit dem Einsatz bestimmter Differenzierungsmaßnahmen zusammenhängen.

6 Diskussion

Ziel der vorliegenden Studie war es zu untersuchen, wie Heterogenitätsmerkmale von Schulklassen und Merkmale der Lehrkraft mit Differenzierungsmaßnahmen im Unterricht zusammenhängen. Die Ergebnisse deuten darauf hin, dass alle Lehrkräfte Differenzierung in unterschiedlichem Umfang einsetzten, die fünf untersuchten Prädiktoren jedoch nur für den Einsatz bestimmter Differenzierungsmaßnahmen eine Rolle spielten.

Der Einsatz von Differenzierung durch Unterstützung durch Mitlernende hing positiv mit der Leistungsheterogenität der Klasse sowie dem Mädchenanteil zusammen. Die Leistungsheterogenität der Klassen nehmen auch die Lernenden war, sodass sie bei hoher Leistungsheterogenität womöglich von sich aus unterstützen. Dass die Unterstützung positiv mit dem Mädchenanteil der Klasse zusammenhängt, deckt sich mit den Annahmen von Thon (2017), die postulierte, dass von Mädchen mehr soziales Verhalten erwartet wird und damit auch mehr gegenseitige Unterstützung. Auch die bei Mädchen höher eingeschätzte Selbststeuerungskompetenz (vgl. Kuhl/Hannover 2012) könnte hier eine Rolle spielen, sodass Lehrkräfte womöglich von Mädchen erwarten, nach Abschluss ihrer Aufgaben selbstständig schwächere Lernende zu unterstützen.

In dieser Studie findet sich nur ein Zusammenhang zwischen den eingesetzten Differenzierungsmaßnahmen und dem Fachstudium Englisch: Lehrkräfte, die Englisch studiert haben, setzten signifikant seltener Differenzierung durch Unterstützung durch Hilfs- und Zusatzmaterialien ein als fachfremd unterrichtende Lehrkräfte. Dieses Ergebnis überrascht zunächst, da wir erwartet hatten, dass im Studium erworbenes fachdidaktisches Wissen die Reduktion komplexer Inhalte und damit Differenzierung erleichtert (vgl. Döhrmann/Kaiser/Blömeke 2008). Womöglich kann der Befund dahingehend interpretiert werden, dass sich studierte Englischlehrkräfte kompetenter fühlen, die Lernenden bei Bedarf selbstständig zu unterstützen, anstatt auf Hilfsmaterialen zu verweisen. Eine alternative Erklärung wäre, dass die fachfremd unterrichtenden Lehrkräfte in stärkerem Maße Fortbildung im Fach Englisch in Anspruch genommen haben. Berufserfahrene Lehrkräfte scheinen häufiger Differenzierung durch unterschiedliche Rollen einzusetzen. Dies ist als komplexe Maßnahme zu betrachten, die nicht nur auf kognitive Lernvoraussetzungen zielt, sondern auch soziale Aspekte berücksichtigt. Sie erfordert ein größeres unterrichtsbezogenes Handlungsrepertoire, dass nach Kunter und Kollegen (2011) erst mit zunehmender Berufserfahrung erworben wird.

Die Befunde der vorliegenden Studie geben einen ersten Eindruck davon, welche Heterogenitätsmerkmale der Klasse und welche Merkmale der Lehrkräfte damit zusammenhängen, wie Lehrkräfte im Unterricht differenzieren. Inwiefern diese tatsächlich adaptiv an die unterschiedlichen Lernvoraussetzungen angepasst waren, also einen Einfluss auf die Tiefenstrukturen ausübt, kann mit dem vorliegenden Studiendesign nicht beantwortet werden. Ein Mehrwert liegt in der Verknüpfung von Beobachtungs- und Fragebogendaten, die einen differenzierten Einblick in die Unterrichtswirklichkeit erlauben. Die Daten wurden im Schuljahr 2003/2004 (vgl. Helmke et al. 2008) erhoben, sodass mögliche Veränderungen im Nachgang des „PISA-Schocks" noch nicht abgebildet sind. Es ist zudem anzunehmen, dass die Heterogenität aller Klassen, insbesondere an Gymnasien, durch weiterhin steigende Übergangsquoten zugenommen hat. Auch in inklusiven Klassen vermuten wir einen größeren Einsatz von Differenzierung. Es erscheint somit notwendig, die Studie mit aktuelleren Videodaten nochmals durchzuführen. Limitierend wirkt auch die Stichprobengröße von $N = 83$ untersuchten Klassen bei $N = 81$ Lehrkräften, von denen nur $N = 76$ den Fragebogen vollständig ausfüllten. Die in dieser Studie ausgewählten Prädiktoren für die Heterogenität der Klasse sowie Lehrkraftmerkmale sind nicht erschöpfend, was die geringen Anteile der Varianzaufklärung verdeutlichen. Es besteht somit weiterer Forschungsbedarf.

Literatur

Baumert, Jürgen/Kunter, Mareike (2006). Stichwort – Professionelle Kompetenz von Lehrkräften. In: Zeitschrift für Erziehungswissenschaft 9, H. 4, S. 469–520.

Baumert, Jürgen/Kunter, Mareike/Brunner, Martin/Krauss, Stefan/Blum, Werner/Neubrand, Michael (2004): Mathematikunterricht aus Sicht der PISA-Schülerinnen und -schüler und ihrer Lehrkräfte. In: Prenzel, Manfred/Baumert, Jürgen/Blum, Werner/Lehmann, Rainer/Leutner, Detlev/Neubrand, Michael/Pekrun, Reinhard/Rolff, Hans-Günter/Rost, Jürgen/Schiefele, Ulrich (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs. Münster u. a.: Waxmann, S. 314–354.

Bohl, Thorsten/Batzel, Andrea/Richey, Petra (2012): Öffnung – Differenzierung – Individualität – Adaptivität: Charakteristika, didaktische Implikationen und Forschungsbefunde verwandter Unterrichtskonzepte zum Umgang mit Heterogenität. In: Bohl. Thorsten/Bönsch, Manfred/Trautmann, Matthias/Wischer, Beate (Hrsg.): Binnendifferenzierung. Teil 1: Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht. Immenhausen bei Kassel: Prolog-Verlag, S. 40–69.

Decristan, Jasmin/Jude, Nina (2017): Heterogenitätskategorie Schulleistung/Leistung von Schülerinnen und Schülern. In: Bohl, Thorsten/Budde, Jürgen/Rieger-Ladich, Markus (Hrsg.): Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen. Bad Heilbrunn: Julius Klinkhardt, S. 109–122.

Döhrmann, Martina/Kaiser, Gabriele/Blömeke, Sigrid (2008): Messung des mathematischen und mathematikdidaktischen Wissens: Theoretischer Rahmen und Teststruktur. In: Blömeke, Sigrid/Kaiser, Gabriele/Lehmann, Rainer (Hrsg.): TEDS-M. Münster u. a.: Waxmann, S. 169–196.

Dumont, Hanna (2018): Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht. Zeitschrift für Erziehungswissenschaft. https://doi.org/10.1007/s11618-018-0840-0.

Helmke, Andreas (2014): Unterrichtsaufzeichnung (22300906-1). Version: 1.0. Video. In DESI-Konsortium. DESI – Deutsch Englisch Schülerleistungen International. Frankfurt a. M.: Forschungsdatenzentrum (FDZ) Bildung am DIPF. https://doi.org./10.7477/6:1:1.

Helmke, Andreas (2015): Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts (6. Auflage). Seelze-Velber: Klett/Kallmeyer.

Helmke, Tuyet/Helmke, Andreas/Schrader, Friedrich-Wilhelm/Wagner, Wolfgang/Nold, Günter/Schröder, Konrad (2008): Die Videostudie des Englischunterrichts. In: DESI-Konsortium (Hrsg.): Unterricht- und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie. Weinheim u. a.: Beltz, S. 345–363.

Hugener, Isabell/Krammer, Katrin/Pauli, Christine (2008): Kompetenzen der Lehrpersonen im Umgang mit Heterogenität: Differenzierungsmaßnahmen im Mathematikunterricht. In: Gläser-Zikuda, Michaela/Seifried, Jürgen (Hrsg.): Lehrerexpertise: Analyse und Bedeutung unterrichtlichen Handelns. Münster u. a.: Waxmann, S. 47–66.

Jennek, Julia/Gronostaj, Anna/Vock, Miriam (2018): Wie Lehrkräfte im Englischunterricht differenzieren. Eine Re-Analyse der DESI-Videos. Unterrichtswissenschaft. https://doi.org/10.1007/s42010-018-0027-7.

Jennek, Julia/Gronostaj, Anna/Vock, Miriam (i.Dr.): Mit Unterschieden in der Klasse umgehen: Vorschlag zur Systematisierung von Methoden der Binnendifferenzierung. In: Stiller, Jurik/Laschke, Christin/Nesbya, Thea/Salascheck, Ulrich (Hrsg.): Berlin-Brandenburger Beiträge zur Bildungsforschung 2019. Berlin: Peter Lang.

Klieme, Eckhard (2012). Deutsch-Englisch-Schülerleistungen-International (DESI). Version: 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. Verfügbar unter http://doi.org/10.5159/IQB_DESI_v1

Klieme, Eckhard/Jude, Nina/Rauch, Dominique/Ehlers, Holger/Helmke, Andreas/Eichler, Wolfgang/Thomé, Günther/Willenberg, Heiner (2008): Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts. In: DESI-Konsortium (Hrsg.): Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie. Weinheim u. a.: Beltz, S. 319–344.

Kuhl, Poldi/Hannover, Bettina (2012): Differenzielle Benotung von Mädchen und Jungen. Der Einfluss der von der Lehrkraft eingeschätzten Kompetenz zum selbstgesteuerten Lernen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 44, H. 3, S. 153–162.

Kunter, Mareike/Kleickmann, Thilo/Klusmann, Uta/Richter, Dirk (2011): Die Entwicklung professioneller Kompetenz von Lehrkräften. In: Kunter, Mareike/Baumert, Jürgen/Blum, Werner/Klusmann, Uta/Krauss, Stefan/Neubrand, Martin (Hrsg.): Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV. Münster u. a.: Waxmann, S. 55–68.

Lipowsky, Frank/Lotz, Miriam (2015): Ist Individualisierung der Königsweg zum Lernen? Eine Auseinandersetzung mit Theorien, Konzepten und empirischen Befunden. In: Mehlhorn, Gerlinde/Schöppe, Karola/Schulz, Frank (Hrsg.): Begabungen entwickeln & Kreativität fördern. München: kopaed, S. 155–219.

Lotz, Miriam/Gabriel, Katrin/Lipowsky, Frank (2013): Niedrig und hoch inferente Verfahren der Unterrichtsbeobachtung. Analysen zu deren gegenseitiger Validierung. Zeitschrift für Pädagogik 59, H. 3, S. 357–380.

Lupatsch, Judith/Hadjar, Andreas (2011): Determinanten des Geschlechtsunterschieds im Schulerfolg: Ergebnisse einer quantitativen Studie aus Bern. In: Hadjar, Andreas (Hrsg.): Geschlechtsspezifische Bildungsungleichheiten. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 177–202.

Marquardt, Christian (2011): Videoanalysen in der Unterrichtsforschung. Momentaufnahme oder typischer Schulalltag? Eine Betrachtung der DESI-Videostudie. Hamburg: Diplomica Verlag.

Meyer, Hilbert. (2013): Was ist guter Unterricht? (9. Auflage). Berlin: Cornelsen.

Pauli, Christine/Reusser, Kurt (2006): Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. Zeitschrift für Pädagogik 52, H. 6, S. 774–798.

Praetorius, Ann-Kathrin (2014): Messung von Unterrichtsqualität durch Ratings. Münster: Waxmann.

Reh, Sabine/Schelle, Carla (2006): Biographieforschung in der Schulpädagogik. Aspekte biographisch orientierter Lehrerforschung. In: Krüger, Heinz-Hermann/Marotzki, Winfried (Hrsg.): Handbuch erziehungswissenschaftliche Biographieforschung. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 291–411.

Schiepe-Tiska, Anja/Reiss, Kristina/Obersteiner, Andreas/Heine, Jörg-Henrik//Seidel, Tina/Prenzel, Manfred (2013): Mathematikunterricht in Deutschland: Befunde aus PISA 2012. In: Prenzel, Manfred/Sälzer, Christine/Klieme, Eckhard/Köller, Olaf (Hrsg.). PISA 2012. Fortschritte und Herausforderungen in Deutschland. Münster u. a.: Waxmann, S. 123–154.

Terhart, Ewald (2011); Lehrerberuf und Professionalität. Gewandeltes Begriffsverständnis – neue Herausforderungen. Zeitschrift für Pädagogik. Beiheft (57), S. 202–224.

Thon, Christine (2017): Kategorie Geschlecht. In: Bohl, Thorsten/Budde, Jürgen/Rieger-Ladich, Markus (Hrsg.): Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen. Bad Heilbrunn: Julius Klinkhardt (S. 77–92).

Trautmann, Matthias/Wischer, Beate (2008): Das Konzept der Inneren Differenzierung – eine vergleichende Analyse der Diskussion der 1970er Jahre mit dem aktuellen Heterogenitätsdiskurs. Zeitschrift für Erziehungswissenschaft 10 (Sonderheft 9), S. 159–172.

Trautmann, Matthias/Wischer, Beate (2011): Heterogenität in der Schule: Eine kritische Einführung. Wiesbaden: VS Verlag für Sozialwissenschaften.

Westphal, Andrea/Becker, Michael/Vock, Miriam/Maaz, Kai/Neumann, Marko/McElvany, Nele (2016): The link between teacher-assigned grades and classroom socioeconomic composition. The role of classroom behavior, motivation, and teacher characteristics. Contemporary Educational Psychology 46, S. 218–227. https://doi.org/10.1016/j.cedpsych.2016.06.004

Soziale Interaktionen in pädagogischen Beziehungen

Erfassung mit State Space Grids

Marion Scherzinger, Benjamin Roth und Alexander Wettstein

Abstract

Der Beziehung zwischen Lehrperson und Schülerinnen und Schülern kommt in Schule und Unterricht eine Schlüsselrolle fürs Lernen und Lehren zu. Die vorliegende Studie geht den Fragen nach, inwieweit globale Beziehungseinschätzungen mit den im Unterricht ablaufenden Interaktionen zusammenhängen und welche Unterschiede sich im Unterricht von Klassen- und Fachlehrpersonen zeigen. Dazu wurden 16 videografierte Unterrichtslektionen von acht fünften und sechsten Primarschulklassen mit State Space Grids, einer Methode zur Erforschung dynamischer Systeme, analysiert (Mikroebene) und mit globalen Fragebogeneinschätzungen der Beziehung zwischen Lehrperson und Lernenden (Makroebene) korreliert. Die Ergebnisse zeigen, dass die Beziehungseinschätzung der Lernenden stärker mit den Interaktionen im Unterricht zusammenhängt als die der Lehrpersonen. Positive Zusammenhänge zeigen sich bei gegenseitig freundlichem Verhalten sowie bei einem beidseitig mittleren kontrollierendem bzw. lenkendem Verhalten. Die Beziehungseinschätzung der Fachlehrpersonen korreliert negativ mit dominantem Verhalten der Lernenden im Unterricht in Kombination mit passivem oder aber kontrollierendem Verhalten auf Lehrpersonenseite. Die Analyse von Interaktionsverläufen im Unterricht mit State Space Grids ermöglicht ein vertieftes Verständnis von Beziehungen zwischen Lehrperson und Schülerinnen und Schülern.

1 Problemstellung

Die Beziehung zwischen Lehrperson und Schülerinnen und Schülern bildet eine wichtige Grundlage für erfolgreiche Lehr-Lern-Prozesse (Pianta/Hamre/Stuhlman 2003; Wentzel 2012). Positive Beziehungen zeichnen sich durch Nähe, Sicherheit, Vertrauen und Unterstützung aus (Wentzel 2012). In der Forschung wird die Beziehung zwischen Lehrperson und Schülerinnen und Schülern meist

global mittels Fragebogen erfasst. Bisher kaum erforscht ist, inwieweit diese globalen Masse mit den im Unterricht ablaufenden Interaktionen zusammenhängen.

2 Theoretischer Hintergrund

Lehrpersonen und Schülerinnen können als Partnerinnen und Partner in einem sozialen System verstanden werden, die sich wechselseitig beeinflussen (Herzog 2009). Die Theorie dynamischer Systeme geht davon aus, dass ein System aus mindestens zwei Einheiten besteht, die sich gegenseitig beeinflussen und dadurch Veränderungen im System hervorrufen (Granic/Hollenstein 2003; van Geert 2009). Dabei wird die Veränderung eines Systems nicht auf isolierte Verhaltensweisen zurückgeführt, sondern das Verhalten wird in Relation zueinander betrachtet (Koopmans/Stamovlasis 2016). Das wechselseitige Verhalten von Interaktionspartnerinnen und -partnern, die soziale Interaktion, bildet den Grundbaustein sozialer Beziehungen (Hinde 1997; Granic/Hollenstein 2003). So entstehen und entwickeln sich auch Beziehungen zwischen Lehrperson und Schülerinnen und Schülern über die Zeit auf Grundlage der Interaktionen im Unterricht und die wahrgenommene Beziehung (Makroebene) beeinflusst wiederum das zwischenmenschliche Verhalten bzw. die Interaktionen (Mikroebene) (Pianta/Hamre/Stuhlman 2003). Die Theorie dynamischer Systeme rückt nicht allein das Verhalten der Lehrperson oder der Schülerinnen und Schüler in den Fokus, sondern das wechselseitige Verhalten, die Interaktionen. Durch die Analyse der Interaktionen im Unterricht kann ein tieferes Verständnis über die Entstehung und Entwicklung von Beziehungen erlangt werden (Pennings/Mainhard 2016). Eine Methode, um Interaktionen über die Zeit zu analysieren und die zugleich in der Theorie dynamischer Systeme verortet werden kann, stellen State Space Grids (SSG) dar (Hollenstein 2013; Lamey et al. 2004). Es handelt sich dabei um orthogonale Raster, welche es ermöglichen aufeinander bezogenes Verhalten zweier Parteien wie auch deren Interaktionsverlauf grafisch darzustellen und zu analysieren. Bisher gibt es wenig Studien, welche State Space Grids zur Erforschung von Unterricht eingesetzt haben. Mainhard und Kollegen (2012) untersuchten mit dieser Methode, die im Unterricht ablaufenden Interaktionen für Communion und Agency. Gemäß der interpersonalen Theorie zeichnet sich jedes interpersonale Verhalten als Kombination von Agency und Communion aus (Wubbels et al. 2015). Agency meint Streben nach Dominanz, Einfluss und Kontrolle, Communion menschliche Wärme bzw. Nähe. Fragebogenstudien haben gezeigt, dass das Verhalten erfolgreicher Lehrpersonen durch eine hohe Agency in Kombination mit einer hohen Communion geprägt ist (ebd.). Die Untersuchung der im Unterricht ablaufenden Interaktionen mit SSG hat ergeben, dass wenn Lernende die Be-

ziehung zu ihrer Lehrperson besser einschätzten, die Lehrpersonen im Unterricht eine hohe, die Lernenden eine mittlere Communion aufwiesen und die Lehrpersonen über eine etwas höhere Agency verfügten als die Lernenden (Mainhard et al. 2012). Die vorliegende Studie setzt hier an und berücksichtigt neben der Beziehungseinschätzung der Lernenden auch jene der Lehrpersonen.

