

JERO
Journal for Educational Research Online
Journal für Bildungsforschung Online
www.j-e-r-o.com

Executive Editors:

Nele McElvany, Institut für Schulentwicklungsforschung, Technische Universität Dortmund

Cornelia Gräsel, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

Editorial Office:

Rahim Schaufelberger, Institut für Schulentwicklungsforschung, Technische Universität Dortmund

Michael Rochnia, Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal

E-mail: editorialoffice@j-e-r-o.com

Editors:

Rolf Becker (Bern), Oliver Böhm-Kasper (Bielefeld), Wolfgang Böttcher (Münster), Martin Bonsen (Münster), Wilfried Bos (Dortmund), Hartmut Ditton (München), Cornelia Gräsel (Wuppertal), Stephan Huber (Zug), Olaf Köller (Kiel), Harm Kuper (Berlin), Katharina Maag Merki (Zürich), Romain Martin (Luxemburg), Nele McElvany (Dortmund), Florian H. Müller (Klagenfurt), R. Daniel Muijs (Southampton), Jost Reinecke (Bielefeld), Kerstin Schneider (Wuppertal), Claudia Schuchart (Wuppertal), Wolfram Schulz (Camberwell), Birgit Spinath (Heidelberg), Petra Stanat (Berlin), Jan Van Damme (Leuven), Stefanie van Ophuysen (Münster), Andreas Voss (Hamburg), Horst Weishaupt (Frankfurt), Ludger Wößmann (München), Olga Zlatkin-Troitschanskaia (Mainz)

Please register as a reader under www.j-e-r-o.com if you want to be informed about the publication of each new issue. There is a spring edition and an autumn edition each year. The subscription of the journal is free of charge.

Please register as an author on the website www.j-e-r-o.com if you want to submit your article considering the submission guidelines.

ISSN 1866-6671

© Waxmann Verlag GmbH 2019
Steinfurter Straße 555, 48159 Münster, Germany
Internet: www.waxmann.com, E-mail: info@waxmann.com

Setting: Stoddart Satz- und Layoutservice, Münster
All rights reserved

Inhalt

<i>Lisa Birnbaum, Elisabeth M. Schüller & Stephan Kröner</i> Überzeugungen zum freizeitlichen Schreiben bei Grundschulkindern	5
<i>Manuela Frommelt, Isabelle Hugener† & Kathrin Krammer</i> Förderung der unterrichtsbezogenen Analysekompetenz in der Ausbildung von Lehrpersonen durch fallbasiertes Lernen mit Videos	37
<i>Charlott Rubach, Rebecca Lazarides & Hendrik Lohse-Bossenz</i> Engagement der Klassenlehrkräfte in der Zusammenarbeit mit Eltern und Motivation Lernender in der Sekundarstufe	61
<i>Jennifer Lambrecht, Katja Bogda, Helvi Koch, Guido Nottbusch & Nadine Spörer</i> Längsschnittliche Effekte der häuslichen und institutionellen Lernumwelt auf den Wortschatz von Grundschulkindern – ein Vergleich	86
<i>Eva-Kristina Franz, Vera Heyl, Albrecht Wacker & Tobias Dörfler</i> Konstruktvalidierung eines Tests zur Erfassung von adaptiver Handlungskompetenz in heterogenen Gruppen	116
<i>Magdalena Muckenthaler, Teresa Tillmann, Sabine Weiß, Andreas Hillert & Ewald Kiel</i> Belastet Kooperation Lehrerinnen und Lehrer? Ein Blick auf unterschiedliche Kooperationsgruppen und deren Belastungserleben.....	147
<i>Ilonca Hardy, Jasmin Decristan & Eckhard Klieme</i> Adaptiver Unterricht in der Lehr-Lernforschung.....	169

Contents

<i>Lisa Birnbaum, Elisabeth M. Schüller & Stephan Kröner</i> Beliefs about leisure time writing of elementary school children	5
<i>Manuela Frommelt, Isabelle Hugener† & Kathrin Krammer</i> Fostering teaching-related analytical skills through case-based learning with classroom videos in initial teacher education	37
<i>Charlott Rubach, Rebecca Lazarides & Hendrik Lohse-Bossenz</i> Class teacher’s engagement in parent-teacher partnership and student’s motivation in secondary schools	61
<i>Jennifer Lambrecht, Katja Bogda, Helvi Koch, Guido Nottbusch & Nadine Spörer</i> Comparing the effect of home and institutional learning environment on children’s vocabulary in primary school	86
<i>Eva-Kristina Franz, Vera Heyl, Albrecht Wacker & Tobias Dörfler</i> Construct validation of a test measuring adaptive teaching competency in heterogeneous groups	116
<i>Magdalena Muckenthaler, Teresa Tillmann, Sabine Weiß, Andreas Hillert & Ewald Kiel</i> Are teachers burdened by collaboration? A view on different collaboration groups and their perceived stress ...	147
<i>Ilonca Hardy, Jasmin Decristan & Eckhard Klieme</i> Adaptive teaching in research on learning and instruction	169

Lisa Birnbaum, Elisabeth M. Schüller & Stephan Kröner

Überzeugungen zum freizeitlichen Schreiben bei Grundschulkindern

Zusammenfassung

Wenn Kinder in ihrer Freizeit schreiben, erwerben sie nicht nur Kompetenzen, sondern üben auch eine kulturelle Aktivität mit Eigenwert aus. Zu den Determinanten freizeitlichen Schreibens von Kindern liegen nur wenige Arbeiten vor. Basierend auf der Theorie des geplanten Verhaltens (Ajzen, 1991) wurde ein Fragebogen zur Erfassung der Determinanten freizeitlichen Schreibens bei Grundschulkindern entwickelt. Ausgehend von einer qualitativen und einer quantitativen Vorstudie wurden Items generiert und in einer Studie mit N = 244 Kindern der dritten Jahrgangsstufe verwendet. Die Überzeugungen bzgl. thematisch kongruenter Folgen, Überzeugungen bzgl. thematisch inkongruenter Folgen, Kontrollierbarkeit und der intrinsische Wert erwiesen sich als relevante Prädiktoren von Schreibhäufigkeit, Schreibdauer und produzierter Textmenge, wobei der intrinsische Wert die Effekte der anderen Prädiktoren mediierte. Insgesamt wurden jeweils 42 % der Varianz in den Kriteriumsitems durch die Skalen erklärt. Für die Konstruktvalidität des Fragebogens spricht, dass Jungen in mehreren Prädiktoren ungünstigere Werte als Mädchen aufwiesen. Perspektiven für die weitere Forschung werden diskutiert.

Schlagworte

Grundschulkindern; Schreiben; Freizeitaktivität

Lisa Birnbaum (corresponding author) · Prof. Dr. Stephan Kröner, Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Empirische Bildungsforschung, Regensburger Str. 160, 90478 Nürnberg, Deutschland
E-Mail: lisa.birnbaum@fau.de
stephan.kroener@fau.de

Dr. Elisabeth M. Schüller, Leibniz-Institut für Bildungsverläufe (LifBi), Wilhelmsplatz 3, 96047 Bamberg, Deutschland
E-Mail: elisabeth.schueller@lifbi.de

Beliefs about leisure time writing of elementary school children

Abstract

When elementary school students write in their leisure time, they acquire competences. Moreover, writing is a cultural activity with value in itself. However, there are only few studies regarding determinants of children's leisure time writing. Therefore, a questionnaire based on the theory of planned behavior (Ajzen, 1991) was developed to examine the determinants of leisure time writing in elementary school students. Departing from a qualitative and a quantitative pilot study, items were generated and used in a study with $N = 244$ children. Beliefs concerning congruent consequences, beliefs concerning incongruent consequences, controllability, and intrinsic value proved to be statistically significant predictors of writing frequency, writing time, and produced amount of text. The intrinsic value mediated the effects of the other predictors. Overall, the scales explained 42 % of the variance in each of the three criterion variables. The questionnaire's construct validity is supported by the boy's lower values in several predictors compared to the girl's values. Perspectives for further research are discussed.

Keywords

Elementary school children; Writing; Leisure time activity

1. Freizeitliches Schreiben: Erwerb von *Literacy* und Teilhabe am kulturellen Leben

Schreiben gehört ebenso wie Lesen zu den grundlegenden Kulturtechniken, die während der Grundschulzeit erworben und in den weiterführenden Schulen, im beruflichen Bereich, aber auch im Alltag und in der Freizeit vorausgesetzt werden. Für eine nachhaltige Förderung dieser Techniken ist es jedoch wichtig, diese nicht nur in der Schule, sondern auch in der Freizeit in vielfältigen Kontexten anzuwenden und dadurch auszubauen (vgl. Bachmann, Bertschi-Kaufmann, Kassis, Schneider & Sieber, 2004). Diese Anwendung kann vom Schreiben von Briefen und E-Mails bis hin zum Verfassen von Tagebucheinträgen und Geschichten reichen. Derartige Schreibaktivitäten in der Freizeit stellen nicht nur wichtige Gelegenheiten zum Erwerb von *Literacy* dar (vgl. European Commission, 2012; UNESCO, 2005a), sie weisen darüber hinaus einen Eigenwert auf, bereichern das tägliche Leben, ermöglichen sozialen Austausch und tragen zur gesellschaftlichen, sozialen und kulturellen Partizipation bei (vgl. Fuchs, 2012; Liebau, 2008; Reinwand, 2012). Die Teilhabe am kulturellen Leben ist eine Grundfreiheit, welche nicht erst erwachsenen Personen, sondern bereits Kindern zusteht (vgl. Art. 27 der allgemeinen Menschenrechte Resolution 217 A (III) der Generalversammlung, 1948; Übereinkommen zur kulturellen Vielfalt der UNESCO, 2005b). Schreibaktivitäten

auch außerhalb von Schule sind vor diesem Hintergrund für jedes Kind wünschenswert.

Ausgehend von diesen Überlegungen zum Eigenwert freizeithlichen Schreibens stehen in der vorliegenden Arbeit motivationale Fragen gegenüber solchen des Erwerbs von Schreibkompetenz im Vordergrund: Es soll untersucht werden, welche Überzeugungen das Ausmaß des freizeithlichen Schreibens von Grundschulkindern erklären. Bislang liegen keine geeigneten Skalen zur systematischen Erfassung dieser Überzeugungen im Grundschulalter vor. Deren Entwicklung stand daher im Zentrum der vorliegenden Arbeit. Da die auf eine Aktivität bezogenen Überzeugungen nicht unabhängig vom jeweiligen Kompetenzniveau sein dürften, gehen wir im Folgenden zunächst kurz auf das Schreiben und seine Entwicklung im Grundschulalter ein.

2. Schreiben im Grundschulalter

Spätestens mit dem Eintritt in die Schule beginnt der Schriftspracherwerb eines Kindes. In dieser Zeit entwickelt er sich graduell aus dem Mündlichen heraus. Dementsprechend verfassen Grundschulkinde r zunächst vor allem solche Texte, die stark an den mündlichen Sprachgebrauch angelehnt sind und bereits bekannte narrative Erzählstrukturen aus vorschulischen und häuslichen Lernumwelten aufweisen (vgl. Lin, Monroe & Troia, 2007; Niklas, 2015). Ein Beispiel hierfür sind Erzählungen, in denen konkrete Erlebnisse zum Ausdruck gebracht werden (Augst, Disselhoff, Henrich & Pohl, 2007; Feilke, 1996). Schriftsprachliche Besonderheiten, die es im Mündlichen nicht gibt, müssen hingegen erst erlernt werden (z. B. Überschriften, vgl. Kruse, Reichardt, Hermann, Heinzl & Lipowsky, 2012).

Auf kognitiver Ebene ergeben sich weitere Herausforderungen für die schreibende Person dadurch, dass im Gegensatz zum Mündlichen meist keine unmittelbare Kommunikation zwischen Sender und Empfänger möglich ist, sondern eine raumzeitliche Trennung überwunden werden muss (und kann). Dies hat zur Folge, dass Schreibende verstärkt darauf achten müssen, das Kommunikationsziel zu erreichen (Becker-Mrotzek, 1994; Pohlmann-Rother, Schoreit & Kürzinger, 2016). Hinzu kommt, dass das Gegenüber nicht in jedem Fall persönlich bekannt ist. Eine derart situationsentbundene schriftliche Kommunikation erfordert in verstärktem Maße, sich in potentielle Rezipientinnen und Rezipienten hineinzusetzen. Schreiben findet nicht nur unter den genannten veränderten Kommunikationsbedingungen statt, sondern erfordert auch andere Versprachlichungsstrategien, da Schriftlichkeit mit einer höheren Komplexität, verstärkten Planungsanforderungen und größerer Elaboriertheit einhergeht als die gesprochene Sprache (vgl. Breidbach, 2007; Günther, 2010). Diese Strategien erfordern wiederum kognitive Ressourcen, welche von motivierten Schülerinnen und Schülern verstärkt genutzt werden (vgl. hierzu das Schreibmodell von Becker-Mrotzek & Böttcher, 2011). Insgesamt ist Schreiben somit ein auf kognitiver, affektiver und psychomotorischer Ebene ein höchst an-

spruchsvoller Prozess (z. B. in andere Hineinversetzen, Hand-Auge-Koordination; vgl. Dockrell, 2009; Hayes, 2000; Kruse et al., 2012).

Bereits bei Kindern der ersten Jahrgangsstufe kann empirisch „die technische Komponente des Schreiben-Könnens (sprachsystematisch) von den sprachlich-inhaltlichen Aspekten (semantisch-pragmatisch)“ getrennt werden (Pohlmann-Rother et al., 2016, S. 128). Dementsprechend versteht man unter Schreibkompetenz das Vermögen, „Texte adressatengerecht zu formulieren und, je nach Zielsetzung, präzise zu informieren, überzeugend zu argumentieren oder Sprache ästhetisch ansprechend und kreativ einzusetzen“ (Harsch, Neumann, Lehmann & Schröder, 2007, S. 54).

Der Erwerb von Schreibkompetenz – sowie von Lesekompetenz – ist entscheidend für Textkompetenz (Waidacher, 2007). Textkompetente Personen sind dazu in der Lage, textgeprägte Sprache im Rahmen der jeweiligen literalen Praxis adäquat zu gebrauchen (vgl. Schmölzer-Eibinger, 2008). Am Anfang der Entwicklung von Textkompetenz steht häufig das Verfassen vorrangig assoziativ erzeugter Texte, gefolgt vom Verfassen von Texten, in welchen die Inhalte sachlogisch verknüpft und mehrere unverbundene Perspektiven umgesetzt werden. Schließlich können die Schreibenden diese Perspektiven beim Verfassen von Texten synthetisieren (Augst et al., 2007).

Da das Schreibalter, also „die Summe der [...] gemachten Schreiberfahrungen“ (Bachmann & Sieber, 2004, S. 199), wichtiger für den Stand der Schreibentwicklung ist als das biografische Alter, sind ausgedehnte und vielfältige Erfahrungen notwendig, um Schreibkompetenz zu erwerben (Bachmann & Sieber, 2004). Darüber hinaus gelten Hintergrundvariablen wie Geschlecht und soziale Herkunftsmerekmale als Determinanten von Schreibalter bzw. Schreibentwicklung. Beispielsweise verfassen Mädchen im Vergleich zu Jungen längere Texte und verfügen über einen umfangreicheren Wortschatz sowie eine bessere Orthographie; analoge Unterschiede finden sich zugunsten von Kindern aus Familien mit höherem sozio-ökonomischem Status, wenn man diese mit solchen aus Familien mit niedrigerem Status vergleicht (vgl. Dummert, Endlich, Schneider & Schwenck, 2014; Steinig, Betzel, Geider & Herbold, 2009).

3. Schreiben in Schule und Freizeit

Angesichts der hohen Komplexität und Bedeutsamkeit des Erwerbs von Schreibkompetenz sollten Kinder hinreichend Gelegenheit zum Schreiben erhalten. Im Unterricht werden vielfältige, meist didaktisch situierte Schreibaufgaben geboten (vgl. Sturm & Weder, 2016). Hierbei macht die Lehrkraft einige Vorgaben: beispielsweise zu Rechtschreibung und Formalien wie der Verwendung von Überschriften und dazu, wie ihre Anmerkungen und Korrekturen umzusetzen sind. Schulische Schreibenlässe beinhalten daher in aller Regel ein gewisses Maß an Strukturierung. Gerade für leistungsschwächere Kinder können Strukturie-

rungshilfen lernförderlich sein und den Lernerfolg ebenso begünstigen wie affektiv-motivationale Aspekte (vgl. Lipowsky, 2009; May, 2001). Zugleich ist es wichtig, dass die Kinder mittelfristig lernen, auch ohne derartige Hilfen eigenständig Texte zu verfassen (Staatsinstitut für Schulqualität und Bildungsforschung [ISB], 2014).

Auch wenn Grundschul Kinder in ihrer Freizeit schreiben, wird dies von ihrer schulischen Praxis beeinflusst – wenn beispielsweise die speziell für die Grundschule entwickelte Gedichtform des Elfhens auch zu Hause in Geburtstagsgedichten für Familienmitglieder angewendet wird. Gleichzeitig können Kinder beim freizeithlichen Schreiben prinzipiell selbst darüber entscheiden, ob, was, wie, wann, wo und mit welchem Medium sie schreiben – und wählen im Optimalfall selbst ihrem Schreibalter und ihren Schreibkompetenzen entsprechende Schreibaufgaben. Auch sind freiwillige, freizeithliche Schreibaktivitäten von Grundschulkindern wahrscheinlich selten primär auf den Erwerb von Schreibkompetenz ausgerichtet. Vielmehr dürften meist die Freude am Schreiben oder andere Beweggründe wie der Austausch mit Freunden ins Zentrum rücken (vgl. Schauble & Glaser, 2009).

4. Das Ausmaß freizeithlicher Schreibaktivitäten bei Kindern

Anders als eine Vielzahl vorliegender Studien zu schulischem Schreiben und Kompetenzerwerb fokussiert die vorliegende Arbeit auf freizeithliche Schreibaktivitäten von Grundschulkindern sowie auf die Überzeugungen, die derartige Aktivitäten determinieren. Es stehen demnach Tätigkeiten im Fokus, die jenseits von Schule und Hausaufgaben stattfinden, beispielweise das Schreiben von Geschichten, Briefen oder Tagebucheinträgen. Auch wenn die Beteiligung an Chats oder das Austauschen von elektronischen Kurznachrichten ebenfalls zu den freizeithlichen Schreibaktivitäten gerechnet werden könnten, fokussiert die vorliegende Arbeit auf das Verfassen längerer Texte. Hier spielen die in Lehrplänen und Bildungsstandards für das Fach Deutsch verankerten Aspekte des Planens, Schreibens und Überarbeitens eine stärkere Rolle (vgl. Beschlüsse der Kultusministerkonferenz, 2004; ISB, 2014). Wie bereits erwähnt, stellen derartige prototypische Schreibaktivitäten wichtige Gelegenheiten zum Erwerb von *Literacy* dar (vgl. European Commission, 2012; UNESCO, 2005a) und tragen letztendlich dazu bei, eine aktive gesellschaftliche, soziale und kulturelle Partizipation zu ermöglichen (vgl. Fuchs, 2012; Liebau, 2008; Reinwand, 2012). Für eine derartige Partizipation – und damit auch für die vorliegende Studie – sind Schreibmaterial oder Schreibmedium sekundär. Ebenso ist es nachrangig, ob die Produkte für die eigene Person gedacht sind oder dem situationsentbundenen Austausch mit anderen dienen sollen.

Informationen zum Ausmaß freizeithlicher Schreibaktivitäten wurden bislang vor allem *en passant* in Untersuchungen mit anderen Fragestellungen und anderen Altersgruppen gewonnen: So zeigte die Untersuchung von Nippold, Duthie und Larsen (2005), dass fast zwei Drittel der befragten Sechstklässlerinnen und Sechstklässler in ihrer Freizeit lesen, aber nur etwa ein Drittel auch selbst schreibt. Aus dem deutschsprachigen Raum ist bekannt, dass Schülerinnen und Schüler der dritten und vierten Jahrgangsstufe neben dem Handschreiben auch digitale Medien zum Schreiben von Wörtern und Texten nutzen – insbesondere beim Kommunizieren mit Freunden (Risch, 2015).

5. Determinanten freizeithlichen Schreibens: ein Rahmenmodell

Wenn es darum geht, das Ausmaß freizeithlicher Schreibaktivitäten zu erklären, so stellen die darauf bezogenen Überzeugungen einen wichtigen Ansatzpunkt dar. Sie stehen daher im Zentrum der vorliegenden Arbeit. Überzeugungen spiegeln die Gesamtheit der Informationen wider, über die eine Person hinsichtlich einer bestimmten Aktivität verfügt (Fishbein & Ajzen, 2010, S. 223). Sie können erklären, wer aus welchen Beweggründen heraus in der Freizeit schreibt. Die Umstände des freizeithlichen Schreibens unterscheiden sich von denen des Schreibens in der Schule und für die Schule: Beispielsweise erfolgt meist keine mit Sanktionen verbundene Bewertung der sich im Ergebnis widerspiegelnden Kompetenzen durch Dritte, und Schreibabsicht sowie Ort können frei gewählt werden (vgl. Schauble & Glaser, 2009). Daher ist zu vermuten, dass die sich in den Überzeugungen widerspiegelnden Beweggründe für freizeithliches Schreiben andere sind als beim Schreiben in und für die Schule.

In der Theorie des planten Verhaltens (TPB; Ajzen, 1991, 2002) nehmen Überzeugungen einen prominenten Platz ein. Diese besagt, dass die Ausführung eines intentionalen Verhaltens durch die drei Konstrukte Einstellung, subjektive Norm und wahrgenommene Verhaltenskontrolle erklärt wird, welche jeweils die Gesamtheit der ihnen zugrundeliegenden Überzeugungen repräsentieren (Ajzen, 1991, 2002): Die Einstellung als Gesamtheit der *verhaltensbezogenen Überzeugungen* bezieht sich darauf, wie eine Person das Verhalten bewertet. Die subjektive Norm als Aggregat der *normativen Überzeugungen* spiegelt wider, inwiefern die befragte Person der Ansicht ist, dass ihr Umfeld das spezifische Verhalten befürwortet oder ablehnt. Die wahrgenommene Verhaltenskontrolle schließlich umfasst die *Kontrollüberzeugungen*. Diese beziehen sich auf die subjektive Einschätzung der eigenen Kompetenz bezüglich des Verhaltens (*Selbstwirksamkeit*, d. h. personenbezogene Kontrollüberzeugungen, vgl. Ajzen, 2002; Kröner, 2013) und auf die Einschätzung, inwieweit die eine Person umgebende Umwelt die Ausführung des Verhaltens erleichtert oder erschwert (*Kontrollierbarkeit*, d. h. umweltbezogene Kontrollüberzeugungen, vgl. Ajzen, 2002; Kröner, 2013).

Die Prädiktorkonstrukte der TPB haben Eingang gefunden in ein Rahmenmodell zur kulturellen Partizipation (Kröner, 2013), welches TPB, soziologische und persönlichkeitspsychologische Ansätze vor dem Hintergrund eines Person-Umwelt-Transaktionsansatzes integriert und sich mehrfach bei der Erklärung verschiedenster Freizeitaktivitäten mit mehr oder weniger engem Kulturbezug bewährt hat (af Ursin, 2016; Penthin, Fritzsche & Kröner, 2017; Schüller, Birnbaum & Kröner, 2017; Staudenmaier, 2012). Die Aktivitäten – im Fall der vorliegenden Arbeit freizeitliche Schreibaktivitäten – werden darin als Feld der Person-Umwelt-Transaktion betrachtet. Beeinflusst werden sie einerseits durch die subjektiven Überzeugungen auf Personenseite (*verhaltensbezogene Überzeugungen* und *personenbezogene Kontrollüberzeugungen*) und andererseits durch die subjektive Widerspiegelung der Umwelt (*normative Überzeugungen* und *umweltbezogene Kontrollüberzeugungen*). Darüber hinaus können Hintergrundvariablen wie das Geschlecht Unterschiede in den spezifischen Überzeugungen erklären (vgl. Ajzen, 2011; Kröner, 2013).

5.1 Forschungsstand zu einzelnen Komponenten des theoretischen Rahmenmodells

5.1.1 Die Personenseite

Positive *verhaltensbezogene Überzeugungen* sind prädiktiv für eine positive Entwicklung der Leistungen im Schreiben (Graham, Berninger & Fan, 2007; Knudson, 1995). Auch in Schreibkompetenzmodellen findet sich die Motivation als wichtige Ressource für Schreibaktivitäten (vgl. Bachmann & Becker-Mrotzek, 2017; Hayes, 2012). Insbesondere ausgeprägte Freude am Schreiben, also ein hoher *intrinsischer Wert*, scheint mit der Produktion längerer Texte einherzugehen (vgl. Steinig et al., 2009). Aus der Forschung zu verschiedenen Aspekten kultureller Freizeitaktivitäten ist außerdem bekannt, dass Kinder mit einer positiven Einstellung zum jeweiligen Verhalten auch über umfangreichere Aktivitäten berichten. Dies gilt insbesondere für den *intrinsischen Wert* als Aspekt der Einstellung (z.B. Durik, Vida & Eccles, 2006; Kröner & Dickhäuser, 2009; Miesen, 2003; Schüller & Kröner, 2017; Schüller et al., 2017). Es ist davon auszugehen, dass diese Befunde auch auf das freizeitliche Schreiben übertragbar sind (vgl. Shanahan, 2006). Graham (2018) weist dementsprechend auf die Bedeutung des *intrinsischen Werts* als Determinante des Schreibens hin. In diesem Sinne gehen auch wir für das freizeitliche Schreiben davon aus, dass der *intrinsische Wert* die zentrale Determinante für das Ausmaß der Aktivität darstellt.

Neben den *verhaltensbezogenen Überzeugungen* gehört die *Selbstwirksamkeit* zu den subjektiven Überzeugungen auf Personenseite. Ein stärker ausgeprägtes Selbstwirksamkeitserleben dürfte nicht nur dazu führen, dass im schulischen Kontext anspruchsvollere, kompetenzförderliche Schreibaufgaben gewählt werden

(Pajares, Johnson & Usher, 2007; Schunk & Swartz, 1993), es sollte auch mit der Entscheidung für Schreibaktivitäten in der Freizeit einhergehen.

5.1.2 Die Umweltseite und ihre subjektive Widerspiegelung

Jenseits der auf die eigene Person bezogenen Überzeugungen sollte auch die Wahrnehmung der sozialen und materiellen Umwelt das Handeln der Kinder und damit auch deren Schreibaktivitäten beeinflussen. Hinsichtlich der sozialen Umwelt sollte eine als positiv wahrgenommene Haltung wichtiger Bezugspersonen zum Schreiben, d. h. positive *normative Überzeugungen*, mit einem hohen Ausmaß an Schreibaktivitäten einhergehen. Da für Kinder im Grundschulalter ihre Eltern die wichtigsten Bezugspersonen darstellen, sind hier auch die stärksten Effekte zu erwarten, wie dies Anderson, Wilson und Fielding (1988) für Freizeitaktivitäten zeigen konnten. Zwar dürften darüber hinaus auch normative Erwartungen hinsichtlich der *Peers* von Bedeutung sein, deren Effekte sollten jedoch angesichts der geringeren Bedeutung der *Peers* im Grundschulalter im Vergleich zu deren Bedeutung in der Adoleszenz schwächer ausfallen (vgl. Kröner & Dickhäuser, 2009). Jenseits der sozialen Umwelt dürfte hinsichtlich der materiellen Umwelt relevant sein, inwiefern die für Schreibaktivitäten benötigten Materialien, geeignete Orte und hinreichende zeitliche Gelegenheiten verfügbar sind (in der TPB gefasst als *Kontrollierbarkeit*; Fritzsche, Pfeiffer & Kröner, 2015). Die Befunde zur Relevanz dieser situativen Bedingungen stehen im Einklang mit einschlägigen Ansätzen wie dem von Becker-Mrotzek und Böttcher (2011).

5.1.3 Geschlechtsunterschiede

Mädchen verfügen über einen größeren Wortschatz als Jungen und verwenden mehr erzählende Elemente (Steinig et al., 2009); analog zu bekannten Geschlechtsunterschieden beim Verfassen von Texten in schulischen Settings zugunsten der Mädchen (Steinig et al., 2009) und bei schulbezogenen Freizeitaktivitäten (Fölling-Albers, Haider & Meidenbauer, 2010) sollten daher auch beim freizeitlichen Schreiben Mädchen in stärkerem Umfang als Jungen aktiv sein. Diese Geschlechtseffekte dürften jedoch vollständig durch Unterschiede in den personen- und umweltbezogenen Überzeugungen vermittelt werden (Ajzen, 2011; Fishbein & Ajzen, 2010, S. 226f.). Damit im Einklang zeigte sich in anderen Studien, dass Mädchen Schreibaktivitäten einen höheren *intrinsischen Wert* zuschreiben als Jungen und dass sie über eine stärker ausgeprägte *Selbstwirksamkeit* berichten (vgl. Pajares et al., 2007; Philipp, 2013; Wigfield & Guthrie, 1997).

5.2 Forschungsanliegen

Wie der Überblick über die Literatur verdeutlicht, bezieht sich die bisherige Forschung zum Schreiben bei Grundschulkindern meist auf die Untersuchung ausgewählter Überzeugungen und Prädiktoren und fokussierte auf Schreibkompetenz als Kriterium. Was dagegen weitestgehend noch aussteht, sind Studien, die auf Prädiktorseite das gesamte Spektrum der Überzeugungen in den Blick nehmen und auf Kriteriumsseite untersuchen, welche Überzeugungen das Ausmaß an freizeithlichen Schreibaktivitäten von Grundschulkindern erklären. Dies ist das Ziel der aktuellen Arbeit. Hierzu waren zunächst Vorarbeiten notwendig, die zeigen, dass sich mit dem gewählten theoretischen Rahmenmodell tatsächlich Skalen entwickeln lassen, mit denen sich eine große Bandbreite von Überzeugungen freizeithlichen Schreibens erfassen lässt. Diese werden im Folgenden kurz berichtet, bevor die Ergebnisse einer empirischen Studie zu folgenden Forschungsfragen vorgestellt werden:

1. Welche Überzeugungen zu freizeithlichen Schreibaktivitäten liegen bei Grundschulkindern vor? Eine Systematisierung derartiger Überzeugungen sollte anhand der aus der TPB übernommenen Prädiktorkonstrukte möglich sein.
2. Können die Überzeugungen das Ausmaß an *freizeithlichen Schreibaktivitäten* bei Grundschulkindern erklären? Es wird erwartet, dass dies tatsächlich der Fall ist, wobei diese Effekte über den *intrinsicischen Wert* mediiert werden sollten. Alle anderen Überzeugungen dürften dementsprechend lediglich vermittelt über den *intrinsicischen Wert* einen Beitrag zur Varianzaufklärung in den Schreibaktivitäten leisten.
3. Können Geschlechtsunterschiede in den Überzeugungen und im Ausmaß der *freizeithlichen Schreibaktivitäten* beobachtet werden? Wir erwarten in allen Variablen höhere Werte bei den Mädchen als bei den Jungen.

6. Empirische Vorarbeiten

6.1 Qualitative Vorstudie

Ausgangspunkt für die Entwicklung der in der vorliegenden Arbeit verwendeten und quantitativ-empirisch ausgewerteten Fragebogenskalen bildete eine qualitativ-empirische Vorstudie (Birnbaum, 2015): In dieser wurden $N = 26$ Kinder der zweiten und dritten Jahrgangsstufe ($n = 12$ weiblich, $n = 14$ männlich) über teilstandardisierte Leitfadeninterviews zum freizeithlichen Schreiben befragt. Der Interviewleitfaden wurde mit Hilfe des auf der TPB fußenden Manuals von Francis et al. (2004) verfasst. Den Kindern wurden Briefpapier, ein Tagebuch und ein Freundschaftsbuch vorgelegt und sie wurden gefragt, ob sie in ihrer Freizeit für sich selbst (Geschichten, Tagebuch) oder für andere (Briefe, Freundschaftsbücher) schreiben. Darauf folgten jeweils vier bis fünf Fragen zu den *verhaltensbezo-*

genen, normativen und Kontrollüberzeugungen bezüglich des freizeithlichen Schreibens (z.B. „Was gefällt dir [nicht] am Schreiben in deiner Freizeit?“). Die Interviews wurden analog zu der bei Kröner et al. (2012) beschriebenen Vorgehensweise transkribiert und inhaltsanalytisch ausgewertet. Kategorien wurden deduktiv-induktiv nach der Vorgehensweise von Mayring (2010) gebildet (Beurteilerübereinstimmung; Cohens $\kappa = .90$). Die der *Einstellung* zuzuordnenden verhaltensbezogenen Überzeugungen wurden in die Unterkategorien *intrinsischer Wert*, *Überzeugungen bzgl. thematisch kongruenter Folgen* und *Überzeugungen bzgl. thematisch inkongruenter Folgen* gegliedert. Bei den *normativen Überzeugungen* wurden aus dem Material heraus Personengruppen kategorisiert. Die auf die wahrgenommene Verhaltenskontrolle bezogenen Kontrollüberzeugungen wurden in die Unterkategorien *Selbstwirksamkeit* und *Kontrollierbarkeit* nach Ajzen (2002) ausdifferenziert (für einen Überblick über das gesamte Kategoriensystem siehe Tabelle A1). Anschließend wurden die gebildeten Kategorien als Grundlage für die Entwicklung von Skalen verwendet; die den einzelnen Kategorien zugeordneten Aussagen dienten dabei als Ausgangspunkt für die Formulierung von Items. Ein erster Test der entwickelten Items und Skalen sowie die Beantwortung der Frage, ob sich die qualitativ-inhaltsanalytisch entwickelte Struktur zu den auf das Schreiben bezogenen Überzeugungen auch quantitativ-empirisch abbilden lässt, war das Ziel der nachfolgenden, quantitativen Vorstudie.

6.2 Quantitative Vorstudie

Das aus der qualitativen Vorstudie resultierende Kategoriensystem und der zugehörige Kodierleitfaden mit Ankerbeispielen wurden als Ausgangslage für die Itementwicklung herangezogen, wobei in der Regel zu jeder Unterkategorie im Kategoriensystem eine eigene Skala entwickelt wurde. Einzige Ausnahme bildeten die *normativen Überzeugungen*, bei der die Skalenentwicklung auf Kategorienebene ansetzte (vgl. Tabelle A1). Dieser Fragebogen wurde mit $N = 190$ Schülerinnen und Schülern der zweiten und dritten Jahrgangsstufe pilotiert (Alter $M = 8.63$, $SD = 0.74$; männlich: 46.0 %). Es wurden vier Antwortmöglichkeiten vorgegeben, die mit (1) „stimmt nicht“, (2) „stimmt kaum“, (3) „stimmt etwa“ und (4) „stimmt genau“ verankert waren.

6.2.1 Kriteriumsvariablen

Als Kriterium diente die über drei kategoriale Items operationalisierte Häufigkeit freizeithlicher Schreibaktivitäten: Die beiden ersten Items, angelehnt an Items aus IGLU (Bos et al., 2005, S. 118f.) erhoben die Häufigkeit, mit der die Kinder Geschichten (Item 1) oder Briefe, E-Mails oder Postkarten (Item 2) schreiben. Hier war das Antwortformat jeweils vierstufig von (1) „nie oder fast nie“, (2)

„ein- bis zweimal pro Monat“, (3) „ein- bis zweimal pro Woche“ bis zu (4) „jeden Tag oder fast jeden Tag“. Im dritten, von Bos et al. (2005, S. 120) unverändert übernommenen Item mit vierstufigem Antwortformat wurde nach der Führung eines Tagebuchs gefragt. Die Anker lauteten: (1) „nein, ich habe noch nie ein Tagebuch geführt“, (2) „früher habe ich mal ein Tagebuch geführt, jetzt nicht mehr“, (3) „ja, hin und wieder, mit Unterbrechung“ und (4) „ja, seit mindestens 2 Monaten regelmäßig“.

6.2.2 Ergebnisse

Da die befragten Schülerinnen und Schüler bestimmten Klassen zugeordnet waren, wurde zunächst überprüft, welcher Anteil der Kriteriumsvarianz auf die Klassenzugehörigkeit zurückgeführt werden kann. Hierfür wurde der Designeffekt (d_{eff}) für die Kriteriumsvariablen berechnet. Dieser betrug $d_{eff} = 1.43$ für das erste Item und jeweils $d_{eff} = 1$ für die anderen beiden Items. Daher erfolgte anstelle von Mehrebenenanalysen lediglich eine Berücksichtigung der hierarchischen Datenstruktur durch Verwendung des Befehls „type = complex“ in Mplus 7.3. (Muthén & Muthén, 1998-2015; Peugh, 2010).

Die internen Konsistenzen der Skalen zum *intrinsicischen Wert* ($\alpha = .89$) und zu den *Überzeugungen bzgl. thematisch kongruenter Folgen* ($\alpha = .80$) sowie zu den *normativen Überzeugungen* ($\alpha = .79$) fielen zufriedenstellend bis gut aus. Für das Kriterium ($\alpha = .37$) und für die Prädiktorskalen *Überzeugungen bzgl. thematisch inkongruenter Folgen* ($\alpha = .61$), *Selbstwirksamkeit* und *Kontrollierbarkeit* ($\alpha = .22$) zeigten sich dagegen sehr niedrige interne Konsistenzen. Dennoch waren die Pearson-Korrelationen der Prädiktoren untereinander statistisch signifikant ($.26 \leq r \leq .63$; Ausnahme: *Überzeugungen bzgl. thematisch kongruenter Folgen* und *Kontrollierbarkeit* korrelierten nicht statistisch signifikant). Die Überzeugungen korrelierten durchweg statistisch signifikant mit dem Item, das nach dem Schreiben von Geschichten fragte. Mit den anderen Items zeigten sich nur einzelne statistisch signifikante Korrelationen (siehe Tabelle 1).

Tabelle 1: Spearman-Korrelation der Überzeugungen mit den drei Kriteriumsitems der quantitativen Vorstudie

Variablen	Item 1: Geschichten schreiben	Item 2: Briefe, E-Mails oder Postkarten schreiben	Item 3: Tagebuch führen
Verhaltensbezogene Überzeugungen			
Intrinsischer Wert	.46 (.04)	.14 (.06)	.30 (.05)
Überzeugungen bzgl. thematisch kongruenter Folgen	.41 (.04)	.46 (.06)	.33 (.05)
Überzeugungen bzgl. thematisch inkongruenter Folgen	.33 (.05)	.14 (.06)	.20 (.06)

Normative Überzeugungen	.27 (.05)	.13 (.06)	-.02 (.08)

Kontrollüberzeugungen			
Selbstwirksamkeit	.22 (.06)	.12 (.06)	.07 (.03)
Kontrollierbarkeit	.34 (.05)	.02 (.07)	.09 (.07)

Anmerkungen. Alle Korrelationen $|r_s| \geq .20$ statistisch signifikant ($p < .01$); Angabe der Standardfehler in Klammern.

6.3 Fazit aus den empirischen Vorarbeiten für die Überzeugungen zum freizeitlichen Schreiben

Die qualitative *Elicitation Study* hat die Bandbreite der für Kinder relevanten Überzeugungen zum freizeitlichen Schreiben aufgezeigt. Das resultierende Kategoriensystem konnte als Basis für die Entwicklung eines inhaltsvaliden Satzes von Items genutzt werden. Bei einer ersten Erprobung der Items zeigte sich, dass einige Skalen noch der Überarbeitung bedurften. Außerdem war die Stichprobe für die Durchführung konfirmatorischer Faktoranalysen zu klein. Die betreffenden Items wurden vor dem Einsatz in der nun folgenden Studie überarbeitet. Einige Items wurden sprachlich angepasst, vereinzelt wurden Items eliminiert, wenn sie bei den Kindern Verständnisprobleme verursacht hatten (z. B. negativ formulierte Items). Dabei wurde darauf geachtet, dass durch die Eliminierung von Items keine inhaltlichen Aspekte verloren gingen.

7. Methode

Die revidierte Fragebogenform wurde $N = 244$ Kindern aus der dritten Jahrgangsstufe vorgelegt (Alter: $M = 8.29$, $SD = 0.53$; männlich: 48.3 %). Dabei entsprach die Vorgehensweise derjenigen in der quantitativen Vorstudie. Die Kinder stammten aus vierzehn Klassen an drei verschiedenen Grundschulen in Bayern, welche sich einerseits hinsichtlich des Anteils an Schülerinnen und Schülern mit

und ohne Migrationshintergrund und andererseits hinsichtlich ihrer städtischen bzw. ländlichen Lage unterschieden. Die Datenerhebung erfolgte durch geschulte Testleiterinnen und Testleiter in den Schulen. Die Befragung wurde an denselben drei bayerischen Grundschulen durchgeführt wie die quantitative Vorstudie – jedoch im darauffolgenden Schuljahr.

7.1 Prädiktorvariablen

Eingesetzt wurden die erfolgreich getesteten Items sowie die überarbeiteten Items aus den Vorarbeiten zu den drei Skalen der *verhaltensbezogenen Überzeugungen*, der Skala zu *normativen Überzeugungen* und den beiden Skalen zu *Kontrollüberzeugungen*.

Die *verhaltensbezogenen Überzeugungen* wurden in drei Subskalen ausdifferenziert erhoben (vgl. entsprechende Unterkategorien in der *Elicitation Study*, Tabelle A1): (1) *intrinsischer Wert* (z.B. „Ich schreibe, weil ich mich dabei richtig in die Geschichten hineindenken kann.“), (2) *Überzeugungen bzgl. thematisch kongruenter Folgen* (z.B. „Ich schreibe Briefe, weil ich dadurch mit anderen in Kontakt bleiben kann.“) und (3) *Überzeugungen bzgl. thematisch inkongruenter Folgen* (z.B. „Oft mache ich lieber etwas Anderes als zu schreiben.“). Hinzu kam eine Skala zu den *normativen Überzeugungen* (z.B. „Wenn ich schreibe, finden meine Eltern das gut.“), sowie je eine Skala zu den beiden Aspekten der *Kontrollüberzeugungen*: Eine Skala bezog sich auf die Überzeugungen hinsichtlich der *Selbstwirksamkeit* (z.B. „Schreiben fällt mir schwer, weil ich mich dabei sehr konzentrieren muss.“) und eine auf die Überzeugungen hinsichtlich der *Kontrollierbarkeit* (z.B. „In meiner Freizeit habe ich Zeit zum Schreiben.“; vgl. Ajzen, 2002).

Das Antwortformat aller Prädiktoritems wurde im Vergleich zu den Vorarbeiten in ein ebenfalls vierstufiges Format abgeändert, welches sich in anderen Studien mit Grundschulkindern bewährt hat (Fritzsche, Kröner, Dresel, Kopp & Martschinke, 2012; Schüller, 2014; Staudenmaier, 2012). Hier waren als Anker in vier Kästchen von links nach rechts (1) ein großgeschriebenes, starke Ablehnung symbolisierendes „NEIN“ enthalten, (2) ein kleingeschriebenes „nein“ als tendenzielle Ablehnung, (3) ein kleingeschriebenes „ja“ als tendenzielle Zustimmung und (4) ein großgeschriebenes „JA“ als starke Zustimmung. Für die Analysen wurden alle Items der *Überzeugungen bzgl. thematisch inkongruenter Folgen* rekodiert. Eine Übersicht mit allen Prädiktoritems findet sich in Tabelle A2.

7.2 Kriteriumsvariablen

Das Kriterium wurde über drei Items erhoben. Das erste Item entstand, indem die drei Items aus der quantitativen Vorstudie zum Geschichtschreiben, dem Schreiben von Briefen und zur Tagebuchführung zu einem Item zusammengefasst

wurden. Die Häufigkeit des freizeithlichen Schreibens wurde dadurch folgendermaßen erhoben: „Wie oft schreibst du in deiner Freizeit eigene Geschichten, Briefe oder Tagebuch zu deinem Vergnügen (also nicht für die Schule)?“ Anker waren hier: (1) „nie oder fast nie“, (2) „ein- bis zweimal pro Monat“, (3) „ungefähr ein- bis zweimal pro Woche“ und (4) „öfter als dreimal pro Woche“ ($Md = 3$, $MQA = 3$). Dieses Item sollte damit erfassen, wie häufig textproduzierende Schreibaktivitäten ausgeübt werden. Darüber hinaus wurden zwei weitere Items neu entwickelt, mit denen erfasst werden sollte, wie lange die Schreibaktivitäten der Kinder jeweils andauern und wie umfangreich die jeweils produzierten Texte sind. Das Item zur Dauer lautete: „Wie viel Zeit verbringst du normalerweise jeden Tag damit, zu deinem Vergnügen (also nicht für die Schule) eigene Geschichten, Briefe oder Tagebuch zu schreiben?“ Anker waren hier: (1) „Ich schreibe fast gar nicht.“, (2) „bis zu 30 Minuten täglich“, (3) „zwischen einer halben und einer Stunde täglich“, (4) „ein bis zwei Stunden täglich“ und (5) „mehr als zwei Stunden täglich“ ($Md = 2$, $MQA = 1$). Das Item zur Textmenge lautete: „Wie viele Seiten schreibst du ungefähr pro Tag, wenn du eigene Geschichten, Briefe oder Tagebuch in deiner Freizeit schreibst?“ Anker waren hier: (1) „weniger als eine Seite“, (2) „ungefähr eine Seite“, (3) „ungefähr drei bis fünf Seiten“, (4) „ungefähr sechs bis acht Seiten“, (5) „mehr als acht Seiten“ ($Md = 2$, $MQA = 2$). Die Items zur Häufigkeit und zur Dauer des freizeithlichen Schreibens ähneln Items, die in der Leseforschung eingesetzt werden (vgl. Becker, McElvany & Kortenbruck, 2010; Schaffner, Schiefele & Ulferts, 2013; Schüller et al., 2017).

8. Ergebnisse¹

Analog zur quantitativen Vorstudie wurde untersucht, ob die auf die Zugehörigkeit der Kinder zu Schulklassen zurückgehende Mehrebenenstruktur der Daten zu berücksichtigen ist. Der Designeffekt lag für die drei Items zu Häufigkeit und Dauer sowie von Textmenge bei freizeithlichen Schreibaktivitäten bei $d_{eff} = 1.93$, $d_{eff} = 1.42$ und $d_{eff} = 1.71$ und damit erneut unterhalb der Grenze, ab der eine Berücksichtigung der Mehrebenenstruktur als zwingend erforderlich betrachtet wird (Peugh, 2010). Die Klassenzugehörigkeit wurde bei den Berechnungen der konfirmatorischen Faktoranalysen und der Strukturgleichungsmodelle wiederum als Clustervariable berücksichtigt.

1 Datendatei, Mplus-Syntax und Mplus-Output sind auf der Plattform des Open Science Framework (OSF) hinterlegt und können unter dem Link <https://osf.io/ha9jv/> heruntergeladen werden.

8.1 Deskriptive Statistiken, Korrelationen und interne Konsistenzen

Die internen Konsistenzen lagen zwischen $\alpha = .76$ und $\alpha = .88$ (vgl. Tabelle 2). Die Korrelationen zwischen den Skalen zu den Überzeugungen und zwischen Überzeugungen und Geschlecht fielen statistisch signifikant aus – mit Ausnahme der Pearson- und der Spearman-Korrelation zwischen Geschlecht und *Selbstwirksamkeit* sowie der Spearman-Korrelation zwischen Geschlecht und *Kontrollierbarkeit*. Die drei Items zu den freizeithlichen Schreibaktivitäten korrelierten statistisch signifikant untereinander: Häufigkeit/Textmenge $r_s = .51$, Häufigkeit/Dauer $r_s = .65$ und Dauer/Textmenge $r_s = .67$ (vgl. Tabelle 2).

8.2 Konfirmatorische Faktoranalysen

Da eine hohe interne Konsistenz noch keine hinreichende Bedingung für die Eindimensionalität einer Skala darstellt, wurden zur Überprüfung der Dimensionalität der Überzeugungsskalen in einem nächsten Schritt konfirmatorische Faktoranalysen durchgeführt. Hierfür wurden drei zunehmend stärker ausdifferenzierte Modelle berechnet und ihre Passung wurde im Rahmen von χ^2 -Differenzentests miteinander verglichen. Als Ausgangspunkt diente ein Modell, in dem lediglich die drei Überzeugungen unterschieden wurden, die den Prädiktor-konstrukten in der TPB nach Ajzen (1991) entsprechen. Ob die Ausdifferenzierung der *Kontrollüberzeugungen* nach Ajzen (2002) auf die *Selbstwirksamkeit* und die *Kontrollierbarkeit* bezogene Überzeugungen für unsere Daten sinnvoll ist, wurde in einem zweiten Modell getestet. Schließlich wurden die *verhaltensbezogenen Überzeugungen* in die drei Subskalen ausdifferenziert, welche sich aus der *Elicitation Study* ergeben hatten: *intrinsischer Wert*, *Überzeugungen bzgl. thematisch kongruenter Folgen* und *Überzeugungen bzgl. thematisch inkongruenter Folgen*. Beide Ausdifferenzierungen resultierten jeweils in einer statistisch signifikant verbesserten Modellpassung (vgl. Tabelle 3). Somit erscheint eine Gliederung der Items zu den Überzeugungen in insgesamt sechs Skalen sinnvoll.

8.3 Erklärung freizeithlicher Schreibaktivitäten durch die Überzeugungen und das Geschlecht

Da die faktorielle Struktur des entwickelten Fragebogens sich als modellkonform erwiesen hatte, wurde zur Gewinnung von Hinweisen zu Kriteriums- und Konstruktvalidität ein Strukturgleichungsmodell mit den drei Variablen zur Erfassung freizeithlichen Schreibens als abhängige Variablen, dem *intrinsischen Wert* als Mediator und den weiteren Überzeugungsskalen als Prädiktoren berechnet. Die Pfade auf die drei abhängigen Variablen Häufigkeit, Dauer und Textmenge

Tabelle 2: Korrelationen (Standardfehler), Mittelwerte, Standardabweichungen, interne Konsistenzen und Anzahl der Items

	1	2	3	4	5	6	7	8	9	10	
Freizeitliche Schreibaktivitäten											
1 Häufigkeit											
2 Dauer	.65 (.02)										
3 Textmenge	.51 (.03)	.67 (.02)									
Verhaltensbezogene Überzeugungen											
4 Intrinsischer Wert	.50 (.03)	.52 (.03)	.55 (.03)								
5 Überzeugungen bzgl. thematisch kongruenter Folgen	.43 (.04)	.44 (.04)	.46 (.04)	.40 (.04)							
6 Überzeugungen bzgl. thematisch inkongruenter Folgen	.33 (.04)	.38 (.04)	.38 (.04)	.50 (.03)	.03 (.06)						
7 Normative Überzeugungen	.14 (.06)	.29 (.05)	.23 (.05)	.43 (.04)	.14 (.06)	.26 (.05)					
Kontrollüberzeugungen											
8 Selbstwirksamkeit	.38 (.04)	.33 (.04)	.36 (.04)	.52 (.03)	.17 (.05)	.56 (.03)	.25 (.05)		.51 (.03)	-.02 (.07)	
9 Kontrollierbarkeit	.39 (.04)	.47 (.03)	.48 (.03)	.70 (.02)	.22 (.05)	.45 (.04)	.45 (.04)	.49 (.03)		-.14 (.07)	
10 Geschlecht	-.25 (.08)	-.23 (.08)	-.24 (.08)	-.25 (.08)	-.32 (.09)	-.13 (.09)	-.14 (.08)	-.02 (.07)	-.13 (.07)		
Zentrale Tendenz ^a	3	2	2	2	2.96	2.70	2.28	3.34	2.75	3.23	0.48
Streuung ^d	3	1	2	2	0.82	0.78	1.00	0.64	0.70	0.72	0.50
Cronbachs α	-	-	-	-	.88	.76	.86	.77	.84	.83	-
Anzahl der Items	1	1	1	1	6	3	3	3	4	4	1

Anmerkungen. Spearman-Korrelationen unterhalb der Diagonalen, Pearson-Korrelationen oberhalb der Diagonalen; alle Korrelationen $|r| \geq .14$ statistisch signifikant ($p < .01$); Angabe der Standardfehler in Klammern; die Angabe zu Cronbachs α entfällt, wenn Einzelitems statt Skalen vorliegen. $N = 244$.
^a*Md* und *MQA* für die Items zu freizeitlichen Schreibaktivitäten (1-3), *M* und *SD* für alle anderen Variablen.

Tabelle 3: Modellvergleich anhand des Satorra-Bentler χ^2 -Differenztests zur Überprüfung der Faktorenstruktur in den Skalen zu den Überzeugungen

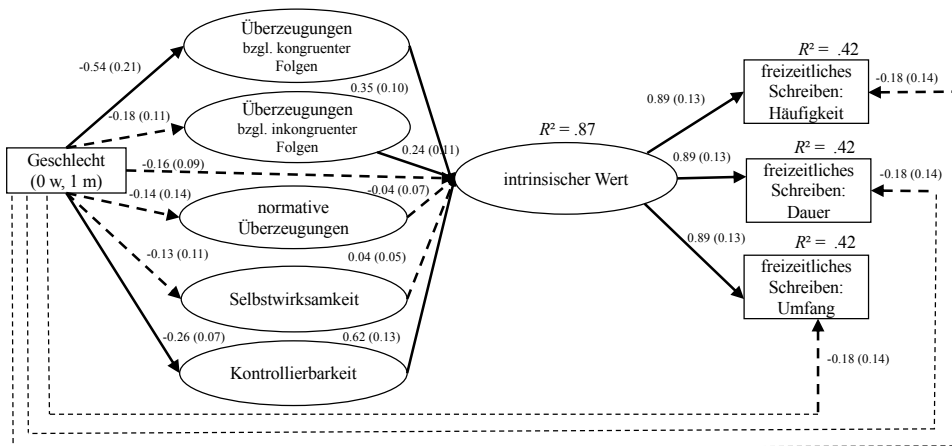
Modell	RMSEA [95 % KI]	CFI	TLI	χ^2	df	Satorra- Bentler- $\Delta\chi^2$ (p)	Δdf
(a) drei Faktoren	.110 [.103, .118]	.710	.677	887.32	227	–	–
(b) vier Faktoren	.097 [.089, .105]	.778	.749	730.09	224	77.23 ($< .001$)	3
(c) sechs Faktoren	.059 [.050, .068]	.921	.07	394.77	215	313.48 ($< .001$)	4

der Schreibaktivitäten wurden in einem ersten Modell frei geschätzt ($\chi^2 = 342.987$, $df = 298$) und in einem zweiten Modell auf Gleichheit restringiert, da diese drei Variablen Aspekte freizeithlicher Schreibaktivitäten messen. Hieraus resultierte keine statistisch signifikant schlechtere Modellpassung ($\Delta\chi^2 = 7.069$, $\Delta df = 4$, $p = .13$). Daher wurde die letztgenannte Modellierung beibehalten. Die Werte bei den berichteten Effekten bezogen auf diese drei Variablen sind durch die Restringierung stets gleich und daher im Folgenden auch nur einmal angegeben. Der ordinalen Datenstruktur des Kriteriums wurde durch die Verwendung des WLSMV-Schätzers Rechnung getragen (Muthén, du Toit & Spisic, 1997). Im Strukturgleichungsmodell wurde zudem das Geschlecht als Kovariate aufgenommen. Dazu wurde es über einen einzelnen Indikator modelliert, ohne einen Messfehler anzunehmen, wie Brown (2015) dies für MIMIC-Modelle (MIMIC = *multiple indicators, multiple causes*) vorschlägt. Derartige Modelle sind für kleinere Stichproben gegenüber Mehrgruppenmodellen zu bevorzugen (Brown, 2015, S. 274), auch wenn damit die Invarianz der Indikatorladungen nicht überprüft werden kann (Brown, 2015, S. 273). Beim vorliegenden Datensatz ergab sich nach dem Fixieren der Beziehung zwischen dem Geschlecht und allen Indikatoren der im Modell enthaltenen latenten Variablen auf Null und der anschließenden Betrachtung der Modifikationsindices lediglich ein statistisch signifikanter Modifikationsindex: Dieser bezog sich auf den Intercept für den Pfad vom Geschlecht auf einen der Indikatoren für die Selbstwirksamkeit, $\chi^2 [1] = 6.79$; $p < .001$. Insgesamt kann damit weitestgehend von Messinvarianz im Hinblick auf die Indikator-Intercepts ausgegangen werden (vgl. zu diesem Vorgehen Brown, 2015, S. 282).

Im Strukturgleichungsmodell wurden die direkten, indirekten und totalen Effekte über den Befehl „*model indirect*“ in Mplus 7.3 angefordert. Die Ermittlung der Konfidenzintervalle aller Koeffizienten erfolgte via Bootstrapping. Dieser Ansatz ist für kleine Stichproben empfehlenswert und setzt keine bestimmte Stichprobenverteilung voraus (Preacher & Hayes, 2004; Shrout & Bolger, 2002). Die Angaben zu den bias-korrigierten Bootstrap-Konfidenzintervallen basieren auf 1000 Bootstrap-Stichprobenziehungen.

Die Modellpassung des Strukturgleichungsmodells fiel zufriedenstellend aus, $\chi^2 = 347.478$, $df = 302$, $RMSEA = .025$ [90 % KI = $.007, .037$], $CFI = .904$, $TLI = .889$. Im Rahmen des MIMIC-Modells wurden jeweils 42 % der Varianz in den manifesten Variablen zu den freizeithlichen Schreibaktivitäten aufgeklärt, sowie 87 % der Varianz des *intrinsicischen Wertes* (siehe Abbildung 1). Insgesamt wiesen drei der fünf Prädiktoren einen statistisch signifikanten Erklärungswert für den Mediator *intrinsicischer Wert* auf. Für die *normativen Überzeugungen* und die *Selbstwirksamkeit* konnten keine Effekte gefunden werden (Abbildung 1). Des Weiteren fanden sich für diejenigen Prädiktoren, die statistisch signifikante, direkte Effekte auf den *intrinsicischen Wert* aufwiesen, auch statistisch signifikante, indirekte, über den *intrinsicischen Wert* vermittelte Effekte auf die drei abhängigen Variablen. Dies gilt für die Überzeugungen bzgl. thematisch kongruenten Folgen ($B = 0.34$, $SE = 0.07$, $BCBS\ KI\ 95\ \% = [0.24, 0.56]$), die Überzeugungen bzgl. thematisch inkongruenten Folgen ($B = 0.24$, $SE = 0.10$, $BCBS\ KI\ 95\ \% = [0.07, 0.44]$) und für die *Kontrollierbarkeit* ($B = 0.51$, $SE = 0.13$, $BCBS\ KI\ 95\ \% = [0.38, 0.82]$)². Die nicht signifikanten Prädiktoren des *intrinsicischen Wertes*, nämlich *normative Überzeugungen* ($B = -0.00$, $SE = 0.10$, $BCBS\ KI\ 95\ \% = [-0.25, 0.56]$) und *Selbstwirksamkeit* ($B = 0.02$, $SE = 0.07$, $BCBS\ KI\ 95\ \% = [-0.12, 0.22]$), wiesen keine statistisch signifikanten indirekten Effekte auf die Variablen zu den freizeithlichen Schreibaktivitäten auf (vgl. Abbildung 1).

Abbildung 1: Strukturmodell zur Vorhersage der freizeithlichen Schreibaktivitäten – *intrinsicischer Wert* als Mediator, Geschlecht als Kovariate.



Anmerkungen. $N = 244$ Grundschulkindern, unstandardisierte Koeffizienten (mit Standardfehler); gestrichelte Linien kennzeichnen nicht signifikante Pfade; alle auf die Kriteriumsitems bezogenen Parameter wurden tripelweise auf Gleichheit restringiert.

2 Es werden die unstandardisierten Koeffizienten des MIMIC-Strukturgleichungsmodells berichtet, da diese im Gegensatz zu standardisierten Koeffizienten bei dichotomen Kovariaten wie dem Geschlecht aussagekräftig sind (Hayes 2013, S. 200).

Hinsichtlich des Geschlechts zeigte sich erwartungsgemäß, dass Jungen statistisch signifikant niedrigere Werte in den Prädiktoren *Überzeugungen bzgl. thematisch kongruenter Folgen* ($B = -0.52, SE = 0.19, BCBS\ KI\ 95\ \% = [-0.98, -0.17]$) und *Kontrollierbarkeit* ($B = -0.25, SE = 0.09, BCBS\ KI\ 95\ \% = [-0.47, -0.10]$) aufwiesen. Für die übrigen Überzeugungen konnten keine statistisch signifikanten Unterschiede zwischen Jungen und Mädchen nachgewiesen werden (vgl. Abbildung 1). Ebenso fiel der direkte Effekt des Geschlechts auf die Kriteriumsvariablen erwartungsgemäß jeweils nicht statistisch signifikant aus, $B = -0.17, SE = 0.13, BCBS\ KI\ 95\ \% = [-0.41, 0.11]$; vgl. Abbildung 1. Es zeigten sich für das Geschlecht die erwarteten indirekten Effekte: Das Geschlecht erklärte zum einen vermittelt über die Überzeugungen bzgl. thematisch kongruenter Folgen und die *Kontrollierbarkeit* den *intrinsischen Wert*, $B = -0.39, SE = 0.12, BCBS\ KI\ 95\ \% = [-0.66, -0.17]$. Zum anderen erklärte es – sowohl über die genannten Überzeugungen als auch über den *intrinsischen Wert* vermittelt – das Ausmaß der freizeitlichen Schreibaktivitäten, jeweils $B = -0.49, SE = 0.15, BCBS\ KI\ 95\ \% = [-0.30, -0.08]$ für alle Indikatoren. Unterschiede in den übrigen Überzeugungen zwischen Jungen und Mädchen blieben aus (vgl. Tabelle 4). Der jeweils statistisch signifikante Gesamteffekt des Geschlechts betrug in Bezug auf den *intrinsischen Wert* $B = -0.55, SE = 0.12, BCBS\ KI\ 95\ \% = [-0.56, -0.15]$ und hinsichtlich der Einzelitems für die freizeitlichen Schreibaktivitäten jeweils $B = -0.66, SE = 0.11, BCBS\ KI\ 95\ \% = [-0.33, -0.18]$.

Tabelle 4: Indirekte Effekte des Geschlechts über die Überzeugungen auf intrinsischen Wert und freizeitliche Schreibaktivitäten

	Über Überzeugungen auf intrinsischen Wert		Über Überzeugungen und intrinsischen Wert auf freizeitliche Schreibaktivitäten	
	<i>B</i> (<i>SE</i>)	BCBS KI 95 %	<i>B</i> (<i>SE</i>)	BCBS KI 95 %
Intrinsischer Wert	–	–	-0.14 (0.09)	[-.33, .01]
Überzeugungen bzgl. thematisch kongruenter Folgen	-0.20 (0.08)	[-.37, -.07]	-0.18 (0.08)	[-.35, -.04]
Überzeugungen bzgl. thematisch inkongruenter Folgen	-0.05 (0.04)	[-.14, .02]	-0.05 (0.04)	[-.14, .02]
Normative Überzeugungen	0.00 (0.02)	[-.05, .03]	0.00 (0.02)	[-.05, .02]
Selbstwirksamkeit	-0.00 (0.01)	[-.03, .02]	-0.00 (0.01)	[-.02, .01]
Kontrollierbarkeit	-0.14 (0.06)	[-.32, -.04]	-0.13 (0.05)	[-.41, -.06]

Anmerkungen. BCBS KI 95 % ist das bias-korrigierte Bootstrap-Konfidenzintervall; Intervalle, die nicht die 0 beinhalten, kennzeichnen, dass die Pfade sich statistisch signifikant von 0 unterscheiden (bei $\alpha = .05$).