3 Fragestellungen

1. Inwieweit hängen die globalen Beziehungseinschätzungen der Lehrpersonen und der Lernenden (makrogenetische Ebene) mit den konkreten Interaktionen im Unterricht (mikrogenetische Ebene) zusammen?
2. Welche Unterschiede zeigen sich zwischen Klassen- und Fachlehrpersonen? Mittelstufenklassen in der Schweiz werden meist von verschiedenen Lehrpersonen unterrichtet. Die Klassenlehrperson trägt die Hauptverantwortung, übernimmt einen Großteil der Lektionen und ist erste Ansprechperson, während Fachlehrpersonen ein oder einzelne Fächer erteilen. Da Fachlehrpersonen die Klassen meist weniger sehen, bieten sich auch weniger Möglichkeiten, eine emotional gefestigte Beziehung aufzubauen (Baker 2006).

4 Methodisches Vorgehen

Die Daten der vorliegenden Studie stammen aus dem Projekt „Wahrnehmung sozialer Interaktionen im Unterricht", in welchem an 83 Klassen in der Schweiz die pädagogischen Beziehung aus Sicht der Lehrpersonen wie auch der Lernenden mit Fragebogen erhoben und an 18 Klassen der Unterricht gefilmt wurde (Wettstein/Ramseier et al. 2016; 2018; Wettstein/Scherzinger/Ramseier 2018; Scherzinger/Wettstein 2019; Wettstein/Scherzinger 2019).

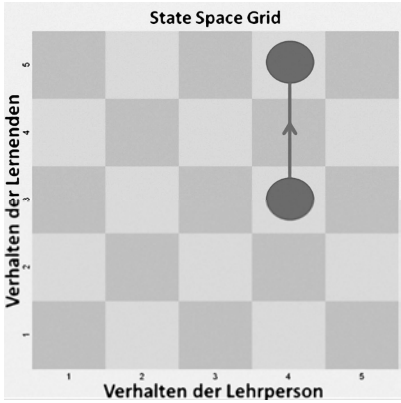
Die Skala „Beziehung" umfasste sechs Items zur Beziehung zwischen Lehrperson und Lernenden (z. B. „Ich mag diese Lehrperson"; vierstufige Antwortskala; Cronbachs $\alpha = .88$; für Details zum Fragebogen vgl. Wettstein/Ramseier et al. 2016). Die Schülerurteile wurden für jede Klasse aggregiert und die Intraklassenkorrelationen für die Skala Beziehung wiesen ein ICC der latenten Variablen von .33 auf (Wettstein/Ramseier/Scherzinger 2018).

In der vorliegenden Studie wurden 16 videografierte Unterrichtslektionen von acht Klassen der Primarschule (5. und 6. Klasse) ausgewertet (Scherzinger et al., submitted). An jeder der acht Klassen ($n = 126$ Schülerinnen und Schüler, 42,1 % weiblich, $M = 11.5$ Jahre) wurde jeweils eine Lektion der Klassenlehrperson und eine Lektion einer Fachlehrperson mit einer GoPro-Kamera gefilmt. Die Klassenlehrpersonen ($n = 8$, 37.5 % weiblich, $M = 41.1$ Jahre alt, $SD = 12.4$)

unterrichten die jeweilige Klasse durchschnittlich M = 21.9 (SD = 2.5) Lektionen pro Woche, die Fachlehrpersonen (n = 8, 75.0 % weiblich, M = 40.6 Jahre alt, SD = 8.1) im Durchschnitt M = 4.9 (SD = 1.6) Lektionen. 14 der 16 Lehrpersonen wiesen mehr als sechs Jahre Berufserfahrung auf.

Die Auswertung der Videos erfolgte mit State Space Grids (Hollenstein 2007; 2013). Mit dieser Methode können Interaktionsverläufe innerhalb eines bestimmten Zeitraums grafisch dargestellt und analysiert werden. Ein State Space Grid bildet alle möglichen Interaktionen (States/Zustände) eines zweidimensionalen Systems, hier der Lehrperson (x-Achse) und der Lernenden (y-Achse), ab (vgl. Abb. 1). Das aufeinander bezogene Verhalten bzw. die Interaktion wird als Punkt abgetragen. Wenn sich im Unterricht das Verhalten einer oder beider Parteien ändert, ändert sich jeweils auch deren Zustand (State), es entsteht ein neuer Punkt, welcher durch eine Linie mit dem vorherigen verbunden wird. Mit State Space Grids können die Häufigkeit und Verweildauer der Interaktionen in den verschiedenen States oder auch die Anzahl Wechsel zwischen den States während einer Lektion ermittelt werden.

Abbildung 1: State Space Grid



Für die Analyse mit State Space wurde mittels deduktiver und induktiver Vorgehensweise ein fünfstufiges Kategoriensystem für Agency (Lenkung, Kontrolle) und Communion (Wärme) entwickelt und das Verhalten im Event-Sampling-Verfahren in MAXQDA 2018 für die Lernenden und die Lehrperson einzeln nach vorab festgelegten Kodierregeln kodiert. Die Kategorien der Dimension Agency reichen von „sehr passiv“ (1, z. B. reagiert nicht, greift nicht ein) bis „sehr aktiv, bestimmend“ bzw. „dominant“ (5, z. B. kontrollierend), diejenigen für Communion reichen von „feindlich“ (1, z. B. beleidigen, bloßstellen) bis „sehr freundlich, kumpelhaft“ (5, z. B. distanzloses Verhalten). Neun Prozent des Datenmaterials wurde durch einen unabhängigen Rater zweitko-

diert. Die prozentuale Übereinstimmung lag bei 87.25 % (Cohen's Kappa $\kappa = 0.65$).

Für jede Lektion wurde anschließend für Agency und für Communion einzeln ein State Space Grid mit dem Programm GridWare 1.15 (Lamey et al. 2004) erstellt, in welchen das zeitlich synchrone Verhalten der Lernenden und der Lehrperson auf einem zweidimensionalen Raster mit 25 States abgebildet wird; x-Achse (Verhalten/Kodierung der Lehrperson), y-Achse (Lernende) (vgl. Abb. 1). Anschließend wurde ermittelt, wie häufig und wie lange sich die Interaktionen während einer Lektion in den jeweiligen States abspielten.

Mit Korrelationsanalysen wurde schließlich untersucht, inwiefern die Ergebnisse der State Space Grids (Mikroebene) mit den globalen Beziehungseinschätzungen aus dem Fragebogen (Makroebene) zusammenhängen und welche Unterschiede sich zwischen Klassen- und Fachlehrpersonen ergeben. Sowohl Klassen- und Fachlehrpersonen wie auch Schülerinnen und Schüler der vorliegenden Studie schätzten die Beziehung im Durchschnitt relativ hoch ein. Es zeigten sich keine statistisch signifikanten Unterscheide in den Beziehungseinschätzungen bei Klassen- und Fachlehrpersonen.

5 Ergebnisse

Fragestellung 1: Zusammenhang Beziehungseinschätzung und pädagogische Interaktionen

Die Beziehungseinschätzung der Schülerinnen und Schülern korrelierte statistisch signifikant positiv mit Interaktionen, bei denen die Lernenden selbst eine hohe Agency (A4) und die Lehrpersonen eine mittlere (A3) besitzen. Negativ hingegen, wenn die Lehrpersonen relativ stark lenken oder das Geschehen kontrollieren (A4), während die Lernenden eher in einer passiven Rolle sind (A2) (vgl. Tab. 1).

Tabelle 1: Korrelationen der prozentualen zeitlichen Anteile an Interaktionen in den einzelnen Zellen (gemittelt über alle 16 Lehrpersonen) mit den Beziehungseinschätzungen der Schülerinnen und Schüler bzw. der Lehrpersonen für Agency und Communion[a b]

Agency der Lernenden	Agency der Lehrperson 1	2	3	4	5	Communion der Lernenden	Communion der Lehrperson 1	2	3	4	5
1	-	-	-	-	-	1	-	-	-	-	-
2	-	-	-	-.21 **-.50***	-	2	-	**-.54*** **-.80****	-.30 -.23	-	-
3	-	-.05 -.05	.11 .37	.17 -.13	-	3	-	-.38 **-.56***	.37 .18	.18 **.54***	-
4	-	-.02 -.18	.35 **.50***	-.14 -.01	-	4	-	-	.10 **.60***	.32 .49	-
5	-.12 -.32	-.37 -.15	-.37 .05	-.48 -.32	-	5	-	-	-	-	-

[a] r Lehrpersonen / r Lernende;

[b] Zellen, in denen weniger als 1 % der Interaktionen stattfinden, sind durch – gekennzeichnet.

* $p < .05$, ** $p < .01$.

Für Communion (vgl. Tab. 1) zeigen sich statistisch signifikant positive Zusammenhänge zwischen der Beziehungseinschätzung der Lernenden und Interaktionen, bei denen entweder sachlich-neutral (C3) oder freundlich (C4) miteinander umgegangen wird. Negative Zusammenhänge zeigten sich hingegen bei einem gegenseitig unfreundlichem Verhalten (C2) im Unterricht.

Fragestellung 2: Pädagogische Interaktionen bei Klassen- und Fachlehrpersonen[1]

Bei Klassenlehrpersonen (vgl. Tab. 2) korrelieren die Beziehungseinschätzung der Schülerinnen und Schüler statistisch signifikant negativ mit Interaktionen, bei denen die Lernenden über eine deutlich niedrige Agency (A2) verfügen als die Klassenlehrpersonen (A4). In der Tendenz zeigen sich zudem negative Zusammenhänge mit einer niedrigen Agency der Klassenlehrpersonen (A2), unabhängig von der Ausprägung der Lernenden-Agency. Positiv hängt die Beziehungseinschätzung hingegen mit einer mittleren Agency (A3) der Lehrperson und einer hohen Agency (A4) der Lernenden zusammen.

Für Communion zeigt sich in der Tendenz ein positiver Zusammenhang zwischen der Beziehungseinschätzung und freundlichem Verhalten im Unter-

1 Die Beziehungseinschätzungen zwischen Klassen- und Fachlehrpersonen unterscheiden sich nicht statistisch signifikant voneinander.

richt und ein negativer Zusammenhang zwischen der Beziehungseinschätzung und gegenseitig unfreundlichem Verhalten.

Tabelle 2: Unterricht Klassenlehrpersonen: Korrelationen der prozentualen zeitlichen Anteile an Interaktionen in den einzelnen Zellen (gemittelt über alle 8 Klassenlehrpersonen) mit den Beziehungseinschätzungen der Schülerinnen und Schüler bzw. der Klassenlehrpersonen für Agency und Communion[a b]

Agency der Lernenden	Agency der Klassenlehrperson 1	2	3	4	5
1	-	-	-	-	-
2	-	-.06 -.19	-	-.36 **-.76***	-
3	-	-.10 -.38	-.33 .38	-.36 -.43	-
4	-	.31 -.09	.51 .62	.29 .04	-
5	.10 -.17	-.05 -.10	-.35 .23	.31 .65	-

Communion der Lernenden	Communion der Klassenlehrperson 1	2	3	4	5
1	-	-	-	-	-
2	-	-.10 -.62	.18 -.19	-	-
3	-	-.49 -.66	-.22 -.12	.24 .71	-
4	-	-	.05 .68	-	-
5	-	-	-	-	-

[a] r Lehrpersonen / r Lernende;

[b] Zellen, in denen weniger als 1 % der Interaktionen stattfinden, sind durch – gekennzeichnet.

* $p < .05$.

Tabelle 3: Unterricht Fachlehrperson: Korrelationen der prozentualen zeitlichen Anteile an Interaktionen in den einzelnen Zellen (gemittelt über alle 8 Fachlehrpersonen) mit den Beziehungseinschätzungen der Schülerinnen und Schüler bzw. der Fachlehrpersonen für Agency und Communion[a b]

Agency der Lernenden	Agency der Fachlehrperson 1	2	3	4	5
1	-	-	-	-	-
2	-	-	-	-	-
3	-	.17 . 24	**.76*** .40	.18 .70	-
4	-	-.57 -.37	.22 .51	-.37 -.04	-
5	**-.76*** **-.90****	-.52 -.22	-.42 -.03	**-.73*** -.61	-

Communion der Lernenden	Communion der Fachlehrperson 1	2	3	4	5
1	-	-	-	-	-
2	-	**-.75*** **-.92****	-.61 -.30	-	-
3	-	-	**.72*** .39	.18 .57	-
4	-	-	.45 .50	.30 .46	-
5	-	-	-	-	-

[a] r Lehrpersonen / r Lernende;

[b] Zellen, in denen weniger als 1 % der Interaktionen stattfinden, sind durch – gekennzeichnet.

* $p < .05$, ** $p < .01$.

Bei Fachlehrpersonen zeigten sich für Agency (vgl. Tab. 3) statistisch signifikant negative Zusammenhänge zwischen der Beziehungseinschätzung der Fachlehrperson und einer sehr hohen Agency (A5) der Lernenden in Kombination mit einer niedrigen (A1) oder auch einer hohen Agency (A4) der Lehrperson. Auch bei den Lernenden korreliert die Beziehungseinschätzung stark negativ mit einer sehr tiefen Agency (A1) der Fachlehrperson und einer sehr hohen (A5) der Lernenden. Für Agency finden sich in der Tendenz positive Korrelationen zwischen der Beziehungseinschätzung und einer mittleren bis hohen Agency der Lernenden (A3, A4) und der Lehrpersonen (A3).

Für Communion zeigt sich, dass die Beziehungseinschätzung aus Sicht der Lernenden wie auch der Fachlehrpersonen negativ mit Interaktionen zusammenhängt, bei denen beide Parteien unfreundlich miteinander umgehen (C2). Die Beziehungseinschätzung der Fachlehrperson korreliert zudem statistisch signifikant positiv mit einem sachlich-nüchternen Umgang im Unterricht (C3). In der Tendenz korreliert die Beziehungseinschätzung negativ mit einer tiefen Communion (C2) und positiv mit einer mittleren (C3) bis hohen Communion (C4).

6 Diskussion

Die Korrelationsanalysen zeigen, dass die globale Beziehungseinschätzung der Schülerinnen und Schüler stärker mit den im Unterricht beobachteten pädagogischen Interaktionen zusammenhängen als die der Lehrpersonen. Dies könnte darauf zurückgeführt werden, dass die Schülerinnen und Schüler während des Unterrichts mehr Zeit zum Beobachten haben, da sie sich in einer handlungsentlasteten Situation befinden und ihre Beziehungseinschätzung stärker vom beobachteten Verhalten im Unterricht abhängt und dieses beeinflusst.

Die Beziehungseinschätzung hängt positiv mit Interaktionen zusammen, in denen die Lehrpersonen über eine mittlere, die Lernenden über eine hohe Agency verfügen und mit Interaktionen, in denen sich mindestens eine der Parteien freundlich verhält. Negative Korrelationen zeigten sich mit pädagogischen Interaktionen, in denen die Lehrpersonen das Geschehen im Klassenzimmer kontrollieren und die Schülerinnen und Schüler sich passiv verhalten wie auch mit Interaktionen, in denen sich die Lehrpersonen unfreundlich und auch die Schülerinnen und Schüler unfreundlich bis sachlich-neutral verhalten. Diese Befunde gehen mit Studien zur Qualität von pädagogischen Interaktionen einher, die zeigen, dass wertschätzende Beziehungen mit der wahrgenommenen Autonomie der Schülerinnen und Schüler zusammenhängen (Rakoczy 2007). Andererseits erleben sich Schülerinnen und Schüler, wenn Lehrpersonen die Interaktionen im Unterricht stark dominieren, nicht in ihrer Autonomie unterstützt, ihre schulische Motivation sowie das Interesse sinken (Seidel/

Prenzel/Rimmele 2003) und sie erleben auch häufiger negative Emotionen (Assor et al. 2005). Wenn Lernende mehr Autonomieunterstützung durch die Lehrperson erfahren, berichten sie auch mehr Freude und Interesse am Unterricht (Park et al. 2012). Es scheint wichtig zu sein, dass sowohl Lehrpersonen als auch Schülerinnen und Schüler im Unterricht über eine ausreichend hohe Autonomie verfügen und den Lernenden auch Autonomie gewährt wird.

Der Vergleich zwischen Klassen- und Fachlehrpersonen ergab einen stärkeren Zusammenhang zwischen den pädagogischen Interaktionen im Unterricht der Fachlehrperson und den Beziehungseinschätzungen insbesondere der Fachlehrpersonen, aber auch der Lernenden. Je schlechter die Fachlehrpersonen die Beziehung einschätzten, desto häufiger zeigten sich im Unterricht Interaktionen, in denen die Lernenden dominierten und die Fachlehrpersonen auch versuchten das Geschehen zu kontrollieren oder aber sich passiv verhielten bzw. sich zurückzogen. Letzteres hängt auch stark negativ mit der Beziehungseinschätzung der Lernenden zusammen. Diese negativen Korrelationen im Bereich einer hohen Lernenden-Agency könnten in Zusammenhang mit auftretenden Unterrichtsstörungen und der Klassenführung stehen. In der Studie von Wettstein, Ramseier und Scherzinger (2018), schätzten die Schülerinnen und Schüler den Unterricht von Fachlehrpersonen hinsichtlich Klassenführung negativer ein und berichteten auch mehr Unterrichtsstörungen als bei Klassenlehrpersonen. Videobeobachtungen des Unterrichts ergaben zudem statistisch signifikante Unterschiede zwischen aggressivem Lernendenverhalten im Unterricht von Klassen- und Fachlehrpersonen (Scherzinger et al. 2018). Es scheint, dass der Unterricht von Fachlehrpersonen stärker durch Unterrichtsstörungen belastet ist. Die Fachlehrpersonen der vorliegenden Studie unterrichten die beobachtete Klasse im Schnitt fünf Lektionen, die Klassenlehrpersonen rund 22 Lektionen pro Woche. Somit haben die Klassenlehrpersonen auch mehr Zeit, die Schülerinnen und Schüler kennenzulernen und eine positive Beziehung aufzubauen.

Die vorliegende explorative Studie weist Stärken und Limitationen auf. Die Erfassung pädagogischer Interaktionen mit State Space Grids ermöglicht in Kombination mit Fragebogendaten ein vertieftes Verständnis für die pädagogische Beziehung und kann Hinweise zu kritischen Interaktionen im Unterricht geben. Als Limitationen können die Stichprobengröße, ein einziger Messzeitpunkt sowie die Korrelationsanalysen genannt werden, welche keine Aussagen zur Kausalrichtung erlauben.

Für die Stärkung der Beziehung zwischen Lehrperson und Lernenden wäre es aufschlussreich, Interaktionen in kritischen Phasen genauer zu untersuchen und Lehrpersonen dafür zu sensibilisieren, sich auch in kritischen Interaktionen nicht negativ reziprok zu verhalten sowie ihre Agency situativ anzupassen und gleichzeitig eine möglichst hohe Communion aufrechtzuerhalten.

Literatur

Assor, A./Kaplan, H./Kanat-Maymon, Y./Roth, G. (2005): Directly controlling teacher behaviors as predictors of poor motivation and engagement in girls and boys: The role of anger and anxiety. In: Learning and Instruction, 15, S. 397–413.

Baker, J. A. (2006): Contributions of teacher–child relationships to positive school adjustment during elementary school. In: Journal of School Psychology, 44, S. 211–229.

Granic, I./Hollenstein, T. (2003): Dynamic systems methods for models of developmental psychopathology. In: Development and Psychopathology, 15, S. 641–669.

Herzog, W. (2009): Schule und Schulklasse als soziale Systeme. In: Becker, R. (Hrsg.): Lehrbuch der Bildungssoziologie. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 155–194.

Hinde, R. A. (1997): Relationships: A dialectical perspective. Hove, UK: Psychology Press.

Hollenstein, T. (2007): State space grids: Analyzing dynamics across development. In: International Journal of Behavioral Development, 31, S. 384–396.

Hollenstein, T. (2013): State space grids: Depicting dynamics across development. Boston, MA: Springer US.

Koopmans, M./Stamovlasis, D. (2016): Introduction to education as a complex dynamical system. In: Koopmans, M./Stamovlasis, D. (Hrsg.): Complex dynamical systems in education: Concepts, methods and applications. Cham: Springer International Publishing, S. 1–8.

Lamey, A./Hollenstein, T./Lewis, M. D./Granic, I. (2004). GridWare (Version 1.1). [Computer software]. http://statespacegrids.org

Mainhard, M. T./Pennings, H. J. M./Wubbels, T./Brekelmans, M. (2012): Mapping control and affiliation in teacher–student interaction with State Space Grids. In: Teaching and Teacher Education, 28, S. 1027–1037.

Park, S./Holloway, S. D./Arendtsz, A./Bempechat, J./Li, J. (2012): What makes students engaged in learning? A time-use study of within- and between-individual predictors of emotional engagement in low-performing high schools. In: Journal of Youth and Adolescence, 41, S. 390–401.

Pennings, H. J. M./Mainhard, T. (2016): Analyzing teacher–student interactions with State Space Grids. In: Koopmans, M./Stamovlasis, D. (Hrsg.): Complex dynamical systems in education: Concepts, methods and applications. Cham: Springer International Publishing, S. 233–271.

Pianta, R. C./Hamre, B./Stuhlman, M. (2003): Relationships between teachers and children. In: Reynolds, W. M./Miller, G. E. (Hrsg.): Handbook of Psychology. Vol. 7: Educational Psychology. Hoboken: Wiley & Sons, S. 199–234.

Rakoczy, K. (2007): Motivationsunterstützung im Mathematikunterricht. Münster: Waxmann.

Scherzinger, M./Roth, B./Wettstein, A. (submitted): Pädagogische Interaktionen als Grundbaustein der Lehrer-Schüler-Beziehung. Die Erfassung mit State Space Grids.

Scherzinger, M./Wettstein, A. (2019): Classroom disruptions, the teacher–student relationship and classroom management from the perspective of teachers, students and external observers: a multimethod approach. In: Learning Environments Research, 22, S. 101–116.

Scherzinger, M./Wettstein, A./Wyler, S. (2018): Aggressive und nicht aggressive Unterrichtsstörungen durch Schülerinnen und Schüler sowie durch Klassen- und Fachlehrpersonen – eine Videostudie. In: Empirische Sonderpädagogik, 4, S. 388–407.

Seidel, T./Prenzel, M./Rimmele, R. (2003): Gelegenheitsstrukturen beim Klassengespräch und ihre Bedeutung für die Lernmotivation – Videoanalysen in Kombination mit Schülerselbsteinschätzungen. In: Unterrichtswissenschaft, 31(2), S. 142–165.

Van Geert, P. (2009): Nonlinear complex dynamical systems in developmental psychology. In: Guastello, S. J./Koopmans, M./Pincus, D. (Hrsg.): Chaos and complexity in psychology: The theory of nonlinear dynamical systems. New York: Cambridge University Press; Cambridge Univ. Press, S. 242–281.

Wentzel, K. R. (2012): Teacher-student relationships and adolescent competence at school. In: Wubbels, T./den Brok, P./van Tartwijk, J./Levy, J. (Hrsg.): Advances in learning environments research: volume 3. Interpersonal relationships in education: An overview of contemporary research. Rotterdam, Boston, Taipei: Sense Publishers, S. 19–36.

Wettstein, A./Ramseier, E./Scherzinger, M. (2018): Eine Mehrebenenanalyse zur Schülerwahrnehmung von Störungen im Unterricht der Klassen- und einer Fachlehrperson. In: Psychologie in Erziehung und Unterricht, 65(1), S. 1–16.

Wettstein, A./Ramseier, E./Scherzinger, M./Gasser, L. (2016): Unterrichtsstörungen aus Lehrer- und Schülersicht. Aggressive und nicht aggressive Störungen im Unterricht aus der Sicht der Klassen-, einer Fachlehrperson und der Schülerinnen und Schüler. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 48, S. 171–183.

Wettstein, A./Scherzinger, M. (2019): Unterrichtsstörungen verstehen und wirksam vorbeugen. Stuttgart: Kohlhammer.

Wettstein, A./Scherzinger, M./Ramseier, E. (2018): Unterrichtsstörungen, Beziehung und Klassenführung aus Lehrer-, Schüler- und Beobachterperspektive. In: Psychologie in Erziehung und Unterricht, 65(1), S. 58–74.

Wubbels, T./Brekelmans, M. (2005): Two decades of research on teacher-student relationships in class. In: International Journal of Educational Research, 43, S. 6–24.

Wubbels, T./Brekelmans, M./den Brok, P./Wijsman, L./Mainhard, T./van Tartwijk, J. (2015): Teacher-student relationships and classroom management. In: Emmer, E. T./Sabornie, E. J. (Hrsg.): Handbook of classroom management. 2nd ed. New York: Routledge, S. 363–386.

Unterrichtsqualität in MINT-Klassen

Befunde aus der quasi-experimentellen Studie ProBiNi

Sonja Nonte, Mareike Haas und Tobias C. Stubbe

Abstract

Profilklassen mit naturwissenschaftlichem Schwerpunkt, sogenannte MINT-Klassen, erfreuen sich zunehmender Beliebtheit. Im Rahmen eines erhöhten Stundenkontingents bieten sie Räume für fächerübergreifenden und projektorientierten Unterricht, der häufig einen hohen Anteil selbstregulierter Lernprozesse aufweist. Empirische Studien zeigen, dass die Beurteilung der Unterrichtsqualität je nach Fach und Klassenzusammensetzung variiert. Bisher gibt es kaum empirische Evidenz hinsichtlich der Unterrichtsqualität von Profilklassen. Ziel des vorliegenden Beitrags ist es, Unterschiede in der Wahrnehmung der Unterrichtsqualität (Klassenführung, kognitive Aktivierung und unterstützendes Verhalten) im Fach Naturwissenschaften zwischen MINT-Klassen und Klassen ohne Profil (Vergleichsklassen) herauszustellen. Anhand der im Rahmen der DFG-Studie ProBiNi (n = 254) erfassten Daten von Schülerinnen und Schülern kann zunächst für das Instrument zur Erfassung von Unterrichtsqualität skalare Messinvarianz angenommen werden. Mittels eines Vergleichs latenter Mittelwertunterschiede zeigen sich unter Kontrolle von Kovariaten keine signifikanten Unterschiede in der Wahrnehmung der Unterrichtsqualität zwischen MINT- und Vergleichsklassen.

1 Einleitung

Neben zahlreichen naturwissenschaftlichen Bildungsangeboten für Kinder und Jugendliche im außerschulischen Bereich (z. B. Mitmach-Museen, Schülerlabore) etablieren sich in den vergangenen Jahren zunehmend auch Klassen mit MINT-Profilen (Forscher-, Entdeckerklassen etc.) an weiterführenden Schulen (Seidel et al. 2016). Mit ein bis zwei zusätzlichen Schulstunden in der Woche eröffnen sie Räume für fächerübergreifendes Lernen im Klassenverband in den Bereichen Mathematik, Informatik, Naturwissenschaft und Technik (MINT)

und ermöglichen so Schülerinnen und Schülern mit naturwissenschaftlichen Themen und Arbeitsweisen in Berührung zu kommen. Ausgehend von einem aktuell vorherrschenden und sich für die Zukunft noch verstärkenden Fachkräftemangel im MINT-Bereich (Anger/Koppel/Plünnecke 2018) kann die Einrichtung von MINT-Profilklassen über eine kontinuierliche Wissensvermittlung hinaus, eine nachhaltige Förderung von Interesse an mathematisch-naturwissenschaftlichen Themen leisten.

2 Theoretischer Hintergrund

Aktuelle Studien heben immer wieder geschlechtsbezogene Disparitäten in mathematisch-naturwissenschaftlichen Leistungen und Interessen hervor. So weisen Mädchen in Mathematik und in Naturwissenschaften ein geringeres Interesse (u. a. Daniels 2008; Schiepe-Tiska/Simm/Schmidtner 2016), eine (tendenziell) niedrigere instrumentelle Motivation (Schiepe-Tiska/Simm/Schmidtner 2016) und signifikant niedrigere Kompetenzwerte (Hammer et al. 2016) auf. Während bisherige Initiativen der Wirtschaft oder der Bundesregierung die Einstellungen gegenüber Naturwissenschaften und Technik nur minimal verändern konnten, müsste vielmehr die subjektive Wahrnehmung des schulischen Fachunterrichts in den Blick genommen werden, so Pfenning und Renn (2012). Enthält der Unterricht Anteile an praktischen und fächerübergreifenden Elementen oder zeigt er Bezüge zur Lebenswelt der Kinder und Jugendlichen, steigert dies das Interesse an Naturwissenschaften der Schülerinnen und Schüler (Höffler/Lüthjohann/Parchmann 2014). Damit kommt der konkreten Umsetzung des Unterrichts eine entscheidende Rolle zu. Unterrichtshandeln umfasst dabei eine Vielzahl unterschiedlicher Verhaltensweisen, die nicht nur hinsichtlich des Inhalts und der Art des Unterrichts, sondern auch nach Fach, Jahrgang und Schulform variieren (Holzberger/Kunter 2016; Klieme/Rakoczy 2008). Eine Beschreibung von Unterrichtsqualität hat sich im deutschsprachigen Forschungsraum über die drei Dimensionen „strukturierte, klare und störungspräventive Unterrichtsführung“, „unterstützendes, schülerorientiertes Sozialklima“ sowie „kognitive Aktivierung“ etabliert (Klieme/Rakoczy 2008, S. 228). Ziel einer gelungenen Klassenführung ist es, die zur Verfügung stehende Zeit effizient für Lernprozesse zu nutzen und durch präventives Handeln Zeitverluste durch Unterrichtsstörungen zu minimieren. Konstruktive Unterstützung beschreibt den Ausprägungsgrad von Unterrichtsaspekten, wie beispielsweise die Fehlerkultur, adaptive Hilfestellung oder das konstruktive Klima in der Klasse. Kognitive Aktivierung beinhaltet hingegen eine auf dem Vorwissen der Schülerinnen und Schüler aufgebaute Integration neuer Wissensinhalte und eine vertiefende Auseinandersetzung mit bestehenden Lerninhalten. Angebot-Nutzungs-Modelle (Helmke 2003) sowie aktuellere Studien wie die von

Praetorius, Pauli, Reusser, Rakoczy und Klieme (2014) verweisen auf die Gelegenheitsstruktur des Unterrichts und die hohe Variabilität insbesondere der kognitiven Aktivierung zwischen verschiedenen Lerngruppen. Diese Befunde deuten darauf hin, dass der Kontext, das Vorwissen, die Herkunft der Schülerinnen und Schüler sowie die motivationalen, sozialen und affektiven Dispositionen eine entscheidende Rolle für Unterrichtsqualität spielen.

3 Forschungsstand

In mathematisch-naturwissenschaftlichen Fächern wurde bereits mehrfach anhand empirischer Studien nachgewiesen, dass kognitive Aktivierung und Klassenführung in einem Zusammenhang mit dem Wissenserwerb stehen (u. a. Baumert et al. 2010). Eine Steigerung der intrinsischen Motivation wird hingegen über eine effektive Klassenführung erzielt (Klieme/Rakoczy 2003). Im Rahmen einer Metaanalyse von Cornelius-White (2007) erwies sich die konstruktive Unterstützung durch die Lehrkraft für die Motivation, Lernleistung und das schulische Verhalten der Schülerinnen und Schüler förderlich. Klieme, Schümer und Knoll (2001) betonen, dass Lehrkräfte unterschiedliche Schwerpunkte im Unterricht setzen. Wie und in welcher Schwerpunktsetzung die verschiedenen Unterrichtsdimensionen zwischen unterschiedlichen Lehrkräften und Fächern variieren, ist bis heute jedoch noch wenig erforscht worden. Erste Tendenzen zeigen, dass Lehrkräfte im Fach Mathematik alle drei Dimensionen gleichermaßen bedienen (Holzberger/Kunter 2016). Studien, die sich mit der Unterrichtsqualität in MINT-Klassen befassen und zudem vergleichend angelegt sind, finden sich bis heute nicht. Dieses Desiderat soll im Rahmen dieses Beitrags aufgegriffen werden.

4 Forschungsfragen

Ausgehend von der Annahme, dass MINT-Klassen besondere Lernumgebungen darstellen, stellt sich die Frage, ob die Unterrichtsqualität von Schülerinnen und Schülern in MINT-Klassen positiver wahrgenommen wird als von Schülerinnen und Schülern in Klassen, in denen ein regulärer Naturwissenschaftsunterricht stattfindet. Um die Ausprägungen auf den drei Dimensionen Klassenführung, kognitive Aktivierung und unterstützendes Verhalten untersuchen zu können, wird zunächst geprüft, ob die Unterrichtsqualität in den MINT- und Vergleichsklassen äquivalent erfasst werden kann. Ist dies der Fall, erfolgt ein Vergleich der Unterrichtsqualitätsdimensionen zwischen MINT- und Vergleichsklassen. In einem letzten Schritt wird zudem untersucht, ob etwaige

Unterschiede auch unter Berücksichtigung von Kovariaten auf Schülerebene (Geschlecht, Fähigkeitsselbstkonzept, Sachinteresse etc.) Bestand haben.

5 Methodisches Vorgehen

Die Datengrundlage bildet die von der Deutschen Forschungsgemeinschaft (DFG) finanzierte Quer- und Längsschnittstudie (DFG-Projektnummer: 312968144) *ProfilBildung an Niedersächsischen Gymnasien und Integrierten Gesamtschulen* (ProBiNi). Über drei Messzeitpunkte (September 2016 bis September 2018) wurden Angaben von Schülerinnen und Schülern, deren Eltern, Lehrkräften sowie Schulleitungen anhand von standardisierten Fragebögen und Onlineumfragen zu Profilierungsmaßnahmen erhoben. Die längsschnittliche Gesamtstichprobe umfasst zwei Gesamtschulen und sieben Gymnasien in Niedersachsen, wobei fünf Klassen über ein musikalisches und fünf Klassen über ein MINT-Profil (eine Schule hat mit zwei Profilklassen teilgenommen) verfügen ($n = 556$). Zwei Gymnasien und eine integrierte Gesamtschule haben an einzelnen Messzeitpunkten teilgenommen und ein Gymnasium und drei integrierte Gesamtschulen haben sich an der ergänzenden Querschnittstudie beteiligt. Fehlende Werte wurden in *R* (R Development Core Team 2017) mit dem Package *mice* zehnfach multipel imputiert (van Buuren/Groothuis-Oudshoorn 2011).