9. Diskussion

In der vorliegenden Arbeit wurde die Struktur von verhaltensbezogenen und normativen Überzeugungen sowie von Kontrollüberzeugungen zu freizeithlichen Schreibaktivitäten bei Grundschulkindern herausgearbeitet (vgl. Ajzen, 1991, 2002). Dabei zeigte sich, dass eine Ausdifferenzierung in insgesamt sechs Überzeugungsskalen sinnvoll ist. Es ergaben sich für jede dieser Skalen statistisch signifikante Korrelationen mit den Indikatoren für das Kriterium freizeithliche Schreibaktivitäten. Dies spricht für die Konstruktvalidität der entwickelten Skalen. Im Strukturgleichungsmodell wiesen mit Ausnahme der *normativen Überzeugungen* und der *Selbstwirksamkeit* die übrigen Skalen zu den Überzeugungen der Kinder einen direkten, statistisch signifikanten Beitrag zur Erklärung des *intrinsic*en Wertes auf und damit mittelbar auch einen Effekt auf die freizeithlichen Schreibaktivitäten. Dass insbesondere dem *intrinsic*en Wert hoher Erklärungswert für *literacy*-bezogene Tätigkeiten zukommt (Durik et al., 2006; Schüller et al., 2017; Steinig et al., 2009), konnte nun auch speziell für freizeithliche Schreibaktivitäten im Grundschulalter belegt werden. Der direkte Effekt des Geschlechts auf die freizeithlichen Schreibaktivitäten fiel bei gleichzeitiger Berücksichtigung der Überzeugungen nicht statistisch signifikant aus, was den Annahmen der TPB sowie des theoretischen Rahmenmodells entspricht (vgl. Ajzen, 2011; Kröner, 2013). Unterschiede zwischen Jungen und Mädchen konnten lediglich in den *Überzeugungen bzgl. thematisch kongruenter Folgen* und in der *Kontrollierbarkeit* beobachtet werden.

9.1 Grenzen der Studie und Implikationen für die Forschung

Trotz der Bestätigung der faktoriellen Struktur der Überzeugungsskalen und des Mediatoreffekts des *intrinsic*en Wertes sind folgende Einschränkungen der vorliegenden Studie zu berücksichtigen: Zunächst fällt auf, dass *Selbstwirksamkeit* und *normative Überzeugungen* keinen statistisch signifikanten Erklärungswert für den *intrinsic*en Wert aufwiesen. Bezüglich der *normativen Überzeugungen* kann dies an der Beschränkung auf den injunktiven Aspekt liegen: Es wurde lediglich gefragt, wie Bezugspersonen es wohl fänden, dass man in der Freizeit schreibt. Nicht erfasst wurde dagegen das bei den relevanten Bezugspersonen wahrgenommene Ausmaß von Schreibaktivitäten, die sogenannte deskriptive Norm. In weiteren Studien sollte daher dieser Aspekt der subjektiven Norm ebenfalls erhoben werden (vgl. Ravis & Sheeran, 2003). So kann untersucht werden, ob für die Kinder die Erwartungen wichtiger sind, welche sie in Bezug auf ihre eigenen Schreibaktivitäten seitens ihrer Eltern und anderer Bezugspersonen wahrnehmen (injunktiver Aspekt) oder ob eher das bei Eltern und anderen Bezugspersonen beobachtete Verhalten entscheidend ist (deskriptiver Aspekt).

Bei der Skala *Selbstwirksamkeit* ist zu berücksichtigen, dass sich die Items relativ allgemein auf das Schreiben und wenig auf konkrete Schreibaktivitäten bezogen (z. B. „Schreiben ist für mich anstrengend.“). Es ist daher zu überlegen, auf welche Weise die Operationalisierung der *Selbstwirksamkeit* in künftigen Studien erfolgen sollte. Möglicherweise könnte anstelle einer speziell für diesen Zweck entwickelten Selbstwirksamkeitsskala auch eine etablierte Skala zum Selbstkonzept, einem der *Selbstwirksamkeit* zumindest sehr eng verwandten Konstrukt, adaptiert werden (z. B. Bos et al., 2005; Marsh, 1990).

Zweitens ist die Erfassung des Kriteriums freizeithliche Schreibaktivitäten zu überdenken: Mit den Aspekten Häufigkeit des Schreibens, Dauer des Schreibens und Textmenge zielten die drei eingesetzten Items grundsätzlich auf verschiedene Facetten des Kriteriumsverhaltens ab. In Bezug auf das Item zur Dauer des Schreibens erscheint es ratsam, für künftige Studien die Formulierung zu überarbeiten. Hier sollte besser gefragt werden: „Wie viel Zeit verbringst du *normalerweise* mit dem Schreiben, *wenn* du in deiner Freizeit eigene Geschichten, Briefe oder Tagebuch schreibst?“ Grund dafür ist, dass bei der derzeitigen Formulierung die Antworten auf die Items zu Häufigkeit und Dauer der Aktivität streng genommen nicht statistisch unabhängig sind. Wer hinsichtlich der Häufigkeit angab, fast nie zu schreiben, konnte in Bezug auf die Dauer nicht sinnvollerweise angeben, täglich 30 Minuten oder gar mehr zu schreiben.³

Ausgangspunkt für die quantitative Vorstudie zur vorliegenden Arbeit war das Bestreben, die Häufigkeit verschiedener, textproduzierender freizeithlicher Schreibaktivitäten mit mehreren Items zu erheben. Im Rahmen der Hauptstudie wurde hingegen die Häufigkeit der Produktion verschiedenster Texte in einem Item zusammengefasst. Hinzu kamen Items zur Dauer der jeweiligen Aktivitäten und zur Textmenge der jeweils produzierten Texte. Allerdings zeigten sich weder für das Geschlecht noch für die Überzeugungen als Determinanten Unterschiede im Erklärungswert für die drei Einzelitems zu den freizeithlichen Schreibaktivitäten. Vor diesem Hintergrund ist zu überlegen, ob eine derartige Differenzierung nach Häufigkeit, Dauer und Textmenge sinnvoll ist und ob nicht künftig wieder auf die Häufigkeit fokussiert und dabei noch stärker als in der Vorstudie nach Textsorten differenziert werden sollte. Hier könnte beispielsweise das Schreiben von Kurznachrichten erhoben werden. Diese wurden in der vorliegenden Untersuchung nicht explizit erfragt, sie stellen jedoch spätestens bei Studien mit Kindern in weiterführenden Schulen einen interessanten Forschungsgegenstand dar, da Jugendliche intensiv auf diesem Weg kommunizieren. Die Fokussierung auf eine nach Textsorten differenzierte Erfassung der Schreibhäufigkeit könnte der Validität der Selbsteinschätzungen der Kinder zuträglich sein. Eventuell gehen die für Häufigkeit, Dauer und Textmenge homogenen Ergebnisse der vorliegenden Arbeit zu einem Gutteil darauf zurück, dass Grundschulkindern unzureichend in der Lage sind, diese drei Aspekte zu differenzieren. Unabhängig davon wäre in je-

3 Eine Inspektion der Kreuztabellen ergab, dass derartige unplausible Antwortmuster in der Tat lediglich bei weniger als 5 % der befragten Grundschulkindern auftraten. Die Syntax für die Kreuztabellen ist bei der Erstautorin erhältlich.

dem Fall wünschenswert, für künftige Studien eine reliable Skala zur Erfassung der Schreibaktivitäten zu konstruieren. Dies würde einer Limitation der vorliegenden Arbeit Rechnung tragen.

Des Weiteren unterschieden die Items für die freizeithlichen Schreibaktivitäten nicht, mit welchem Medium geschrieben wird. Mittlerweile wird jedoch auch im Grundschulalter nicht mehr nur mit Stift und Papier geschrieben, sondern – insbesondere in der Freizeit – auch mit mobilen Geräten und Computern. Deshalb sollte die Erhebung von Schreibaktivitäten in künftigen Studien nach Schreibmedien aufgeschlüsselt werden. Dabei könnten dann die unterschiedlichen Anforderungen des Tastaturschreibens und Handschreibens berücksichtigt werden: So gibt es Personen, denen Tastaturschreiben leichter fällt als Handschreiben und umgekehrt. Damit einhergehend könnten auch die einzelnen Überzeugungen je nach Person und Medium unterschiedlich ausfallen (vgl. Schneider & Anskait, 2017, S. 284; Stevenson & Just, 2014). Auch wenn angesichts der allgemeinen Formulierung unserer Kriteriumsitems die Befunde zum Erklärungswert der einzelnen Überzeugungen und zur zentralen Rolle des *intrinsischen Werts* einigermaßen robust gegenüber dem verwendeten Medium sein dürften, lohnt sich doch eine differenziertere Erhebung in künftigen Studien. Dies gilt insbesondere für solche Kinder, die – beispielsweise aufgrund besonderer Umstände, wie sie mit einem Migrationshintergrund einhergehen können – Schwierigkeiten beim Schreiben haben (vgl. Dockrell, 2009; Steinig et al., 2009).

Eine vierte Einschränkung der vorliegenden Studie bezieht sich darauf, dass aufgrund der Stichprobengröße Geschlechtseffekte im Rahmen eines MIMIC-Modells modelliert werden mussten (vgl. Brown, 2015, S. 242). Dabei konnte keine vollständige Überprüfung der Messinvarianz für das Geschlecht erfolgen, wie sie im Rahmen eines Mehrgruppenstrukturgleichungsmodells möglich gewesen wäre. Dies und die Beantwortung der Frage der Generalisierbarkeit auf größere geographische Einheiten bleiben daher Folgestudien vorbehalten.

Als fünfter Aspekt ist zu nennen, dass es neben dem Geschlecht weitere Faktoren geben könnte, welche die Überzeugungen zu freizeithlichen Schreibaktivitäten beeinflussen können. Insbesondere die individuelle Schreibkompetenz und der Bildungshintergrund der Eltern gehören zu den Variablen, die in künftigen Studien mit erhoben werden sollten. So ist aus der Literatur bekannt, dass sich bereits im Grundschulalter sozial bedingte sprachliche Unterschiede beim Verfassen von Texten zeigen (Steinig et al., 2009, S. 344). Aufgrund des theoretischen Rahmenmodells ist jedoch zu erwarten, dass diese Variablen ähnlich wie das Geschlecht lediglich indirekte, über die Überzeugungen der Kinder vermittelte Effekte auf die Freizeitaktivitäten aufweisen. Zudem kann argumentiert werden, dass die wahrgenommene eigene Schreibkompetenz sich in der *Selbstwirksamkeit* wiederfindet und die Effekte des Bildungshintergrunds der Eltern und anderer wichtiger Bezugspersonen sich in den *normativen Überzeugungen* widerspiegeln. Angesichts des für den sozialen Hintergrund anzunehmenden starken Effekts auf sprachliche Variablen (Steinig et al., 2009) erscheint es jedoch lohnenswert,

Bildungshintergrund und individuelle Schreibkompetenzen in künftigen Studien als Prädiktoren freizeithlicher Schreibaktivitäten aufzunehmen.

Sechstens fokussierte die vorliegende Arbeit ausschließlich auf den statistischen Erklärungswert der Überzeugungen als Determinanten freizeithlicher Schreibaktivitäten. Hierfür können jedoch – wie bei Korrelationsstudien unvermeidbar – unterschiedliche Zusammenhangsmuster verantwortlich sein, die in der vorliegenden Arbeit nicht untersucht wurden. Insbesondere könnte das Ausmaß der freizeithlichen Schreibaktivitäten selbst wiederum auf die individuellen Überzeugungen wirken. Derartige Zusammenhänge sollten künftig im Rahmen von Längsschnittstudien und experimentellen Designs untersucht werden.

9.2 Zur Relevanz der einzelnen Überzeugungen

Die Hypothese, dass der *intrinsische Wert* die zentrale Determinante freizeithlicher Schreibaktivitäten im Grundschulalter darstellt und die Effekte der übrigen Skalen vermittelt, konnte durch die vorhandenen Daten gestützt werden. Dies steht im Einklang sowohl mit theoretischen Erwartungen als auch mit anderen Studien zu außerschulischen Aktivitäten (Graham et al., 2007; Kröner & Dickhäuser, 2009; Schüller, 2014; Steinig et al., 2009). Kinder im Grundschulalter tun in ihrer Freizeit insbesondere das, was ihnen Spaß macht (vgl. Durik et al., 2006; Schüller et al., 2017). Hier zeigt sich deutlich der Kontrast zum schulischen Schreiben, das oft nicht primär aus intrinsischen Beweggründen ausgeführt wird, sondern mehr oder weniger aufgrund vorgegebener Aufgaben. Zudem zielt es insbesondere auf den Erwerb von Schreibkompetenzen ab (Pohlmann-Rother et al., 2016; Sturm & Weder, 2016). Wenn daher schreibbezogene Freizeitangebote für Kinder als Angebot zur Förderung kultureller Partizipation Sinn ergeben sollen, so sollten sie von den Kindern mit einem hohem *intrinsischen Wert* verbunden werden (Baurmann & Pohl, 2011; vgl. Liebau, 2008; Reinwand, 2012). Dies kann dazu beitragen, dass Kinder im Grundschulalter nicht durch die kognitiven Herausforderungen vom Schreiben in der Freizeit abgehalten werden (Dockrell, 2009; Kruse et al., 2012; Steinig et al., 2009).

Davon ausgehend stellt sich die Frage, was dazu beiträgt, dass freizeithlichen Schreibaktivitäten ein hoher *intrinsischer Wert* beigemessen wird und wie dieser erlebbar gemacht werden könnte. Hier kann eine Betrachtung der anderen, durch den *intrinsischen Wert* vermittelten schreibbezogenen Überzeugungen hilfreich sein: Wenn man auf diejenigen Überzeugungen fokussiert, für die sich in dieser Studie ein statistisch signifikanter Effekt auf den *intrinsischen Wert* ergeben hat, betrifft dies zunächst die Schreibumgebung, die Zeit und Raum zum Schreiben bieten sollte, in der geeignete Materialien verfügbar und Methoden zugänglich sein sollten (*Kontrollierbarkeit*, Pohlmann-Rother et al., 2016; Reichardt, Kruse & Lipowsky, 2014). Hier könnten die Eltern oder andere Betreuungspersonen bei der Vorbereitung behilflich sein. Diese sollten darauf achten, dass die von ih-

nen geschaffene Umgebung dem Schreibalter des Kindes und seiner individuellen Schreibkompetenz entspricht (Augst et al., 2007; Dockrell, 2009; Feilke, 1996).

Angesichts des Effektes von *Überzeugungen bzgl. thematisch kongruenten Folgen* auf den *intrinsicischen Wert* gilt es darüber hinaus darauf zu achten, dass die Kinder sich in einer Weise schriftlich mit anderen austauschen können, die verstärkende Wirkungen mit sich bringt. Dies kann beispielsweise erfolgen, indem sie sich gegenseitig Briefe schreiben. Eventuell kann es auch motivierend wirken, sich gegenseitig Rückmeldung zu einem Text einzuholen, indem bestimmte Methoden wie Textlupe und Schreibkonferenz auf die eigenständige Anwendung in der Freizeit zugeschnitten werden (vgl. Reichardt et al., 2014). Angesichts der Relevanz von *Überzeugungen bzgl. thematisch inkongruenten Folgen* sollten Zeitfenster für das freizeithliche Schreiben geschaffen werden, in denen dieses nicht mit weiteren Aktivitäten konkurriert.

9.3 Zur Rolle des Geschlechts für die Entscheidung zu freizeithlichem Schreiben

Im Rahmen unserer Analysen ergaben sich für Jungen statistisch signifikant geringere Werte in den Skalen *Überzeugungen bzgl. thematisch kongruenter Folgen* und *Kontrollierbarkeit* als für Mädchen. Zugleich zeigte sich kein direkter Erklärungswert des Geschlechts für die Kriteriumsvariablen im Strukturgleichungsmodell. Dies steht sowohl im Einklang mit den Ausführungen von Ajzen (1991, 2011) zur umfassenden Rolle der Überzeugungen, die den TPB-Prädiktorkonstrukten zugrunde liegen als auch mit den Befunden aus der Lese- und Schreibforschung zu einem geringen Aktivitäts- und Kompetenzniveau bei Jungen (z. B. Bos, Tarelli, Bremerich-Vos & Schwippert, 2012; Gracia, 2015).

10. Fazit

Mit der vorliegenden Arbeit liegt erstmals ein Überblick über den Erklärungswert von systematisch entwickelten Skalen zu Überzeugungen von Grundschulkindern zum freizeithlichen Schreiben vor. Diese Skalen können künftig im Rahmen von Längsschnittstudien eingesetzt werden. Sie ermöglichen es, zu untersuchen, wie sich die Überzeugungen freizeithlichen Schreibens im Laufe des Grundschulalters und darüber hinaus verändern und welche wechselseitigen Effekte sich zwischen diesen Überzeugungen sowie jeweils mit dem Ausmaß der Schreibaktivitäten von Grundschulkindern zeigen. Auch die Integration von Skalen zu Überzeugungen und freizeithlichen Schreibaktivitäten in Längsschnittstudien zur Schreibkompetenz ist von Interesse, da für deren Entwicklung außerschulische Schreibaktivitäten und die darauf bezogenen Überzeugungen bedeutsam sein dürften (vgl. Bachmann et al., 2004, S. 242). Aus derartigen Studien lassen sich Hinweise für pädagogische

Fachkräfte und Eltern dazu ableiten, wie sie freizeithliche Schreibaktivitäten von Kindern fördern und unterstützen können. Dies ist wie eingangs erwähnt auch unabhängig vom Kompetenzerwerb erstrebenswert (Liebau, 2008; Reinwand, 2012), da es sich um eine Aktivität mit Eigenwert handelt, die darüber hinaus kulturelle Teilhabe ermöglicht.

Förderhinweis

Die Datenerhebung der vorliegenden Arbeit wurde durch eine Forschungsbeihilfe der Staedtler-Stiftung an den Drittautor gefördert (Förderkennzeichen: DS/eh S22-S26/10).

Literatur

- af Ursin, P.-K. (2016). *Explaining cultural participation in childhood. Applying the theory of planned behavior to German and Finnish primary school children*. Dissertation. University of Turku, Turku. Zugriff am 11.01.2018 unter <https://www.utu.fi/fi/sivustot/cyri/yhteystiedot/PublishingImages/Annales%20B%20423%20af%20Ursin%20B5.pdf>
- Ajzen, I. (1991). The theory of planned behavior. *Organisational behavior and human decision processes*, 50(2), 179–211.
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32(4), 665–683.
- Ajzen, I. (2011). The theory of planned behaviour. Reactions and reflections. *Psychology & Health*, 26(9), 1113–1127.
- Anderson, R. C., Wilson, P. T. & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 23(3), 285–303.
- Augst, G., Disselhoff, K., Henrich, A. & Pohl, T. (2007). *Text – Sorten – Kompetenz. Eine echte Longitudinalstudie zur Entwicklung der Textkompetenz im Grundschulalter* (Theorie und Vermittlung der Sprache, Bd. 48). Frankfurt a. M.: Lang.
- Bachmann, T. & Becker-Mrotzek, M. (2017). Schreibkompetenz und Textproduktion modellieren. In M. Becker-Mrotzek, J. Grabowski & T. Steinhoff (Hrsg.), *Forschungshandbuch empirischer Schreibdidaktik* (S. 25–54). Münster: Waxmann.
- Bachmann, T., Bertschi-Kaufmann, A., Kassis, W., Schneider, H. & Sieber, P. (2004). Resultate und Konsequenzen. In A. Bertschi-Kaufmann, W. Kassis & P. Sieber (Hrsg.), *Mediennutzung und Schriftlernen. Analysen und Ergebnisse zur literalen und medialen Sozialisation* (Lesesozialisation und Medien, S. 239–248). Weinheim: Juventa.
- Bachmann, T. & Sieber, P. (2004). Wechselwirkungen zwischen literaler Praxis und Schreibkompetenz. In A. Bertschi-Kaufmann, W. Kassis & P. Sieber (Hrsg.), *Mediennutzung und Schriftlernen. Analysen und Ergebnisse zur literalen und medialen Sozialisation* (Lesesozialisation und Medien, S. 199–218). Weinheim: Juventa.
- Baurmann, J. & Pohl, T. (2011). Schreiben – Texte verfassen. In A. Bremerich-Vos, D. Granzer, U. Behrens & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Deutsch konkret. Aufgabenbeispiele, Unterrichts Anregungen, Fortbildungsideen* (3. Aufl., S. 75–103). Berlin: Cornelsen.

- Becker, M., McElvany, N. & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy. A longitudinal study. *Journal of Educational Psychology*, 102(4), 773–785.
- Becker-Mrotzek, M. (1994). Schreiben als Handlung. Das Verfassen von Bedienungsanleitungen. In G. Brünner & G. Graefen (Hrsg.), *Texte und Diskurse. Methoden und Forschungsergebnisse der Funktionalen Pragmatik* (S. 158–178). Opladen: VS.
- Becker-Mrotzek, M. & Böttcher, I. (2011). *Schreibkompetenz entwickeln und beurteilen. Praxishandbuch für die Sekundarstufe I und II* (3. Aufl.). Berlin: Cornelsen.
- Beschlüsse der Kultusministerkonferenz. (2004). *Bildungsstandards im Fach Deutsch für den Primarbereich*. München: Luchterhand.
- Birnbaum, L. (2015). *Überzeugungen zum freizeitlichen Schreiben von Grundschulkindern. Eine qualitative Interviewstudie*. Unveröffentlichte Masterarbeit, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberg.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R., Voss, A., Walther, G., Buddeberg, I., Gneckow, S., Hügler, U. & Kowalski, K. (2005). *IGLU. Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Zugriff am 08.05.2015 unter <http://www.waxmann.com/?eID=texte&pdf=2828Volltext.pdf&typ=zusatztext>
- Breidbach, S. (2007). *Bildung, Kultur, Wissenschaft. Reflexive Didaktik für den bilingualen Sachfachunterricht*. Münster: Waxmann.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (second edition). New York, NY: Guilford.
- Dockrell, J. (2009). Causes of delay and difficulties in the production of written text. In R. Beard, D. Myhill, J. Riley & M. Nystrand (Hrsg.), *The SAGE handbook of writing development* (S. 489–505). Los Angeles, CA: Sage.
- Dummert, F., Endlich, D., Schneider, W. & Schwenck, C. (2014). Entwicklung schriftsprachlicher und mathematischer Leistungen bei Kindern mit und ohne Migrationshintergrund. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 46(3), 115–132.
- Durik, A. M., Vida, M. & Eccles, J. S. (2006). Task values and ability beliefs as predictors of high school literacy choices. A developmental analysis. *Journal of Educational Psychology*, 98(2), 382–393.
- European Commission. (2012). *EU High Level Group of Experts on Literacy. Final report, September 2012*. Luxembourg: Publications Office of the European Union.
- Feilke, H. (1996). Die Entwicklung der Schreibfähigkeiten. In H. Günther & O. Ludwig (Hrsg.), *Schrift und Schriftlichkeit. Ein interdisziplinäres Handbuch internationaler Forschung* (Handbücher zur Sprach- und Kommunikationswissenschaft/HSK, 10.2, S. 1178–1191). Berlin: de Gruyter.
- Fishbein, M. & Ajzen, I. (2010). *Predicting and changing behavior. The reasoned action approach*. New York, NY: Psychology Press.
- Fölling-Albers, M., Haider, T. & Meidenbauer, K. (2010). Schule ist auch nach der Schule – Schulbezogene Aktivitäten, Gespräche und Gedanken in der unterrichtsfreien Zeit. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 30(4), 406–420.
- Francis, J. J., Eccles, M. P., Johnston, M., Walker, A., Grimshaw, J., Foy, R., Kaner E. F. S., Smith, L. & Bonetti, D. (2004). *Constructing questionnaires based on the theory of planned behaviour. A manual for health service researchers*. Newcastle upon Tyne, England: Centre for Health Services Research, University of Newcastle upon Tyne.
- Fritzsche, E. S., Kröner, S., Dresel, M., Kopp, B. & Martschinke, S. (2012). Confidence scores as measures of metacognitive monitoring in primary students? (Limited) va-

- lidity in predicting academic achievement and the mediating role of self-concept. *Journal for Educational Research Online*, 2(4), 120–142.
- Fritzsche, E. S., Pfeiffer, W. & Kröner, S. (2015). Musical elite gymnasia as learning environments and settings for personality development in secondary students? The case of musical self-concept. *Psychological Test and Assessment Modeling*, 2(57), 236–261.
- Fuchs, M. (2012). Kulturelle Bildung als Menschenrecht? In H. Bockhorst, V.-I. Reinwand & W. Zacharias (Hrsg.), *Handbuch Kulturelle Bildung* (Bd. 30, S. 91–94) München: kopaed.
- Gracia, P. (2015). Parent-child leisure activities and cultural capital in the United Kingdom: The gendered effects of education and social class. *Social Science Research*, 52, 290–302.
- Graham, S. (2018). A writer(s)-within-community model of writing. In C. Bazerman, A. N. Applebee, V. W. Berninger, D. Brandt, S. Graham, J. V. Langer, P. K. Matsuda, S. Murphy, D.W. Wells Rowe, M. Schleppegrell & K. C. Campbell Wilcox (Hrsg.), *The lifespan development of writing* (S. 272–325). Urbana, IL: National Council of English.
- Graham, S., Berninger, V. & Fan, W. (2007). The structural relationship between writing attitude and writing achievement in first and third grade students. *Contemporary Educational Psychology*, 32(3), 516–536.
- Günther, H. (2010). *Beiträge zur Didaktik der Schriftlichkeit* (Kölner Beiträge zur Sprachdidaktik Reihe A, Bd. 6). Duisburg: Gilles & Francke.
- Harsch, C., Neumann, A., Lehmann, R. & Schröder, K. (2007). Schreibfähigkeit. In B. Beck & K. Eckhard (Hrsg.), *Sprachliche Kompetenzen: Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistungen International)* (S. 42–62). Weinheim: Beltz.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis. A regression-based approach*. New York, NY: Guilford.
- Hayes, J. R. (2000). A new framework for understanding cognition and affect in writing. In R. Indrisano & J. R. Squire (Hrsg.), *Perspectives on writing. Research, theory, and practice* (1. Aufl., S. 6–44). Newark, DE: International Reading Association.
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388.
- Knudson, R. (1995). Writing experiences, attitudes, and achievement of first to sixth graders. *Journal of Educational Research*, 89(2), 90–97.
- Kröner, S. (2013). Kulturelle Partizipation bei Jugendlichen als Feld der Person-Umwelt-Transaktion. *Zeitschrift für Erziehungswissenschaft*, 16 (Suppl. 3), 233–256.
- Kröner, S. & Dickhäuser, O. (2009). Die Rolle von Eltern, Peers und intrinsischem Wert für die rezeptive hochkulturelle Praxis von Gymnasiasten der Sekundarstufe II. *Zeitschrift für Pädagogische Psychologie*, 23(1), 53–63.
- Kröner, S., Schüller, E. M., Penthin, M., Fritzsche, E. S., Friedrich, M. C. G. & Krol, M. M. (2012). Elternvertreter mit Migrationshintergrund an allgemeinbildenden Schulen. *Zeitschrift für Erziehungswissenschaft*, 15(4), 707–726.
- Kruse, N., Reichardt, A., Hermann, M., Heinzl, F. & Lipowsky, F. (2012). Zur Qualität von Kindertexten. Entwicklung eines Bewertungsinstruments in der Grundschule. *Didaktik Deutsch*, 8(32), 87–110.
- Liebau, E. (2008). *Mit Kunst und Kultur Schule gestalten – Teilhabe ermöglichen*. Vortrag anlässlich BKJ-Fachtagung „Kultur macht Schule – Bildung gemeinsam gestalten!“, Berlin.
- Lin, S.-J. C., Monroe, B. W. & Troia, G. A. (2007). Development of writing knowledge in grades 2–8. A Comparison of typically developing writers and their struggling peers. *Reading & Writing Quarterly*, 23(3), 207–230.

- Lipowsky, F. (2009). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (2. überarbeit. Aufl., S. 69–105). Berlin: VS.
- Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, 32(4), 623–636.
- May, P. (2001). *Lernförderlicher Unterricht. Teil I. Untersuchung zur Wirksamkeit von Unterricht und Förderunterricht für den schriftsprachlichen Lernerfolg*. Frankfurt a. M.: Lang.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz.
- Miesen, H. W. J. M. (2003). Predicting and explaining literary reading. An application of the theory of planned behavior. *Poetics*, 31(3-4), 189–212.
- Muthén, B. O., du Toit, S. H. C. & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes*. Zugriff am 23.07.2018 unter https://www.statmodel.com/download/Article_075.pdf
- Muthén, L. K. & Muthén, B. O. (2015). *Mplus User's Guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Niklas, F. (2015). Die familiäre Lernumwelt und ihre Bedeutung für die kindliche Kompetenzentwicklung. *Psychologie in Erziehung und Unterricht*, 62(2), 106.
- Nippold, M. A., Duthie, J. K. & Larsen, J. (2005). Literacy as a leisure activity. Free-time preferences of older children and young adolescents. *Language, Speech, and Hearing Services in School*, 36(2), 93–102.
- Pajares, F., Johnson, M. J. & Usher, E. L. (2007). Sources of writing self-efficacy beliefs of elementary, middle, and high school students. *Research in the Teaching of English*, 42(1), 104–120. Zugriff am 23.07.2018 unter <http://www.jstor.org/stable/40171749>
- Penthin, M., Fritzsche, E. S. & Kröner, S. (2017). Bereichsspezifische Determinanten außerschulischer musikalischer Aktivitäten von Grundschulkindern. *Beiträge empirischer Musikpädagogik* (8), 1–30. Zugriff am 11.01.2018 unter <https://www.b-em.info/index.php/ojs/article/view/147/300>
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 4(1), 85–112.
- Philipp, M. (2013). *Wie motiviert lese und schreibe ich? Und wenn ja: auf wie viele Arten und Weisen? Zur Klärung und Klärungsbedürftigkeit der vielschichtigen Lese- und Schreibmotivation*. Zugriff am 11.08.2016 unter http://leseforum.ch/myUploadData/files/2013_2_Philipp.pdf
- Pohlmann-Rother, S., Schoreit, E. & Kürzinger, A. (2016). Schreibkompetenzen von Erstklässlern quantitativ-empirisch erfassen – Herausforderungen und Zugewinn eines analytisch-kriterialen Vorgehens gegenüber einer holistischen Bewertung. *Journal for Educational Research Online*, 8(2), 107–135.
- Preacher, K. J. & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717–731.
- Reichardt, A., Kruse, N. & Lipowsky, F. (2014). Textüberarbeitung mit Schreibkonferenz oder Textlupe – Zum Einfluss der Schreibumgebung auf die Qualität von Schülertexten. *Didaktik Deutsch*, 19(36), 64–85.
- Reinwand, V.-I. (2012). *Literaturvermittlung als kulturelle Bildung*. Zugriff am 15.07.2016. unter <http://www.bpb.de/gesellschaft/kultur/kulturelle-bildung/137304/literaturvermittlung-als-kulturelle-bildung?p=all>
- Resolution 217 A (III) der Generalversammlung. (1948). *Allgemeine Erklärung der Menschenrechte*. Zugriff am 11.08.2016 unter <https://www.un.org/depts/german/menschenrechte/aemr.pdf>

- Risch, C. (2015). Eine Untersuchung zur Nutzung neuer Medien für das Lernen in der Grundschule. In A. Wildemann & S. Fornol (Hrsg.), *Sprachliches und mediales Lernen in der Grundschule* (S. 21–44). Aachen: Shaker.
- Rivis, A. & Sheeran, P. (2003). Descriptive norms as an additional predictor in the theory of planned behaviour. A meta-analysis. *Current Psychology*, 22(3), 218–233.
- Schaffner, E., Schiefele, U. & Ulferts, H. (2013). Reading amount as a mediator of the effects of intrinsic and extrinsic reading motivation on reading comprehension. *Reading Research Quarterly*, 48(4), 369–385.
- Schauble, L. & Glaser, R. (Hrsg.). (2009). *Innovations in learning. New environments for education*. New York, NY: Routledge.
- Schmölzer-Eibinger, S. (2008). *Lernen in der Zweitsprache. Grundlagen und Verfahren der Förderung von Textkompetenz in mehrsprachigen Klassen* (Europäische Studien zur Textlinguistik, Bd. 5). Tübingen: Narr.
- Schneider, H. & Anskait, N. (2017). Einsatz digitaler Schreibwerkzeuge. In M. Becker-Mrotzek, J. Grabowski & T. Steinhoff (Hrsg.), *Forschungshandbuch empirische Schreibdidaktik* (S. 283–298). Münster: Waxmann.
- Schüller, E. M. (2014). *Lesen als Freizeitbeschäftigung von Grundschulkindern. Entwicklung von Skalen auf Grundlage der Theorie des geplanten Verhaltens*. Dissertation. Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberg. Zugriff am 10.05.2015 unter <http://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/5488>
- Schüller, E. M., Birnbaum, L. & Kröner, S. (2017). What makes elementary school students read in their leisure time? Development of a comprehensive questionnaire. *Reading Research Quarterly*, 52(2), 161–175.
- Schüller, E. M. & Kröner, S. (2017). A pink dog with 20 legs – An explorative study on the beliefs of elementary school children explaining engagement in artistic leisure activities. *Journal for Educational Research Online*, 9(2), 183–204.
- Schunk, D. H. & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, 18(3), 337–354.
- Shanahan, T. (2006). Relations among oral language, reading, and writing development. In C. MacArthur, S. Graham & J. Fitzgerald (Hrsg.), *Handbook of writing research* (S. 171–186). New York, NY: Guilford.
- Shrout, P. E. & Bolger, N. (2002). Mediation in experimental and nonexperimental studies. New procedures and recommendations. *Psychological Methods*, 7(4), 422–445.
- ISB – Staatsinstitut für Schulqualität und Bildungsforschung. (2014). *LehrplanPLUS Grundschule. Fachlehrplan Deutsch 3/4*. Zugriff am 04.06.2018 unter <https://www.lehrplanplus.bayern.de/fachlehrplan/grundschule/3/deutsch>
- Staudenmaier, R. (2012). *Sportliche Aktivitäten von Grundschulkindern. Entwicklung von Skalen vor dem Hintergrund der Theorie des geplanten Verhaltens*. Unveröffentlichte Dissertation. Friedrich-Alexander-Universität Erlangen-Nürnberg, Nürnberg.
- Steinig, W., Betzel, D., Geider, F. J. & Herbold, A. (2009). *Schreiben von Kindern im diachronen Vergleich. Texte von Viertklässlern aus den Jahren 1972 und 2002*. Münster: Waxmann.
- Stevenson, N. C. & Just, C. (2014). Why teach handwriting before keyboarding? *Early Childhood Education Journal*, 42(1), 49–56.
- Sturm, A. & Weder, M. (2016). *Schreibkompetenz, Schreibmotivation, Schreibförderung. Grundlagen und Modelle zum Schreiben als soziale Praxis*. Seelze: Kallmeyer.
- UNESCO. (2005a). *Education for All. Literacy for life*. Zugriff am 10.05.2015 unter <http://www.uis.unesco.org/Library/Documents/gmr06-en.pdf>

- UNESCO. (2005b). Übereinkommen über den Schutz und die Förderung der Vielfalt kultureller Ausdrucksformen. Zugriff am 14.02.2014 unter <http://www.unesco.de/infotehek/dokumente/uebereinkommen/konvention-kulturelle-vielfalt.html>
- Waidacher, G. (2007). Multimodale Textkompetenz. In S. Schmölder-Eibinger & G. Weidacher (Hrsg.), *Textkompetenz. Eine Schlüsselkompetenz und ihre Vermittlung* (S. 39–55). Tübingen: Narr.
- Wigfield, A. & Guthrie, J. T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology, 89*(3), 420–432.