Die Ergebnisse in diesem Beitrag basieren auf Daten des ersten und zweiten Messzeitpunktes. Nicht berücksichtigt wurden die Daten derjenigen Schülerinnen und Schüler, die nur am ersten Messzeitpunkt teilgenommen haben. Die Teilstichprobe setzt sich aus den fünf teilnehmenden Klassen, die über ein naturwissenschaftliches Profil (MINT) verfügen ($n = 126$), sowie aus deren direkten Parallelklassen zusammen, die kein Profil aufweisen (Vergleichsklassen) ($n = 128$). Die Klassen stammen aus insgesamt fünf Schulen: zwei Gesamtschulen und drei Gymnasien. Zu Beginn der 6. Jahrgangsstufe sind die Schülerinnen und Schüler im Durchschnitt $M = 11.51$ Jahre alt ($SD = 0.49$). Der Anteil an Mädchen liegt bei 40.5 % in den MINT-Klassen und bei 50.0 % in den Vergleichsklassen.

Instrumente

Für die Erfassung der *Unterrichtsqualität* im Fach Naturwissenschaften wurden die Schülerinnen und Schüler zu Beginn der 6. Jahrgangsstufe um ihre retrospektive Einschätzung der drei Dimensionen kognitive Aktivierung (KA), Klassenführung (KF) und unterstützendes Verhalten (UV) des vergangenen Schuljahres gebeten (Baumert et al. 2009; Fauth/Decristan/Riesner/Klieme/Büttner 2014). Sie sollten je fünf bis sieben Aussagen pro Dimension auf einer vierstufi-

gen Likert-Skala (1 = „stimmt gar nicht“ bis 4 = „stimmt genau“ (KA, UV) bzw. 1 = „in jeder Stunde“ bis 4 = „nie oder fast nie“ (KF)) beantworten. Beispielitems sind „Wir bearbeiten im Naturwissenschaftsunterricht Aufgaben, über die ich sehr gründlich nachdenken muss.“ (KA), „Es ist laut und unruhig.“ (KF) und „Unsere Naturwissenschaftslehrkraft ist auch dann nett zu mir, wenn ich einen Fehler mache.“ (UV). Das Messmodell ist in Abbildung 1 dargestellt.

Mittels eines standardisierten Kompetenztests wurden zu Beginn der 6. Jahrgangsstufe die *naturwissenschaftlichen Fähigkeiten* erhoben. Da der zeitliche Rahmen nicht ausreichend war, um unterschiedliche Teilbereiche der Naturwissenschaftskompetenz zu erfassen, fand eine Beschränkung auf die Biologiekompetenz statt. Biologie wurde gewählt, da dieser Teilbereich in der Grundschule besonders umfangreich und die ProBiNi-Studie zu Beginn der 5. Jahrgangsstufe gestartet ist. Der Test umfasst insgesamt 42 Items aus TIMSS 2011. Skaliert wurde der Test als dichotomes Raschmodell (1PL) mit dem Package *TAM* (Robitzsch/Kiefer/Wu 2017) in *R* (R Development Core Team 2017). Einschränkend ist zu berücksichtigen, dass die WLE-Reliabilität mit .57 gering ausfällt.

Die *kognitiven Fähigkeiten* der Schülerinnen und Schüler wurden anhand des Subtests Figurenanalogien (N2) für die 6. Klasse aus dem KFT-NF 2 (Heller/Perleth 2000) erfasst. Zu diesem Zwecke wurden die Rohwerte nach Vorgabe des Manuals in Normwerte transformiert.

Das *naturwissenschaftliche Sachinteresse* wurde ebenfalls zu Beginn der 6. Jahrgangsstufe über eine Skala mit vier Items erfasst (Daniels 2008). Aussagen wie „Die Beschäftigung mit naturwissenschaftlichen Themen ist für mich sehr wichtig, unabhängig von der Schule oder anderen Personen.“ beantworteten die Schülerinnen und Schüler auf einer vierstufigen Likert-Skala (1 = „stimmt gar nicht“ bis 4 = „stimmt genau“; Cronbachs $\alpha = .87$).

Das *naturwissenschaftliche Fähigkeitsselbstkonzept* wurde anhand einer Skala, die aus acht Items besteht, mit Hilfe einer vierstufigen Likert-Skala (1 = „stimmt gar nicht“ bis 4 = „stimmt genau“) zu Beginn der 6. Jahrgangsstufe, erhoben (z. B. „Ich gehöre im naturwissenschaftlichen Unterricht zu den guten Schülerinnen und Schülern.“; Cronbachs $\alpha = .86$) (Rost/Sparfeldt 2002).

Der *Sozialindex* umfasst 47 dichotome und dichotomisierte Items, die über den Schüler- und Elternfragebogen zum ersten Messzeitpunkt erhoben wurden. Er schließt zahlreiche Angaben zum ökonomischen, kulturellen und sozialen Kapital ein, wie etwa die Anzahl an Bücher im Haushalt, den Bildungshintergrund der Eltern (ISCED) sowie kulturelle Aktivitäten. Wie bei der Biologiekompetenz wurde ein dichotomes Raschmodell (1PL) berechnet. Die WLE-Reliabilität des Indexes weist einen zufriedenstellenden Wert von .78 auf.

Methode

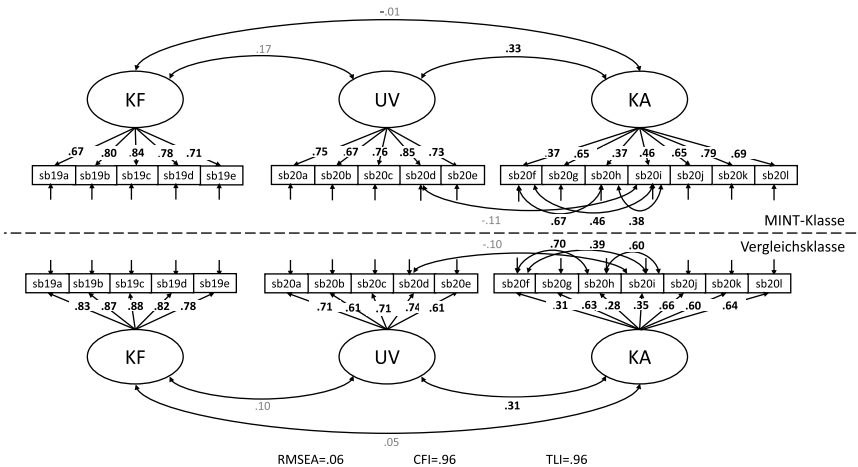
Um Unterschiede in der Wahrnehmung der Unterrichtsqualität im Naturwissenschaftsunterricht von Schülerinnen und Schülern in MINT- und Vergleichsklassen zu ermitteln, werden Mehrgruppenanalysen in *Mplus 7.11* (Muthén/Muthén 2012) umgesetzt. Da es sich bei den verwendeten Items um kategoriale Variablen handelt, die zudem nicht normalverteilt sind, wird der robuste WLSMV-Schätzer gewählt. Anhand des in Mplus implementierten Verfahrens für geclusterte Stichproben wird mit dem Befehl „*type is complex*" im Zusammenhang mit der Angabe der Cluster-Variable (*idclass*) eine Korrektur des Standardfehlers vorgenommen. Für die Beurteilung der Modellgüte werden die Fit-Werte RMSEA: ≤ .06, CFI: ≥ .95 und TLI: ≥ .95 herangezogen (Hu/Bentler 1999). Zunächst wird die Güte des Messinstruments (Unterrichtsqualität) anhand einer konfirmatorischen Faktorenanalyse (CFA) geprüft. Anschließend wird das Vorliegen von Messinvarianz überprüft (*multiple group confirmatory factor analysis*, Wu, Li/Zumbo 2007). Dabei wird untersucht, ob das latente Konstrukt in den Gruppen (MINT-Klassen versus Vergleichsklassen) äquivalent erfasst werden kann. Um latente Mittelwertdifferenzen auf signifikante Unterschiede zu prüfen, muss nach Meredith (1993) mindestens strenge (skalare) Messinvarianz vorliegen. Zur Beurteilung der Modellgüte werden zusätzlich zum X^2-Differenztest und den gängigen Fit-Statistiken der CFI-Differenzentest ($\Delta CFI \leq .01$, Cheung/Rensvold 2002) verwendet. Damit Residualvarianzen der kategorialen abhängigen Variablen geschätzt werden können, wird die *theta parametrisation* gewählt (vgl. Schulte/Nonte/Schwippert 2013). Abschließend werden unter Berücksichtigung skalarer Messinvarianz Kovariaten in das Mehrgruppenmodell aufgenommen, um Aussagen unter Kontrolle von Individualmerkmalen treffen zu können. Hier wurden nur Angaben von Schülerinnen und Schülern verwendet, die an beiden Messzeitpunkten teilgenommen haben. Alle Kovariaten werden zu Zwecken der Vergleichbarkeit z-standardisiert.

6 Ergebnisse

Zunächst wird im Rahmen einer CFA überprüft, ob sich das theoretisch postulierte Modell der Unterrichtsqualität auch empirisch abbilden lässt. Nachdem auf der Grundlage von Modifikationsindizes entschieden wurde, Kovarianzen zwischen Items der latenten Dimension kognitive Aktivierung zuzulassen, weisen die Fit-Werte auf eine moderate Modellanpassung hin (vgl. Abb. 1). Die Zusammenhänge zwischen der Klassenführung und dem unterstützenden Verhalten sowie der Klassenführung und der kognitiven Aktivierung fallen sehr gering aus und sind zudem nicht signifikant. Vergleichsweise hoch fallen die

Zusammenhänge hingegen zwischen dem unterstützenden Verhalten und der kognitiven Aktivierung aus (MINT-Klasse: $r = .31$; $p \leq .001$; Vergleichsklasse: $r = .33$; $p \leq .01$).

Abbildung 1: Konfirmatorische Faktorenanalyse für die drei Basisdimensionen guten Unterrichts ($n = 254$)



Anmerkungen: standardisierte Pfadkoeffizienten. KF = Klassenführung; UV = Unterstützendes Verhalten; KA = Kognitive Aktivierung; fett gedruckte Werte sind signifikant ($p \leq .05$)

In einem nächsten Schritt wird die Überprüfung von Messinvarianz schrittweise über die Vergleichsgruppen hinweg im Rahmen von konfirmatorischen Faktorenanalysen im Mehrgruppenansatz vorgenommen. Der X^2-Differenztest, die absoluten (CFI, TLI) und inkrementellen Fit-Werte (X^2-Test, RMSEA) sowie der ΔCFI verdeutlichen, dass das Modell ein Mindestmaß an skalarer faktorieller Messinvarianz aufweist. Die drei Dimensionen von Unterrichtsqualität werden somit in MINT- und Vergleichsklassen äquivalent erfasst, sodass ein Vergleich der latenten Mittelwertdifferenzen zulässig ist.

Unter Annahme skalarer Invarianz nehmen Schülerinnen und Schüler in den MINT-Klassen einen signifikant höheren Grad an Klassenführung ($\Delta M = 0.57$; $SE = 0.24$; $p < .05$) sowie kognitiver Aktivierung ($\Delta M = 0.18$; $SE = 0.08$; $p < .05$) wahr als Schülerinnen und Schüler in den Vergleichsklassen. Keine Unterschiede können hinsichtlich der wahrgenommenen Unterstützung durch die Lehrkraft beobachtet werden ($\Delta M = 0.34$; $SE = 0.19$; $p \geq .05$). Unter Kontrolle von Individualmerkmalen unterscheidet sich die Wahrnehmung der Unterrichtsqualität von Schülerinnen und Schülern in MINT- und Vergleichsklassen hingegen in keiner der drei Dimensionen (vgl. Tab. 1). Es können jedoch differenzielle Effekte in Bezug auf die Unterrichtsqualitätsdimensionen beobachtet werden. So zeigt sich etwa in den Vergleichsklassen ein systemati-

scher Zusammenhang des Geschlechts mit der Wahrnehmung der kognitiven Aktivierung. Differenzielle Effekte werden an anderer Stelle beschrieben (Stubbe/Nonte/Haas/Krieg 2019). Insgesamt ist der Anteil aufgeklärter Varianz für die einzelnen Dimensionen gering und variiert zudem stark, wie beispielsweise der Grad an kognitiver Aktivierung in den MINT-Klassen (R^2 = .02) und in den Vergleichsklassen (R^2 = .12).

Tabelle 1: Unterrichtsqualität differenziert nach Klassenprofilen unter Aufnahme von Kovariaten (Mehrgruppenmodell; unstandardisierte Koeffizienten; *n* = 228)

	Kognitive Aktivierung	Klassen-führung	Unterstützendes Verhalten
Vergleichsklasse	.00	.00	.00
Sozialindex	-.04	.17	**.17** *
Mädchen	**.17** **	.15	.19
Biologiekompetenz	-.04	.21	.10
Kognitive Fähigkeiten	-.01	.01	-.06
Fähigkeitsselbstkonzept	-.01	.15	**.15** *
Sachinteresse	**.08** *	-.09	.09
R^2	.13	.07	.17
MINT-Klasse	.21	.48	-.07
Sozialindex	.01	**.19** *	.00
Mädchen	.08	.26	**.71** ***
Biologiekompetenz	.03	.20	.21
Kognitive Fähigkeiten	.01	-.08	.13
Fähigkeitsselbstkonzept	-.01	.08	**.36** ***
Sachinteresse	.04	.10	-.01
R^2	.03	.09	.23

Anmerkungen: Alle Kovariaten wurden z-standardisiert. ***$p \leq .001$; **$p \leq .01$; *$p \leq .05$;
Fit-Werte: RMSEA = .05; CFI = .95; TLI = .94

7 Zusammenfassung und Diskussion

Der Beitrag befasst sich mit der Frage, ob sich die Wahrnehmung der Unterrichtsqualität von Schülerinnen und Schülern in MINT-Klassen systematisch von der Wahrnehmung von Schülerinnen und Schülern in den Vergleichsklassen unterscheidet. Da für das Gesamtmodell die Annahme skalarer Messinvarianz bestätigt werden konnte, erfolgte in einem nächsten Schritt die Bestimmung der latenten Mittelwertunterschiede zwischen MINT-Klassen und Vergleichsklassen. Dabei wurden der Grad der Klassenführung sowie die kognitive Aktivierung von Schülerinnen und Schülern in MINT-Klassen im Vergleich zu

Schülerinnen und Schülern in Klassen ohne Profil unter Annahme skalarer Messinvarianz signifikant positiver wahrgenommen. Kein Unterschied konnte hinsichtlich des unterstützenden Verhaltens der Lehrkraft beobachtet werden. Die Unterschiede in der kognitiven Aktivierung und der Klassenführung konnten nicht mehr beobachtet werden, wenn Kovariaten wie beispielsweise das Fähigkeitsselbstkonzept, die Biologiekompetenz und das Geschlecht in das Modell aufgenommen wurden. Damit scheinen individuelle Lernvoraussetzungen von Schülerinnen und Schülern für die Wahrnehmung der Unterrichtsqualität bedeutsamer als die projektorientierte und fächerübergreifende Umsetzung des Unterrichts im Rahmen von MINT-Klassen. Einschränkend muss festgehalten werden, dass in dieser Untersuchung die konkrete Umsetzung von MINT-Profilen (Forscher-, Entdeckerklassen etc.) nicht berücksichtigt werden konnte. Dies liegt unter anderem in der kleinen Stichprobe begründet, die keine weitere Differenzierung zulässt. Aufgrund der geringen Anzahl an MINT- und Vergleichsklassen ist die Durchführung von Mehrebenenanalysen zudem nicht sinnvoll, auch wenn die Unterschiede zwischen den Klassen bedeutsam sein dürften, was sich nicht zuletzt anhand der Variation des Grades an aufgeklärter Varianz zeigt. Möglicherweise findet sich hier einmal mehr ein Hinweis für die Bedeutung von Kompositionseffekten auf Klassenebene (Dumont/Neumann/Maaz/Trautwein 2013). So kann die kognitive Leistungsfähigkeit, das Interesse und das Fähigkeitsselbstkonzept für die Wahrnehmung der Unterrichtsqualität gruppenübergreifend zwar nicht direkt, jedoch vermittelt über die Klassenkomposition verantwortlich sein. Diese Effekte werden vermutlich stark konfundiert sein und sollten in zukünftigen Studien, die größere Stichprobenumfänge realisieren können, systematisch untersucht werden. Des Weiteren wurden die Schülerinnen und Schüler erstmalig zu Beginn der sechsten Jahrgangsstufe retrospektiv zur Unterrichtsqualität in der fünften Jahrgangsstufe befragt. Entwicklungen über die Zeit lassen sich anhand dieser Daten daher nicht abbilden. Mit Hilfe der Daten des dritten Messzeitpunktes können die hier präsentierten Ergebnisse auf differenzielle Effekte im Zusammenhang mit MINT-Klassen und Unterrichtsqualitätsdimensionen auch im Längsschnitt untersucht werden. Interessant ist dabei unter anderem, ob der Besuch einer MINT-Klasse im Zeitverlauf mit dem naturwissenschaftlichen Interesse der Schülerinnen und Schüler zusammenhängt.

Literatur

Anger, C., Koppel, O. & Plünnecke, A. (2018). MINT-Frühjahrsreport 2018. MINT – Offenheit, Chancen, Innovationen. Zugriff am 23.05.2019 unter https://www.iwkoeln.de/fileadmin/user_upload/Studien/Gutachten/PDF/2018/MINT-Fr%C3%BChjahrsreport_2018neu.pdf

Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M. & Tsai, Y.-M. (2009). Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV). Dokumentation der Erhebungsinstrumente. Berlin: Max-Planck-Institut für Bildungsforschung.

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

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9(2), 233–255.

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. Review of Educational Research, 77(1), 113–143.

Daniels, Z. (2008). Entwicklung schulischer Interessen im Jugendalter (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 69). Münster: Waxmann.

Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. Psychologie in Erziehung und Unterricht, 60, 163–183.

Fauth, B., Decristan, J., Riesner, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. Learning and Instruction, 29, 1–9.