Anhang

Tabelle A1: Kategoriensystem der qualitativen Elicitation Study zu den Überzeugungen freizeithlichen Schreibens

Kategoriensystem	$N_{\text{kodierte Aussagen}}$	$N_{\text{Interviews}}$	Ankerbeispiel
1. Verhaltensbezogene Überzeugungen	153	26	
1.1 Intrinsischer Wert	76	23	„Wenn ich mal angefangen habe, kann ich nicht mehr aufhören [...], weil es dann Spaß macht.“
1.2 Überzeugungen bzgl. thematisch kongruenter Folgen	31	16	„Da kann man Erlebtes aufschreiben und immer wieder nachlesen.“
1.3 Überzeugungen bzgl. thematisch inkongruenter Folgen	46	21	„Manchmal finde ich, es ist gut zum Zeitvertreib.“
2. Normative Überzeugungen	58	23	„Meine Eltern. Wenn ich schreibe [...], sagen die, es ist gut.“
3. Kontrollüberzeugungen	88	26	
3.1 Selbstwirksamkeit	63	26	„weil [im] Schreiben bin ich eigentlich relativ gut.“
3.2 Kontrollierbarkeit	25	16	„Wenn ich meine Ruhe habe, [...] dann kann ich mich leicht konzentrieren.“
Gesamt	299	26	

Anmerkung. Die Zahl der Interviews, die einer übergeordneten Kategorie zugeordnet wurden, kann sich von der Summe der ihren untergeordneten Kategorien zugeordneten Interviews unterscheiden, da für die übergeordneten Kategorienebenen Personen nur einmal gezählt wurden, die sich entsprechend mehrerer der untergeordneten Kategorien geäußert haben.

Tabelle A2: Formulierungen aller verwendeten Prädiktoritems

Intrinsischer Wert

- Ich schreibe, weil ich mich dabei richtig in die Geschichten hineindenken kann.
- Ich schreibe, weil ich selbst entscheiden kann, was ich schreiben will.
- Ich schreibe, weil ich ganz verschiedene Geschichten erzählen kann (z. B. spannende oder lustige Geschichten).
- Schreiben macht mir Spaß.
- Schreiben gefällt mir.
- Oft habe ich richtig Lust darauf, zu schreiben.

Überzeugungen bzgl. thematisch kongruenter Folgen

- Ich schreibe Briefe, weil ich dadurch mit anderen in Kontakt bleiben kann.
- Ich schreibe Briefe, weil ich mich dann mit anderen austauschen kann.
- Ich schreibe Briefe, weil andere sich darüber freuen.

Überzeugungen bzgl. thematisch inkongruenter Folgen

- Oft treffe ich mich in meiner Freizeit lieber mit Freunden, als etwas zu schreiben.
- Ich spiele lieber draußen, als etwas zu schreiben.
- Lieber beschäftige ich mich mit meinen Spielsachen, als etwas zu schreiben.

Normative Überzeugungen

- Wenn ich schreibe, finden meine Eltern das gut.
- Wenn ich schreibe, finden meine Verwandten (zum Beispiel Großeltern, Onkel, Tanten) das gut.
- Wenn ich schreibe, finden andere Erwachsene das gut.

Selbstwirksamkeit

- Schreiben ist für mich anstrengend.
- Schreiben fällt mir schwer, weil ich mich dabei sehr konzentrieren muss.
- Ich schreibe nicht gerne, weil das sehr lange dauert.
- Ich schreibe nicht gerne, weil mir dann die Hand weh tut.

Kontrollierbarkeit

- In meiner Freizeit habe ich Zeit zum Schreiben.
- In meiner Freizeit nehme ich mir Zeit zum Schreiben.
- In meiner Freizeit kann ich in Ruhe schreiben.
- Zu Hause kann ich gut schreiben, weil ich da meine Ruhe habe.

Anmerkung. Zur verwendeten Formulierung der Kriteriumsitems und Vorschlägen zu deren Optimierung siehe Diskussionsteil.

Manuela Frommelt, Isabelle Hugener[†] & Kathrin Krammer

Fostering teaching-related analytical skills through case-based learning with classroom videos in initial teacher education

Abstract

The ability to analyze teaching is deemed to be crucial to successful adaptive classroom teaching. It requires teachers to use their professional knowledge for identifying significant teaching situations, for interpreting these situations in terms of student learning processes, and for considering possibilities for improvements. In our intervention study “Video Analysis in Teacher Education” (VideA), we applied a quasi-experimental pretest/posttest control-group design (N = 128 student teachers) for pursuing the question of whether teaching-related analytical skills can be fostered through analyzing one’s own or other teachers’ classroom videos. The intervention focused on three basic features of effective teaching (goal clarity, learning support, learning climate). For measuring the participants’ analytical skills, we used open questions and standardized rating items, all of which related to video clips. Content analyses of the comments on the open questions showed a significant increase in the ability to identify classroom situations that are relevant to student learning in both video groups whereas there was no change as regards the ability to suggest improvements in teaching and the ability to give reasons for why the identified situations were thought to be relevant. Moreover, our results indicate that there is a positive correlation between the ability to identify significant classroom situations and the ability to interpret them by referring to pertinent knowledge.

Keywords

Analytical skills; Learning with classroom videos; Teacher education; Open and standardized measurement formats

Manuela Frommelt, M.Sc. (corresponding author), currently not affiliated with an institution.
e-mail: manuela.frommelt@gmail.com

Prof. Dr. Isabelle Hugener[†]

Prof. Dr. Kathrin Krammer, University of Teacher Education Lucerne, Pfistergasse 20, 6000 Lucerne, Switzerland
e-mail: kathrin.krammer@phlu.ch

Förderung der unterrichtsbezogenen Analysekompetenz in der Ausbildung von Lehrpersonen durch fallbasiertes Lernen mit Videos

Zusammenfassung

Die unterrichtsbezogene Analysekompetenz gilt als bedeutsam für erfolgreiches adaptives Handeln im Unterricht. Sie zeigt sich darin, dass Lehrpersonen ihr professionelles Wissen nutzen, um relevante Situationen im Unterricht zu identifizieren, deren Bedeutung in Bezug auf das Lernen der Schülerinnen und Schüler zu interpretieren sowie begründete Optimierungsmöglichkeiten zu entwickeln. Im Rahmen der Interventionsstudie Videos in der Ausbildung von Lehrpersonen (VideA) wurde mit einem quasi-experimentellen Prä-Post-Kontrollgruppendesign (N = 128 Studierende) untersucht, ob sich die unterrichtsbezogene Analysekompetenz durch die Analyse von eigenen bzw. fremden Unterrichtsvideos fördern lässt. Inhaltlicher Fokus der Intervention waren die drei Unterrichtsqualitätsmerkmale „Zielklarheit“, „Lernbegleitung“ und „Lernatmosphäre“. Zur Erfassung der Analysekompetenz wurden sowohl offene Fragen als auch standardisierte Ratingitems zu Videosequenzen eingesetzt. Die inhaltsanalytische Auswertung der Kommentare zu den offenen Fragen zeigt einen signifikanten Zuwachs beim Identifizieren lernrelevanter Situationen im Unterricht in beiden Videogruppen, hingegen keine Veränderung beim Vorschlagen von Optimierungsmöglichkeiten und beim Begründen von deren Bedeutung für die Lernenden. Die Ergebnisse deuten zudem darauf hin, dass die Fähigkeit des Erkennens und des wissensbasierten Interpretierens von lernrelevanten Situationen im Unterricht einen positiven Zusammenhang aufweisen.

Schlagworte

Analysekompetenz; Lernen mit Unterrichtsvideos; Lehrerinnen- und Lehrerbildung; Offenes und standardisiertes Erhebungsinstrument

1. Introduction

Teaching is a complex process in which teachers regularly find themselves in ambiguous and unpredictable situations that require the ability to act flexibly and adaptively. It is therefore important that they are able to identify significant situations, to interpret their relevance to student learning processes (Berliner, 2001; Sherin, 2007), and, if need be, to infer possibilities for improvements. The ability to analyze teaching is considered to be a crucial prerequisite for successful classroom teaching (Sherin, Jacobs, & Philipp, 2011). In order to enhance this ability, teacher preparation programs have increasingly made use of classroom videos (Gaudin & Chaliès, 2015). First studies have already yielded empirical evidence that points to the overall effectiveness of video-based learning as regards the pro-

motion of analytical skills in initial teacher education (e.g., Hellermann, Gold, & Holodynski, 2015; Stürmer, Könings, & Seidel, 2013). The more specific question concerning the effects on particular facets of student teachers' analytical skills is still in need of clarification, however. The intervention study VideA (Video Analysis in Teacher Education) addresses this question and explores the effects of case-based learning with regard to whether student teachers' own, or other teachers' classroom videos are used for fostering different facets of teaching-related analytical skills. Furthermore, the study pursues the question of whether there is a correlation between the ability to identify features of effective teaching and the ability to interpret teaching situations by referring to these features.

2. Theoretical background

2.1 Analytical skills of teachers

The ability to analyze classroom teaching is an essential aspect of a teacher's expertise (Seidel & Stürmer, 2014). It is grounded in professional knowledge (Blömeke, Gustafsson, & Shavelson, 2015; König et al., 2014; Plöger & Scholl, 2014; Stürmer et al., 2013) that consists of general pedagogical knowledge, content knowledge, and pedagogical content knowledge (Shulman, 1987). First empirical findings support the assumption that teaching-related analytical skills and especially the facet pertaining to the development of alternative teaching strategies are associated with more student-centred practices (Sun & van Es, 2015) and student performance (Kersting, Givvin, Sotelo, & Stigler, 2010; Kersting, Givvin, Thompson, Santagata, & Stigler, 2012; Roth et al., 2011). It is therefore of crucial importance to foster analytical skills already in initial teacher education.

Research on analytical skills has its origins in Goodwin's (1994) concept of professional vision that was transferred to classroom teaching by Sherin (e.g., 2007) and includes the two fundamental processes of noticing and interpreting significant classroom situations (Seidel & Stürmer, 2014). Noticing (or selective attention) refers to the identification of aspects that are relevant to learning while interpretation is understood as knowledge-based reasoning about classroom practices in terms of the impact of teachers' decision on student learning. As for the second ability, Seidel and Stürmer (2014) distinguish between three empirically separable but closely interrelated dimensions: 1) the ability to describe features that are relevant to learning; 2) the ability to explain such features; and 3) the ability to predict their effects on student learning. In addition to the two processes of noticing/identifying and interpreting, analytical skills include the ability to use one's professional knowledge for developing possible ways of improving learning processes and for deciding what to do next in a particular teaching situation (Blömeke et al., 2015; Plöger & Scholl, 2014).

In our study, the promotion of teaching-related analytical skills focuses on the ability of knowledge-based identification and interpretation of significant classroom situations and the ability to make reasoned suggestions for improvements in teaching. As regards the contents, the training concentrates on three basic features of teaching, namely goal clarity, learning support, and learning climate, all of which are essential conditions for effective teaching irrespective of subject and grade (Hattie, 2009; Seidel & Shavelson, 2007). These three features belong to a teacher's general pedagogical knowledge and correspond to the curriculum of the seminar. "Goal clarity" requires transparency about learning objectives and demands of the lesson as well as a clear lesson structure. "Learning support" consists in process-oriented support of learning processes by means of asking questions that stimulate thinking, adaptive scaffolding and feedback, thus encouraging reflection. The third feature concerns the provision of a positive and supportive "learning climate", in which teachers take students seriously and humor and appreciation are essential (Seidel & Stürmer, 2014; Stürmer & Seidel, 2017).

2.2 Development of teaching-related analytical skills in (student) teachers

Findings from research on expertise indicate that experts tend to center on events that are significant in terms of student learning while novices are often inclined to pay particular attention to the teacher's actions (Berliner, 2001). Studies that explored video-based learning found that experienced teachers, owing to their sophisticated professional knowledge, discern the complexities of classroom teaching in more detail and are better able to keep track of what is going on than novices (Stürmer et al., 2013). This allows them to identify relevant situations and to develop alternative teaching strategies in a flexible way. Furthermore, there is evidence showing that novices, as opposed to experts who are better able to interpret possible effects of teaching on students' learning, mainly stick to descriptions of classroom situations (Seidel & Prenzel, 2007).

The study conducted by Gold, Hellermann, and Holodynski (2016) on student teachers', teacher candidates', and practicing teachers' abilities to identify and interpret teaching situations in terms of classroom management did not find significant differences between the three samples. All the same, the findings show that the student teachers proposed less alternative teaching strategies than the other groups. Moreover, their comments included the highest number of descriptions. This is consistent with the findings published by Schäfer and Seidel (2015) who looked at how student teachers identified and interpreted relevant classroom situations with respect to the features "goal clarity" and "learning climate". Although the participants were able to identify several relevant events, they also paid attention to comparatively irrelevant aspects. Furthermore, they had difficulties in knowledge-based reasoning about classroom practices.

In recent years, teacher preparation programs have increasingly made use of case-based learning with classroom videos for fostering analytical skills (Gaudin & Chaliès, 2015). Classroom videos are a suitable means because they capture the complexity of teaching in particularly authentic ways, allow repeated observations of complex situations, and facilitate multi-perspective analyses (Santagata, 2014). Also, video-based learning opportunities are considered to be more motivating than text-based case studies (Moreno & Valdez, 2007). Empirical research on video-supported learning in teacher education has provided evidence of positive effects on teaching-related analytical skills in both student teachers (e.g., Hellermann et al., 2015; Stürmer et al., 2013) and practicing teachers (e.g., Sherin & van Es, 2009; Tripp & Rich, 2012). The findings available so far indicate that learning with videos fosters the ability to identify relevant classroom situations (Barnhart & van Es, 2015; Star & Strickland, 2008) and to interpret them by referring to pertinent knowledge (Stürmer et al., 2013; Yeh & Santagata, 2015) as well as the ability to propose suggestions for improvements (Santagata & Guarino, 2011).

Case-based learning with videos can center on clips from one's own or another teacher's classroom. The suitability of the two types of video has been investigated in several studies. Seidel, Stürmer, Blomberg, Kobarg, and Schwindt (2011) showed that analyses of one's own teaching led to a higher degree of immersion and that they are perceived to be more authentic and more motivating than analyses of other teachers' videos. All the same, a potential disadvantage of videos that document the teachers' own practice is that they are less critically commented on than videos from unknown classrooms. This is in line with the findings proposed by Kleinknecht and Schneider (2013) who demonstrated, that videos of other teachers prompted more suggestions for possible improvements in teaching than the participants' own videos. What is still largely unresolved, by contrast, is the question of whether videos from one's own classroom or other teachers' videos are more suitable for fostering specific facets of a teacher's analytical skills. Moreover, there are hardly any studies with larger samples and control-group design that evaluate the effectiveness of video-based learning in teacher education (Gaudin & Chaliès, 2015). One study that is already available is the one by Hellermann et al. (2015) who investigated the effects of the two types of video on student teachers' analytical skills concerning classroom management. In comparison with the untreated control group, both exclusive learning with other teachers' videos and combined learning with other teachers' and the participants' own videos had positive effects on the development of the participants' analytical skills. The largest increase was achieved in the group that had worked with both types of video.

Irrespective of the type of video, the instructional setting is a decisive factor in the video-supported promotion of analytical skills (Blomberg, Renkl, Sherin, Borko, & Seidel, 2013; Brophy, 2004). Successful learning with videos requires a clear focus of the analyses and a well-structured learning environment (Borko, Koellner, Jacobs, & Seago, 2011). This ensures that (student) teachers focus their attention on relevant aspects and facilitates a profound analysis of the recorded classroom situations. The "Lesson Analysis Framework" devised by Santagata and

Guarino (2011) is designed to assist student teachers in learning to notice and interpret relevant situations and to develop suggestions for improvements in teaching. The VideA study used this framework for the analysis of the student teachers' own and other teachers' videos (Krammer, Hugener, Frommelt, Furrer Auf der Maur, & Biaggi, 2015).

The VideA results available so far show that the participants considered both types of video to be effective in terms of the development of their professional competence and that they were motivated and liked dealing with authentic examples. Working with one's own video received the highest ratings (Krammer et al., 2015). In summary, the analysis of both types of videos enhanced the ability to interpret teaching situations on the basis of pertinent knowledge (Krammer et al., 2016), which was measured by means of the standardized video-based tool "Observer" (Seidel & Stürmer, 2014; Stürmer & Seidel, 2017).

2.3 Measurement of analytical skills

In teacher education, classroom videos serve not only to foster teaching-related analytical skills but also to measure them. Video clips situate teaching in a concrete context, capture its complexity, and make it possible to measure the analytical skills of teachers in a contextual frame (Blömeke, 2013). Ratings of and comments on videotaped teaching situations give an indication of how well-developed the target skills are. Instruments commonly used for this purpose are either open-ended question formats with a subsequent coding of the answers (e.g., Sherin & van Es, 2009) or standardized tools with items whose rating can be compared to an expert rating (e.g., Gold & Holodynski, 2017; Meschede, Steffensky, Wolters, & Möller, 2015; Seidel & Stürmer, 2014). Rating items of standardized, closed formats direct the teachers' attention to specific aspects whereas open questions facilitate the measurement of the ability to identify teaching situations that are relevant to learning and the ability to propose suggestions for improvements. So far, the two types of instrument have only rarely been used in combination, which is why their synergistic potential has not been fully explored yet (Gold et al., 2016; Schäfer & Seidel, 2015).

3. Research questions

Our contribution investigates the effects of the use of classroom videos in initial teacher education in terms of specific facets of teaching-related analytical skills. We aim to find out to what extent student teachers are able to apply their theoretical knowledge of three basic features of effective teaching – goal clarity, learning support, and learning climate – to real-life situations if they are asked to write com-

ments on video clips and to suggest improvements in teaching. The results of our study are supposed to answer the following research questions:

1. Are the participating student teachers able to identify the three basic features of effective teaching (goal clarity, learning support, learning climate)?
2. Do the student teachers propose suggestions for improvements that relate to how student learning could be additionally enhanced through teaching strategies that relate to the three basic features of effective teaching?
3. Do the student teachers provide knowledge-based reasons for why they think that the identified features and their suggestions for improvements are significant for student learning?

We expect student teachers to be able to notice the three basic features of effective teaching after the intervention. As novices are less able to interpret effects of teaching on students' learning and propose alternative teaching strategies (e.g., Gold et al., 2016; Seidel & Prenzel, 2007) and due to student teachers' still limited professional knowledge in the first year of their preparation program, we assume student teachers' learning gain in providing reasons for the relevance of the identified features and to suggest alternative teaching strategies to be lower. However, due to the provided learning opportunities during the intervention we hypothesize positive changes in the three examined facets of the student teachers' analytical skills.

The results of our study are also supposed to answer the question if the three intervention groups differ in their abilities to identify the three basic features of effective teaching, propose suggestions for improvements and provide reasons for why they think that the identified features and their suggestions for improvements are significant for student learning. We expect student teachers in the three intervention groups to benefit from the intervention and show positive developments in their analytical skills. Students in the video groups are predicted to profit more than the control group. As working with one's own videos met with the highest degree of acceptance in our study and as it is considered to lead to a higher degree of immersion (Seidel et al., 2011), we expect this video group to show the highest change in identifying and interpreting. In contrast, because of less emotional involvement and more critical stance when analyzing other teachers' videos (Kleinknecht & Schneider, 2013), student teachers' in this intervention group may focus more on proposing suggestions for improvement, which could lead to the highest change between the pre- and the post-test in this facet of student teacher's analytical skills.

Moreover, we are interested in the way in which the results of this methodological approach with open-ended questions are associated with the results generated by means of the standardized tool "Observer". Thus, the following research question was addressed in an explorative way:

4. Do the facets of analytical skills (identifying, providing reasons) as measured in comments on open questions correlate with the ability to interpret (describing, explaining, predicting) teaching situations by referring to theoretical knowledge as measured by the standardized tool "Observer"?

4. Design and method

4.1 Project design

The intervention was embedded in a compulsory seminar and involved three conditions with nine seminar groups in total. Three of these groups worked with videos from their own practice, which had been recorded during an internship, and with the teaching materials they had used, while three other groups dealt with videos and materials from other teachers. So as to evaluate the effects of case-based video analyses, we included a control group with three further seminar groups that attended to written documents and materials instead of videos. Both the participating student teachers and the facilitators were randomly assigned to one of the three conditions. In preparation of the intervention, the facilitators had been introduced to the contents and to the procedure of case analysis in a three-day workshop.

For all three conditions, the examples of classroom teaching and the supplementary materials had been specifically selected to ensure that they provided a solid basis for an analysis with respect to the three focused basic features of effective teaching. They had to contain the realization of at least one of the three focused features in an explicit and clearly observable way. In all conditions, the participants had to analyze the examples (their own or other teachers' videos; written materials) for 90 minutes per week. The analyses were structured according to the "Lesson Analysis Framework" mentioned above (Santagata & Guarino, 2011). The theoretical basics were made available in a script that included descriptions, indicators, and examples of the three features (details in Biaggi, Krammer, & Hugener, 2013; Krammer & Hugener, 2014; Krammer et al., 2015).

In order to check the implementation, in each seminar group the intervention was filmed once. These videotaped seminar sessions were inspected by two raters. The results confirmed that all facilitators followed the procedure of the "Lesson Analysis Framework" and focused on the three basic features of effective teaching.

4.2 Sample

In total, 163 student teachers from the University of Teacher Education Lucerne, Switzerland, took part in the intervention. As stated above, participation was compulsory. 35 participants had to be excluded from the analyses because of missing data, which resulted in a final sample consisting of 128 student teachers. 46 of them worked with their own videos, 40 with videos from other teachers, and 42 with written teaching/learning materials. All participants were in their second semester and studied for a teaching certificate at kindergarten and primary-school

level (71.1 %) or at lower secondary-school level (28.9 %). On average, they were 21.75 years old ($SD = 2.10$), and 78.9 % of them were female.¹

4.3 Instruments

Before and after the intervention, the participants completed an online video-survey that consisted of two parts. The first part measured different facets of teaching-related analytical skills (identification of key features of effective teaching, proposing suggestions for improvements, providing reasons) by means of written comments on three video clips while the second part consisted in a shortened version of the standardized tool “Observer” (Seidel & Stürmer, 2014).

4.3.1 Written comments

The open question format asked the student teachers to write comments on videotaped classroom teaching. The task consisted of three video clips from different school levels and subjects that had been specifically selected so that they illustrated the features “goal clarity”, “learning support”, and “learning climate”. The two questions to be answered were the following: “1. Please describe what you have noticed in these teaching situations. Give reasons for why you think that your observations are important in terms of student learning. 2. Please make suggestions for how student learning could be additionally enhanced and give reasons for your suggestions”.

Before the intervention, the average length of the comments on the two questions amounted to 82 words per video clip. After the intervention, the comments contained an average of 98 words per video clip and were thus significantly longer than at the beginning of the intervention (Wilcoxon test: $z = 3.42, p < .01$).

4.3.2 “Observer”

In the standardized survey format, which also included three video clips, the participants had to rate a range of items in terms of the three basic features of effective teaching on a four-level scale. Their ratings were thereafter compared to an expert rating (three experts, Cohen’s $\kappa = .79$) so that the percentage of agreement could be determined. As several studies have shown, the tool “Observer” is a well-suited means for measuring aspects of professional vision respectively the ability of knowledge-based interpretation of teaching situations in a valid and reliable

¹ In first year of the full time preparation program, all student teachers follow the same curriculum, which mainly covers educational psychology, general pedagogy and subject specific pedagogy. The three intervention groups were comparable with respect to age, sex and the school level they were being prepared for.

way (Seidel & Stürmer, 2014). In our study, we reached good EAP/PV reliabilities for knowledge-based interpretation ($\alpha = .81$) and its three dimensions “describing” ($\alpha = .77$), “explaining” ($\alpha = .81$), and “predicting” ($\alpha = .85$).

4.4 Data analyses

4.4.1 Coding of the comments

The content analysis of the comments on the video clips was based on a category system that was grounded in both theory and data and had been newly developed for the purposes of the study. The theoretical foundation consisted in the script concerning the three basic features of effective teaching that had been created by the research team and served as a reference in the video analyses of the intervention. The categories had been derived from comments of the pilot study² and were thereafter validated by applying them to comments of the main sample. The focus of the analysis was twofold: 1) content of the comments (aspects of basic features of effective teaching that the participants had identified and/or mentioned in their suggestions for improvements) and 2) the quality of the reasons that the participants had provided for their suggestions.

The coding procedure first divided the comments into thematic units that were defined by one of the three basic features of effective teaching (goal clarity, learning support, learning climate) and, if existing, reasons for their significance for student learning. Thus, each thematic unit was coded in terms of content and quality of reasons. The content of the comments was separately categorized for each identified aspect of the three basic features of effective teaching and for each suggestion for improvements. The aspects that the student teachers could theoretically have noticed in the three video clips had beforehand been determined by three experts (see Table 1). All of them had expertise in educational psychology and were familiar with both the intervention and the associated materials. The coders decided for each of the participants' observations whether they corresponded to one of the predetermined aspects. Statements that did not relate to one of the three features but referred to aspects like clothing, language, or the furniture of the classroom were assigned to the category “Further aspects”.

For evaluating the quality of the reasons, the research team defined three categories that were similar to the “levels of interpretation” proposed by Kersting, Givvin, Thompson, Santagata, and Stigler (2012). The first category applies to comments that merely describe the teaching situation whereas the second category contains comments that, at least in part, provide an interpretation of the teaching situation. The reasons rudimentarily relate to student learning processes. Essential concepts of educational psychology are mentioned but not further explained (e.g.,

2 In the pilot study all instruments and intervention materials had been tested with 110 student teachers.

Table 1: Categories relating to the content of the comments (with examples)

Content	Category/Aspect	Example
Goal clarity	Clarification of goals	“Definition of clear goals so that the students know why they have to learn a certain thing (transparency).”
	Clarification of expectations	“In this circle, the task was explained and looked at in more detail so that the students ask as less questions as possible while working and can fully concentrate on the task.”
	Use of media enhances understanding	“The teacher could have asked the students to create a transparency so that all students can see the posters in the classroom.”
	Contextualization: verbal	“The teacher established a close connection with everyday life by planting the calculations also rhetorically in the restaurant.”
	Contextualization: media	“Getting started together arouses the students’ interest and curiosity. The prepared environment with objects from the restaurant activates the students’ prior knowledge.”
	Structure of the teaching	“The teaching situation is very structured. There is an introduction to the topic, a phase of student work, and at the end of the lesson the teacher documents the results.”
Learning support	Questions: degree of open-endedness	“Open-ended question format; that is to say, the students were free to find their own way of solving the problem.”
	Factual-constructive feedback	“Furthermore, I noticed that the teacher gave factual and constructive feedbacks when the students asked him whether their solution was correct. Each time, the teacher answered and stimulated further thoughts, but without giving the solution away.”
	Reflection on the procedure	“The students present their own solutions. This shows that there isn’t just one way of doing it.”
Learning climate	Teacher takes the students seriously	“The teacher made time for each group and carefully looked at the calculations as well. The students can see this as a confirmation, and it shows that the teacher takes the students and their suggestions seriously.”
	Humor	“The teacher uses humor in a conscious and purposeful way.”
Further aspects	--	“The teacher said ‘um’ too often.”

metacognition, motivation). The third category consists of comments providing elaborate reasons and comprehensive descriptions of the significance for student learning processes (see Table 2).

Table 2: Categories relating to the quality of the reasons (with examples)

Category	Example
No reasons	“The teacher presents a real-world problem to the students that can be transferred to everyday life.”
Rudimentary reasons	“They dealt with a real-world problem that was connected to a brain-teaser. This arouses the students’ interest and motivation.”
Elaborate reasons	“By telling this little story, the teacher confronts the students with a real-world problem. I can imagine that the children now ask themselves why it isn’t possible to transport light. They try to understand this. By doing so, they activate their prior knowledge, which allows them to integrate new knowledge in a better way.”

4.4.2 Reliability of the coding

The comments were coded by six trained persons who had no information about the intervention groups or the time-points of the measurements. In a training, coders were introduced to the coder manual and its application based on data of the pilot study. The manual contained the description of the categories and examples. To achieve a high consensus about the different codes difficult examples were discussed and added to the category system in order to specify the definitions. Before the coding and after the first half of the comments had been coded, we conducted a reliability check on all categories. For this purpose, we randomly selected three to five comments per category, each of which consisted of at least twenty thematic units. The percentage of agreement served as the reliability measure and had to meet at least 85 % according to Hugener, Pauli, and Reusser (2006). The first check compared the coding with the model coding by the research team while the second check pertained to inter-rater agreement. Both checks resulted in reliability values of at least 85 % for each facet of teaching-related analytical skills.

4.4.3 Statistical analyses

The quantitative analyses of the comments were separately conducted for the identified features and for the suggestions for improvements. We determined the relation between the number of features/suggestions mentioned in the comments and the total number possible as defined by the category system. As for the quality of the reasons, we determined the proportions of the three reasoning categories to the total number of the identified features of effective teaching and the suggestions for improvements. This procedure was chosen in order to prevent bias because of the varying length of the comments.

For investigating the development of the ability to identify significant teaching situations, the ability to provide suggestions for improvements, and the ability to provide adequate reasons in the comments, we conducted repeated-measures multivariate analyses of variance. The relations between the ability to identify basic

features of effective teaching and to provide reasons for their relevance as measured in the comments and the ability to provide knowledge-based interpretations of teaching situations as measured with the standardized tool “Observer” were calculated by means of Pearson correlations.

5. Results

5.1 Identification of basic features of effective teaching

The first research question focused on the identification of features of teaching that are conducive to learning. Before the intervention, the student teachers had, on average, been able to identify 20 % of those 23 aspects that were included in the category system and thus theoretically observable in the video clips. After the intervention, the proportion of identified aspects amounted to an average of 24 % (see Table 3). The results of the repeated-measures analyses of variance show that the participants of the total sample achieved a significant increase with a medium effect size in their ability to identify basic features of effective teaching (see Table 4). The increase is statistically significant for all three features (goal clarity: $t(127) = -2.36, p < .05$; learning support: $t(127) = -3.18, p < .01$; learning climate: $t(127) = -2.81, p < .01$). The highest increase was attained in the feature “learning support”. Apart from aspects concerning the three key features of effective teaching, the student teachers mentioned two or three further aspects on average and did so before and after the intervention (see Table 3).