Hammer, S., Reiss, K., Lehner, M. C., Heine, J.-H., Sälzer, C. & Heinze, A. (2016). Mathematische Kompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), PISA 2015. Eine Studie zwischen Kontinuität und Innovation (S. 219–247). Münster: Waxmann.

Heller, K. A. & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Beltz Test.

Helmke, A. (2003). Unterrichtsqualität erfassen, bewerten, verbessern. Seelze-Velber: Kallmeyer.

Holzberger, D. & Kunter, M. (2016). Unterricht aus der Perspektive der Pädagogischen Psychologie und der empirischen Unterrichtsforschung. In J. Möller, M. Köller & T. Riecke-Baulecke (Hrsg.), Basiswissen Lehrerbildung. Schule und Unterricht – Lehren und Lernen (S. 39–52). Seelze-Velber: Kallmeyer.

Höffler, T. N., Lüthjohann, F. & Parchmann, I. (2014). Welche Wirkungen erzielt ein Anfangsunterricht. Zeitschrift für Didaktik der Naturwissenschaften, 20(1), 87–99.

Hu, L. T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6(1), 1–55.

Klieme, E. & Rakoczy, K. (2003). Unterrichtsqualität aus Schülerperspektive: kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland (S. 333–359). Opladen: Leske & Budrich.

Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. Zeitschrift für Pädagogik, 54(2), 222–237.

Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung im internationalen Vergleich. In J. Baumert & E. Klieme (Hrsg.), TIMSS – Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente (S. 43–57). Bonn: Bundesministerium für Bildung und Forschung.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. Psychometrika, 58(4), 525–543.

Muthén, L. K. & Muthén, B. O. (2012). Mplus user's guide (7. Aufl.). Los Angeles, CA: Muthén & Muthén.

Pfenning, U. & Renn, O. (2012). Wissenschafts- und Technikbildung auf dem Prüfstand. Zum Fachkräftemangel und zur Attraktivität der MINT-Bildung und -berufe im europäischen Vergleich. Baden-Baden: Nomos.

Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K. & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. Learning and Instruction, 31, 2–12.

R Development Core Team (2017). R: A language and environment for statistical computing [Computer software]. Vienna: R Foundation for Statistical Computing. Zugriff am 23.05.2019 unter http://www.R-project.org

Robitzsch, A., Kiefer, T. & Wu, M. (2017). TAM: Test analysis modules [Computer software]. Zugriff am 23.05.2019 unter https://CRAN.R-project.org/package=TAM

Rost, D. H. & Sparfeldt, J. R. (2002). Facetten des schulischen Selbstkonzepts: Ein Verfahren zur Messung des differentiellen Selbstkonzepts schulischer Leistungen und Fähigkeiten (DISK-Gitter). Diagnostica, 48(3), 130–140.

Schiepe-Tiska, A., Simm, I. & Schmidtner, S. (2016). Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), PISA 2015. Eine Studie zwischen Kontinuität und Innovation (S. 99–132). Münster: Waxmann.

Schulte, K., Nonte, S. & Schwippert, K. (2013). Die Überprüfung von Messinvarianz in international vergleichenden Schulleistungsstudien am Beispiel der Studie PIRLS. Zeitschrift für Bildungsforschung, 3(2), 99–118.

Seidel, T., Reinhold, S., Holzberger, D., Mok, S. Y., Schiepe-Tiska, A. & Reiss, K. (2016). Wie gelingen MINT-Schulen? Anregungen aus Forschung und Praxis. Münster: Waxmann.

Stubbe, T. C., Nonte, S., Haas, M. & Krieg, M. (2019). Musik- und MINT-Profile an Niedersächsischen Gymnasien und integrierten Gesamtschulen. Ergebnisse der Studie ProBiNi. Manuskript in Vorbereitung.

van Buuren, S. & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. Journal of Statistical Software, 45(3), 1–67.

Wu, A., Li, Z. & Zumbo, B. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. Practical Assessment, Research & Evaluation, 12(3), 1–26.

Gerechtigkeit von Leistungsbewertung im Urteil von Oberstufenschülerinnen und -schülern

Carolin Körner, Sabine Gruehn und Sylvia Rahn

Abstract

Während zu Leistungsbeurteilung in einer professionalisierungstheoretischen Perspektive als Bestandteil diagnostischer Kompetenz von Lehrkräften bereits zahlreiche Studien existieren, beschäftigen sich nur wenige Studien mit Leistungsbeurteilung als (normatives) Merkmal von Unterrichtsqualität. Der vorliegende Beitrag befasst sich mit der zuletzt genannten Perspektive und geht der Frage nach, ob das Gerechtigkeitsempfinden von Schülerinnen und Schülern fachspezifisch variiert und ob diese schülerperzipierte Gerechtigkeit mit der Fachnote des letzten Zeugnisses sowie der Unterrichtszufriedenheit der Lernenden in Verbindung steht. Dazu werden die im Rahmen der DFG-geförderten Studie „Determinanten des Schülerfeedbacks" erhobenen quantitativen Daten von 2234 Schülerinnen und Schülern der elften Jahrgangsstufe deutscher Gesamtschulen und Berufskollegs mittels Varianz- und Regressionsanalysen ausgewertet. Die Ergebnisse zeigen deutliche und signifikante Zusammenhänge zwischen dem Gerechtigkeitsempfinden und der Zufriedenheit mit dem Unterricht sowie auch kleinere Effekte zwischen dem Gerechtigkeitsempfinden und der letzten Zeugnisnote. Für das Geschlecht der Schülerinnen und Schüler sowie das Fach lassen sich hingegen keine entsprechenden Effekte nachweisen.

1 Problemstellung

Seitdem die neueren Leistungsvergleichsstudien wie beispielsweise PISA (Baumert et al. 2001) herausstellen konnten, dass z. T. deutliche Diskrepanzen zwischen den gemessenen Testleistungen und den von Schülerinnen und Schülern berichteten Zeugnisnoten bestehen, werden Fragen einer gerechten Leistungsbeurteilung wieder verstärkt diskutiert (z. B. Dalbert 2013; Lin-Klitzing/Di Fuccia/Gaube 2016; Oelkers 2016).

Mehrere Studien weisen darauf hin, dass Gerechtigkeits- und Ungerechtigkeitserfahrungen im Schulalltag in vielerlei Hinsicht eine bedeutende Auswirkung auf das Leben der Schülerinnen und Schüler haben: Beispielsweise sind Gerechtigkeitserfahrungen wichtig für das Selbstkonzept der Schülerinnen und Schüler und ihre Selbstwirksamkeitserwartungen (Israelashvili 1997; Beutel/Beutel 2010), aber auch für ihre Motivation und Anstrengungsbereitschaft (Beutel 2005; Resh/Sabbagh 2009).

Die bisherige Forschung zur Gerechtigkeit von Leistungsbeurteilung stützt sich zumeist auf Erhebungen durch externe Beobachterinnen und Beobachter, auf die Befragung von Lehrkräften oder Eltern. Offen geblieben ist bisher dabei, wie die Schülerinnen und Schüler selbst die Gerechtigkeit ihrer Noten bewerten (Beutel/Vollstaedt 2002, S. 592). Aus diesem Grund widmet sich die vorliegende Untersuchung maßgeblich der Sicht der Schülerinnen und Schüler.

Zunächst werden in diesem Beitrag zwei Perspektiven, die in der Literatur zur Beurteilung der Gerechtigkeit von Leistungsbeurteilung in den Blick genommen werden, herausgearbeitet und im Hinblick auf ihre Bedeutung für die eigene Untersuchung beleuchtet. Vor diesem Hintergrund werden die Forschungsfragen benannt sowie die Stichprobe und das methodische Vorgehen dargelegt. Im Anschluss daran erfolgt die Präsentation der Ergebnisse. Der Beitrag endet mit einer Diskussion über die Bedeutung der Ergebnisse im Hinblick auf zukünftige Forschung in diesem Feld sowie im Hinblick auf unterrichtliches Handeln von Lehrkräften.

2 Theoretischer Hintergrund

Zur Beurteilung der Gerechtigkeit von Leistungsbewertungen können zwei Perspektiven eingenommen werden: So kann die Gerechtigkeit von Leistungsbeurteilungen einerseits aus einer professionalisierungstheoretischen Perspektive als Bestandteil diagnostischer Kompetenz von Lehrkräften betrachtet werden. Andererseits lässt sie sich unterrichtstheoretisch als (normatives) Merkmal von Unterrichtsqualität konzipieren, in welcher der Perspektive der Schülerinnen und Schüler eine besondere Bedeutung zukommt (Kyburz-Graber 2002).

Bezüglich der erstgenannten Perspektive wird der diagnostischen Kompetenz von Lehrkräften seit einigen Jahren verstärktes Interesse beigemessen. Dieses resultiert nicht nur aus den wiederholt belegten unbefriedigenden Leistungen eines Großteils der deutschen Schülerschaft (Bos et al. 2003), sondern auch aus der geringen Treffsicherheit von Lehrkräften bei der Identifikation von Risikoschülerinnen und -schülern (Baumert et al. 2001). Dabei ist die Beurteilung der Leistungen von Schülerinnen und Schülern eine wesentliche Aufgabe von Lehrkräften, die aufgrund ihrer gesellschaftlichen Funktionen der Qualifikation, Allokation und Selektion in besonderer Weise der Vergleichbar-

keit von Leistungen verpflichtet ist. Gerade da Schulnoten wesentlich über die Lebensverläufe junger Menschen entscheiden (Sacher 2009), wird dieser Perspektive, die eine objektiv gerechtere Leistungsbewertung zum Ziel hat, in der Forschung bereits viel Aufmerksamkeit geschenkt (für einen Überblick s. Südkamp/Praetorius 2017).

Innerhalb der Forschung zur zweiten, unterrichtstheoretischen Perspektive lässt sich mit Meyer (2004) eine gerechte Leistungsbewertung als ein Qualitätsmerkmal des Unterrichts beschreiben. Unter seinen 10 Merkmalen guten Unterrichts findet sich u. a. auch der Aspekt der transparenten Leistungserwartungen. Diese speisen sich nach Meyer aus den sechs Aspekten (1) verständliche Aufgabenstellung, (2) Orientierung am Lernstand des Lernenden sowie (3) an den Bildungsstandards, (4) Offenlegung der Bezugsnorm bei der Leistungsbeurteilung, (5) Förderhaltung der Lehrkraft sowie (6) zügige Rückmeldungen. Sind diese Aspekte erfüllt, führt dies – neben einer stärkeren Identifikation der Schülerin/des Schülers mit der gestellten Aufgabe und ihrer/seiner erhöhten Leistungsbereitschaft – zu höherer Beurteilungsgerechtigkeit (ebd., S. 116). Auch Dalbert (2013) stellt in ihren Gerechtigkeitsdimensionen heraus, dass Transparenz und Offenheit als Teil der interaktionalen Gerechtigkeit zentrale Voraussetzungen für hohe Gerechtigkeitskognitionen sind. Im Zuge der interaktionalen Gerechtigkeit betont Dalbert außerdem die interpersonale Gerechtigkeit. Hiernach beurteilen Menschen einen Vorgang dann als gerecht, wenn sie mit Würde, Respekt und Höflichkeit behandelt werden (vgl. ebd.). Dieser Aspekt findet sich auch in dem *unterstützenden, schülerorientierten Sozialklima* in den *Grunddimensionen der Unterrichtsqualität* von Klieme und Rakoczky (2008) wieder.

In der Forschung zur Unterrichtsqualität hat sich – nicht zuletzt initiiert durch die Arbeit von Gruehn (2000) – eingebürgert, die Unterrichtsqualität nicht nur über Beobachtungsdaten zu erheben, sondern auch die Perspektive der Schülerinnen und Schüler einzubeziehen. Diese ist vor allem bei jenen Qualitätsmerkmalen unentbehrlich, die am besten durch Schülerinnen und Schüler beurteilt werden können. Zu diesen zählen – neben z. B. dem angemessenen Lerntempo – auch transparente Leistungserwartungen, die sich in der Wahrnehmung einer gerechten Leistungsbeurteilung durch die Schülerinnen und Schüler wiederfinden. Diesem Kriterium ist allerdings in den vorliegenden Studien zur Wahrnehmung der Unterrichtsqualität durch Schülerinnen und Schüler (zusammenfassend s. Göllner et al. 2016) bislang kaum Aufmerksamkeit geschenkt worden, sodass sich die Frage stellt, wie bedeutsam überhaupt dieses Kriterium für die Einschätzung der Unterrichtsqualität in der Wahrnehmung der Schülerinnen und Schüler insgesamt ist.

3 Fragestellungen

Bislang gibt es nur wenige empirische Befunde zu der Frage, wie gerecht Schülerinnen und Schüler ihre Leistungsbeurteilung empfinden. Bezüglich der oben beschriebenen ersten, professionalisierungstheoretischen Perspektive wird in empirischen Studien immer wieder angemerkt, dass schulische Leistungsüberprüfungen den testtheoretischen Gütekriterien – allen voran der Objektivität – nur unzureichend genügen (z. B. Brüggelmann et al. 2006), sodass vermutet werden kann, dass dieses auch zu einer geringen Gerechtigkeitsbeurteilung durch die Schülerinnen und Schüler führt. Die wenigen vorliegenden Studien (Dalbert/Sallay 2004) zeigen jedoch, dass Schülerinnen und Schüler die Gerechtigkeit ihrer Leistungsbewertung im Allgemeinen als hoch einschätzen. Dabei wurde das Gerechtigkeitsempfinden jedoch nicht fächerdifferenziert erhoben, sodass zu prüfen ist, ob es je nach Unterrichtsfach unterschiedlich ausfällt. So könnte es z. B. sein, dass in Fächern, in denen vermeintlich objektivere Bewertungsmaßstäbe zur Verfügung stehen, wie z. B. in Mathematik, die Leistungsbewertung als gerechter wahrgenommen wird als in anderen Fächern.

Ebenfalls wurde bislang kaum untersucht, ob der Leistungsstand des Schülers/der Schülerin einen Einfluss auf das Gerechtigkeitsempfinden hat. Mit der Theorie der Kausalattribution, nach der Menschen Erfolg sich selbst, Misserfolg hingegen eher externalen Faktoren zuschreiben, kann argumentiert werden, dass Schülerinnen und Schüler dazu tendieren, schlechte Noten eher als Ergebnis einer ungerechten Leistungsbewertung zu interpretieren. In der Tat konnten Maes & Kals (2004) zeigen, dass ein höheres Gerechtigkeitsempfinden mit guten Noten einhergeht. Da zudem die Noten in verschiedenen Fächern geschlechtsspezifisch ausfallen (vgl. Bernhart/Konrad 2014), kann darüber hinaus vermutet werden, dass auch das Gerechtigkeitsempfinden geschlechtsspezifisch variiert.

Vor dem Hintergrund des referierten Forschungsstands und den aufgezeigten Forschungslücken widmet sich die folgende Studie zwei Fragen:

1. Wie gerecht empfinden Schülerinnen und Schüler die Leistungsbewertung im Fach Deutsch und Mathematik? Lassen sich hierbei Geschlechtseffekte erkennen?
2. In welchem Zusammenhang steht die von Schülerinnen und Schülern wahrgenommene Gerechtigkeit von Leistungsbeurteilungen mit a) der Fachnote im letzten Zeugnis sowie b) der Unterrichtszufriedenheit der Schülerinnen und Schüler insgesamt?

4 Methodisches Vorgehen

Zur Erforschung der durch die Schülerinnen und Schüler subjektiv erlebten Gerechtigkeit bezüglich ihrer Leistungsbewertung sowie zur Einschätzung der Unterrichtszufriedenheit wurde ein standardisierter Fragebogen eingesetzt. Die Unterrichtszufriedenheit von Schülerinnen und Schülern als ein globales Maß der Unterrichtsqualität zu verwenden, erscheint vor dem Hintergrund der bisherigen Forschung zu Einschätzungen der Unterrichtsqualität durch Schülerinnen und Schüler als gerechtfertigt, haben diesbezügliche Studien doch ergeben, dass das Urteil der Schülerinnen und Schüler bis zu einem gewissen Grad als zuverlässig und gültig gelten kann (Göllner et al. 2016; Gruehn 2000; Prenzel/Lankes 2013).

Stichprobe

Die verwendeten Daten wurden im Rahmen des von der DFG geförderten Projekts „Determinanten des Schülerfeedbacks: Eine fächervergleichende Analyse von Schüleraussagen zur Unterrichtsqualität an allgemeinbildenden und beruflichen Gymnasien" erhoben. Dieses am Ende des Schuljahres 2012/13 durchgeführte Projekt umfasste sowohl eine Schülerinnen- und Schüler- als auch eine Lehrkräftebefragung. Für die vorliegende Studie wurden nur die Daten der Schülerinnen und Schüler herangezogen. Insgesamt nahmen 2234 Schülerinnen und Schüler (49,4 % Schülerinnen) aus 10 Gesamtschulen und 39 Berufskollegs in NRW teil; die Schülerinnen und Schüler befanden sich zum Zeitpunkt der Befragung am Ende der 11. Jahrgangsstufe.

Beschreibung des Instruments

Die Unterrichtszufriedenheit (Eigenentwicklung) wurde mit 3 Items (Beispielitem „Mit diesem Unterricht bin ich sehr zufrieden"; $\alpha_{D=}.81$; $\alpha_{M=}.83$), das Gerechtigkeitsempfinden in Anlehnung an Seeber/Squarra (2003) sowie LASSO (Saldern/Littig 1987) mit fünf Items (Beispielitem „Meine Leistungen werden durch meine Lehrerin/meinen Lehrer richtig eingeschätzt."; $\alpha_{D=}.86$; $\alpha_{M=}.89$) auf einer vierstufigen Likert-Skala mit den Antwortvorgaben „trifft nicht zu" bis „trifft zu" erhoben. Die internen Konsistenzen liegen für beide Skalen in jeweils beiden Fächern im guten Bereich. Zudem wurde die letzte Zeugnisnote (= Halbjahresnote in Klasse 11) in den Fächern Deutsch und Mathematik erfragt.

Auswertungsmethoden

Mit Hilfe von T-Tests und Varianzanalysen wurde geprüft, ob es Unterschiede im fachspezifischen Gerechtigkeitsempfinden zwischen Schülerinnen und Schülern sowie zwischen unterschiedlich leistungsstarken Schülerinnen und Schülern gibt. Da aufgrund der Stichprobengröße bereits kleine Unterschiede schon signifikant werden, wurde neben einem strengen Signifikanzniveau von $p = .001$ zusätzlich die gemeinsame Varianz zwischen den jeweils betrachteten Variablen als ein Maß der praktischen Bedeutsamkeit herangezogen. Das strenge Signifikanzniveau dient zudem auch dazu, die hierarchische Struktur der Stichprobe (Auswahl von Schulen statt Schülerinnen und Schülern, die zu einer geschichteten Stichprobe von Schülerinnen und Schülern in Klassen in Schulen führt) zu berücksichtigen.

Zur Überprüfung der Zusammenhangsstruktur zwischen der durch die Schülerinnen und Schüler wahrgenommenen Gerechtigkeit und der Zufriedenheit mit dem Unterricht sowie der letzten Zeugnisnote wurden Korrelations- und Regressionsanalysen berechnet.

5 Ergebnisse

Um zu klären, wie gerecht Schülerinnen und Schüler die Leistungsbewertung im Fach Deutsch und Mathematik empfinden, werden zunächst die Mittelwerte und Standardabweichungen des selbstberichteten Gerechtigkeitsempfindens berichtet: So fällt einerseits auf, dass beide Werte deutlich über dem Skalenmittel (= 2.5) liegen (vgl. Tab. 1), sowie andererseits, dass der Mittelwert im Fach Deutsch signifikant höher ist ($t(4475) = 12.033$, $p < .001$).

Das Gerechtigkeitsempfinden wurde außerdem mittels einer 2x2-faktoriellen Varianzanalyse mit Messwiederholung untersucht. Es ergab sich ein signifikanter Haupteffekt für den Faktor Fach ($F(1,2210) = 153.912$, $p < .001$, $\eta^2 = .065$), jedoch keine signifikanten Effekte für das Geschlecht ($F(1,2210) = 1.719$, $p = .190$, $\eta^2 = .001$) und die Fach-Geschlecht-Interaktion ($F(1,2210) = .515$, $p = .473$, $\eta^2 = .000$).

Die erste Forschungsfrage kann damit so beantwortet werden, dass die Schülerinnen und Schüler die Leistungsbeurteilung analog zu den Befunden von Dalbert/Sallay (2004) im Allgemeinen als relativ gerecht einschätzen und das Unterrichtsfach einen signifikanten, wenngleich eher geringen, das Geschlecht der Schülerinnen und Schüler hingegen keinen Einfluss auf die Bewertung der Benotungsgerechtigkeit aufweisen.

Tabelle 1: Mittelwerte und Standardabweichungen der Skalen Gerechtigkeitsempfinden und Unterrichtszufriedenheit sowie der Zeugnisnote

	Mathematik				Deutsch			
	männlich		weiblich		männlich		weiblich	
	MW	SD	MW	SD	MW	SD	MW	SD
GE*	2.81	.49	2.82	.44	2.96	.42	2.99	.40
ZN	2.84	1.09	2.81	1.06	2.98	.83	2.74	.76
UZ*	2.82	1.00	2.75	.97	2.93	.86	2.96	.86

GE = Gerechtigkeitsempfinden, ZN = Zeugnisnote, UZ = Unterrichtszufriedenheit
* Die Antwortskala reicht von 1–4.

Da sich die zweite Forschungsfrage auf den Zusammenhang der letzten Zeugnisnote und der Unterrichtszufriedenheit mit dem fachspezifischen Gerechtigkeitsempfinden bezieht, werden auch diese beiden weiteren Variablen zunächst deskriptiv betrachtet.

Die Notenmittelwerte fallen für beide Fächer gleich aus, jedoch ist die Standardabweichung im Fach Mathematik deutlich größer als im Fach Deutsch (vgl. Tab. 1), was sich in der Verteilung der Noten über die verschiedenen Notenstufen bemerkbar macht: Während im Fach Deutsch fast die Hälfte aller Schülerinnen und Schüler die Note *befriedigend* auf dem letzten Zeugnis stehen hatte, hatte dies im Fach Mathematik nur etwa ein Drittel, dafür gab es in Mathematik sowohl mehr Einsen als auch mehr Vieren und Fünfen als im Fach Deutsch. Bezüglich möglicher Geschlechtsunterschiede kann konstatiert werden, dass sich für beide Fächer keine signifikanten geschlechtsspezifischen Unterschiede in der Notenverteilung finden lassen, obwohl sich im Fach Deutsch die Notenmittelwerte zwischen den Geschlechtern deutlich unterscheiden ($t(2119) = .083$, $p < .001$).

Zur weiteren Untersuchung können außerdem die Interkorrelationen zwischen den Variablen Zeugnisnote, Unterrichtszufriedenheit und Gerechtigkeitsempfinden betrachtet werden. Zwischen den Variablen Unterrichtszufriedenheit und Gerechtigkeitsempfinden lassen sich in beiden Fächern hohe Korrelationen finden ($r_{D=}.49^{**}$; $r_{M=}.53^{**}$), zwischen der Zeugnisnote und dem Gerechtigkeitsempfinden außerdem im Fach Mathematik ($r_{M=}-.40^{**}$)[1]. Die Korrelation zwischen diesen Variablen ist im Fach Deutsch deutlich geringer ($r_{D=}-.24^{**}$). Für beide Fächer besteht zudem ein ebenfalls geringer Zusammenhang zwischen der Note und der Zufriedenheit mit dem Unterricht ($r_{D=}-.20^{**}$; $r_{M=}-.29^{**}$).

1 Das negative Vorzeichen resultiert aus der Polarisierung der Noten (1 = sehr gut, 6 = ungenügend) im Kontrast zu den verwendeten Skalen (trifft nicht zu = 1, trifft zu = 4)

Die Ergebnisse der linearen Regression zeigen, dass ein deutlicher Zusammenhang zwischen einer als gerecht empfundenen Leistungsbewertung und der Zufriedenheit mit dem Unterricht sowie der Zeugnisnote besteht, wobei letzterer eine deutlich geringere Bedeutung zukommt (s. Tab. 2). Für das Geschlecht ließ sich kein signifikanter Regressionskoeffizient ermitteln, so dass es im Modell nicht weiter berücksichtigt wird. Darüber hinaus sind keine Unterschiede zwischen den beiden Fächern feststellbar.

Tabelle 2: Prädiktoren des Gerechtigkeitsempfindens

	Mathematik		Deutsch	
	Modell 1	Modell 2	Modell 1	Modell 2
Prädiktoren	β	β	β	β
UZ	.53***	.49***	.49***	.48***
ZN		-.09***		-.09***
(korr.) R^2	.28	.29	.24	.25

UZ = Unterrichtszufriedenheit, ZN = Zeugnisnote
***$p < .001$

Beim Vergleich von Modell 1 und 2 der schrittweisen Regression wird allerdings deutlich, dass die Hinzunahme der Zeugnisnote im zweiten Modell kaum zur weiteren Aufklärung der Varianz im Gerechtigkeitsempfinden beiträgt – ihr spezifischer Varianzanteil beträgt jeweils nur 1 %. Die leichte bis marginale Veränderung des Regressionskoeffizienten der Unterrichtszufriedenheit im zweiten Modell verweist zudem auf eine eher geringe Konfundierung der beiden Prädiktorvariablen. Eine Varianzzerlegung zeigt, dass 20 % bzw. 21 % der Unterschiede im Gerechtigkeitsempfinden in Mathematik bzw. Deutsch allein auf Differenzen in der Unterrichtszufriedenheit zurückgehen – nur 8 % (Mathematik) bzw. 4 % (Deutsch) werden gemeinsam von beiden Prädiktoren erklärt.

6 Diskussion

Dieser Beitrag hat gezeigt, dass das Gerechtigkeitsempfinden im Hinblick auf die Leistungsbewertung – in Übereinstimmung mit den bisherigen Befunden – auch bei Elftklässlerinnen und Elftklässlern eher hoch ausfällt, das Geschlecht hierbei keinen Einfluss hat, wohl aber das Unterrichtsfach: In Deutsch wird die Leistungsbewertung etwas gerechter beurteilt als in Mathematik. Darüber hinaus lässt sich ein substantieller Zusammenhang mit der Zufriedenheit mit dem Mathematik- bzw. Deutschunterricht und in geringerem Ausmaß auch mit der Fachnote belegen. Ob dabei die Unterrichtszufriedenheit das Gerechtigkeits-

empfinden beeinflusst oder ob eine umgekehrte Wirkungsrichtung angenommen werden muss, kann mit den vorliegenden Daten aufgrund ihres Querschnittsdesigns nicht bestimmt werden. Festhalten lässt sich allerdings, dass die Gerechtigkeit der Leistungsbeurteilung ein ganz wesentlicher Bestandteil der allgemeinen Unterrichtszufriedenheit von Schülerinnen und Schüler ist – die Korrelation liegt in einer Größenordnung, wie sie auch für andere Prozessmerkmale des Unterrichts (z. B. Motivierungsqualität oder Erklärkompetenz, vgl. Bertram, im Druck) gefunden wurde. Dies bedeutet, dass eine gerechte Leistungsbewertung auch aus der Sicht von Schülerinnen und Schülern ein bedeutsames Merkmal von Unterrichtsqualität darstellt.

Auch wenn Elftklässlerinnen und Elftklässler sowohl in Deutsch als auch in Mathematik die Leistungsbewertung ihrer Fachlehrkraft im Durchschnitt als eher gerecht einschätzen, so zeigen die Daten auch, dass sich leistungsschwächere Schülerinnen und Schüler – vor allem im Fach Mathematik – ungerechter bewertet fühlen als leistungsstärkere. Wenngleich davon auszugehen ist, dass Kausalattributionen hier eine Rolle spielen, könnten doch auch weitere Faktoren diesen Zusammenhang erklären. So könnte ein weiterer Grund auch darin liegen, dass sich leistungsschwächere Schülerinnen und Schüler aufgrund nicht angemessener Bewertungsmaßstäbe der Lehrkräfte ungerecht bewertet fühlen: So ist z. B. gut erforscht, dass Schülerinnen und Schüler die Orientierung an der kriterialen Bezugsnorm am gerechtesten bewerten, wobei sie aber diese – dem Differenzprinzip entsprechend – um die individuelle Bezugsnorm ergänzt wissen möchten (Dalbert 2013). Lehrkräfte hingegen wählen oft die soziale Bezugsnorm als Vergleichsmaßstab, bei der vor allem leistungsschwache Schülerinnen und Schüler schlecht bewertet werden, auch wenn sie Leistungsfortschritte erzielt haben sollten. Darüber hinaus könnten auch Beurteilungsfehler der Lehrkraft (z. B. der Halo-Effekt) ein Grund dafür sein, dass sich leistungsschwächere Schülerinnen und Schüler ungerechter bewertet fühlen, zumal Noten nicht nur auf schriftlichen, sondern auch auf mündlichen Leistungen basieren, die aber im Vergleich zu den schriftlichen oft nicht reliabel und valide gemessen werden (können) und dadurch anfällig für Beurteilungsfehler sind. Hier fehlen weiterhin Studien, die diese Fragen beantworten können.

Literatur

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W. et al. (Hrsg.) (2001): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.

Bernhart, D. & Konrad, K. (2014): Leistungsbewertung bei erweitertem Lern- und Leistungsbegriff. Qualitativ hochwertige Formen der Bewertung umsetzen. In G. Birkigt (Hrsg.), Leistung zwischen Anspruch und Wirklichkeit. Wohin geht Schule? (S. 177–204). Stuttgart: Raabe.

Beutel, S.-I. (2005): Zeugnisse aus Kindersicht. Kommunikationskultur an der Schule und Professionalisierung der Leistungsbeurteilung. Weinheim und München: Juventa.

Beutel, S.-I. & Beutel, W. (2010): Beteiligt oder bewertet? Leistungsbeurteilung und Demokratiepädagogik. Schwalbach: Wochenschau Verlag.

Beutel, S.-I. & Vollstaedt, W. (2002): Kinder als Experten für Leistungsbewertung. Zeitschrift für Pädagogik, 4(48), 591–613.

Bos, W., Lankes, E.-M., Schwippert, K., Valtin, R., Voss, A., Badel, I. et al. (2003): Lesekompetenzen deutscher Grundschülerinnen und Grundschüler am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valentin (Hrsg.), Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich (S. 69–134). Münster: Waxmann.

Brüggelmann, H., Backhaus, A., Brinkmann, E., Coelen, H., Franzkowiak, T., Knorre, S. et al. (2006): Sind Noten nützlich – und nötig? Ziffernzensuren und ihre Alternativen im empirischen Vergleich. Eine wissenschaftliche Expertise des Grundschulverbandes erstellt von der Arbeitsgruppe Primarstufe an der Universität Siegen. Frankfurt a. M.: Grundschulverband – Arbeitskreis Grundschule e.V.

Dalbert, C. (2013): Gerechtigkeit in der Schule. Wiesbaden: Springer VS.

Dalbert, C. & Sallay, H. (Hrsg.) (2004): The Justice Motive in Adolescence and Young Adulthood. Origins and consequences. New York: Routledge.

Göllner, R. V., Wagner, W. V., Klieme, E. V., Lüdtke, O. V., Nagengast, B. V. & Trautwein, U. V. (2016): Erfassung der Unterrichtsqualität mithilfe von Schülerurteilen. Chancen, Grenzen und Forschungsperspektiven. In Bundesministerium für Bildung und Forschung (Hrsg.), Forschungsvorhaben in Ankopplung an Large-Scale-Assessments (S. 63–82). Bielefeld: wbv.

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

Israelashvili, M. (1997): Situational determinants of school students feeling of injustice. Elementary School Guidance and Counseling, 31, 283–292.

Klieme, E., Rakoczky, K. (2008): Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. Zeitschrift für Pädagogik, 54(2), 222–237.

Kyburz-Graber, R. (2002): Die Beurteilung von Lehrerinnen und Lehrern – Ein Beitrag zur Qualitätsentwicklung der Schule? In H. Rhyn (Hrsg.), Beurteilung macht Schule. Leistungsbeurteilung von Kindern, Lehrpersonen und Schule (S. 37–52). Bern: Haupt.

Lin-Klitzing, S., Di Fuccia, D. & Gaube, T. (Hrsg.) (2016): Leistungsstandards und Leistungsbewertung an Gymnasien und Universitäten. Beiträge zur (nicht) vorhandenen Passung (Gymnasium – Bildung – Gesellschaft). Bad Heilbrunn: Verlag Julius Klinkhardt.

Maes, J. & Kals, E. (2004): Two facets of the belief in a just world and achievement behaviour at school. In C. Dalbert & H. Sallay (Hrsg.), The Justice Motive in Adolescence and Young Adulthood. Origins and consequences (S. 135–152). New York: Routledge.

Meyer, H. (2004): Was ist guter Unterricht? (2., durchges. Aufl.). Berlin: Cornelsen Scriptor.

Oelkers, J. (2016, September): Leistungsbeurteilung: Ein Widerspruch zur Bildungsgerechtigkeit? 7. Kieler Schulleitersymposium, Kiel.

Prenzel, M. & Lankes, E.-M. (2013): Was können Schülerinnen und Schüler über ihren Unterricht sagen? Ein Blick in die Schülerfragebogen von internationalen Vergleichsstudien. In N. McElvany & H. G. Holtappels (Hrsg.), Empirische Bildungsforschung. Theorien, Methoden, Befunde und Perspektiven (S. 93–107). Münster: Waxmann.

Resh, N. & Sabbagh, C. (2009): Justice in teaching. In Saha, L. J. & Dworkin, A. G. (Hrsg.), International Handbook of Research on Teachers and Teaching. Part Two. New York: Springer.

Sacher, W. (2009): Leistungen entwickeln, überprüfen und beurteilen. Bewährte und neue Wege für die Primar- und Sekundarstufe (5., überarbeitete und erweiterte Auflage). Bad Heilbrunn: Verlag Julius Klinkhardt. Verfügbar unter https://www.content-select.com/index.php?id =bib_view&eaN = 9783781552036

Saldern, M. von & Littig, K. E. (1987): Landauer Skalen zum Sozialklima. Weinheim.

Seeber, S. & Squarra, D. (2003): Lehren und lernen in beruflichen Schulen. Schülerurteile zur Unterrichtsqualität (Berufliche Bildung im Wandel, Bd. 3). Frankfurt a. M.: Lang.

Südkamp, A. & Praetorius, A.-K. (Hrsg.) (2017): Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 94). Münster: Waxmann.

Schriftliches Argumentieren im Deutschunterricht

Entwicklung eines Kompetenztests für Lernende der Sekundarstufe I

Julian Kempf und Dorit Bosse

Abstract

Präsentiert wird die Entwicklung und der Einsatz eines Kompetenztests zum schriftlichen Argumentieren im Fach Deutsch für Schülerinnen und Schüler der Jahrgangsstufen sieben bis zehn. Bisher liegt zu dem Themenbereich kein standardisierter deutschsprachiger Test vor. Für einen qualitativ hochwertigen Unterricht ist entscheidend, die Kompetenzen der Lernenden zu berücksichtigen. Der entwickelte Test kann dazu beitragen, Lernvoraussetzungen und -erfolge zu identifizieren, um die Gestaltung des Unterrichts daran auszurichten. Im Beitrag wird aufgezeigt, wie der Test für ein Deutsch-Schweizer Forschungsprojekt entwickelt wurde. Entsprechend wurden für die Testkonstruktion die für den Bereich „Schriftliches Argumentieren" relevanten Bildungsstandards der beiden Länder berücksichtigt. Die Testgüte wird bezogen auf offene, halboffene und geschlossene Aufgaben präsentiert ($N = 1760$). Die Auswertung der offenen Aufgabe erfolgte analytisch mittels dichotomer Kodierungen mit einer guten bis sehr guten Interraterreliabilität. Die Raschskalierung des Tests zeigt eine gute Reliabilität mit überwiegend akzeptablen Item-Fits.