Table 3: Identification of basic features of effective teaching and further aspects ($N = 128$)

	Measure- ment	Descriptive statistics							
		Absolute values				Proportions in percentages ^a			
		<i>M</i>	<i>SD</i>	Min	Max	<i>M</i>	<i>SD</i>	Min	Max
All features of effective teaching (23) ^b	t_1	4.53	2.07	1	12	0.20	0.09	0.04	0.52
	t_2	5.55	2.47	1	12	0.24	0.11	0.04	0.52
Goal clarity (10) ^b	t_1	2.20	1.25	0	6	0.22	0.12	0.00	0.60
	t_2	2.50	1.18	0	6	0.25	0.12	0.00	0.60
Learning support (8) ^b	t_1	1.59	1.15	0	5	0.20	0.14	0.00	0.63
	t_2	2.05	1.52	0	7	0.26	0.19	0.00	0.88
Learning climate (5) ^b	t_1	0.74	0.86	0	4	0.15	0.17	0.00	0.80
	t_2	1.00	0.76	0	3	0.20	0.15	0.00	0.60
Further aspects	t_1	2.88	1.98	0	12	-	-	-	-
	t_2	3.05	2.08	0	9	-	-	-	-

^aProportion of identified aspects of effective teaching (three video clips) relative to expert analysis in percentages. ^bNumber of aspects of effective teaching identified by experts (three video clips).

A comparison between the three groups shows that the groups that had worked with their own or other teachers' videos identified significantly more aspects of effective teaching after the intervention than before the intervention. The control group, by contrast, which had solely dealt with written materials, did not do significantly better after the intervention (see Table 4). The results of the repeated-measures analysis of variance revealed, however, that the development of the ability to identify features of effective teaching did not significantly differ between the three groups ($F(2, 125) = 0.85, p = .43, \eta^2 = .01$).

Table 4: Development of the ability to identify aspects of features of effective teaching by intervention group ($N = 128$)

Sample	<i>N</i>	<i>F</i>	<i>df</i>	η^2	Pretest		Posttest		$\Delta t_1, t_2$	
					<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
All intervention groups	128	20.40***	1,127	0.14	0.20	0.09	0.24	0.11	0.04	0.11
Other teachers' videos	40	11.39**	1, 39	0.23	0.19	0.11	0.26	0.12	0.06	0.12
Participants' own videos	46	7.27*	1, 45	0.14	0.20	0.08	0.23	0.10	0.04	0.10
Teaching/learning materials	42	3.11	1, 41	0.07	0.20	0.08	0.24	0.10	0.03	0.12

* $p < .05$, ** $p < .01$, *** $p < .001$.

5.2 Suggestions for improvements in teaching

The second research question focused on the participants' suggestions for improvements in teaching through strategies that relate to the three basic features of effective teaching. On average, the student teachers mentioned 11 % of the 16 possibilities included in the category system before the intervention and 12 % after the intervention (see Table 5). Before the intervention, the participants had most often proposed suggestions that concerned the feature "learning support", but there was no increase after the intervention ($t(127) = 1.21, p = .23$). The same applies to the feature "goal clarity" in which we found no increase either ($t(127) = 0.97, p = .34$). Only as regards the feature "learning climate", the student teachers provided more suggestions for improvements after the intervention than they had done before ($t(127) = 4.74, p < .001$). Before and after the intervention, the participants moreover mentioned two to three suggestions for improvements concerning other aspects on average.

Table 5: Suggestions for improvements in terms of the three basic features of effective teaching and further aspects ($N = 128$)

	Measurement	Descriptive statistics							
		Absolute values				Proportions in percentages ^a			
		<i>M</i>	<i>SD</i>	Min	Max	<i>M</i>	<i>SD</i>	Min	Max
All features of effective teaching (16) ^b	t_1	1.78	1.21	0	6	0.11	0.08	0.00	0.38
	t_2	1.91	1.28	0	6	0.12	0.08	0.00	0.38
Goal clarity (9) ^b	t_1	0.70	0.87	0	4	0.08	0.10	0.00	0.44
	t_2	0.60	0.85	0	3	0.07	0.09	0.00	0.33
Learning support (4) ^b	t_1	0.82	0.68	0	3	0.21	0.17	0.00	0.75
	t_2	0.73	0.66	0	2	0.18	0.16	0.00	0.50
Learning climate (3) ^b	t_1	0.27	0.48	0	2	0.09	0.16	0.00	0.67
	t_2	0.57	0.71	0	3	0.19	0.24	0.00	1.00
Further aspects	t_1	2.20	1.42	0	8	-	-	-	-
	t_2	2.52	1.53	0	9	-	-	-	-

^aProportion of suggested improvements (three video clips) relative to expert analysis in percentages.

^bNumber of expert suggestions for improvements (three video clips).

The results of the repeated-measures analysis of variance confirm that the student teachers of the total sample did not make more suggestions for improvements after the intervention than before the intervention (see Table 6). Furthermore, Table 6 shows that there were no differences in the development between the three intervention groups ($F(2,125) = 0.03, p = .97, \eta^2 = .00$). Thus, in none of the groups we could find an increase in the ability to make suggestions for improvements in teaching.

Table 6: Development of the ability to suggest improvements by intervention group ($N = 128$)

Sample	<i>N</i>	<i>F</i>	<i>df</i>	η^2	Pretest		Posttest		$\Delta t_1, t_2$	
					<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
All intervention groups	128	0.71	1,127	0.01	0.11	0.08	0.12	0.08	0.01	0.10
Other teachers' videos	40	0.06	1,39	0.00	0.11	0.09	0.12	0.09	0.00	0.12
Participants' own videos	46	0.42	1,45	0.01	0.10	0.07	0.10	0.06	0.01	0.08
Teaching/learning materials	42	0.44	1,41	0.01	0.13	0.07	0.14	0.09	0.01	0.10

5.3 Quality of the reasons

Our analyses concerning the third research question and thus the quality of the reasons for the significance of the identified features of effective teaching and proposed suggestions for improvements show that before the intervention the student teachers on average did not provide reasons for 62 % of the identified features of effective teaching and their suggestions for improvements. For 35 % of the aspects, the comments contained rudimentary reasons, and 3 % of the aspects were backed by elaborate reasons (see Table 7). After the intervention, the proportion of aspects without reasons was as high as before the intervention and amounted to 62 % again ($t(127) = -0.04, p = .97$). The same can be stated regarding rudimentary reasons for the significance of identified features of effective teaching: There was no significant increase after the intervention ($t(127) = 1.04, p = .30$). Elaborate reasons, by contrast, were provided for 5 % of the aspects, which marks a significant increase (Wilcoxon Test: $z = 2.70, p < .02$).³

Table 7: Descriptive statistics: Quality of reasons ($N = 128$)

		Descriptive statistics							
		Absolute values				Proportion in percentages ^a			
		<i>M</i>	<i>SD</i>	Min	Max	<i>M</i>	<i>SD</i>	Min	Max
No reasons	t_1	3.77	2.02	0	11	0.62	0.26	0.00	1.00
	t_2	4.50	2.49	0	13	0.62	0.26	0.00	1.00
Rudimentary reasons	t_1	2.35	1.96	0	9	0.35	0.24	0.00	0.88
	t_2	2.53	2.01	0	8	0.32	0.23	0.00	1.00
Elaborate reasons	t_1	0.20	0.67	0	5	0.03	0.09	0.00	0.71
	t_2	0.42	0.78	0	5	0.05	0.10	0.00	0.56

^aProportions of reasons provided for identified aspects of features of effective teaching and suggestions for improvements in percentages. 100 % refer to the total of identified aspects of features of effective teaching and suggestions for improvements.

The results of the repeated-measures analysis of variance show that there was no difference in the development of the quality of the reasons between the three intervention groups ($F(2,125) = 1.08, p = .34, \eta^2 = .02$). In all three groups, the quality of the reasons remained the same (see Table 8).

3 Because of the dependence of the three variables that relate to the quality of the reasons, the significance level was adjusted through Bonferroni correction.

Table 8: Development of the quality of the reasons (rudimentary and elaborate) by intervention group ($N = 128$)

Sample	N	F	df	η^2	Pretest		Posttest		$\Delta t_1, t_2$	
					M	SD	M	SD	M	SD
All intervention groups	128	0.01	1, 127	0.00	0.38	0.26	0.38	0.26	0.00	0.34
Other teachers' videos	40	0.63	1, 39	0.02	0.39	0.24	0.35	0.25	-0.04	0.32
Participants' own videos	46	0.99	1, 45	0.02	0.39	0.29	0.45	0.29	0.06	0.40
Teaching/learning materials	42	0.39	1, 41	0.01	0.36	0.25	0.33	0.21	-0.03	0.30

Note. Owing to the small number of elaborate reasons, rudimentary reasons and elaborate reasons have been merged.

5.4 Correlations between different facets of analytical skills

The fourth question dealt with correlations between the ability to identify features of effective teaching and to provide reasons for their relevance as reflected in the comments and the ability of knowledge-based interpretation of teaching situations as measured by the tool “Observer”. The reported results exclusively relate to one of the three video clips of the “Observer” because this clip was part of the open question format as well. After the intervention, the agreement between the student teachers' ratings and the experts' rating of the video clip under consideration amounted to 30 % (in total). Moreover, the values of the three dimensions “describing,” “explaining,” and “predicting” are very similar (see Table 9).

The results show that knowledge-based interpretation and its three dimensions significantly correlate with the ability to identify features of effective teaching. The higher the ability to identify relevant teaching situations is, the higher the values for knowledge-based interpretations in the “Observer” are. We also found positive correlations between the reasons provided in the comments and knowledge-based interpretations in the “Observer”, but they were not statistically significant.

Table 9: Descriptive statistics: Knowledge-based interpretation in the tool “Observer” and correlations with comments on one video clip (identification and reasons) ($N = 128$)

	Proportions “Observer” in percentages t_2		Comments t_2	
	<i>M</i>	<i>SD</i>	Identification <i>M</i> = .20, <i>SD</i> = .15 ^b	Reasons ^a <i>M</i> = .23, <i>SD</i> = .31 ^b
Knowledge-based interpretation (in total)	0.30	0.26	.33**	.14
Describing	0.28	0.26	.27**	.12
Explaining	0.29	0.27	.27**	.08
Predicting	0.32	0.30	.35**	.17

^aOwing to the small number of elaborate reasons, rudimentary reasons and elaborate reasons have been merged. ^bIn percentages.

** $p < .01$.

6. Discussion

The aim of our article was to investigate the effects of learning with classroom videos on different facets of teaching-related analytical skills in initial teacher education. Furthermore, we were interested in how the ability to identify basic features of effective teaching and to give reasons for their relevance as reflected in the comments and the ability of knowledge-based interpretation of teaching situations as measured by the tool “Observer” correlate. In what follows, we summarize and discuss our findings and conclude with some remarks on the limitations of our study and an outlook.

6.1 Development of different facets of teaching-related analytical skills

With regard to the effects of case-based learning with classroom videos, the findings vary between the different facets of teaching-related analytical skills. After the intervention, the student teachers of the two video groups were significantly better able to identify teaching situations that are relevant to learning than the control group. This means that the structured analysis of written materials did not in the same way lead to a measurable increase in the ability to identify the three focused features of effective teaching and might be attributed to the media-specific advantages of classroom videos to foster analytical skills of teachers (e.g., authenticity, capture complexity of teaching). Because there were no differences between the two video groups, we can conclude that structured analyses of classroom vide-

os in general can contribute to enhancing the ability to identify essential features of effective teaching, irrespective of whether the student teachers deal with their own or with other teachers' videos.

As for the suggestions for improvements in teaching, there was no significant increase in any of the three intervention groups. When the participants suggested alternative teaching strategies, they mostly related to the feature "learning support", no matter whether the suggestions were made before or after the intervention. The highest increase occurred in the feature "learning climate". This finding may indicate that it is easier to acquire the ability to suggest improvements pertaining to this feature than to acquire this ability in connection with the other two features.

Apart from the focused features of effective teaching, the participants' comments also included several other aspects of classroom teaching. The findings by Schäfer and Seidel (2015) make it plausible to assume that the category "Further aspects" may at least in part contain irrelevant features. Additional analyses of the comments are needed for clarifying this hypothesis, however.

In general, the student teachers had difficulty in providing reasons for the relevance of identified features of effective teaching and for their suggestions for improvements both before and after the intervention. The proportion of features that are supplemented with elaborate reasons is very low in both measurements. Thus, establishing connections with student learning processes and linking them to basic knowledge of educational psychology seems to be a demanding task for student teachers. This is consistent with the findings published by Schäfer and Seidel (2015) who also reached the conclusion that student teachers find it difficult to give knowledge-based reasons for their observations.

Summing up our findings, we can conclude that case-based learning with classroom videos at the beginning of a teacher preparation program was not to the same extent effective in fostering all three facets of teaching-related analytical skills. The fact that there was neither an increase in the ability to make suggestions for improvements nor an increase in the ability to provide reasons could point to the limits of the learning gains to be expected within one single semester.

Another explanation might be the timing of the intervention. It took place in the first year of the preparation program. Since this is quite early, it could account for why the student teachers had not been fully able to apply their newly acquired knowledge flexibly to authentic teaching situations (Berliner, 2001; Steffensky, Gold, Holodynski, & Möller, 2015). The finding that it is easier for student teachers to identify relevant teaching situations than to interpret their significance in terms of the impact of teachers' decisions on student learning and to propose reasoned suggestions for improvements could therefore indicate that the second and the third facet rest on more closely interconnected knowledge structures than identifying alone. This assumption can be supported by referring to the model of teaching-related analytical skills devised by Plöger and Scholl (2014). Their model is grounded in research on expertise and places the ability to make reasoned suggestions for alternative teaching strategies on higher levels that require a well-de-

veloped perception of complexities and the integration of pertinent knowledge. Furthermore, also the study by Gold et al. (2016) showed that student teachers were less capable of suggesting alternative strategies than teacher candidates and experienced teachers. Thus, it is vital that teacher preparation programs not only impart knowledge of educational psychology, but also provide student teachers with opportunities to relate this knowledge to authentic examples, as this is practiced in video-supported case studies.

Furthermore, our findings point to the importance of the facilitators' role in case-based work with classroom videos. Other studies particularly emphasize the aspect of how the analysis is guided (e.g., Borko et al., 2011; Santagata & Guarino, 2011; van Es, Tunney, Goldsmith, & Seago, 2014). This could point to further explanations for the differences in the development of the three facets. It is possible, for instance, that in our intervention – although the procedure was standardized – the ability to give reasons and the ability to suggest improvements were less intensely practiced than the ability to identify situations that are relevant to learning. The implementation check confirmed a comparable application of the “Lesson Analysis Framework”, but the facilitators' support activities were not analyzed in detail. Hence, this point requires further clarification.

6.2 Correlations between different facets of analytical skills

The results from the complementary survey with the standardized tool “Observer” show that case-based learning with classroom videos can foster the ability to interpret teaching situations on the basis of pertinent knowledge (Krammer et al., 2016). Those facets that were measured with both instruments (comments and “Observer”) correlate positively, but only the correlation between identification and knowledge-based interpretation was statistically significant. König et al. (2014) were also able to find a positive, though not significant correlation between the identification and the interpretation of general pedagogical aspects of classroom teaching. Therefore, the authors conclude that the two facets are only loosely associated. Similarly, loose connections were found by Gold et al. (2016), who conducted their study on professional vision concerning classroom management. The number of identified events (open question format) did not significantly correlate with the results from a standardized scale that measured the abilities relating to describing and interpreting. Owing to the differences in the instruments, the focused contents, and the design of the studies, these findings are comparable to ours only within limits, however.

That the correlations between knowledge-based interpretation in the “Observer” and the ability to give reasons in comments were not significant could indicate that the latter is a distinct facet that requires more flexibly structured knowledge or even other skills than the former. Since while the answers in standardized instruments are limited to a particular range of options, an open format involves unguided recalling of relevant theoretical knowledge and associating this knowledge with

a given teaching situation. Without this flexibility in the use of pertinent knowledge it is not possible to generate hypotheses concerning the effects of a teacher's actions on student learning and thus not possible to provide well-founded reasons.

6.3 Limitations and outlook

The only minor developments of student teachers' different facets of teaching-related analytical skills, which can also be ascribed to the design of the study, might be due to the timing of the second measurement, which took place at the end of the semester and thus simultaneously with the exam period. It is likely that the participants devoted less priority to the second survey than to their exams, which raises doubts about whether we could document their actual progress.

Our findings concerning the correlations between the different facets of teaching-related analytical skills rest on the analysis of one single video. Hence, we should emphasize that the reported results are only explorative in nature and that they merely provide first indications of correlations between the abilities under discussion. In order to clarify the question of whether and how these abilities correlate, more research is necessary.

Moreover, our data do not provide information about the effects of video analyses on the participants' actual behavior in the classroom or on the learning gains of the students taught by the participants. Pursuing these questions would yield knowledge on the effects of a video-based enhancement of teaching-related analytical skills on practical teaching and thus clarify whether this approach is functional. This, in turn, would provide valuable inputs when it comes to improving the design of teacher preparation programs.

Despite the limitations mentioned above, the study presented in this article makes a contribution to the research on effects of video-based interventions on the development of different facets of teaching-related analytical skills at the beginning of initial teacher education. So as to document and monitor this development over the course of the whole preparation program, a longitudinal study would be required. Furthermore, this study provides valuable information for the design of video-based teacher education as our findings show that students' own videos and videos from other teachers can be used to foster their ability to identify relevant classroom situations. In contrast, however, interpreting classroom situations and proposing suggestions of improvement seem to be challenging for student teachers in the first year of their preparation program. Thus, the question arises how to design propitious settings for fostering these two facets of analytical skills.

Acknowledgements

The research project VideA was conducted at the University of Teacher Education Lucerne from 2012 to 2017 and funded by the Swiss National Science Foundation (Project No. 100016_140863 and 100019_159632). The project was directed by Prof. Dr. Isabelle Hugener[†] and Prof. Dr. Kathrin Krammer, and assisted by

Manuela Frommelt, Sandro Biaggi, and Gabriela Fürrer Auf der Maur. Cooperating research partners were Prof. Dr. Christine Pauli, University of Fribourg, Prof. em. Dr. Kurt Reusser, University of Zurich, Prof. Dr. Tina Seidel, Technical University of Munich, and Prof. Dr. Kathleen Stürmer, University of Tübingen. We also thank Jonna Truniger for editing the manuscript. This article is published in memory of Prof. Dr. Isabelle Hugener who contributed substantially to its completion.

References

- Barnhart, T., & van Es, E. (2015). Studying teacher noticing: Examining the relationship among pre-service science teachers' ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education*, *45*, 83–93.
- Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research*, *35*, 463–482.
- Biaggi, S., Krammer, K., & Hugener, I. (2013). Vorgehen zur Förderung der Analysekompetenz in der Lehrerbildung mit Hilfe von Unterrichtsvideos – Erfahrungen aus dem ersten Studienjahr. *SEMINAR*, *19*(2), 26–34.
- Blomberg, G., Renkl, A., Sherin, M. G., Borko, H., & Seidel, T. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for Educational Research Online*, *5*(1), 90–114.
- Blömeke, S. (2013). Moving to a higher state of confusion. Der Beitrag der Videoforschung zur Kompetenzforschung. In U. Riegel & K. Macha (Eds.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (pp. 25–44). Münster, Germany: Waxmann.
- Blömeke, S., Gustafsson, J. E., & Shavelson, R. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, *223*, 3–13.
- Borko, H., Koellner, K., Jacobs, J., & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM*, *43*, 175–187.
- Brophy, J. (Ed.). (2004). *Using video in teacher education*. Oxford, United Kingdom: Elsevier.
- Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41–67.
- Gold, B., Hellermann, C., & Holodynski, M. (2016). Professionelle Wahrnehmung von Klassenführung – Vergleich von zwei videobasierten Erfassungsmethoden. In D. Prinz & K. Schwippert (Eds.), *Der Forschung – Der Lehre – Der Bildung: Aktuelle Entwicklungen der Empirischen Bildungsforschung* (pp. 103–118). Münster, Germany: Waxmann.
- Gold, B., & Holodynski, M. (2017). Using digital video to measure the professional vision of elementary classroom management: Test validation and methodological challenges. *Computers & Education*, *107*, 13–30.
- Goodwin, C. (Ed.). (1994). Professional vision. *American Anthropologist*, *96*, 606–633.
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Routledge.
- Hellermann, C., Gold, B., & Holodynski, M. (2015). Förderung von Klassenführungsfähigkeiten im Lehramtsstudium: Die Wirkung der Analyse eigener und fremder Unterrichtsvideos auf das strategische Wissen und die professionelle Wahrnehmung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *47*, 97–109.
- Hugener, I., Pauli, C., & Reusser, K. (2006). Videoanalysen. In E. Klieme, C. Pauli & K. Reusser (Eds.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie "Unterrichtsqualität, Lernverhalten und ma-*

- thematisches Verständnis*". Materialien zur Bildungsforschung, Bd. 15. Frankfurt a. M., Germany: GFFP.
- Kersting, N. B., Givvin, K. B., Sotelo, F. L., & Stigler, J. W. (2010). Teachers' analyses of classroom video predict student learning of mathematics: Further explorations of a novel measure of teacher knowledge. *Journal of Teacher Education*, *61*, 172–181.
- Kersting, N. B., Givvin, K. B., Thompson, B. J., Santagata, R., & Stigler, J. W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, *49*, 568–589.
- Kleinknecht, M., & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching? *Teaching and Teacher Education*, *33*, 13–23.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education*, *38*, 76–88.
- Krammer, K., & Hugener, I. (2014). Förderung der Analysekompetenz angehender Lehrpersonen anhand von eigenen und fremden Unterrichtsvideos. *Journal für Lehrerinnen- und Lehrerbildung*, *14*(1), 25–32.
- Krammer, K., Hugener, I., Biaggi, S., Frommelt, M., Furrer Auf der Maur, G., & Stürmer, K. (2016). Videos in der Ausbildung von Lehrkräften: Förderung der professionellen Unterrichtswahrnehmung durch die Analyse von eigenen und fremden Videos. *Unterrichtswissenschaft*, *44*, 357–372.
- Krammer, K., Hugener, I., Frommelt, M., Furrer Auf der Maur, G., & Biaggi, S. (2015). Case-based learning in initial teacher education: Assessing the benefits and challenges of working with student videos and other teachers' videos. *Orbis Scholae*, *9*(2), 119–137.
- Meschede, N., Steffensky, M., Wolters, M., & Möller K. (2015). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht. Theoretische Beschreibung und empirische Erfassung. *Unterrichtswissenschaft*, *43*, 317–335.
- Moreno, R., & Valdez, A. (2007). Immediate and delayed effects of using a classroom case exemplar in teacher education: The role of presentation format. *Journal of Educational Psychology*, *99*, 194–206.
- Plöger, W., & Scholl, D. (2014). Analysekompetenz von Lehrpersonen – Modellierung und Messung. *Zeitschrift für Erziehungswissenschaft*, *17*, 85–112.
- Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Videobased lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, *48*, 117–148.
- Santagata, R. (2014). Video and teacher learning: Key questions, tools, and assessments guiding research and practice. *Beiträge zur Lehrerinnen- und Lehrerbildung*, *32*, 196–209.
- Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM*, *43*, 133–145.
- Schäfer, S., & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. *Journal for Educational Research Online*, *7*(2), 34–58.
- Seidel, T., & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen – Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. *Zeitschrift für Erziehungswissenschaft*, *10* (Special Issue 8), 201–216.
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the last decade: Role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*, 454–499.

- Seidel, T., & Stürmer, K. (2014). Modeling the structure of professional vision in pre-service teachers. *American Educational Research Journal*, *51*, 739–771.
- Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, *27*, 259–267.
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron, & S. Derry (Eds.), *Video research in the learning sciences* (pp. 383–395). Hillsdale, NJ: Erlbaum.
- Sherin, M. G., Jacobs V. R., & Philipp R. A. (Eds.). (2011). *Mathematics teacher noticing. Seeing through teachers' eyes*. New York, NY: Routledge.
- Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, *60*, 20–37.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*(1), 1–22.
- Star, J. R., & Strickland, S. K. (2008). Learning to observe: using video to improve pre-service mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, *11*, 107–125.
- Steffensky, M., Gold, B., Holodynski, M., & Möller, K. (2015). Professional vision of classroom management and learning support in science classrooms – Does professional vision differ across general and content-specific classroom interactions? *International Journal of Science and Mathematics Education*, *13*, 351–368.
- Stürmer, K., Könings, K. D., & Seidel, T. (2013). Declarative knowledge and professional vision in teacher education: Effect of courses in teaching and learning. *British Journal of Educational Psychology*, *83*, 467–483.
- Stürmer, K., & Seidel, T. (2017). A standardized approach for measuring teachers' professional vision: The Observer research tool. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks* (pp. 359–380). Cham, Switzerland: VS.
- Sun, J., & van Es, E. A. (2015). An exploratory study of the influence that analyzing teaching has on preservice teachers' classroom practice. *Journal of Teacher Education*, *66*, 201–214.
- Tripp, T. R., & Rich, P. J. (2012). Using video to analyze one's own teaching. *British Journal of Educational Technology*, *43*, 678–704.
- van Es, E. A., Tunney, J., Goldsmith, L. T., & Seago, N. (2014). A framework for the facilitation of teachers' analysis of video. *Journal of Teacher Education*, *65*, 340–356.
- Yeh, C., & Santagata, R. (2015). Preservice teachers' learning to generate evidence-based hypotheses about the impact of mathematics teaching on learning. *Journal of Teacher Education*, *66*, 21–34.

Charlott Rubach, Rebecca Lazarides & Hendrik Lohse-Bossenz

Engagement der Klassenlehrkräfte in der Zusammenarbeit mit Eltern und Motivation Lernender in der Sekundarstufe

Zusammenfassung

Die gelingende Zusammenarbeit von Eltern und Lehrkräften gilt als eine wichtige Voraussetzung für den schulischen Bildungserfolg Lernender und wirkt sich zudem positiv auf die Beziehung von Lehrkräften und Lernenden aus. Der vorliegende Beitrag untersucht, inwiefern das Engagement von Klassenlehrkräften in der Zusammenarbeit mit Eltern, operationalisiert über angebotene formelle und informelle Kontakte, sowie die von Lernenden wahrgenommene Lehrer-Schüler-Beziehung mit der intrinsischen Motivation Lernender in Zusammenhang stehen. Ausgewertet wurden Daten von 881 Schülerinnen und Schülern aus 39 neunten und zehnten Klassen aus 13 Schulen der Sekundarstufe (Gymnasien und Integrierte Sekundarschulen). Manifest-latente Mehrebenenmodelle zeigen signifikant positive Zusammenhänge zwischen der von Lehrkräften angebotenen Vielfalt formeller Kontakte und der intrinsischen Motivation der Lernenden auf Klassenebene. Auf Individualebene zeigen sich signifikante positive Zusammenhänge zwischen der von Lernenden berichteten Lehrer-Schüler-Beziehung und der intrinsischen Motivation.

Schlagworte

Intrinsische Motivation; Längsschnittliche Mehrebenenanalyse; Vielfältige Zusammenarbeit von Eltern und Lehrkräften; Lehrer-Schüler-Beziehung

Dr. Charlott Rubach (corresponding author) · Jun.-Prof. Dr. Rebecca Lazarides, Arbeitsbereich für Schulpädagogik m.d.S. Schul- und Unterrichtsentwicklung, Universität Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Deutschland
E-Mail: rubach@uni-potsdam.de
rebecca.lazarides@uni-potsdam.de

Jun.-Prof. Dr. Hendrik Lohse-Bossenz, Institut für Psychologie, Pädagogische Hochschule Heidelberg, Keplerstraße 87, 69120 Heidelberg, Deutschland
E-Mail: hendrik.lohse-bossenz@ph-heidelberg.de

Class teacher's engagement in parent-teacher partnership and student's motivation in secondary schools

Abstract

Parent-teacher partnerships are important prerequisites for successful school-related learning processes of students and is positive related to the relationship between teachers and students. This two-wave longitudinal study examines the relationship between parental work of class teachers, measured by their formal and informal communication practices, the teacher-student relationship, and the students' intrinsic motivation. The analyses are based on questionnaire data from N = 881 students of 39 classrooms (9th and 10th grade) at German secondary schools (two school types: Gymnasium and Integrated Secondary School). Manifest-latent multilevel analyses show a significant and positive relationship between formal communication practices and students' intrinsic motivation at the class level. At the individual level, student-reported teacher-student relationship is significantly and positively related to students' intrinsic motivation.

Keywords

Intrinsic motivation; Longitudinal multilevel analysis; Parent-teacher partnership; Teacher-student relationship

1. Einleitung

Die Motivation von Schülerinnen und Schülern gilt als ein zentraler Faktor für erfolgreiche Lernprozesse im schulischen Kontext (Schiefele & Schaffner, 2015). Zahlreiche Studien zeigen, dass eine hohe intrinsische Motivation sowohl die Ausdauer im Lernprozess als auch die Freude am Lernen begünstigt (Rohlf's, 2011; Wild & Lorenz, 2010). Besonders am Ende der Sekundarstufe I ist die Förderung der Lernmotivation in der Schule hochrelevant, da Lernende wichtige berufliche und bildungsbezogene Entscheidungen treffen müssen. Diese sind maßgeblich durch ihre Motivation beeinflusst (Lauer mann, Tsai & Eccles, 2017; Nagy, Trautwein, Baumert, Köller & Garrett, 2006). Der *Stage-Environment Fit-Hypothese* von Eccles et al. (1993) folgend führt die fehlende Passung zwischen schulischen Lehr-Lernarrangements und den Bedürfnissen Lernender zur kontinuierlichen Abnahme der Motivation nach dem Übergang in die Sekundarstufe I. Auf Grund der hohen Relevanz der Motivation Jugendlicher für ihre berufs- und bildungsbezogenen Entscheidungsprozesse ist es notwendig, Möglichkeiten der schulischen Förderung der Motivation in dieser Entwicklungsphase zu untersuchen.

Die Motivation eines Menschen, also die „aktive Ausrichtung des momentanen Lebensvollzugs auf einen positiv bewerteten Zielzustand“ (Rheinberg & Vollmeyer, 2012, S. 15) kann unter anderen in die Lern- und der Leistungsmotivation unter-

schieden werden (Schlag, 2013). Die Lernmotivation ist definiert als die selbstinitiierte Intention eines Individuums, bestimmte Inhalte und Fertigkeiten zu erlernen, während die Leistungsmotivation als Absicht des Erreichens oder Überragens von individuellen oder sozialen Erwartungen beschrieben werden kann (Schiefele & Schaffner, 2015). Als Teilkonzept der Lernmotivation wird in der vorliegenden Studie die schulbezogene intrinsische Motivation fokussiert, welche nach Ryan und Deci (2000, S. 55) darin besteht eine Handlung auszuüben, „because it is inherently interesting or enjoyable“. Ein zentraler Bedingungsfaktor für die positive Entwicklung der Lernmotivation und dem damit verbundenen Bildungserfolg von Schülerinnen und Schülern ist die effektive Unterstützung durch Lehrkräfte sowie durch Eltern (Schiefele, 2004; Wild & Hofer, 2002). Neben der separaten Unterstützung durch Eltern und Lehrkräfte spielt zudem die Kooperation von Eltern und Lehrkräften eine Rolle in der Unterstützung der Motivation der Lernenden (Fan & Williams, 2010; Harris & Goodall, 2007). Allerdings wird in Studien zur Zusammenarbeit von Eltern und Lehrkräften zumeist nur die Perspektive der Eltern fokussiert (Buhl & Hilkenmeier, 2017; Domina, 2005; Schmitt & Kleine, 2010). Die Berücksichtigung der Perspektive von Lehrkräften, um das Engagement der Lehrkräfte in der Zusammenarbeit mit Eltern zu erfassen, kann jedoch von hoher Relevanz sein, um aufzuzeigen, wie das kooperationspezifische Engagement der Lehrkräfte mit der Motivation von Lernenden zusammenwirkt (Deng et al., 2017). Die vorliegende Studie untersucht daher längsschnittlich die Zusammenhänge zwischen dem Engagement der Lehrkräfte in der Zusammenarbeit mit Eltern, der Lehrer-Schüler-Beziehung und der intrinsischen Motivation von Schülerinnen und Schülern am Ende der Sekundarstufe I.

2. Theoretischer Hintergrund

Das Modell der überlappenden Sphären von Epstein (1995, 2010) postuliert, dass verschiedene Beziehungsformen zwischen Schule, Familie und Umwelt den Bildungserfolg von Schülerinnen und Schülern prägen. Hierbei werden inter- und intrainstitutionelle Beziehungsformen unterschieden: Intrainstitutionelle Beziehungen finden innerhalb einer Institution statt, wie beispielsweise die Beziehung zwischen Lehrkräften und Lernenden in der Institution Schule. Interinstitutionelle Beziehungen beschreiben das Zusammenwirken zwischen verschiedenen Institutionen, beispielsweise die Zusammenarbeit zwischen Eltern und Lehrkräften. Obwohl beschrieben wird, dass solche Beziehungsformen den Bildungserfolg der Lernenden fördern, werden im Modell der überlappenden Sphären von Epstein (1995, 2010) jedoch keine konkreten Wirkmechanismen zwischen unterschiedlichen Beziehungsformen und Aspekten des Bildungserfolges beschrieben. Zum Bildungserfolg von Lernenden zählen dabei nicht nur Leistungen und Kompetenzen, sondern auch schulbezogene Merkmale wie die Lern- und Leistungsmotivation von Schülerinnen und Schülern (Grundmann, 2009; Maaz,

Neumann & Baumert, 2014). Als zweiter theoretischer Bezugsrahmen wird das Modell zum elterlichen Schulengagement von Hoover-Dempsey und Sandler (1995) genutzt. Das Modell beschreibt unter anderem, dass die Kommunikation zwischen Eltern und Lehrkräften die Motivation Lernender in der Schule positiv beeinflusst. Anzumerken ist, dass ausschließlich das kooperationspezifische Engagement der Eltern fokussiert wird, obwohl davon ausgegangen werden kann, dass sowohl Eltern als auch Lehrkräfte aktive Beteiligte in der gemeinsamen Zusammenarbeit sind (Bauch & Goldring, 1998). In der hier vorliegenden Studie wird daran anknüpfend das Engagement der Lehrkräfte in der Zusammenarbeit untersucht. Basierend auf beiden theoretischen Modellen wird der Frage nachgegangen, inwiefern die intrinsische Motivation von Schülerinnen und Schülern in der Sekundarstufe I mit dem kooperationspezifischen Engagement der Lehrkräfte (interinstitutionelle Beziehungen) und der Lehrer-Schüler-Beziehung (intrainstitutionelle Beziehungen) in Zusammenhang steht.

2.1 Zusammenarbeit zwischen Eltern und Lehrkräften und Motivation von Schülerinnen und Schülern

Eine gelungene Erziehungs- und Bildungspartnerschaft zwischen Lehrkräften und Eltern wird häufig als Ziel der Kooperation von Eltern und Lehrkräften beschrieben, wobei darunter die zielgerichtete, strukturierte und partnerschaftliche Beziehung zwischen Elternhaus und Schule zur effizienten Unterstützung von Schülerinnen und Schülern verstanden wird (Stange, 2012a). Forschungsarbeiten verweisen darauf, dass Formen der Zusammenarbeit zwischen Lehrkräften und Eltern sowohl von den Eltern als auch seitens der Schule initiiert werden können (Fan & Williams, 2010). In der vorliegenden Arbeit werden die von Lehrkräften initiierten Kontakte, also das Engagement der Lehrkräfte in der Zusammenarbeit zwischen Elternhaus und Schule, fokussiert. Diese Kontakte können theoriebasiert in formelle und informelle Kontakte unterschieden werden (Hoover-Dempsey & Walker, 2002; Sacher, 2008). Dabei wird ein vielfältiges Angebot von formellen und informellen Kontakten in der Kooperation von Eltern und Lehrkräften als ein Qualitätsmerkmal verstanden (Vodafone Stiftung Deutschland, 2013). Zu den formellen Kontakten, welche meist durch Schulgesetze, Vorschriften oder die Schule selbst vorgegeben werden, gehören beispielsweise Elternabende und Elterngespräche sowie Informationsbriefe zu den Schulleistungen der Lernenden oder zu Prüfungs- und Leistungsanforderungen in der Schule. Formelle Kontakte sind verpflichtend und daher häufig in einem gewissen Maße ritualisierte Formen der Kommunikation zwischen Eltern und Lehrkräften (vgl. Hoover-Dempsey & Walker, 2002). Informelle Kontaktformen der Zusammenarbeit sind im Gegensatz zu formellen Kontakten nicht verpflichtend für Lehrkräfte und meist nicht schulintern gefordert. Zu informellen Formen zählen unter anderem Hausbesuche, E-Mails, Telefonate oder Unterrichtshospitationen. Da das Angebot informeller Kontakte freiwillig ist, wird angenommen, dass diese eine gute, also qualitäts-

voll wahrgenommene Gesprächskultur und Atmosphäre zwischen den Kooperationsakteuren voraussetzen (Sacher, 2008). Das bedeutet, dass die bestehende Zusammenarbeit von den Beteiligten als funktionierend interpretiert wird.

Empirische Analysen zum Zusammenhang des kooperationspezifischen Engagements der Lehrkräfte, spezifisch zur Wirksamkeit des Angebots formeller und informeller Kontakte, und der intrinsischen Motivation der Lernenden sind lückenhaft. Allgemein wurde herausgestellt, dass das kooperationspezifische Engagement von Lehrkräften positiv mit der Leistungs- und Persönlichkeitsentwicklung von Schülerinnen und Schülern der Sekundarstufe zusammenhängt (Deng et al., 2017). Fan und Williams (2010) verdeutlichen mit dem Fokus auf die intrinsische Motivation von Lernenden, dass durch die Schule initiierte Kontakte, wie beispielsweise Informationsgespräche mit den Eltern (formelle Kontakte), in positivem Zusammenhang zur intrinsischen Motivation Lernender im Fach Englisch stehen. Auch Kraft und Dougherty (2013) konnten aufzeigen, dass von Lehrkräften initiierte Kontakte, hier Telefonate mit Eltern (informelle Kontakte) mit der Motivation der Schülerinnen und Schüler in Zusammenhang stehen, wenn Lernende diese Kontakte wahrnehmen. Weiterhin kann angenommen werden, dass Lehrkräfte, die den Willen zeigen Eltern einzubeziehen, nicht nur individuell Lernende fördern, sondern auch den Bildungserfolg auf Klassenebene bedingen. Kraft und Dougherty (2013) konnten darlegen, dass Lehrkräfte ihr Klassenmanagement und die Unterstützung der Klasse positiver einschätzen, wenn es zum Austausch mit Eltern in informellen Kontakten kam. Da diese Ergebnisse aus qualitativen Interviews resultieren, schließt sich durch die vorliegende Studie die quantitative Überprüfung der Ergebnisse an. Zwar kann vereinzelt die Relevanz des kooperationspezifischen Engagements von Lehrkräften herausgestellt werden, jedoch wird, wie schon vorher verwiesen, dieses meist aus der Perspektive der Eltern als Kooperationspartner erfasst (Fan & Williams, 2010). Dabei ist es nicht nur wichtig, die Perspektive der Lehrkräfte zu berücksichtigen, sondern das kooperationspezifische Engagement auch differenziert zu erfassen (Busse & Helsper, 2008). Eine bedeutsame Rolle nimmt hier die differenzierte Operationalisierung von formellen und informellen Kontakten ein. Beide Kontaktformen wirken sich positiv auf die Motivation von Lernenden aus, werden aber bisher in keiner Studie gleichzeitig als mögliche Prädiktoren berücksichtigt. So könnten differenzielle Auswirkungen auf die Motivation der Lernenden durch diese Kontakte nachgewiesen werden.

Weiterhin wird in den Qualitätsmerkmalen schulischer Elternarbeit (Vodafone Stiftung Deutschland, 2013) die vielfältige Kommunikation von Eltern und Lehrkräften als Voraussetzung einer gelungenen Erziehungs- und Bildungspartnerschaft herausgestellt. Vielfältige Kommunikation zwischen Eltern und Lehrkräften kann auf Grundlage theoretischer Modelle (Hoover-Dempsey & Walker, 2002; Sacher, 2008) im Sinne informeller und formeller Kontakte erfasst werden. Jedoch wurde bisher nicht untersucht, in welchem Zusammenhang die vielfältige Kommunikation zwischen Eltern und Lehrkräften zur Motivation von Lernenden steht. In der vorliegenden Studie wird daher untersucht, welche Zusammenhänge zwischen der

intrinsischen Motivation in der Klasse und dem Angebot vielfältiger formeller und informeller Kontakte durch die Klassenlehrkraft bestehen.

2.2 Lehrer-Schüler-Beziehung und Motivation von Schülerinnen und Schülern

Die positive Entwicklung der intrinsischen Motivation von Lernenden im Schulkontext wird nicht nur durch die Zusammenarbeit von Lehrkräften und Eltern begünstigt, sondern auch durch eine emotional unterstützende Lehrer-Schüler-Beziehung (Looser, 2011; Meyer & Turner, 2002; Wild, 2001). Die vorliegende Studie untersucht, welche Rolle die Lehrer-Schüler-Beziehung für die intrinsische Motivation der Schülerinnen und Schüler spielt, wenn gleichzeitig die Zusammenarbeit zwischen Lehrkräften und Eltern berücksichtigt wird. Die Lehrer-Schüler-Beziehung wird dabei in der vorliegenden Arbeit als eine durch Zugewandtheit und emotionale Unterstützung geprägte Beziehung zwischen Lehrkräften und Lernenden definiert (Knierim, Raufelder & Wettstein, 2016).

In der Selbstbestimmungstheorie (Deci & Ryan, 2002; Ryan & Deci, 2000) wird die soziale Eingebundenheit, die sich auch im Kontext der Lehrer-Schüler-Beziehung manifestiert, als relevant für die intrinsische Motivation Lernender hervorgehoben. Empirische Arbeiten (Wentzel, 1998; Wild, 2001) verdeutlichen, dass eine emotional unterstützende Lehrer-Schüler-Beziehung ein wichtiger Prädiktor der intrinsischen Motivation von Schülerinnen und Schülern in der Sekundarstufe I ist. Ähnliche Befunde werden auch in einer Meta-Analyse von Cornelius-White (2007) berichtet – die Beziehung zu den Lehrkräften ist prädiktiv für die Motivation (siehe auch Martin & Dowson, 2009). Wenige Studien untersuchen die Zusammenhänge zwischen der emotional unterstützenden Lehrer-Schüler-Beziehung und der Motivation sowohl auf Individual- als auch auf Klassenebene. Bieg, Backes und Mittag (2011) zeigen beispielsweise in ihrer querschnittlich angelegten Studie, dass die durch Sekundarschülerinnen und -schüler wahrgenommene emotional unterstützende Lehrer-Schüler-Beziehung mit der intrinsischen Motivation der Lernenden auf Klassenebene und Individualebene korreliert. Im hierarchischen Modell ist die emotional unterstützende Lehrer-Schüler-Beziehung ausschließlich prädiktiv für die intrinsische Motivation auf Individualebene (ebd.). Die längsschnittliche Studie von Fauth, Decristan, Rieser, Klieme und Büttner (2014) verdeutlicht damit einhergehend, dass die unterstützende Lehrer-Schüler-Beziehung mit dem Interesse der Lernenden auf Individualebene in Zusammenhang steht. Allerdings ist dieser Zusammenhang auch auf Klassenebene signifikant, wenn weitere Unterrichtsqualitätsmerkmale (Kognitive Aktivierung, Klassenführung) berücksichtigt werden. Hier zeigt sich eine heterogene Befundlage zum Zusammenhang zwischen der Lehrer-Schüler-Beziehung und der Motivation auf Klassenebene. Angenommen werden könnte, dass die emotional unterstützende Beziehung zur Lehrkraft vor allem auf Individualebene wirkt, das heißt, dass die individuelle Beziehung zwischen Lehrkräften und den Lernenden von großer

Bedeutung für die intrinsische Motivation einzelner Lernender ist, aber die in der Klasse wahrgenommene emotional unterstützende Beziehung zu Lehrkräften für das durchschnittliche Niveau der intrinsischen Motivation in der Schulklasse eher nachrangig ist. Hier könnten andere Faktoren wie beispielsweise ein anregendes und herausforderndes Lernumfeld relevant sein (Fauth et al., 2014).

2.3 Zusammenarbeit von Eltern und Lehrkräften, Lehrer-Schüler-Beziehung und intrinsische Motivation

Schlussfolgernd kann die intrinsische Motivation sowohl im Zuge der Zusammenarbeit zwischen Eltern und Lehrkräften, spezifisch dem kooperationspezifischen Engagement von Klassenlehrkräften, und zudem durch die Lehrer-Schüler-Beziehung gefördert werden. Hiermit lassen sich theoretische Annahmen von Epstein (1995, 2010) bestätigen, dass sowohl inter- als auch intrainstitutionelle Beziehungen positiv auf den Bildungserfolg von Schülerinnen und Schülern, im Detail die intrinsische Motivation, wirken. Im Modell wird auch postuliert, dass sich inter- und intrainstitutionelle Beziehungen gegenseitig bedingen, sodass angenommen werden kann, dass die Zusammenarbeit von Eltern und Lehrkräften positiv mit der Lehrer-Schüler-Beziehung zusammenhängt (Epstein & Becker, 1982). Bestätigend gaben Lehrkräfte in empirischen Untersuchungen an, dass sie durch den Austausch mit Eltern wahrnehmen, dass sie individuell und klassenbezogen bessere Unterstützungsangebote machen können (Kraft & Dougherty, 2013). Topor, Keane, Shelton und Calkins (2010) sowie Wyrick und Rudasill (2009) zeigen, dass die Lehrer-Schüler-Beziehung in Grundschulklassen mit Aspekten der Zusammenarbeit von Eltern und Lehrkräften (elterliches Engagement in der Schule) zusammenwirkt. Daran anschließend konnte empirisch nachgewiesen werden, dass hier nicht nur von einer Korrelation ausgegangen werden kann, sondern dass das kooperationspezifische Engagement von Klassenlehrkräften über die Lehrer-Schüler-Beziehung die Leistungsentwicklung von Lernenden positiv beeinflussen kann (Deng et al., 2017).

Unklar ist bislang, welche Effekte sich auf die intrinsische Motivation Lernender zeigen, wenn intrainstitutionelle Beziehungen (Lehrer-Schüler-Beziehung) und interinstitutionelle Beziehungen (Zusammenarbeit zwischen Lehrkräften und Eltern) gemeinsam berücksichtigt werden. In der Studie von Kraft und Dougherty (2013) zeigte sich, dass durch individuelle Telefonate zwischen Eltern und Lehrkräften sowohl die Beziehung zwischen Lehrkräften und Lernenden durch Lehrkräfte positiver beurteilt wurde, als auch die intrinsische Motivation der Lernenden stieg. Inwiefern die Effekte bei der differenzierten Analyse von Kooperationsformen nachgewiesen werden kann, steht an dieser Stelle noch aus.

2.4 Schulform, Migrationshintergrund, Note und Geschlecht – Ihre Bedeutung für die Zusammenarbeit zwischen Eltern und Lehrkräften und schulbezogene intrinsische Motivation

Empirische Forschungsarbeiten verweisen darauf, dass die Berücksichtigung individueller und kontextualer Merkmale Lernender und ihrer Lernumwelt unverzichtbar sind, um die Zusammenhänge zwischen der Zusammenarbeit von Eltern und Lehrkräften und dem Bildungserfolg der Schülerinnen und Schüler differenziert zu untersuchen (u. a. Domina, 2005; Graham, 2015; Ulich, 2001).

Einerseits sollte die Schulform als Kontextmerkmal berücksichtigt werden. Sacher (2004) verdeutlicht, dass Eltern an Hauptschulen, im Vergleich zu Eltern an Realschulen und Gymnasien, häufiger von Gesprächen mit Lehrkräften berichten. Von Rosenblandt und Thebis (2003) verweisen darauf, dass Eltern an Gymnasien im Vergleich zu Eltern anderer Schulformen eine höhere Zufriedenheit in der Zusammenarbeit mit Lehrkräften wahrnehmen.

Auch der Migrationshintergrund der Familie hat sowohl eine bedeutsame Rolle für die intrinsische Motivation der Lernenden (Stanat, Segeritz & Christensen, 2010) als auch für die Gestaltung der Zusammenarbeit zwischen Eltern und Lehrkräften (u. a. Domina, 2005; Gomolla, 2009; Graham, 2015). Auf Grund mangelnder Sprachkenntnisse seitens der Eltern und seitens der Lehrkräfte kann es zu Schwierigkeiten in der Kommunikation kommen (Gomolla, 2009).

Andererseits gilt der Notendurchschnitt der Schülerinnen und Schüler als Faktor, der die Zusammenarbeit zwischen Lehrkräften und Eltern prägt. Studien zeigen beispielsweise, dass die Zusammenarbeit von Eltern und Lehrkräften auf Grund von den Leistungen der Lernenden häufig anlassbezogen stattfindet (Neuenschwander et al., 2004) und die Kontaktaufnahme von Eltern und Lehrkräften geringer ist, wenn die Leistungen der Schülerinnen und Schüler gut sind (Menheere & Hooge, 2010). Ulich (2001) sowie Schmitt und Sixt (2014) verdeutlichen weiterhin, dass die Beziehung zwischen Lehrkräften und Lernenden dann positiver wahrgenommen wird, wenn Schülerinnen und Schüler einen besseren Notendurchschnitt haben.

Empirische Befunde belegen auch, dass das Geschlecht sowie der individuelle Notendurchschnitt mit der intrinsischen Motivation der Lernenden korrelieren – Mädchen und Lernende mit besserem Notendurchschnitt berichten häufig eine höhere schulbezogene Motivation (Freudenthaler, Spinath & Neubauer, 2008; Hattie, 2009).

In der vorliegenden Studie werden basierend auf dieser Befundlage sowohl die Schulform, der Migrationshintergrund der Lernenden (operationalisiert über die Muttersprache), der Notendurchschnitt und das Geschlecht der Schülerinnen und Schüler in den Analysen berücksichtigt.

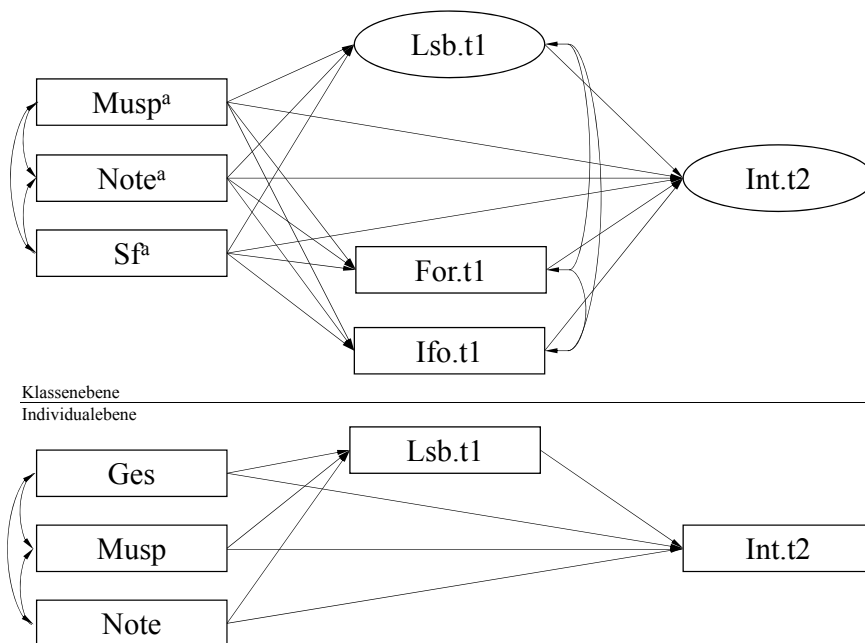
2.5 Die vorliegende Studie

Die vorliegende Arbeit ergänzt den Forschungsstand dahingehend, dass Zusammenhänge zwischen verschiedenen von Klassenlehrkräften berichteten Formen der Zusammenarbeit zwischen Lehrkräften und Eltern (Angebot vielfältiger formeller und informeller Kontakte), der von Lernenden berichteten emotional unterstützenden Lehrer-Schüler-Beziehung und der intrinsischen Motivation der Lernenden über zwei Messzeitpunkte untersucht werden.

Es werden folgende Hypothesen überprüft:

1. Das von Klassenlehrkräften berichtete formelle und informelle Kooperationsengagement in der Zusammenarbeit mit Eltern zu Beginn des Schuljahres steht in signifikant positivem Zusammenhang mit der intrinsischen Motivation der Lernenden zur Mitte des Schuljahres.
2. Es wird ein positiver Zusammenhang zwischen der durch Schülerinnen und Schüler berichteten Lehrer-Schüler-Beziehung zu Beginn des Schuljahres und der intrinsischen Motivation der Schülerinnen und Schüler in der Mitte des Schuljahres erwartet. Dabei wird angenommen, dass sich substantielle Effekte vor allem auf Individualebene zeigen.

Abbildung 1: Darstellung der in der Studie untersuchten Pfade



Anmerkungen. Musp = Muttersprache der Lernenden, Note = Durchschnittsnote, Ges = Geschlecht der Lernenden, Sf = Schulform, Lsb = Lehrer-Schüler-Beziehung, Int = Intrinsische Motivation, For = vielfältiges Angebot formeller Kontakte, Ifo = vielfältiges Angebot informeller Kontakte, t1 = Messzeitpunkt 1, t2 = Messzeitpunkt 2; a = das Konstrukt ist auf Klassenebene aggregiert.

3. Methodenteil

3.1 Stichprobe

Datengrundlage der Fragebogenstudie sind zwei Messzeitpunkte der längsschnittlichen Move-Studie (Lazarides & Rubach, 2015-2017). Die Studie untersucht die Rolle von Unterrichtsmerkmalen, Lehrpersonen und Eltern für die Entwicklung schulischer Motivation im Jugendalter. Die zwei Messzeitpunkte fanden im Herbst 2015 sowie im Frühjahr 2016 statt. Dabei wurden nur jene Lernende ($N = 881$, 52.7 % weiblich) und Schulklassen ($N = 39$; durchschnittliche Klassengröße: 23 Lernende) in die Analysen einbezogen, für die zu beiden Messzeitpunkten auch die Angaben der Klassenkräfte ($N = 39$, 42.7 % weiblich) vorlagen. Die Teilstichprobe umfasst Daten von Lernenden aus 12 Berliner Schulen, davon sind fünf Gymnasien (11 Klassen) und sieben Integrierte Sekundarschulen.¹ Die befragten Schülerinnen und Schüler waren in den neunten (54.5 %) und zehnten Jahrgangsstufen (45.5 %) und durchschnittlich 14.59 Jahre alt ($SD = 0.88$). Circa ein Viertel (24.4 %) der Schülerinnen und Schüler gaben an, dass Deutsch nicht ihre (einzige) Muttersprache ist. Der rekodierte Notendurchschnitt der Lernenden lag auf einer Skala von 1 bis 6 bei $M = 4.06$ ($SD = 0.79$).

3.2 Instrumente

3.2.1 Engagement der Klassenlehrkräfte in der Zusammenarbeit mit Eltern

Das Engagement der Klassenlehrkräfte in der Zusammenarbeit von Lehrkräften und Eltern wurde zum ersten Messzeitpunkt anhand von neun Items erfasst (Sacher, 2008). Die Lehrkräfte sollten angeben, welche der aufgezählten Kontaktformen sie selbst im letzten Schulhalbjahr für Eltern in der gemeinsamen Kooperation anboten (0 = „nein“, 1 = „ja“). Da die Items theoriebasiert in formelle (sechs Items) und informelle Kontakte (drei Items) unterschieden werden können (vgl. Sacher, 2008; Hoover-Dempsey & Walker, 2002), wurden die zugehörigen Items zu formellen Kontakten beziehungsweise informellen Kontakten aufsummiert. Diese zwei Summenscores (vielfältiges Angebot formeller Kontakte und informeller Kontakte) wurden als Prädiktoren berücksichtigt. Ein hoher Summenscore stellt so beispielsweise eine hohe Vielfalt im Angebot formeller beziehungsweise informeller Kontakte dar. Die Formen für formelle Kontakte sind Elternabend, Elternsprechstunde, Elternrundbriefe, Leistungsinformationen,

1 Zusätzlich durchgeführte Analysen verweisen darauf, dass sich ausschließlich der Notendurchschnitt der Lernenden zwischen den Schulformen unterscheidet, sodass an Gymnasien ($\bar{x} = 4.34$, $SD = 0.75$) im Vergleich zu Integrierten Sekundarschulen ($\bar{x} = 3.90$, $SD = 0.77$) ein höherer Notendurchschnitt von den Schülerinnen und Schülern berichtet wurde ($t(664) = 7.031$, $p = .000$, Tabelle A1 im Anhang).

Information zur Prüfungspraxis und Leistungsanforderungen sowie Lern- und Beurteilungsgespräche. Formen für informelle Kontakte sind Hausbesuche, Unterrichtshospitationen und Schulveranstaltungen sowie eingeholtes elterliches Feedback.

3.2.2 Lehrer-Schüler-Beziehung

Die Skala zur Lehrer-Schüler-Beziehung (Kunter et al., 2002) erfasst mit vier Items die von Schülerinnen und Schülern wahrgenommene Beziehung zu ihren Lehrkräften zum ersten Messzeitpunkt. Die Reliabilität ist gut (T_1 : $\alpha = .82$). Die Skala ist fünfstufig (1 = „stimme gar nicht zu“ bis 5 = „stimme voll und ganz zu“). Ein Beispielim lautet: „Ich komme mit den meisten meiner Lehrerinnen und Lehrer gut aus“.

3.2.3 Intrinsische Motivation

Die Skala zur Erfassung der intrinsischen Motivation erfasst mit vier Items, orientiert an der adaptierten und ergänzten Version des Academic Self-Regulation Questionnaire (SRQ-A; Ryan & Connell, 1989) von Müller, Hanfstingl und Andreitz (2007), die intrinsische Motivation der Jugendlichen zum zweiten Messzeitpunkt. Die Reliabilität ist gut $\alpha = .90$. Die Skala ist fünfstufig (1 = „stimme gar nicht zu“ bis 5 = „stimme voll und ganz zu“). Ein Beispielim lautet: „Ich arbeite und lerne für die Schule, weil es mir Spaß macht“.

3.2.4 Kontrollvariablen

Mit Hilfe der Variable Muttersprache wird ein möglicher Migrationshintergrund der Lernenden abgebildet. Die Schülerinnen und Schüler gaben an, ob Deutsch ihre Muttersprache ist (0 = „Deutsch nicht als (einzige) Muttersprache“, 1 = „Deutsch als Muttersprache“). Die Schulform wird mit Hilfe eines dichotomen Items erfasst (0 = „Integrierte Sekundarschule“, 1 = „Gymnasium“). Die Durchschnittsnote als Kontrollvariable wurde aus der von den Lernenden angegebenen Englisch-, Mathematik- und Deutschnote des letzten Schulhalbjahres gebildet. Die Notenwerte (Range von 1–6 mit 1 = „sehr gut“ bis 6 = „ungenügend“) wurden rekodiert. Bei der Variable Geschlecht gaben die Lernenden an, ob sie männlich (mit 1 kodiert) oder weiblich (mit 0 kodiert) sind.

3.3 Statistische Analyse

Auf Grund der hierarchischen Datenstruktur (Hox, 2010) wurden manifest-latente Mehrebenen-Pfadmodelle mit *Mplus* 8.1 geschätzt (Muthén & Muthén, 2016, type = twolevel). Im ersten Schritt der Analysen wurden Korrelationsanalysen der Konstrukte auf Individual- und Klassenebene durchgeführt, um die Voraussetzung des Zusammenhangs der untersuchten Konstrukte nachzuweisen.² Auf der Klassenebene wurden die Angaben der Klassenlehrkräfte einbezogen sowie die aggregierten Angaben der Lernenden pro Klasse. Bei der Mehrebenen-Pfadanalyse wurden die Kovariaten auf der Individualebene, welche zum ersten Messzeitpunkt erfasst wurden, mittels Groupmean Centering zentriert (Lüdtke, Robitzsch, Trautwein & Kunter, 2009). In allen Berechnungen wurde der Maximum-Likelihood-Robust-Schätzer verwendet (Muthén & Muthén, 2017). Zur Untersuchung der Varianzaufklärung der untersuchten Konstrukte auf den beiden Ebenen wurden Interklassenkorrelationen berechnet (Hox, 2010). Die ICC-Werte der untersuchten Konstrukte lagen oberhalb des Grenzwertes (Varnell, Murray, Janega & Blitstein, 2004): Lehrer-Schüler-Beziehung (T_1): ICC = .036, intrinsische Motivation (T_2): ICC = .028. Auf Grund des Stichprobenumfangs, der Anzahl der Klassen und der geringen ICC-Werte der untersuchten Konstrukte wurden die Modelle mit manifesten Konstrukten geschätzt (Marsh et al., 2009).

Die Teilstichprobe zum ersten Messzeitpunkt umfasste $N = 1209$ Lernende. Zum zweiten Messzeitpunkt umfasste die Teilstichprobe $N = 881$ Schülerinnen und Schüler (Dropout 27,1 %). Mittels des MCAR-Tests (MCAR = Missing Completely At Random; Little & Rubin, 2002) wurde geprüft, inwiefern ein systematischer Stichprobenausfall zwischen den Messzeitpunkten hinsichtlich der abhängigen Variable (intrinsische Motivation) vorlag. Der MCAR-Test war nicht signifikant ($\chi^2(2) = 3.62, p > .05$). Auf Grund des nichtsystematischen Stichprobenausfalls wurden fehlende Werte durch das in *Mplus* implementierte Full-Information-Maximum-Likelihood-Verfahren (FIML) theoriebasiert geschätzt. Darüber hinaus wurde getestet, inwiefern sich die untersuchte abhängige Variable (intrinsische Motivation zum zweiten Messzeitpunkt) zwischen den Lernenden, für deren Klassenlehrkräfte Daten vorlagen ($N = 881$) von den Lernenden unterschieden, von deren keine Daten der Klassenlehrkräfte vorlagen ($N = 481$). Die Ergebnisse zeigen, dass es keinen signifikanten Unterschied zwischen den Gruppen gibt: $F(1, 969) = 2.03, p > .05$.

Die untersuchten Modelle wurden mit Hilfe folgender Modell-Fit-Indizes beurteilt: χ^2 -Wert (Cheung & Rensvold, 2002), CFI- und TLI-Wert (Hu & Bentler, 1999), RMSEA und SRMR (Hu & Bentler, 1999).

2 Mögliche Zusammenhänge zwischen nominalskalierten Konstrukten und dichotomen Variablen wurde mit Hilfe der punktbiserialen Korrelation bestimmt (Field, 2009).

4. Ergebnisse

4.1 Deskriptive Ergebnisse

Tabelle 1 zeigt die Ergebnisse der manifesten Korrelationen zwischen den untersuchten Konstrukten sowie deren Mittelwerte und Standardabweichungen auf Individual- und Klassenebene.

Die Mittelwerte der Studie lassen darauf schließen, dass Lehrkräfte durchschnittlich drei verschiedene formelle Kontakte für Eltern anbieten und zwei unterschiedliche informelle Kontakte. Im Folgenden werden die Ergebnisse der Korrelationsanalyse auf Individual- und Klassenebene dargelegt. Auf Individualebene korreliert die intrinsische Motivation zum zweiten Messzeitpunkt sowohl mit der Lehrer-Schüler-Beziehung als auch mit dem Notendurchschnitt zum ersten Messzeitpunkt. Die Lehrer-Schüler-Beziehung steht in signifikant positiven Zusammenhang mit der Durchschnittsnote zum ersten Messzeitpunkt (siehe Tabelle 1).