1 Problemstellung

Seit langem ist bekannt, dass für einen qualitativ hochwertigen Unterricht und den Lernerfolg die Berücksichtigung des Vorwissens entscheidend ist (McNamara/Kintsch 1996; Snow/Corno 1986; Sweller 1994). Damit ein Unterricht gelingen kann, der an den Lernvoraussetzungen der Schülerinnen und Schüler ansetzt und deren individuelle Lernfortschritte berücksichtigt, gilt es, die entsprechenden Leistungsstände zu identifizieren (Hosenfeld et al. 2002). Darüber hinaus können Kompetenztests als Mittel zur Überprüfung und Weiterent-

wicklung von Unterrichtsqualität eingesetzt werden. Je besser dies Lehrkräften gelingt, desto differenzierender gehen sie beim Unterrichten vor (vgl. Richter et. al 2014; für den Mathematikunterricht Westphal et al. 2016).

Im Folgenden werden die Entwicklung und der Einsatz eines Kompetenztests zur Erfassung von Fähigkeiten im schriftlichen Argumentieren bei Lernenden der Jahrgangsstufen sieben bis zehn dargestellt. Dieser Test wurde im Rahmen der binationalen Studie „Coaching im Praktikum" (COPRA)[1] entwickelt. Die Studie untersucht im Rahmen eines Mixed-Methods-Ansatzes die Wirkungen von Unterstützungselementen, insbesondere des fachspezifischen Unterrichtscoachings, in Schulpraktika (vgl. Arnold/Gröschner/Hascher 2014; Staub 2015) bezogen auf Deutschunterricht in der Sekundarstufe I. Das Unterrichtsthema wurde gewählt, da es zahlreiche Anknüpfungspunkte an die Inhalte der Deutschcurricula in der Schweiz und in Deutschland bietet, besondere gesellschaftliche Relevanz hat und breite Gestaltungsfreiräume in der Unterrichtsplanung und -gestaltung lässt. Um die Wirkung der Unterstützungselemente in Schulpraktika auf der Ebene der Leistungen der Schülerinnen und Schüler erfassen zu können, wurde der Kompetenztest für den Einsatz in beiden Ländern entwickelt.

2 Theoretischer Hintergrund

Argumentative Fähigkeiten sind in nahezu allen Bereichen des öffentlichen wie auch privaten Lebens von hoher Relevanz und gelten als Schlüsselkompetenz für demokratisches Zusammenleben und wissenschaftlichen Erkenntnisgewinn. Sie werden immer dann benötigt, wenn Personen unterschiedliche Standpunkte vertreten und sich hierüber austauschen.

In der Argumentationsforschung gibt es zwei Ansätze, den „arguing-to-learn approach" und den „learning-to-argue approach" (Rapanta/Garcia-Mila/Gilabert 2013, S. 486). Bezug nehmend auf ersteren Ansatz belegen zahlreiche internationale Studien, dass Lernende im Rahmen angepasster Lernsettings und adäquater Unterstützung bedeutsame Lernerfolge im Sinne von Wissenserwerb über bestimmte Themen durch gemeinsames Argumentieren erzielen (z. B. Andriessen 2006; Leitão 2003; Nussbaum 2005). Die Ziele des Deutschunterrichts sind dem zweiten Ansatz zuzuordnen, womit das Fach jenen Fächern zuarbeitet, in denen argumentatives Auseinandersetzen dem fachlichen Lernzuwachs dient (Budke/Meyer 2015).

1 Es handelt sich um ein vom SNF und der DFG gefördertes Lead Agency Projekt gemeinsam mit Fritz Staub, Tina Hascher, Sibylle Rahm und Alois Niggli, Laufzeit 04/2016 – 09/2019; GZ für den Standort Kassel: BO 2533/4-1; s. auch www.copra-online.net.

Bei argumentativen Kompetenzen, fokussiert auf das schriftliche Argumentieren, handelt es sich um eine Bündelung unterschiedlicher Fähigkeiten und Fertigkeiten unter Berücksichtigung der kontextuellen, funktionalen und strukturellen Ebene (Bosse/Kempf 2019). Argumentationskompetenz umfasst aus Perspektive des Faches Deutsch die Fähigkeit, Informationen einzuordnen, komplexe Sachverhalte zu analysieren, logische Schlüsse zu ziehen und eigene Argumentationen adressatengerecht zu formulieren (Winkler 2003). Das schriftliche Argumentieren verlangt im speziellen die rezeptive und produktive Auseinandersetzung mit Texten, die durch Begründungen einer Position zu überzeugen versuchen, was zugleich auch die gezielte Auseinandersetzung mit Sprache und Kommunikation beinhaltet (Becker-Mrotzek et al. 2010). Fachlich kann das schriftliche Argumentieren im Fach Deutsch in den Kompetenzbereichen *Beherrschung der deutschen Sprache*, *Erwerb grundlegender Lese- und Schreibfähigkeiten* sowie S*ituationsangemessener, sachgerechter und zielgerichteter Gebrauch von Wort und Schrift* verortet werden.

Konkreter werden die Erwartungshorizonte, welche argumentativen Fähigkeiten Lernende zu den unterschiedlichen Zeitpunkten ihrer schulischen Laufbahn beherrschen sollten, in den Bildungsstandards aufgeführt. Während eine Vielzahl der dort beschriebenen Kompetenzen standardisiert überprüft werden können, trifft dies für den Bereich der Argumentationskompetenz nur bedingt zu. Zur Erfassung argumentativer Fähigkeiten existieren zwar zahlreiche Studien (Budke/Meyer 2015; Krelle 2014; Nussbaum 2008; Rapanta/Garcia-Mila/Gilabert 2013), dennoch liegt kein Instrument vor, mit dem sich argumentative Fähigkeiten im Fach Deutsch sowohl standardisiert als auch zeitökonomisch erfassen lassen. Hieraus resultiert die Notwendigkeit der Entwicklung eines Testinstruments.

Wie Domenech, Krah und Hollmann (2017) darstellen, entwickeln sich argumentative Fähigkeiten während der Sekundarstufe I im Durchschnitt linear, aber individuell und je nach Geschlecht unterschiedlich. Zudem konnte gezeigt werden, dass die gemessene Argumentationskompetenz die Leistungen im Fach Deutsch bedeutsam vorhersagte (ebd.). Im Rahmen des Tests werden somit die Variablen *Alter*, *Geschlecht* und *letzte Zeugnisnote im Fach Deutsch* mit erhoben.

3 Fragestellungen

- Können die Anforderungen, die im Bereich des schriftlichen Argumentierens an Schülerinnen und Schüler der Jahrgangsstufen sieben bis zehn im Fach Deutsch gestellt werden, mittels eines zeitökonomischen Kompetenztests erfasst werden?

- Welche Einflüsse zeigen sich zwischen den gemessenen Kompetenzen und weiteren Merkmalen der Schülerinnen und Schüler?

4 Methodisches Vorgehen

Testinstrument

Der Entwicklung des Tests zum Thema schriftliches Argumentieren lagen mehrere Anforderungen zu Grunde, die sich zum einen aus der Konzeption der Studie, zum anderen aus der Perspektive der weiteren Verwendbarkeit des Testinstruments zur Kompetenzmessung als eine zentrale Voraussetzung von Unterrichtsqualität ergaben. Konkret sollte der Test ein breites Spektrum argumentativer Kompetenz abdecken, um zum einen sowohl den Grundkompetenzen für die Schulsprache am Ende des 8. Schuljahres der Nationalen Bildungsstandards der EDK (2011) als auch den Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss der KMK (2003) gerecht zu werden; zum anderen, um den Einsatz von der siebten bis zur zehnten Jahrgangsstufe zu ermöglichen. Eine weitere Anforderung war der zeitökonomische Einsatz, d. h. der Zeitrahmen einer Schulstunde von 45 Minuten sollte für die gesamte Durchführung nicht überschritten werden. Leitend für die Testkonstruktion waren zudem Forderungen der Deutschdidaktik, im Deutschunterricht bei dem Thema *Argumentation* jenseits des schulischen Textformats *Erörterung* auch Bezüge zur außerschulischen Schreibwirklichkeit herzustellen. Damit sind Textsorten mit argumentativem Charakter wie Leserbriefe, Zeitungskommentare oder Filmkritiken gemeint, die im Gegensatz zur klassischen Erörterung einen expliziten Situations- und Adressatenbezug haben (Ludwig/Spinner 2000).

Das Vorgehen bei der Testentwicklung verlief anhand folgender Schritte:

1. Sichtung und Sammlung von relevanten Inhalten für das Thema „schriftliches Argumentieren“ in den Schweizer und deutschen Bildungsstandards sowie in Schulbüchern und Tests.
2. Abgleich, Analyse und Komprimierung der gesammelten Inhalte innerhalb eines Kompetenzrasters sowie Erstellung von Aufgaben und Auswertungsstrategie.
3. Mehrere Pilotphasen, mit jeweils anschließender Modifikation von Aufgaben und Auswertungsstrategie auf Basis der Ergebnisse mit Kooperationspartnerinnen und -partnern aus Erziehungswissenschaft, Fachdidaktik und der Schulpraxis.

Bei dem entwickelten Instrument handelt es sich somit um einen Leistungstest zum Thema *schriftliches Argumentieren* für Schülerinnen und Schüler der Jahrgangsstufen sieben bis zehn mit einer Bearbeitungsdauer von 35 Minuten. Enthalten sind eine offene, eine halboffene sowie 24 single-choice Aufgaben, welche auf die Kenntnis der Struktur und der Bestandteile von Argumenten, die Analyse argumentativer Texte, das selbstständige Verfassen einer Argumentation sowie Aspekte logischen Schlussfolgerns, der Sprachoberfläche und kritischen Denkens fokussieren, um sowohl das deklarative als auch das prozedurale Wissen überprüfen zu können. Tabelle 1 stellt beispielhaft dar, welche Kompetenzen mit Bezug zum schriftlichen Argumentieren in den Bildungsstandards genannt werden und wie diese als Teilbereiche in den Test einfliessen.

Tabelle 1: Kompetenzbereiche in den Bildungsstandards und im Test

EDK (2011)	**KMK (2003)**	**Teilbereiche Testentwicklung**
Lesen		
argumentative Texte verstehen, deren Themen und Wortschatz dem Weltwissen der Schülerinnen und Schüler entsprechen	Verfahren zur Textaufnahme kennen und nutzen: z. B. Aussagen erklären und konkretisieren, Stichwörter formulieren	argumentative Texte verstehen
zentrale inhaltliche Elemente erkennen	wesentliche Elemente eines Textes erfassen	zentrale inhaltliche Elemente erkennen und verstehen
explizite Informationen erkennen	Information und Wertung in Texten unterscheiden	Information und Wertung erkennen und unterscheiden
einige strukturierende Textelemente erkennen	sprachliche Gestaltungsmittel in ihren Wirkungszusammenhängen erkennen	Organisation und Struktur eines Textes nachvollziehen, Intention erfassen
die Organisation eines Textes nachvollziehen	analytische Methoden anwenden	
Schreiben		
der jeweiligen Kommunikationssituation angemessene argumentative Texte schreiben (unter Einbezug von Mustertexten)	Thesen formulieren, Gegenargumente formulieren, überdenken und einbeziehen	Thesen und Argumente formulieren, Gegenargumente einbeziehen
Textteile thematisch verknüpfen	Argumente gewichten und Schlüsse ziehen, begründet Stellung nehmen	begründet Stellung nehmen und logische Schlüsse ziehen
gebräuchliche Verknüpfungswörter verwenden, um den Textzusammenhang herzustellen		sprachliche Gestaltungsmittel in ihren Wirkungszusammenhängen kennen und anwenden

Die im Test eingesetzten Aufgaben basieren überwiegend auf eigenen Entwicklungen, Abbildung 1 zeigt beispielhaft eine Aufgabe aus dem Kompetenzbereich *begründet Stellung nehmen und logische Schlüsse ziehen*, die versucht auch Aspekte dialogischer Interaktion im Rahmen des schriftlichen Tests zu erfassen. Enthalten sind jedoch auch eine Adaption der Leseaufgabe *Graffiti* aus der ersten PISA-Studie (OECD 2000) sowie die Schreibaufgabe *Leserbrief* aus den bundesweiten Vergleichsarbeiten VERA im Fach Deutsch für die 8. Klasse (IQB 2011). Letztere Aufgabe stellt ein kontroverses Thema (Verhalten von Schülerinnen und Schülern während einer Busfahrt) dar und fordert dazu auf, dieses schriftlich zu diskutieren. So muss zunächst der Leserbrief einer Lokalzeitung gelesen werden, in dem sich eine ältere Dame über das unhöfliche Verhalten von Jugendlichen beschwert. Anschließend soll ein eigener Leserbrief als Antwort verfasst und dabei argumentativ auf den vorliegenden Leserbrief Bezug genommen werden.

Abbildung 1: Aufgabe 6 des entwickelten Tests zu *Induktion und Deduktion beim Argumentieren*

6) Lies die beiden folgenden Dialoge durch, bevor Du die nachfolgende Frage beantwortest:

Marjem: „So, jetzt ist es endgültig: Jeder, der letztes Jahr in Französisch eine sehr gute Note hatte, darf in diesen Sommerferien mit nach Frankreich zum Austausch."

Urs: „Super! Ich habe letztes Jahr eine sehr gute Note in Französisch bekommen. Dann könnte ich ja mitfahren."

Sina: „Hast Du schon gehört? Paul muss heute keine Hausaufgaben machen."

Philipp: „Echt? Warum das denn?"

Sina: „Er hat heute Geburtstag und unser Lehrer meinte, dass er dann lieber feiern soll."

Philipp: „Wow, in eurer Klasse wäre ich gerne! Da muss man keine Hausaufgaben machen, wenn man Geburtstag hat."

Wer schließt von einem Einzelfall auf eine allgemeine Regel?
Setze bitte *nur ein* Kreuz.

O Marjem

O Urs

O Sina

O Philipp

Stichprobe

Die Datenbasis bildet der Prätest der Baseline-Erhebung des SNF-DFG Lead Agency Projekts COPRA mit $N = 1760$ deutschen (39 %) und Schweizer (61 %) Schülerinnen und Schüler aus den Jahrgangsstufen sieben bis zehn (Alter $M = 14.18$, $SD = 1.16$, 50 % weiblich). Im Rahmen der Studie wurde der Test jeweils vor und nach einer sechsstündigen Unterrichtseinheit eingesetzt, die von Studierenden im Schulpraktikum zu den oben dargestellten Kompetenzbereichen gehalten wurde. Die Schulform der Stichprobe war somit nicht vorgeschrieben, sondern ergab sich aus den Praktikumsschulen, an denen die Studierenden unterrichteten. Durchgeführt wurden die Tests von geschulten Testleiterinnen und Testleitern, die allgemeine Fragen zur Beantwortung der Testaufgaben, nicht aber zu den Inhalten des Tests beantworteten und instruiert waren, den Lehrenden keinen Einblick in die Testhefte zu gewähren. Zudem wurde erklärt, dass der Test ein breites Spektrum argumentativer Fähigkeiten erfasst und in seinen Anforderungen nicht mit einer Klassenarbeit vergleichbar sei.

Analysen

Zur Aufbereitung der Daten erfolgte zunächst eine Rekodierung aller Antworten in die Formate *gelöst* oder *nicht gelöst*, wobei nicht oder mehrfach beantwortete Aufgaben ebenfalls als *nicht gelöst* gewertet wurden. Die Kodierung der halboffenen Frage erfolgte zunächst von zwei unabhängigen Kodierern auf Basis des vorliegenden Auswertungsmanuals (OECD 2000) anhand einer Teilstichprobe von $n = 89$, wobei sich eine sehr gute Interraterreliabilität zeigte (Cohen's $\kappa = .82$). Auch für die Kodierung der offenen *Aufgabe 10 Leserbrief* wurde auf das bestehende Auswertungsmanual (IQB 2011) zurückgegriffen, um die formalen, strukturellen und inhaltlichen Kriterien der Leserbriefe anhand eines analytischen, auf dichotome Kodierungen ausgelegten Auswertungsschemas vorzunehmen (zum Vergleich mit dem holistischen Auswertungsschema s. Schipolowski/Böhme 2016). Auf Basis konsensueller Validierungen der ersten Kodierungen durch vier Kodiererinnen und Kodierer wurden einige Kriterien spezifiziert oder gesplittet, um den Interpretationsspielraum bei der Auswertung der Texte zu verringern. Die Interraterreliabilität für die zweite Teilstichprobe zeigte sich zwar zufriedenstellend auf Ebene der einzelnen Kriterien und hoch für die Summe der dichotomen Kodierungen, zugleich zeigte sich jedoch Optimierungspotential hinsichtlich einzelner Kategorien, da einfallsreiche und inhaltlich differenzierte Leserbriefe durch das vorliegende Manual nicht hinreichend gewürdigt werden konnten. Nach der entsprechenden Ergänzung um weitere, vor allem inhaltliche Aspekte, enthält das adaptierte Kodierschema ein formales sowie drei strukturelle und acht inhaltliche Kriterien. Auf Basis einer Teilstichprobe von $n = 311$ zeigt sich die Interraterreliabilität auf Ebene der

einzelnen Kriterien akzeptabel bis sehr gut (.61 < Cohen's κ < .82) und gut für die Summe der dichotomen Kodierungen (ordinal Krippendorff's α = .76; zur Berechnung siehe Hayes/Krippendorff 2007).

Zur Ermittlung von Testgüte und Personenfähigkeit wurden eindimensionale Partial-Credit-Modelle mit ConQuest2 (Wu et al. 2007) berechnet. Bei der *Aufgabe 10 Leserbrief* sowie bei *Aufgabe 9 Kennen und Erkennen der einzelnen Bestandteile einer Argumentation* wurden die gelösten Antworten summiert und somit auch in der Skalierung gestuft, während alle weiteren Items dichotom einflossen.