Tabelle 1: Interkorrelationen zwischen den untersuchten Konstrukten auf Individualebene ($N = 881$; unter der Diagonale) und auf Klassenebene ($n = 39$; über der Diagonale)

	<i>M (SD)</i>	1	2	3	4	5	6	7	8
1 Musp.t1	–		.31*	–	.34*	-.43	-.42*	.22	-.54*
2 Note.t1	4.06 (0.79)	-.01		–	-.51**	-.10	.04	-.21	-.27
3 Ges.t1	–	.06	-.12*		–	–	–	–	–
4 Sf.t1	–	–	–	–		-.06	.01	.01	.05
5 Lsb.t1	3.35 (0.04)	-.04	.21**	-.03***	–		-.05	.19	-.08
6 For.t1	3.04 (2.33)	–	–	–	–	–	–	-.53**	.84**
7 Ifo.t1	1.92 (1.03)	–	–	–	–	–	–	–	-.25
8 Int.t2	2.59 (0.05)	-.09	.13*	.07***	–	.33**	–	–	–

Anmerkungen. Musp = Muttersprache der Lernenden (1 = Deutsch als Muttersprache), Note = Durchschnittsnote aus Mathematik-, Deutsch- und Englischnote (rekodiert), Ges = Geschlecht der Lernenden (1 = männlich), Sf = Schulform (1 = Gymnasium), Lsb = Lehrer-Schüler-Beziehung, For = vielfältiges Angebot formeller Kontakte, Ifo = vielfältiges Angebot informeller Kontakte, Int = intrinsische Motivation, t1 = Messzeitpunkt 1, t2 = Messzeitpunkt 2.

* $p < .05$, ** $p < .01$., *** $p < .001$

Die Korrelationsmatrix der aggregierten Konstrukte auf Klassenebene (siehe Tabelle 1) verdeutlicht einen signifikanten und positiven Zusammenhang zwischen der auf Klassenebene aggregierten intrinsischen Motivation zum zweiten Messzeitpunkt und dem vielfältigen Angebot formeller Kontakte durch die Klassenlehrkräfte. Bei der Betrachtung korrelativer Zusammenhänge mit den integrierten Kontrollvariablen zeigen die Analysen, dass in Klassen mit einem hohen

Anteil Lernender nicht deutscher Muttersprache eine signifikant höhere intrinsische Motivation der Jugendlichen besteht als in Klassen mit einem geringen Anteil Lernender nicht deutscher Muttersprache. Die Ergebnisse zeigen weiterhin, dass es einen signifikanten Zusammenhang zwischen dem Angebot formeller Kontakte von Klassenlehrkräften und dem Anteil Lernender in der Schulklasse, die Deutsch nicht als (einzige) Muttersprache haben, gibt. In Klassen, in denen eine höhere Variabilität von Muttersprachen existiert, werden vielfältigere formelle Kontakte von den Klassenlehrkräften in der Kooperation mit Eltern angeboten.

4.2 Zusammenhänge zwischen dem Engagement der Klassenlehrkräfte in der Zusammenarbeit mit Eltern, der Lehrer-Schüler-Beziehung und der intrinsischen Motivation

In der vorliegenden Arbeit wurde ein Mehrebenen-Pfadmodell modelliert.³ Im Folgenden werden im ersten Abschnitt Ergebnisse der Individualebene berichtet. Im zweiten Abschnitt erfolgt die Beschreibung der Resultate auf Klassenebene.

Bezugnehmend auf die Frage, inwieweit die Lehrer-Schüler-Beziehung prädictiv für die intrinsische Motivation ist, zeigt sich auf Individualebene, dass die von Lernenden berichtete Lehrer-Schüler-Beziehung zum ersten Messzeitpunkt in signifikant positivem Zusammenhang mit der intrinsischen Motivation zum zweiten Messzeitpunkt steht. Bei simultaner Berücksichtigung der Kontrollvariablen hängt ausschließlich der Notendurchschnitt der Lernenden (rekodiert) signifikant positiv mit von den Lernenden berichteten Lehrer-Schüler-Beziehung zusammen.

Hauptfokus der Arbeit stellt die Untersuchung der Forschungsfrage dar, inwieweit die Vielfalt von angebotenen formellen und informellen Kontakten in der Kooperation mit Eltern mit der intrinsischen Motivation der Schülerinnen und Schüler auf Klassenebene in Zusammenhang stehen. Auf Klassenebene steht das vielfältige Angebot an formellen Kontakten durch die Klassenlehrkraft in positivem Zusammenhang zur intrinsischen Motivation der Lernenden zum zweiten Messzeitpunkt. Das Angebot vielfältiger informeller Kontakte der Klassenlehrkräfte in der Kooperation mit Eltern trägt nicht signifikant zur Aufklärung der intrinsischen Motivation der Schülerinnen und Schüler zum zweiten Messzeitpunkt bei. Zudem verdeutlichen die Ergebnisse, dass die durchschnittlich in der Klasse wahrgenommene Lehrer-Schüler-Beziehung zum ersten Messzeitpunkt nicht signifikant mit der intrinsischen Motivation der Klasse zum zweiten Messzeitpunkt zusammenhängt. Weiterhin wurde der Frage nachgegangen, inwieweit das kooperationspezifische Engagement der Klassenlehrkräfte mit der auf Klassenebene aggregierten Lehrer-Schüler-Beziehung in Zusammenhang steht. Weder

3 Die Modellierung erfolgte schrittweise. Zunächst wurde nur die Lehrer-Schüler-Beziehung als Prädiktor auf beiden Ebenen berücksichtigt, im zweiten Schritt wurde das kooperationspezifische Engagement der Klassenlehrkräfte einbezogen und im dritten Schritt wurden die Kovariaten ins Modell aufgenommen. Die Effekte der geschätzten Pfade waren in allen Schritten stabil.

Tabelle 2: Mehrebenen-Pfadmodell der längsschnittlichen Zusammenhänge zwischen der Vielfalt formeller und informeller Kontakte in der Zusammenarbeit mit Eltern, der Lehrer-Schüler-Beziehung und der intrinsischen Motivation

	L1: Lsb.t1		L1: Int.t2		L2: For.t1		L2: Ifo.t1		L2: Lsb.t1		L2: Int.t2	
	β	SE	β	SE	β	SE	β	SE	β	SE	β	SE
<i>Individualebene</i>												
Lsb.t1			.31**	.04								
Note	.21**	.04	.07	.05								
Musp	-.04	.04	-.08	.05								
Ges	.00	.04	.09	.05								
<i>Klassenebene</i>												
Lsb.t1											-.21	.47
For.t1											.86*	.27
Ifo.t1											.24	.25
Sf.t1					.15	.19	-.21	.21	-.05	.24	-.07	.27
Note ^a					.27	.21	-.43	.25	.01	.37	-.23	.31
Musp ^a					-.51**	.14	.36*	.15	-.43	.28	-.25	.36
R ²	4.5		12.3		22.1		16.2		18.6		86.6	

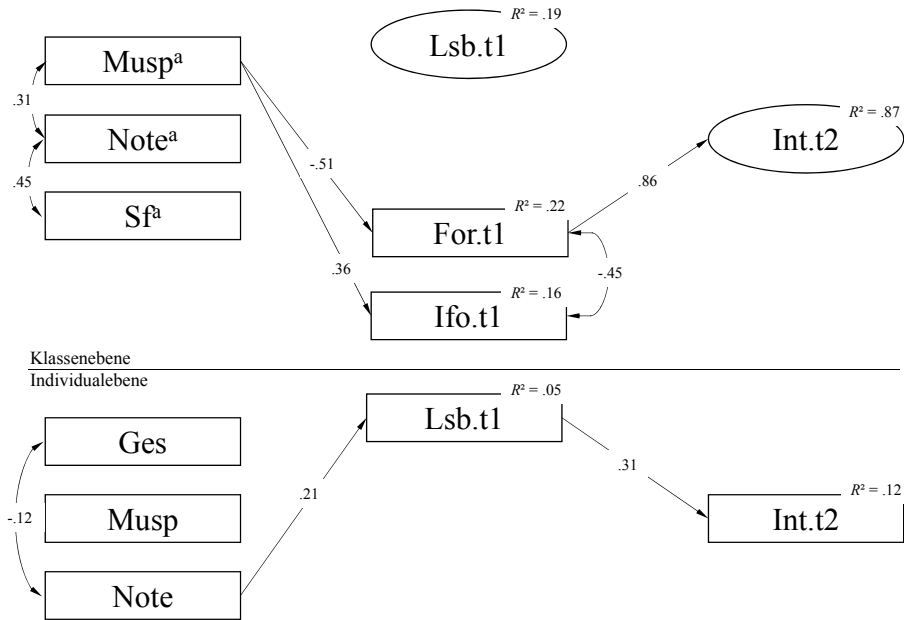
Anmerkungen. Musp = Muttersprache der Lernenden (1 = Deutsch als Muttersprache), Note = Durchschnittsnote (rekodiert), Ges = Geschlecht der Lernenden (1 = männlich), Sf = Schulform (1 = Gymnasium), Lsb = Lehrer-Schüler-Beziehung, Int = Intrinsische Motivation, For = vielfältiges Angebot formeller Kontakte, Ifo = vielfältiges Angebot informeller Kontakte, L1 = Individualebene, L2 = Klassenebene, t1 = Messzeitpunkt 1, t2 = Messzeitpunkt 2.

^aDas Konstrukt ist auf Klassenebene aggregiert.

* $p < .05$, ** $p < .01$.

das vielfältige Angebot formeller noch informeller Kontakte steht in signifikantem Zusammenhang zur aggregierten Lehrer-Schüler-Beziehung. Bei der Analyse der Zusammenhangsstrukturen beider Formen des kooperationspezifischen Engagements kann nachgewiesen werden, dass das vielfältige Angebot formeller Kontakte signifikant negativ mit dem vielfältigen Angebot informeller Kontakte in Zusammenhang steht. Dies verdeutlicht, dass Lehrkräfte entweder eher vielfältige formelle oder eher informelle Kontakte anbieten. Zudem korrelieren sowohl das formelle als auch informelle Engagement der Klassenlehrkräfte mit dem Anteil an Lernenden in der Klasse, die Deutsch nicht als Muttersprache haben. Im Detail zeigen die Ergebnisse, dass Lehrkräfte vielfältigere formelle Kontakte anbieten, wenn der Anteil an Lernenden, die nicht Deutsch als Muttersprache haben, höher ist. Dem gegenüber bieten Lehrkräfte vielfältiger informelle Kontakte an, wenn der Anteil der Lernenden, die Deutsch nicht als Muttersprache haben, niedriger in der Klasse ist.

Abbildung 2: Standardisierte, signifikante ($p < .05$) Pfade des Mehrebenen-Pfadmodells – Zusammenhänge zwischen der Vielfalt formeller und informeller Kontakte in der Zusammenarbeit mit Eltern, der Lehrer-Schüler-Beziehung und der intrinsischen Motivation



Anmerkungen. Musp = Muttersprache der Lernenden, Note = Durchschnittsnote, Ges = Geschlecht der Lernenden, Sf = Schulform, Lsb = Lehrer-Schüler-Beziehung, Int = Intrinsische Motivation, For = vielfältiges Angebot formeller Kontakte, Ifo = vielfältiges Angebot informeller Kontakte, t1 = Messzeitpunkt 1, t2 = Messzeitpunkt 2; a = das Konstrukt ist auf Klassenebene aggregiert.

Im berichteten Modell werden 4.5 % der Varianz der Lehrer-Schüler-Beziehung und 12.3 % der intrinsischen Motivation zum zweiten Messzeitpunkt auf Individualebene aufgeklärt. Auf Klassenebene werden 22.1 % der Varianz der angebotenen formellen Kontakte, 16.2 % der angebotenen informellen Kontakte, 18.6 % der Lehrer-Schüler-Beziehung und 86.6 % der durch die Klasse angegebenen intrinsischen Motivation zum zweiten Messzeitpunkt aufgeklärt.

5. Diskussion

Ziel der vorliegenden Studie war es – basierend auf dem theoretischen Modell der überlappenden Sphären von Epstein (2010) – differenzielle Effekte von inter- und intrainstitutionellen Beziehungen auf die intrinsische Motivation unter Berücksichtigung der Individual- und Klassenebene zu untersuchen. Intrainstitutionelle Beziehungen wurden über die wahrgenommene Lehrer-Schüler-Beziehung operationalisiert. Interinstitutionelle Beziehungen wurden durch das von Klassenlehrkräften berichtete Engagement in der Zusammenarbeit mit Eltern er-

fasst. Die Ergebnisse der Studie zeigen, dass das von den Klassenlehrkräften berichtete vielfältige Angebot formeller Kontakte in positivem Zusammenhang zur intrinsischen Motivation zur Mitte des Schuljahres steht. Des Weiteren zeigen die Ergebnisse, dass eine positive Lehrer-Schüler-Beziehung ausschließlich auf Individualebene mit der intrinsischen Motivation einhergeht.

5.1 Zusammenfassung und Einordnung

Insgesamt lässt sich festhalten, dass das vielfältige Angebot von formellen Kontakten durch die Klassenlehrkräfte positiv zur Erklärung der intrinsischen Motivation auf Klassenebene in der Sekundarstufe beiträgt – vielfältige formelle Kontaktangebote wirken auf Klassenebene motivationsförderlich. Damit können Forschungsergebnisse bestätigt werden, die aufzeigen, dass die Zusammenarbeit zwischen Eltern und Lehrkräften die Motivation Jugendlicher begünstigt (Fan & Williams, 2010; Hoover-Dempsey et al., 2001). Zudem wird der aktuelle Forschungsstand dahingehend erweitert, als dass Kontakte die durch die Schule initiiert werden, nicht nur mit der intrinsischen Motivation in Englisch und Mathematik (Fan & Williams, 2010; Fan, Williams & Wolters, 2012), sondern auch mit der allgemeinen schulbezogenen intrinsischen Motivation in positiven Zusammenhang stehen. Eine mögliche Erklärung für die Wirksamkeit formeller Kontakte ist der professionelle Kontext dieser Kontakte in der Schule, der solche Kontakte durch Schulordnungen vorstrukturiert. Formelle Kontakte sind meist ritualisierte Kontakte, in denen Lehrkräfte sich sicher in ihrer professionellen Rolle fühlen (Sacher, 2014). Daran anschließend stellt sich die Frage, welche medienrenden Prozesse zwischen dem Angebot vielfältiger formeller Kontakte und der intrinsischen Motivation wirken. Zukünftige Studien sollten sich diesen Fragen verstärkt widmen. Methodisch sollte dabei berücksichtigt werden, dass in der vorliegenden Studie ausschließlich das Konstrukt „vielfältige Kommunikation“ als Prädiktor in den Analysen berücksichtigt wurde. Da hier ein Summenscore aus verschiedenen Kontaktformen gebildet wurde, ist es durchaus auch relevant zu untersuchen, inwieweit spezifische Kontaktformen auf die Lernmotivation von Schülerinnen und Schülern wirken.

Die Ergebnisse der vorliegenden Studie zeigen des Weiteren, dass das vielfältige Angebot informeller Kontakte keinen signifikanten Beitrag zur Aufklärung der intrinsischen Motivation auf Klassenebene leistet. Hier stellt sich die Frage nach Unterschieden zwischen dem vielfältigen Angebot informeller und formeller Kontakte. In Anlehnung an das Angebots-Nutzungs-Modell für die Elternberatung im schulischen Kontext (Hertel, Bruder, Jude & Steinert, 2013) könnte die Rolle der Eltern in der Zusammenarbeit ausschlaggebend für die fehlende Signifikanz dieser Zusammenhänge sein, da angenommen werden könnte, dass Eltern häufiger formelle als informelle Kontakte nutzen (Rubach & Lazarides, 2017).

Die Ergebnisse der Studie zeigen zudem, dass intrainstitutionelle Beziehungen – hier die Beziehung zwischen Lehrkräften und Lernenden – einen besonderen

Beitrag für die schulbezogene intrinsische Motivation auf der Individualebene leisten. Die Ergebnisse schließen damit an die bestehende Befundlage an, die zeigt, dass die Beziehung zu den Lehrkräften mit der Motivation von Schülerinnen und Schülern in Zusammenhang steht (vgl. Looser, 2011; Meyer & Turner, 2002; Wild & Hofer, 2000). Auf Klassenebene zeigte sich hingegen ein nicht signifikanter Effekt der wahrgenommenen Lehrer-Schüler-Beziehung auf die intrinsische Motivation, so dass Hypothese 2 nur teilweise bestätigt werden kann. Auch Bieg et al. (2011) verdeutlichten, dass die emotionale Unterstützung der Lehrkräfte keinen statistisch bedeutsamen Beitrag zur Aufklärung der intrinsischen Motivation auf Klassenebene zeigte. Daran anschließend könnte angenommen werden, dass in der Entwicklungsphase der Adoleszenz andere Faktoren relevant für die Motivation auf Klassenebene sind. Diese Annahme bestätigt sich durch Ergebnisse der Analyse von Hoferichter, Raufelder, Eid und Bukowski (2014), welche zeigten, dass nur eine spezifische Gruppe Jugendlicher ausschließlich durch Lehrkräfte motiviert wird, da in der Adoleszenz vor allem auch Peers zur Motivationsförderung beitragen. Zudem sollte bedacht werden, dass hier die generelle Lehrer-Schüler-Beziehung erfasst wurde, und nicht die Beziehung zur jeweiligen Klassenlehrkraft. Zukünftig könnte in Anlehnung an die vorliegende Studie auch darüber nachgedacht werden, die Lehrer-Schüler-Beziehung als Summenscore zu operationalisieren und über konkrete Verhaltensweisen von Lehrkräften und Lernenden zu erfassen. So wäre die Möglichkeit gegeben, Aussagen darüber zu treffen, welche konkreten Situation bzw. inwieweit die Vielfältigkeit bestimmter Handlungen mit der Motivation von Lernenden in Zusammenhang steht.

Die vorliegende Arbeit berücksichtigte auch individuelle Merkmale der Lernenden und Merkmale des Lernkontextes. Die Ergebnisse zeigen, dass Lehrkräfte in Klassen mit einem hohen Anteil an Lernenden, die nicht (nur) Deutsch als Muttersprache angaben, von einem vielfältigeren Angebot formeller Kontakte berichteten. Eine mögliche Interpretation dieser Befunde ist, dass das Engagement der Lehrkräfte in der Zusammenarbeit mit Eltern an die Bedarfe der Eltern im Sinne differenzierender Elternarbeit angepasst wird (Sacher, 2012). Formelle Kontakte sind für Familien mit Migrationshintergrund auf Grund möglicher sprachlicher Barrieren eventuell effektiver und verständlicher als informelle Kontakte (vgl. Sacher, 2014). Zudem benötigen Familien mit Kindern, die nicht (nur) Deutsch als Muttersprache haben, eventuell auf Grund differenzieller kultureller Hintergründe, eher Informationen durch die Schule als Familien, in denen Kinder ausschließlich Deutsch als Muttersprache haben.

Überraschend zeigte sich in dieser Studie, dass das Engagement der Klassenlehrkräfte in der Zusammenarbeit mit Eltern nicht in signifikantem Zusammenhang zum Notendurchschnitt der Klasse stand. Dies verdeutlicht, dass Leistungsprobleme in Klassen für Lehrkräfte nicht mit dem Angebot formeller oder informeller Kontakte in Zusammenhang stehen und nicht als deren Ursache betrachtet werden können. Allerdings bedürfen diese Interpretationen weiterführender Untersuchungen. An dieser Stelle muss kritisch angemerkt werden, dass das arithmetische Mittel in der vorliegenden Studie aus den Notenangaben der

Lernenden gebildet wurde, obwohl diese Variable durch eine ordinale Skalierung abgebildet wird. Zwar ist es eine gängige Praxis in der Bildungsforschung, Notenwert als intervallskalierte Konstrukte zu behandeln (Tachtsoglou & König, 2017), sinnvoller für zukünftige Studien wäre jedoch die Messung der Leistungen durch standardisierte Leistungstest.

5.2 Limitationen und Ausblick

Methodische Limitationen der vorliegenden Studie sind das kleine Sample der Studie mit 39 Klassen sowie der relativ kurze Abstand zwischen den beiden Messzeitpunkten. Untersuchungen mit größeren Stichproben sowie über mehrere Schuljahre hinweg würden eine differenziertere Analyse der untersuchten Zusammenhänge erlauben. Außerdem wird durch die geringen ICC-Werte sichtbar, dass in der vorliegenden Studie nur wenig Varianz auf Klassenebene aufgeklärt werden konnte. Daher wäre ein weiteres Vorgehen, die Prozesse auf Individualebene zu untersuchen: Lehrkräfte sollten hier zu individuellen Kontakten mit Eltern bezüglich einzelner Schülerinnen und Schüler befragt werden. So könnte der Frage nachgegangen werden, welche Rolle das kooperationspezifische Engagement der Klassenlehrkräfte für die intrinsische Motivation auf Individualebene spielt. Zudem wären weiterführende Studien sinnvoll, die der Frage nachgehen, inwiefern das kooperationspezifische Engagement der Lehrkräfte auch in anderen Kontexten, beispielsweise der Grundschule, prädiktiv für die Varianzaufklärung der intrinsischen Motivation der Lernenden durch die Klassenzugehörigkeit ist. Da es sich in dieser Studie ausschließlich um Selbstberichtsdaten handelt, könnte in zukünftigen Studien auch die Perspektive des objektiven Beobachters berücksichtigt werden, umso das Kooperationsengagement von Lehrkräften zu erfassen.

In Bezug auf inhaltliche Limitationen muss angemerkt werden, dass in dieser Studie ausschließlich die durch die Klassenlehrkräfte initiierte Zusammenarbeit mit Eltern durch die Abfrage einzelner angebotener Kontaktformen erfasst wurde. Über die Häufigkeit und Qualität der Kontakte kann in dieser Studie keine Auskunft gegeben werden. Im Zuge dessen muss angemerkt werden, dass in der vorliegenden Studie theoretisch angenommene Formen des kooperationspezifischen Engagements von Klassenlehrkräften erfasst wurden, jedoch im Zuge der Veränderung von Kommunikationsprozessen von Eltern und Lehrkräften nun weitere Formen existieren, die hier keine Berücksichtigung fanden (Stange, 2012b). Diese sollten in zukünftigen Untersuchungen zusätzlich erfasst werden, um ein noch differenziertes Bild zum kooperationspezifischen Engagements von Klassenlehrkräften zu erhalten.

Die Ergebnisse verweisen darauf, dass eine differenzierte Untersuchung der Zusammenarbeit von Eltern und Lehrkräften, die auch die Vielfalt angebotener Kontakte in den Blick nimmt, zu einem besseren Verständnis möglicher schulbezogener Faktoren beitragen kann, die die intrinsische Motivation Lernender am Ende der Sekundarstufe begünstigen. Mit Blick auf die praktische Relevanz zeigen unse-

re Ergebnisse, dass vielfältige formelle Kontakte zwischen Eltern und Lehrkräften zu einer freudvollen und selbstinitiierten Auseinandersetzung mit schulischen Lerngegenständen beitragen. Wichtig ist, dass sich Lehrkräfte ihrer Rolle als Akteure der Motivationsförderung bewusst sind und diese Möglichkeiten in der Zusammenarbeit mit Eltern effektiv nutzen.

Literatur

- Bauch, P. A. & Goldring, E. B. (1998). Parent–teacher participation in the context of school governance. *Peabody Journal of Education*, 73(1), 15–35.
- Bieg, S., Backes, S. & Mittag, W. (2011). The role of autonomous motivation for teaching, teachers' care and autonomy support in students' self-determined motivation. *Journal for Educational Research Online*, 3, 122–140.
- Buhl, H. M. & Hilkenmeier, J. (2017). Editorial. Gespräche zwischen Eltern und Lehrkräften. *Psychologie in Erziehung und Unterricht*, 64, 241.
- Busse, S. & Helsper, W. (2008). Schule und Familie. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 469–494). Wiesbaden: VS.
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 233–255.
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective. A meta-analysis. *Review of Educational Research*, 77, 113–143.
- Deci, E. L. & Ryan, R. M. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 3–33). Rochester, NY: University of Rochester.
- Deng, L., Zhou, N., Nie, R., Jin, P., Yang, M. & Fang, X. (2017). Parent-teacher partnership and high school students' development in mainland China. The mediating role of teacher-student relationship. *Asia Pacific Journal of Education*, 38, 15–31.
- Domina, T. (2005). Leveling the home advantage. Assessing the effectiveness of parental involvement in Elementary School. *Sociology of Education*, 78, 233–249.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C. & Iver, D. M. (1993). Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48, 90–101.
- Epstein, J. L. (1995). School/family/community partnerships: Caring for the children we share. *Phi Delta Kappan*, 76, 701–712.
- Epstein, J. L. (2010). School/family/community partnerships. Caring for the children we share. *Phi Delta Kappan*, 92, 81–96.
- Epstein, J. L. & Becker, H. J. (1982). Teachers' reported practices of parent involvement: Problems and possibilities. *The Elementary School Journal*, 83, 103–113.
- Fan, W. & Williams, C. M. (2010). The effects of parental involvement on students' academic self-efficacy, engagement and intrinsic motivation. *Educational Psychology*, 30, 53–74.
- Fan, W., Williams, C. M. & Wolters, C. A. (2012). Parental involvement in predicting school motivation. Similar and differential effects across ethnic groups. *The Journal of Educational Research*, 105(1), 21–35.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Student ratings of teaching quality in primary school. Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9.
- Field, A. (2009). *Discovering statistics using SPSS. (and sex and drugs and rock 'n' roll)* (3. Aufl.). Los Angeles, CA: Sage.

- Freudenthaler, H. H., Spinath, B. & Neubauer, A. C. (2008). Predicting school achievement in boys and girls. *European Journal of Personality*, 22, 231–245.
- Gomolla, M. (2009). Elternbeteiligung in der Schule. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Elternbeteiligung* (S. 21–49). Wiesbaden: VS.
- Graham, D. (2015). Patterns of parent involvement: A longitudinal analysis of family-school partnerships in the early years of school in Australia. *Australian Journal of Early Childhood*, 14, 119–128.
- Grundmann, M. (2009). Sozialisation – Erziehung – Bildung. Eine kritische Begriffsbestimmung. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 61–83). Wiesbaden: Springer VS.
- Harris, A. & Goodall, J. (2007). *Engaging parents in raising achievement. Do parents know they matter? A research project commissioned by the Specialist Schools and Academies Trust*. Zugriff am 25.06.2016 unter <http://dera.ioe.ac.uk/6639/1/DCSF-RW004.pdf>
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hertel, S., Bruder, S., Jude, N. & Steinert, B. (2013). Elternberatung an Schulen im Sekundarbereich. Schulische Rahmenbedingungen, Beratungsangebote der Lehrkräfte und Nutzung von Beratung durch die Eltern. In E. Klieme & N. Jude (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (Zeitschrift für Pädagogik. Beiheft, Bd. 59, S. 40–62). Weinheim: Beltz.
- Hoferichter, F., Raufelder, D., Eid, M. & Bukowski, W. M. (2014). Knowledge transfer or social competence? A comparison of German and Canadian adolescent students on their socio-motivational relationships in school. *School Psychology International*, 35, 627–648.
- Hoover-Dempsey, K. V., Battiato, A. C., Walker, J. M. T., Reed, R. P., DeJong, J. M. & Jones, K. P. (2001). Parental involvement in homework. *Educational Psychologist*, 36(3), 195–209.
- Hoover-Dempsey, K. V. & Sandler, H. M. (1995). Parental involvement in children's education: Why does it make a difference? *Teachers College Record*, 97, 310–331.
- Hoover-Dempsey, K. V. & Walker, J. M. T. (2002). Family-school communication. *Elementary School Journal*, 106(2), 105–130.
- Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (Quantitative methodology series, 2. Aufl.). New York, NY: Routledge.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55.
- Knierim, B., Raufelder, D. & Wettstein, A. (2016). Die Lehrer-Schüler-Beziehung im Spannungsfeld verschiedener Theorieansätze. *Psychologie in Erziehung und Unterricht*, 84, 35.
- Kraft, M. A. & Dougherty, S. M. (2013). The effect of teacher-family communication on student engagement. Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness*, 6, 199–222.
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Eckhard, K., Neubrand, M., Prenzel, M., Schiefel, U., Schneider, W., Stanat, P., Tillmann, K. J. & Weiß, M. (2002). *PISA 2000. Dokumentation der Erhebungsinstrumente* (Materialien aus der Bildungsforschung, Bd. 72). Berlin: MPIB.
- Lauermann, F., Tsai, Y.-M. & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy-value theory of achievement-related behaviors. *Developmental Psychology*, 53, 1540–1559.
- Little, R. J. A. & Rubin, D. B. (2002). *Statistical analysis with missing data* (Wiley series in probability and statistics, 2. Aufl.). Hoboken, NJ: Wiley-Interscience.

- Looser, D. (2011). *Soziale Beziehungen und Leistungsmotivation. Die Bedeutung von Bezugspersonen für die längerfristige Aufrechterhaltung der Lern- und Leistungsmotivation* (Beiträge der Schweizer Bildungsforschung, Bd. 2). Opladen: Budrich.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Kunter, M. (2009). Assessing the impact of learning environments. How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, 34, 120–131.
- Maaz, K., Neumann, M. & Baumert, J. (2014). Editorial. In K. Maaz, M. Neumann & J. Baumert (Hrsg.), *Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter. Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive* (Zeitschrift für Erziehungswissenschaft – Sonderheft, Bd. 24, S. 1–7). Wiesbaden: Springer VS.
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B. & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, 44, 764–802.
- Martin, A. J. & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement. Yields for theory, current issues, and educational practice. *Review of Educational Research*, 79, 327–365.
- Menheere, A. & Hooge, E. H. (2010). Parental involvement in children's education. A review study about the effect of parental involvement on children's school education with a focus on the position of illiterate parents. *Journal of the European Teacher Education Network*, 144–157.
- Meyer, D. K. & Turner, J. C. (2002). Discovering emotion in classroom motivation research. *Educational Psychologist*, 37, 107–114.
- Müller, F. H., Hanfstingl, B. & Andreitz, I. (2007). *Skalen zur motivationalen Regulation beim Lernen von Schülerinnen und Schülern. Adaptierte und ergänzte Version des Academic Self-Regulation Questionnaire (SRQ-A) nach Ryan & Connell*. Klagenfurt: Alpen-Adria-Universität.
- Muthén, L. K. & Muthén, B. O. (2016). *Mplus 8.1* [Computer software]. Los Angeles, CA: Muthén & Muthén.
- Muthén, L. K. & Muthén, B. O. (2017). *Mplus User's Guide* (8. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Nagy, G., Trautwein, U., Baumert, J., Köller, O. & Garrett, J. (2006). Gender and course selection in upper secondary education. Effects of academic self-concept and intrinsic value. *Educational Research and Evaluation*, 12, 323–345.
- Neuenschwander, M. P., Balmer, T., Gasser, A., Goltz, S., Hirt, U., Ryser, H., Wartenweiler, H. (2004). *Eltern, Lehrpersonen und Schülerleistungen. Schlussbericht*. Bern: Stelle für Forschung und Entwicklung, Lehrerinnen- und Lehrerbildung Bern.
- Rheinberg, F. & Vollmeyer, R. (2012). *Motivation*. Stuttgart: Kohlhammer.
- Rohlf, C. (2011). *Bildungseinstellungen. Schule und formale Bildung aus der Perspektive von Schülerinnen und Schülern*. Wiesbaden: VS.
- Rosenblatt, B. von & Thebis, F. (2003,). *Schule aus der Sicht von Eltern. Das Eltern-Forum als neues Instrument der Schulforschung und mögliche Form der Elternmitwirkung. Eine Studie der Infratest Bildungsforschung*. München: Infratest.
- Rubach, C. & Lazarides, R. (2017). Eine Untersuchung kooperativer Beziehungen zwischen Eltern und Lehrkräften an Berliner Schulen. In J. Stiller & C. Laschke (Hrsg.), *Berlin-Brandenburger Beiträge zur Bildungsforschung 2017. Herausforderungen, Befunde und Perspektiven interdisziplinärer Bildungsforschung* (S. 177–204). Frankfurt a. M.: Lang.