5 Ergebnisse

Testgüte

Die Ergebnisse der getesteten Schülerinnen und Schüler zeigen nach Ausschluss von fünf Items aufgrund ihrer Trennschärfe ($r_{i(t-i)} < .30$) für die Skalierung von 19 dichotomen und zwei gestuften Items eine gute Reliabilität von .86. Die Passung (wMNSQ) der dichotom skalierten Items liegt zwischen 0.82 und 1.04 in einem akzeptablen Bereich. Die gestuften Items weisen einen wMNSQ zwischen 1.04 und 1.33 auf. Der letztere Wert liegt über 1.2 und weist somit auf eine zu geringe Passung zwischen der einzelnen Aufgabe und dem Gesamttest hin. Konkret repräsentiert der Wert die Punktzahl *null* für die *Aufgabe 10 Leserbrief* und sagt aus, dass Lernende trotz ausreichender Fähigkeiten die Aufgabe nicht gelöst haben. Die Varianz der Fähigkeitsverteilung beträgt 1.21. Abbildung 2 zeigt die geschätzte benötigte Personenfähigkeit, um die einzelnen Aufgaben lösen zu können. Deutlich wird, dass nur die Schülerinnen und Schüler mit den höchsten Fähigkeiten die volle Punktzahl bei den gestuften Items erreichten. Item zehn stellt hierbei den beschriebenen Leserbrief dar, wobei die dritte und höchste Stufe erreicht wurde, wenn mindestens fünf von maximal neun Aspekten als erfüllt kodiert wurden.

Abbildung 2: Personenfähigkeit (Itemschwierigkeit) im Test nach dichotomen und gestuften Items; jedes „X“ repräsentiert 11.5 Fälle

```
Personen-        dichotome  gestufte
fähigkeit        Items      Items
                 |
               X|
               X|
               X|
  2            X|
              XX|            10.3
              XX|
             XXX|
            XXXX|            9.5
           XXXXX|
         XXXXXXX|
  1       XXXXXX|
         XXXXXXX|            10.2
         XXXXXXX|19
        XXXXXXXX|21          9.4
       XXXXXXXXX|20
      XXXXXXXXXX|
  0     XXXXXXX|4, 17        9.3
        XXXXXXXX|14          10.1
         XXXXXXX|1, 3
        XXXXXXXX|
         XXXXXXX|2, 11
         XXXXXXX|12, 15      9.2
           XXXXX|8
 -1        XXXXX|
            XXXX|7
            XXXX|
            XXXX|16, 18
             XXX|
              XX|
              XX|5, 13       9.1
 -2           XX|
               X|
               X|6
               X|
               X|
                |
               X|
 -3             |
```

Einflüsse

Signifikante, jedoch schwache Einflüsse auf die im Rasch-Modell ermittelte Personenfähigkeit als abhängige Variable zeigen die letzte Zeugnisnote im Fach Deutsch ($F(1, 1048) = 65.53$, $p < .001$, $\eta^2 = .06$ und das Alter ($F(1, 1048) = 109.61$, $p < .001$, $\eta^2 = .08$). Das Geschlecht erklärt keinen relevanten Anteil der Varianz ($F(1, 1048) = 4.03$, $p < .05$, $\eta^2 = .00$).

6 Diskussion

Im Beitrag wurde für den Bereich des schriftlichen Argumentierens im Fach Deutsch aufgezeigt, welche Kompetenzerwartungen an Schülerinnen und Schüler gestellt werden und eine Möglichkeit zu deren zeitökonomischer standardisierter Erhebung vorgestellt. Das entwickelte Testinstrument weist statistisch betrachtet eine ausreichende Güte auf, um schriftliche Argumentationskompetenz zu erfassen.

Als Grund für das Passungsproblem von Personenfähigkeit und Lösung der *Aufgabe 10 Leserbrief* liegt nahe, dass die Lernenden im Rahmen eines unbenoteten Tests mit freiwilliger Teilnahme und überwiegend single-choice-Aufgaben nicht immer die Motivation aufbrachten, eine Schreibaufgabe zu absolvieren. Dies spiegelt sich auch darin wider, dass lediglich 60 % der Leserbriefe kodiert werden konnten, da die Aufgabe von 19 % gar nicht und von weiteren 21 % in nicht auswertbarer Art und Weise (z. B. weniger als 20 Wörter oder nicht lesbar) bearbeitet wurde. Die niedrige Varianz der Fähigkeitsverteilung kann auf den breiten Kompetenzbereich, den der Test abdeckt, zurückgeführt werden.

Dennoch sollten weitere Aspekte bei dem Einsatz des beschriebenen Leistungstests berücksichtigt und noch genauer untersucht werden. So ist grundsätzlich zu diskutieren, inwieweit sich ein auf dialogische Interaktion ausgerichteter Kommunikationsbereich wie das Argumentieren überhaupt mittels eines Tests, der individuell bearbeitet wird, prüfen lässt. Die Befürchtung, dass der Test für Schülerinnen und Schüler der Jahrgangsstufen sieben bis zehn eine zu breite Altersspanne abdecken soll, da sich wie beschrieben die argumentativen Fähigkeiten während der Sekundarstufe I stetig weiterentwickeln, konnte anhand der Ergebnisse etwas entschärft werden. Dies könnte jedoch auch durch die unterschiedlichen Schulformen, in denen getestet wurde, resultieren und muss differenzierter analysiert werden. Für eine Gymnasialklasse im Jahrgang zehn sollte jedoch beispielsweise ein Test mit schwierigeren Aufgaben gewählt werden.

Die letzte Deutschnote hat einen hoch signifikanten Einfluss auf das Testergebnis, erklärt jedoch lediglich 6 % der Varianz. Diesbezüglich sollen weitere Analysen auf Basis der Ergebnisse der Interventionsphase des Projekts anhand des Pre-post-Designs zeigen, ob die im Unterricht behandelten Kompetenzfacetten des Bereichs „Schriftliches Argumentieren“ tatsächlich auch mit dem Instrument erfasst werden. So sind der interaktive Adressatenbezug oder die Interdependenz zwischen mündlichem und schriftlichem Argumentieren gewichtige Aspekte argumentativer Kompetenz, die sich im Rahmen des Tests jedoch nicht direkt erfassen lassen. Zur Fundierung oder Modifikation des Tests sollte zudem überprüft werden, inwiefern die gemessenen Kompetenzen von den Einflussgrößen der Lese- und Schreibfähigkeit abhängen, denn aktuell

bleibt noch offen, inwieweit etwa Leseschwächen oder Schwierigkeiten beim Schreiben Einfluss auf die Testbearbeitung hatten.

Als praktische Implikationen können nach Abschluss des Forschungsprojekts zum einen der Test selbst sowie ergänzend didaktisches Material in Aussicht gestellt werden, was interessierten Lehrkräften als Mittel eines kompetenzorientiert ausgerichteten Unterrichts im Bereich des schriftlichen Argumentierens dienen kann.

Literatur

Andriessen, Jerry (2006): Arguing to learn. In: Sawyer, Robert Keith (Hrsg.): The Cambridge handbook of the learning sciences. New York: Cambridge University Press, S. 443–460.

Arnold, Karl-Heinz/Gröschner, Alexander/Hascher, Tina (Hrsg.) (2014): Schulpraktika in der Lehrerbildung / Pedagogical field experiences in teacher education, Theoretische Grundlagen, Konzeptionen, Prozesse und Effekte / Theoretical foundations, programmes, processes, and effects. Münster/New York: Waxmann.

Becker-Mrotzek, Michael/Schneider, Frank/Tetling, Klaus (2010): Argumentierendes Schreiben – lehren und lernen. www.schulentwicklung.nrw.de/cms/upload/netzwerk_NfUE/deutsch/argumentieren_einfuehrung_lang.pdf (Abfrage: 13.12.2018).

Bosse, Dorit/Kempf, Julian (2019): Argumentationskompetenz im Deutschunterricht der Oberstufe und deren Förderung durch digitales Lernen. In: Steffens, Ulrich/Messner, Rudolf (Hrsg.): Konzepte und Bedingungen qualitätsvollen Unterrichts – Grundlagen der Qualität von Schule (Band 3). Münster: Waxmann, S. 403–417.

Budke, Alexandra/Meyer, Miriam (2015): Fachlich argumentieren lernen – die Bedeutung der Argumentation in den unterschiedlichen Schulfächern. In: Budke, Alexandra/Kuckuck, Miriam/Meyer, Michael/Schäbitz, Frank/Schlüter, Kirsten/Weiss, Günther (Hrsg.): Fachlich argumentieren lernen. Didaktische Forschungen zur Argumentation in den Unterrichtsfächern. Münster: Waxmann 2015, S. 9–28.

Domenech, Madeleine/Krah, Antje/Hollmann, Jelena (2017). Entwicklung und Förderung der Argumentationskompetenz in der Sekundarstufe I: Die Relevanz familiärer Ressourcen. In: Bildung und Erziehung, 70, 1, S. 91–108.

EDK (2011): Grundkompetenzen für die Schulsprache. Nationale Bildungsstandards. Freigegeben von der EDK-Plenarversammlung am 16. Juni 2011.

Hayes, Andrew F./Krippendorff, Klaus (2007): Answering the call for a standard reliability measure for coding data. In: Communication Methods and Measures 1, S. 77–89.

Hosenfeld, Ingmar/Helmke, Andreas/Schrader, Friedrich-Wilhelm (2002): Diagnostische Kompetenz: Unterrichts- und lernrelevante Schülermerkmale und deren Einschätzung durch Lehrkräfte in der Unterrichtsstudie SALVE. Zeitschrift für Pädagogik, Beiheft 45, S. 65–82.

IQB (2011): VERA. Beispielaufgaben Deutsch Sek I. Aufgabenbereich: VERA-8 Deutsch Schreibaufgaben. www.iqb.hu-berlin.de/vera/aufgaben/de1 (Abfrage: 25.11.2018).

KMK (2003): Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. Beschluss der Kultusministerkonferenz vom 4.12.2003. München: Wolters Kluwer.

Krelle, Michael (2014): Mündliches Argumentieren in leistungsorientierter Perspektive. Eine empirische Analyse von Unterrichtsdiskussionen in der neunten Jahrgangsstufe. Baltmannsweiler: Schneider-Verlag Hohengehren.

Leitão, Selma (2003): Evaluating and Selecting Counterarguments: Studies of Children's Rhetorical Awareness. In: Written Communication 20, 3, S. 269–306.

Ludwig, Otto/Spinner, Kasper H. (2000). Mündlich und schriftlich argumentieren. Praxis Deutsch 160, 16–22.

McNamara, Danielle/Kintsch, Walter (1996): Learning from texts: Effects of prior knowledge and text coherence. In: Discourse Processes 22, 247–288.

Nussbaum, E. Michael (2008): Collaborative discourse, argumentation, and learning: Preface and literature review. In: Contemporary Educational Psychology 33, 3, S. 345–359.

Nussbaum, E. Michael (2005): The effect of goal instructions and need for cognition on interactive argumentation. In: Contemporary Educational Psychology 30, 3, S. 286–313.

OECD (2000): Lesekompetenz. Sammlung freigegebener PISA-Aufgaben.

Rapanta, Chrysi/Garcia-Mila, Merce/Gilabert, Sandra (2013): What Is Meant by Argumentative Competence? An Integrative Review of Methods of Analysis and Assessment in Education. In: Review of Educational Research 83, 4, S. 483–520.

Richter, Dirk/Böhme, Kathrin/Becker, Michael/Pant, Hans Anand/Stanat, Petra (2014): Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten. Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen von Schülerinnen und Schülern. In: Zeitschrift für Pädagogik 60, 2, S. 225–244.

Schipolowski, Stefan/Böhme, Katrin (2016): Assessment of writing ability in secondary education: comparison of analytic and holistic scoring systems for use in large-scale assessments. In: L1-Educational Studies in Language and Literature 16, S. 1–22.

Snow, Richard E./Corno, Lucia (1986): Adapting teaching to individual differences among learners. In: Wittrock, Merlin Carl (Hrsg.): Handbook of research on teaching. London: Macmillan, S. 605–629.

Staub, Fritz. C. (2015): Fachspezifisches Unterrichtscoaching. In: Rolff, Hans-Günter (Hrsg.): Handbuch Unterrichtsentwicklung. Weinheim u. a.: Beltz, S. 476–489.

Sweller, John (1994): Cognitive load theory, learning difficulty, and instructional design. In: Learning and Instruction 4, 4, S. 295–312.

Westphal, Andrea/Gronostaj, Anna/Vock, Miriam/Emmrich, Rico/Harych, Peter (2016): Differenzierung im gymnasialen Mathematik- und Deutschunterricht – vor allem bei guten Diagnostiker/innen und in heterogenen Klassen? In: Zeitschrift für Pädagogik 26, S. 131–148.

Winkler, Iris (2003): Argumentierendes Schreiben im Deutschunterricht: Theorie und Praxis. Frankfurt a. M.: Lang.

Wu, Margaret L./Adams, Russ J./Wilson, Mark R./Haldane, S. A. (2007): ACER ConQuest Version 2.0. Generalised item response modelling software. Camberwell, Victoria: ACER Press.

Verzeichnis der Autorinnen und Autoren

Frederik Ahlgrimm, Dr. (HU Berlin)
Anna F. Altmann, Dr. (Universität Freiburg)
Katharina Asen-Molz (Universität Regensburg)
Frederike Bartels, Prof. Dr. (Universität Vechta)
Thorsten Bohl, Prof. Dr. (Universität Tübingen)
Dorit Bosse, Prof. Dr. (Universität Kassel)
Nina Brendel, Prof. Dr. (Universität Potsdam)
Kai S. Cortina, Prof. Dr. (University of Michigan)
Markus Dresel, Prof. Dr. (Universität Augsburg)
Kerstin Drossel, Dr. (Universität Paderborn)
Timo Ehmke, Prof. Dr. (Universität Lüneburg)
Christina Ehras (Universität Regensburg)
Birgit Eickelmann, Prof. Dr. (Universität Paderborn)
Michael Elmer (Universität Regensburg)
Manuela Endberg, Dr. (TU Dortmund)
Tobias Feldhoff, Prof. Dr. (Universität Mainz)
Marike Feierabend, (Universität Vechta)
Eva-Kristina Franz, Dr. (PH Freiburg)
Mario Frei (Universität Regensburg)
Lisa Gaier (Universität Regensburg)
Maria Gastl-Pischetsrieder (Universität Regensburg)
Miriam M. Gebauer, PD Dr. phil. (TU Dortmund)
Silvia Greiten, Dr. (Universität Wuppertal)
Anna Gronostaj, Dr. (Die Deutsche Schulakademie)
Jana Groß Ophoff, Dr. (Universität Tübingen)
Sabine Gruehn, Prof. Dr. (Universität Münster)
Eileen Gunga (Universität Regensburg)
Mareike Haas (Universität Göttingen)
Andreas Hartinger, Prof. Dr. (Universität Augsburg)
Sarah-Larissa Hecker (Universität Bielefeld)
Jana Heinze (Universität Regensburg)
Frank Hellmich, Prof. Dr. (Universität Paderborn)
Fabian Hoya, Dr. (Universität Paderborn)
Julia Jennek, Dr. (Universität Potsdam)
Julian Kempf, (Universität Kassel)
Stephanie Klein (Universität Bielefeld)
Marc Kleinknecht, Prof. Dr. (Universität Lüneburg)
Gabriele Klewin, Dr. (Universität Bielefeld)
Barbara Koch-Priewe. Prof Dr. (Universität Bielefeld)
Anne Köker, Dr. phil. (Universität Bielefeld)

Carolin Körner (Universität Münster)
Karsten Krauskopf, Dr. rer. nat. (Universität Potsdam)
Poldi Kuhl, Prof. Dr. (Universität Lüneburg)
Rebecca Lazarides, Prof. Dr. (Universität Potsdam)
Sabine K. Lehmann-Grube, Dr. (Universität Augsburg)
Svenja Lemmrich (Universität Lüneburg)
Timo Leuders, Prof. Dr. (PH Freiburg)
Alfred Lindl (Universität Regensburg)
Marwin Felix Loeper (Universität Paderborn)
Ramona Lorenz, Dr. (TU Dortmund)
Daniel Mays, Prof. Dr. (Universität Siegen)
Samuel Merk, Jun.-Prof. Dr. (Universität Tübingen)
Karina Meyer (Universität Göttingen)
Franziska Müller (Universität Wuppertal)
Renate Murmann (Universität Regensburg)
Sonja Nonte, Dr. (Universität Göttingen)
Matthias Nückles, Prof. Dr. (Universität Freiburg)
Andreas Obersteiner, Prof. Dr. (PH Freiburg)
Udo Ohm, Prof. Dr. (Universität Bielefeld)
Saskia Opalinski (PH Freiburg)
Marcus Pietsch, Dr. (Universität Lüneburg)
Christopher Neil Prilop (Universität Lüneburg)
Sylvia Rahn, Prof. Dr. (Universität Wuppertal)
Kristina Reiss, Prof. Dr. (TU München)
Simone Röhrl (Universität Regensburg)
Benjamin Roth, Dr. (Universität München)
Charlott Rubach, Dr. (Universität Potsdam)
Anna-Maria Ruck (Universität Regensburg)
Katja Scharenberg, Jun.-Prof. Dr. (PH Freiburg)
Marion Scherzinger, Dr. (PH Bern)
Claudia Schmaltz, Dr. (PH Freiburg)
Sabine Schöb, Dr. (Universität Tübingen)
Tobias Schroedler, Dr. (Universität Hamburg)
Ilse Stangen (Universität Hamburg)
Jan Christoph Störtländer, Dr. (Universität Bielefeld)
Tobias C. Stubbe, Prof. Dr. (Universität Göttingen)
Marcus Syring, Dr. (Universität Tübingen)
Matthias Trautmann, Prof. Dr. (Universität Siegen)
Miriam Vock, Prof. Dr. (Universität Potsdam)
Kira Elena Weber, Dr. (Universität Lüneburg)
Matthias Weich (Universität Regensburg)
Alexander Wettstein, Prof. Dr. (PH Bern)
Angelika Wildgans-Lang (TU München)
Ariane S. Willems, Prof. Dr. (Universität Göttingen)
Nicole Zaruba (Universität Potsdam)