- Ryan, R. M. & Connell, J. P. (1989). Perceived locus of causality and internalization. Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57, 749–761.
- Ryan, R. M. & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.
- Sacher, W. (2004). *Elternarbeit in den bayerischen Schulen. Repräsentativ-Befragung zur Elternarbeit im Sommer 2004*. SUN – Schulpädagogische Untersuchungen Nürnberg, Nr. 23. Nürnberg.
- Sacher, W. (2008). *Elternarbeit: Gestaltungsmöglichkeiten und Grundlagen für alle Schularten*. Bad Heilbrunn: Klinkhardt.
- Sacher, W. (2012). Erziehungs- und Bildungspartnerschaften in der Schule: zum Forschungsstand. In W. Stange, R. Krüger, A. Henschel & C. Schmitt (Hrsg.), *Erziehungs- und Bildungspartnerschaften. Grundlagen und Strukturen von Elternarbeit* (S. 232–243). Wiesbaden: VS.
- Sacher, W. (2014). *Elternarbeit als Erziehungs- und Bildungspartnerschaft. Grundlagen und Gestaltungsvorschläge für alle Schularten* (2., vollst. überarb. Aufl.). Bad Heilbrunn: Klinkhardt.
- Schiefele, U. (2004). Förderung von Interessen. In G. W. Lauth, M. Grünke & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen* (2. Aufl., S. 134–144). Göttingen: Hogrefe.
- Schiefele, U. & Schaffner, E. (2015). Motivation. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (Springer-Lehrbuch, 2., vollst. überarb. und aktualisierte Aufl., S. 153–175). Berlin: VS.
- Schlag, B. (2013). *Lern- und Leistungsmotivation*. Wiesbaden: Springer VS.
- Schmitt, M. & Kleine, L. (2010). The influence of family-school relations on academic success. *Journal for Educational Research Online*, 2, 145–167.
- Schmitt, M. & Sixt, M. (2014). Social relations and academic success in the German educational system. *Journal for Educational Research Online*, 6, 66–88.
- Stanat, P., Segeritz, M. & Christensen, G. (2010). Schulbezogene Motivation und Aspiration von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (S. 31–58). Münster: Waxmann.
- Stange, W. (2012a). Erziehungs- und Bildungspartnerschaften – Grundlagen, Strukturen, Begründungen. In W. Stange, R. Krüger, A. Henschel & C. Schmitt (Hrsg.), *Erziehungs- und Bildungspartnerschaften. Grundlagen und Strukturen von Elternarbeit* (S. 12–39). Wiesbaden: VS.
- Stange, W. (2012b). Methoden-Muster: Partizipation und Verhandlung – Elternbeteiligung und -mitbestimmung. In W. Stange, R. Krüger, A. Henschel & C. Schmitt (Hrsg.), *Erziehungs- und Bildungspartnerschaften. Grundlagen und Strukturen von Elternarbeit* (S. 438–450). Wiesbaden: VS.
- Tachtsoglou, S. & König, J. (2017). *Statistik für Erziehungswissenschaftlerinnen und Erziehungswissenschaftler*. Wiesbaden: VS.
- Topor, D. R., Keane, S. P., Shelton, T. L. & Calkins, S. D. (2010). Parent involvement and student academic performance: A multiple mediational analysis. *Journal of Prevention & Intervention in the Community*, 38, 183–197.
- Ulich, K. (2001). *Einführung in die Sozialpsychologie der Schule* (Studium Paedagogik). Weinheim: Beltz.
- Varnell, S. P., Murray, D. M., Janega, J. B. & Blitstein, J. L. (2004). Design and analysis of group-randomized trials. A Review of recent practices. *American Journal of Public Health*, 94, 393–399.
- Vodafone Stiftung Deutschland. (2013). *Qualitätsmerkmale schulischer Elternarbeit. Ein Kompass für die partnerschaftliche Zusammenarbeit von Schule und Elternhaus*. Zugriff am 21.06.2019 unter https://www.vodafone-stiftung.de/uploads/tx_newsjson/vfst_qm_elternarbeit_web.pdf

- Wentzel, K. R. (1998). Social relationships and motivation in middle school. The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*, 202–209.
- Wild, E. (2001). Familiäre und schulische Bedingungen der Lernmotivation von Schülern. *Zeitschrift für Pädagogik, 47*, 481–499.
- Wild, E. & Hofer, M. (2000). Elterliche Erziehung und die Veränderung motivationaler Orientierungen in der gymnasialen Oberstufe und der Berufsschule. In U. Schiefele & K.-P. Wild (Hrsg.), *Interesse und Lernmotivation. Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 31–52). Münster: Waxmann.
- Wild, E. & Hofer, M. (2002). Familien mit Schulkindern. In M. Hofer, E. Klein-Allermann, P. Noack & M. Hassebrauck (Hrsg.), *Familienbeziehungen. Eltern und Kinder in der Entwicklung: ein Lehrbuch* (2. Aufl., S. 216). Göttingen: Hogrefe.
- Wild, E. & Lorenz, F. (2010). *Elternhaus und Schule. StandardWissen Lehramt*. Stuttgart: UTB.
- Wyrick, A. J. & Rudasill, K. M. (2009). Parent involvement as a predictor of teacher-child relationship quality in third grade. *Early Education and Development, 20*, 845–864.

Anhang

Tabelle A1: Mittelwerte der getesteten Variablen an Gymnasien und Integrierten Sekundarschulen

	Gymnasium <i>M (SD)</i>	Integrierte Sekundarschule <i>M (SD)</i>	Signifikanter Unterschied
Notendurchschnitt	4.34 (0.75)	3.90 (0.77)	$t(664) = 7.031, p = .000^a$
Lehrer-Schüler-Beziehung	3.36 (0.84)	3.30 (0.89)	$t(659) = 0.721, p = .471$
vielfältiges Angebot formeller Kontakte	3.15 (2.31)	2.86 (2.40)	$t(596) = 1.467, p = .143$
vielfältiges Angebot informeller Kontakte	1.88 (1.21)	2.00 (0.88)	$t(638) = -1.317, p = .189$
Intrinsische Motivation	2.56 (0.88)	2.61 (0.91)	$t(580) = -0.704, p = .482$

^aDie Bonferroni-Korrektur ergibt ein Signifikanzniveau von $p = .01$.

Jennifer Lambrecht, Katja Bogda, Helvi Koch, Guido Nottbusch & Nadine Spörer

Comparing the effect of home and institutional learning environment on children's vocabulary in primary school

Abstract

It is assumed that additionally to the family background and child characteristics, the children's learning environments are crucial for the acquisition of early competencies. This study aimed to compare the effects of home and institutional learning environment on young children's vocabulary and to test necessary conditions for a potential compensatory effect of the institutional learning environment. Using longitudinal data from $N = 557$ preschool children (German National Educational Panel Study), we analysed to what extent family background and children's characteristics predicted home and institutional learning environments and to what extent these learning environments predicted vocabulary in pre-school and primary school. In order to test if both learning environments predict vocabulary separately, we used almost identical indicators to operationalize them. The effects were estimated within a structural equation model. The study revealed that both, home and institutional learning environment, had small and separate effects on children's vocabulary. The home learning environment was more closely related to the family background, while the institutional learning environment was more closely related to the children's characteristics. This evokes new possibilities to discuss compensatory effects.

Jennifer Lambrecht, M.A. (corresponding author) · Katja Bogda · Prof. Dr. Nadine Spörer, Structural Unit Educational Science, Chair of Psychology for Primary Education, University of Potsdam, Karl-Liebknecht-Str. 24–25, 14476 Potsdam, Germany

e-mail: jlambrec@uni-potsdam.de
katja.bogda@uni-potsdam.de
nadine.spoerer@uni-potsdam.de

Dr. Helvi Koch, Structural Unit Educational Science and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany
e-mail: helvi.koch@fu-berlin.de

Prof. Dr. Guido Nottbusch, Structural Unit Educational Science, Chair for Primary Education/German, University of Potsdam, Karl-Liebknecht-Str. 24–25, 14476 Potsdam, Germany
e-mail: gnott@uni-potsdam.de

Keywords

Home learning environment; Institutional learning environment; Vocabulary; Compensatory effect

Längsschnittliche Effekte der häuslichen und institutionellen Lernumwelt auf den Wortschatz von Grundschulkindern – ein Vergleich

Zusammenfassung

Es wird angenommen, dass für den frühen Kompetenzerwerb eines Kindes neben dem familiären Hintergrund und Merkmalen des Kindes die Lernumgebungen eine wichtige Rolle spielen. Die vorliegende Studie verfolgte das Ziel, die Effekte der häuslichen und der institutionellen Lernumgebung von Kindergartenkindern auf den frühen Wortschatzerwerb zu vergleichen und notwendige Bedingungen für einen kompensatorischen Effekt der institutionellen Lernumgebung zu überprüfen. Anhand längsschnittlicher Daten von $N = 557$ Kindergartenkindern aus dem deutschen Nationalen Bildungspanel (NEPS) wurde untersucht, in welchem Ausmaß der familiäre Hintergrund und die Merkmale eines Kindes die häusliche und institutionelle Lernumgebung präzisieren und in welchem Ausmaß diese wiederum den Wortschatz in der Vorschule und in der ersten Klasse vorhersagen. Um zu überprüfen, ob die beiden Lernumgebungen einen jeweils eigenständigen Beitrag zur Prädiktion des Wortschatzes leisten, wurden nahezu identische Indikatoren zur Operationalisierung verwendet. Mittels Strukturgleichungsmodellierung wurden die Effekte geschätzt. Die Studie zeigte, dass die häusliche und institutionelle Lernumgebung eines Kindes kleine sowie voneinander abgrenzbare Effekte auf den Wortschatz hatten. Dabei war die häusliche Lernumwelt stärker mit dem familiären Hintergrund assoziiert, während die institutionelle Lernumgebung stärker durch Merkmale des Kindes selbst prädiiziert wurde. Dies eröffnet neue Möglichkeiten der Diskussion kompensatorischer Effekte.

Schlagworte

Institutionelle Lernumgebung; Häusliche Lernumgebung; Wortschatz; Kompensatorischer Effekt

1. Theoretical background

The ability to comprehend written texts is a universal cultural technique that is directly linked to educational success in modern societies (e.g., Bynner, Londra, & Jones, 2004). The acquisition of this ability is a complex process that involves different precursor literacy skills, such as the knowledge of letters and vocabulary (Whitehurst & Lonigan, 1998). These skills develop in the early years of education

and are not only linked to children's genetic predispositions but also to their home (e.g., Lehl, Ebert, Roßbach, & Weinert, 2012; Melhuish, Phan, Sylva, Sammons, Siraj-Blatchford, & Taggart, 2008; Weinert & Ebert, 2013) as well as institutional learning environments (Becker, 2010; Melhuish, 2010; Melhuish et al., 2008). The aim of the present study is to investigate the effects of the home and institutional learning environment on vocabulary as an indicator of precursor literacy skills¹.

1.1 Home learning environment

The home learning environment (HLE) provided by the family, refers to all aspects, offers, and activities that might enhance a child's learning and, thereby, the acquisition of skills and competencies. As the child's development of academic competencies can be regarded as domain-specific (Wellman & Gelman, 1998), there is a well-established distinction between the home literacy and home numeracy environment in the English-speaking scientific community (e.g., LeFevre, Skwarchuk, Smith-Chant, Fast, Kamawar, & Bisanz, 2009). Taking this domain-specificity and prior research findings into account, Niklas (2015) postulated a theoretical framework on the relation between family background, HLE, and a child's competencies.

He grounded his theoretical framework mainly on two approaches, the Ecological Theory (Bronfenbrenner, 1979) and the Theory of Social Capital (Bourdieu, 1983). According to Bronfenbrenner (1979, 1986), an individual is surrounded by proximal and distal system components interacting with each other. The distal systems (macro- and exosystem) represent aspects of society or neighborhood whereas the proximal system (mesosystem) consists of the child's home and learning institutions (e.g., kindergarten, school). Interactions between the child and people involved in these environments (e.g., parents, educators) have a direct impact on the child's development. In his model, Niklas (2015) distinguished the distal component family background (e.g., socioeconomic status, education, migration background) and the proximal component HLE. The HLE is constituted by three important dimensions: cultural capital, cultural practice, and implicit learning (Niklas, 2015). Cultural capital includes cultural possessions, like books and artworks (Bourdieu, 1983). Cultural practice, in turn, refers to the family's cultural activities and interactions, such as reading together, library visits or the grammatical complexity of spoken language (e.g., Hoff, 2010; Niklas, Möllers, & Schneider, 2013; Retelsdorf & Möller, 2008). Implicit learning refers to assumed effects of parental cultural practice with the child that is not intended as direct teaching (Niklas, 2015).

Consistent with Niklas' theoretical framework (2015), on the one hand, studies already indicated that a more stimulating HLE promotes the acquisition of precursor literacy skills (e.g., de Jong & Leseman, 2001). On the other hand, it is

1 In line with Niklas and Schneider (2017) we use the term 'skill' (e.g., vocabulary) with regard to early forms of later literacy 'competencies' (e.g., reading).

well-documented that family background influences the child's academic competencies (e.g., Baumert & Maaz, 2006; Bos, Schwippert, & Stubbe, 2007; McElvany, Becker, & Lüdtke, 2009; OECD, 2013; Sylva et al., 2013; Tarelli, Valtin, Bos, Bremerich-Vos, & Schwippert, 2013; Weinert & Ebert, 2013). Children growing up in more privileged families without migration background have higher literacy achievements in school. However, recent research suggests that this is an indirect effect. For instance, Lehl, Ebert, Roßbach, and Weinert (2012) found that this relation could be partly explained by indicators of the HLE (see Foster, Lambert, Abbott-Shim, McCarty, & Franze, 2005; Niklas et al., 2013). Thus, children who grew up in more privileged families were also provided with a more stimulating HLE. With regard to the different dimensions of the HLE, Niklas et al. (2013) documented close relations between family background and cultural capital as well as cultural practice, whereas the relation to implicit learning needs further investigation.

With reference to prior research, the various conceptualizations and operationalizations of implicit learning within the HLE have to be considered (Burgess, Hecht, & Lonigan, 2002). Ebert et al. (2013), for example, conceptualized implicit learning activities in a longitudinal study as a literacy-specific characteristic of a family and asked parents to rate ten literacy-related activities. They found a positive relation between the amount of these activities and children's initial vocabulary but no influence on vocabulary growth. Dumais (2006) investigated the effects of extracurricular activities during kindergarten and primary school on the gain of reading skills between first and third grade. Here, a positive relation, indicating the relevance of implicit learning by engagement in extracurricular activities, was found. Melhuish et al. (2008) included fourteen activities, whereby half of them were domain-unspecific (e.g., visiting friends or painting) and half of them were literacy- or numeracy-related activities (e.g., going to library). Their results revealed that domain-unspecific activities were not significantly related to under- or overachievement (e.g., vocabulary) of children, whereas the literacy-specific activities played a significant role. However, the operationalization did not allow separating cultural activities and implicit learning activities.

1.2 Institutional learning environment

In today's society, children typically attend preschool before entering school. In Germany, for example, 95 % of three to five year olds attend a preschool or day-care (Autorengruppe Bildungsberichterstattung, 2016, p. 59). Therefore, children are not only provided with an HLE but also with an institutional learning environment (ILE) prior to school enrolment. The ILE concept, similar to the HLE, unites offers and activities in an educational institution that might enhance a child's early skills. In empirical studies (e.g., Kuger & Kluczniok, 2008), three quality dimensions were identified: structural quality (e.g., socioeconomic group composition, availability of material), orientation quality (e.g., educators' attitudes), and process

quality (e.g., child-educator interaction). They can be conceptualized globally as well as literacy-specific (Kluczniok & Roßbach, 2014; Roßbach, 2005).

Studies revealed that attending preschool has a robust positive influence on a child's early literacy skills (e.g., Burger, 2010; Melhuish et al., 2008). It is assumed that especially children growing up in a less stimulating HLE benefit more from attending preschool in contrast to their peers growing up in a more stimulating HLE (Autorengruppe Bildungsberichterstattung, 2016; Becker, 2010). However, the state of research is not consistent: Ebert et al. (2013) found a positive influence of ILE (e.g., number of different kinds of books, promotion of language) on precursor literacy skills among children with a migration background but not with reference to native speaking German children. In contrast, Becker's study (2010) revealed that all children benefited from a higher number of books and toys within kindergarten regarding vocabulary. Referring to the process quality of the ILE, Niklas and Tayler (2018) as well as Schmerse et al. (2018) found a positive relation to verbal abilities (e.g., vocabulary), while Weinert and Ebert (2013), who analysed literacy-specific facets of process and structural quality (e.g., availability of books), did not find a positive effect on early literacy skills.

The question to what extent the ILE is able to compensate for family background and/or a less stimulating HLE is internationally relevant and was investigated in different studies with different approaches. Based on a systematic review, Burger (2010) concluded that a compensatory effect of the ILE on the cognitive development of children is supported by a number of studies. At the same time, however, another number of studies indicated no compensatory, but an additional effect for all students. A review of international empirical studies on compensatory effects for the cognitive development of children revealed that, within the regular ILE, compensatory effects are unlikely and costly since specialized programmes are needed for compensation (Roßbach, Kluczniok, & Kuger, 2008). The authors recommended to further investigate the learning environment of a child and pointed out, that the majority of studies were conducted in the United Kingdom and middle Europe, while the empirical evidence was rare for the German educational system. Recently, in a study with a German sample, Schmerse et al. (2018) found contra-indicators for a compensatory effect. Here, children growing up in a high-quality HLE benefited more from the ILE than their disadvantaged peers at age four.

To sum up, empirical evidence in terms of a compensatory effect of the ILE for children from disadvantaged families is hard to interpret, as the studies were conducted in different settings and point to different results. Moreover, only a few studies investigated literacy-related competencies. Against the backdrop of those inconsistent results, it may be appropriate to take a step back and investigate the conditions that have to be fulfilled to presume a compensatory effect.

To presume a compensatory effect of the ILE, three conditions have to be accomplished: The effect of family background on a child's skills or competencies is (at least partly) mediated by the child's learning environments (compensatory effect condition 1). The ILE of a child is not determined by family background (com-

pensatory effect condition 2). HLE and ILE have an independent effect on a child's skills or competencies (compensatory effect condition 3).

According to these conditions, recent research findings can now be systematized. In synopsis, condition 1 is likely to be fulfilled, as several studies indicated that a child's learning environment is relevant, which is a pre-condition to fulfil condition 1. However, the results differ across studies (Burger, 2010). Condition 2 is likely to be fulfilled, yet several studies suggest that children from disadvantaged families access institutions of lower quality compared to their peers from more privileged families (Becker & Biedinger, 2016; Cloney, Cleveland, Hattie, & Tayler, 2016; Lehl, Kuger, & Anders, 2014). According to Hogrebe (2016) this might be due to a lower mobility of the parents belonging to a minority group based on ethnic and social segregation. These results suggest that ILE is not determined, yet to some extent predicted by family background.

With respect to condition 3, the independent effect of HLE and ILE on a child's skills or competencies, we can state that most studies taking HLE and ILE into account found effects of both learning environments on a child's vocabulary (e.g., Schmerse et al., 2018). So far, however, the effects of both learning environments were hardly comparable as they were operationalized based on different theoretical frameworks resulting in different indicators (e.g., number of books at home vs. number of children with a migration background at the institution). This raises the question how similar aspects of HLE and ILE contribute to a child's early literacy skills (e.g., number of books at home vs. number of books in the institution). Compensatory effect condition 3 has not yet been examined under this perspective.

2. The present study

To investigate whether the three conditions for a compensatory effect are fulfilled and thus, to verify if HLE and ILE predict vocabulary independently and if they are both predicted by family background in the same way, it is necessary to conceptualize both learning environments in a way that makes them comparable. Prior theoretical models focussing either the home or institutional learning environment display different conceptions and operationalizations of both learning environments. Due to these different research approaches of the HLE and ILE, there is no integrative theoretical model taking both learning environments and their longitudinal effects on a child's early skills into account. Furthermore, the indicators of both learning environments differed. Therefore, it is an open question, if the ILE can compensate for the HLE. It remains unclear (a) whether family background predicts the quality of HLE and ILE in the same way, (b) whether HLE and ILE predict early literacy skills equally, and (c) whether these effects last beyond the transition from preschool to primary school. With reference to the attempt that every student succeeds (ESSA, 2015), it is necessary to parallelize the HLE and ILE regarding their conception as well as operationalization and to investigate wheth-

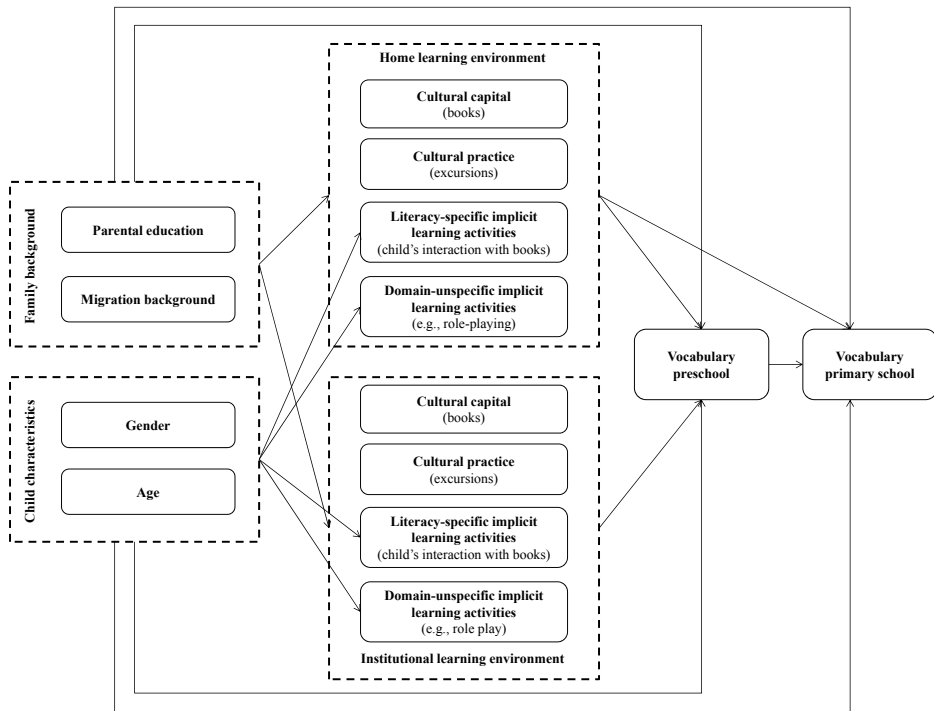
er both learning environments influence a child's competence development independently and additively or not.

Moreover, in consequence of focusing on parent-child interactions in the former research framework of HLE by Niklas (2015) and child-educator interactions with respect to process quality in the ILE, children's implicit learning activities by themselves (e.g., reading alone) and general activities (e.g., role-playing) that might be meaningful for developing precursor literacy skills were not taken into account yet. Defining these child-driven activities in a model also contributes to a better distinction between the different dimensions of the learning environments. For instance, cultural practice and implicit learning (Niklas, 2015) have a substantial overlap. Therefore, it is difficult to disentangle both constructs regarding their operationalization. Parental reading to a child or visiting a library could be ascribed to cultural practice as well as to implicit learning since both dimensions focus on the adult-child interactions. However, they do not cover child-driven activities with, for example, books. With reference to these child-driven implicit learning activities, it needs to be assumed that not only family background but additionally child characteristics are influencing variables. Regarding the child-driven choice of learning activities it is presumed that inter-individual systematic differences among children depending on the child's characteristics, such as gender and age, predict HLE, ILE, and vocabulary.

To address these desiderates, we propose a joint model of HLE and ILE based on the theoretical framework by Niklas (2015) that includes comparable dimensions of both learning environments to predict precursor literacy skills (see Figure 1).² Here, we focus on vocabulary as a precursor literacy skill since early vocabulary significantly predicts later reading development (Muter, Hulme, Snowling, & Stevenson, 2004; Speece, Ritchey, Cooper, Roth, & Schatschneider, 2004). We assume that the HLE directly predicts a child's vocabulary in preschool as well as in primary school. The ILE is directly linked to vocabulary in preschool, but is indirectly predictive for a child's vocabulary in first grade via the gained vocabulary in preschool. In our model, both the HLE and ILE are conceptualized to be four-dimensional. The dimensions cultural capital (availability of books at home resp. in preschool) and cultural practice (child visits a library with parents resp. educators) are domain-specific, and therefore, directly linked to literacy. The third dimension refers to a child's literacy-specific implicit learning activities and includes activities which are directly related to text material (child's interaction with books at home resp. in preschool). Finally, the fourth dimension covers a child's domain-unspecific learning activities which are not exclusively related to literacy but may enhance literacy-related skills (e.g., role-playing at home resp. in preschool).

2 Another possibility would have been to adapt and modify the model proposed by Kluczniok, Lehl, Kuger, and Roßbach (2013), who developed a model integrating HLE and ILE by adapting the model of kindergarten quality by Tietze et al. (2013). However, to analyse potential compensatory effects we presume that the theoretical basis of Niklas' model (2015) allows a more adequate interpretation in terms of social disparities.

Figure 1: Model of relations between family background, child characteristics, HLE, ILE, and vocabulary acquisition



In line with prior research, we suppose that the family background does not only predict precursor literacy skills but also HLE and ILE. As a consequence, the model presumes that a part of the effects of family background on vocabulary is indirect via HLE and ILE.

In addition to existing models, we further assume that a child's characteristics influence implicit learning activities inherent in HLE and ILE components. As prior HLE models typically focused on parent-child interactions, the influence of family (and therefore parents') characteristics on these interactions were taken into account. However, when we integrate learning activities that a child can choose self-directed at home and in preschool, one can suppose that children select different activities biased by gender and age (Siraj-Blatchford, 2010). Thus, it is finally assumed that effects of child characteristics on vocabulary are indirect via both learning environments.

Based on our model, the main goal of the present study was to analyse the conditions of compensatory effects of the ILE. To gain a better understanding of not only cross-sectional relations but indications of longitudinal effects, we aimed to predict vocabulary in preschool and primary school. Therefore, we addressed the following questions within one model in order to test the necessary conditions for a compensatory effect:

1. To what extent do family background and child characteristics predict vocabulary in preschool and primary school?

To test the compensatory effect condition 1 (that the effect of family background on vocabulary is mediated by the learning environments of a child), we analysed if vocabulary is not only predicted by family background and child characteristics but also indirectly via HLE and ILE.

2. To what extent do family background and child characteristics predict HLE and ILE?

To test the compensatory effect condition 2 (that ILE is not determined by family background), we focussed on the relation between family background and ILE within our integrative model.

3. To what extent do HLE and ILE predict vocabulary in preschool and primary school?

To test compensatory effect condition 3 (that HLE and ILE have an independent effect on vocabulary), we compared their specific contribution to predict vocabulary simultaneously within our integrative model.

3. Method

3.1 Participants and design

The used sample was a subsample of the second starting cohort of the longitudinal German National Educational Panel Study (NEPS). The aim of NEPS is to analyse educational transitions from early childhood to late adulthood with a multi-cohort sequence design (see Blossfeld, Roßbach, & Maurice, 2011). The NEPS consists of six panel cohorts with random samples for each cohort covering crucial educational transitions over the life course. Cohort 2 focuses on the transition from preschool to primary school. It was selected using indirect sampling based on the link between probable kindergartens and primary schools to enhance longitudinal and contextual analyses (Aßmann et al., 2011). At the first measurement point (t1) in kindergarten in 2011, $N = 2,996$ children were included in the sample. Due to the drop out caused by the transition from kindergarten to primary school, $N = 557$ children could be tested again in first grade in 2013 (t2). Since we were interested in longitudinal relations, we included data from those children, their parents, and their educators, who took part in NEPS at both measurement points. The children came from 159 preschools in 274 preschool groups. They were $M = 5.00$ ($SD = 0.32$) years old at t1 and $M = 6.98$ ($SD = 0.31$) years old at t2. Nearly 52 % were girls and 12 % of the children had a migration background, i.e. at least one parent and/or the child's grandparent was not born in Germany. About 60 % of the children had at least one parent with a higher education entrance qualification. This indicates a selective sample, restricting the representativeness.

3.2 Instruments

3.2.1 Children's vocabulary

Children's vocabulary was individually assessed at t1 and t2 using a German adaptation of the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981; Roßbach, Tietze, & Weinert, 2005). The test consisted of 77 items at t1 ($M = 51.61$; $SD = 11.31$) and 66 items at t2 ($M = 41.34$; $SD = 9.25$). The children heard a word and were then asked to find the corresponding picture out of four alternatives. The internal consistency of the test was sufficient at both measurement points (Cronbach's $\alpha_{t_1} = .76$; $\alpha_{t_2} = .72$). To make both tests comparable, the sum scores were z-standardized.

3.2.2 Family background

The family background was assessed at t1 via standardized interviews with the children's parents (LifBi, 2016). We used the Comparative Analyses of Social Mobility in Industrial Nations (CASMIN; Lüttinger & König, 1988; see also Lechert, Schroedter, & Lüttinger, 2006) as an indicator for parental education. We took the highest CASMIN from either, father or mother. Moreover, we determined the child's migration background. A child was ascribed a migration background if at least one parent and/or the child's grandparent was not born in Germany.

3.2.3 Learning environments

HLE was assessed at t1 via standardized interviews with the children's parents, while ILE was assessed at t1 via standardized questionnaires with the children's educators (LifBi, 2016). We chose the variables to operationalize HLE and ILE with respect to comparability between the learning environments.

3.2.4 Cultural capital

With respect to the HLE, parents were asked how many books they have at home (1 = "0 to 10 books" to 6 = "more than 500 books"). Within the ILE, we used the availability of books within preschool (Are picture books available in the preschool for 0 = "not available" to 3 = "almost all of children" so they can play with them at the same time) as an indicator for cultural capital.

3.2.5 Cultural practice

Cultural practice was captured by the frequency of parents visiting a library with their child (1 = “never” to 8 = “several times a day”) for the HLE and the frequency of visiting a library with the preschool group during the last twelve months (1 = “never” to 6 = “daily”) for the ILE.

3.2.6 Implicit learning activities (ILA)

Both the parents (HLE) and the educators (ILE) were asked how often the child is busy with different activities (1 = “never” to 8 = “several times a day”). One out of ten ILA was literacy-specific (interaction with books), while nine items were domain-unspecific for literacy (e.g., doing jigsaw puzzles, doing crafts, role-playing, see Table 1) in the HLE and in the ILE.

3.3 Statistical analyses

Before answering our research questions, we ran a set of preliminary analyses. All preliminary analyses were conducted with SPSS 23. In a first step, we handled missing data with the SPSS default algorithm for regression imputation. The amounts of data missing were rather small, ranging from none with respect to child characteristics, to 6 % (HLE items) and 7 % (ILE items) up to 10 % (family background variables). With regard to the clustered structure of the data (children in groups, groups in preschools), please note that no whole groups were investigated. In the present data set, only 6 % of the preschool groups had five or more children nested within them. Therefore, data were analysed at the individual level.

In a second step, we analysed the distribution of HLE and ILE variables and their interrelations. Table 1 shows the descriptive statistics for all variables used (a table showing the correlations between all variables analyzed can be found in the appendix). All HLE and ILE variables had a median above the theoretical median and were non-normally distributed.

Table 1: Descriptive statistics of the child characteristics, family background, early literacy skills, HLE, and ILE variables

	<i>M</i>	<i>SD</i>	Range	%	<i>Md</i>	<i>Sk</i>	<i>Ku</i>
<i>Child characteristics</i>							
Age	5.01	0.32	4.25–5.92		5	0.79	-0.63
Gender							
Girls				52			
Boys				48			
<i>Family background</i>							
Parental education			1–8		6	-0.26	-0.93
Migration background							
No				88			
Yes				12			
<i>Early literacy skills</i>							
Vocabulary t1	51.61	11.31	0–77		54	-1.36	2.12
Vocabulary t2	41.34	9.25	0–66		43	-0.58	-0.08
<i>Home learning environment</i>							
Cultural capital			1–6		4	-0.55	-0.61
Cultural practice			1–8		2	0.60	-0.60
Literacy-specific ILA			1–8		8	-2.31	5.64
Domain-unspecific ILA							
Comparing/collecting things			1–8		7	-1.41	3.16
Number games or counting			1–8		6	-1.15	1.71
Doing jigsaws			1–8		6	-0.75	0.57
Building games			1–8		6	-0.80	0.11
Doing crafts			1–8		7	-1.11	1.16
Role-playing			1–8		7	-1.35	1.74
Sports			1–8		7	-0.99	2.18
Music, singing or dancing			1–8		6	-0.94	1.16
Nature and gardening			1–8		6	-0.96	1.16
<i>Institutional learning environment</i>							
Cultural capital			0–3		1	0.90	1.37
Cultural practice			1–6		2	0.69	-0.80
Literacy-specific ILA			1–8		7	-1.21	1.10
Domain-unspecific ILA							
Comparing/collecting things			1–8		6	-0.69	0.13
Number games or counting			1–8		6	-0.79	0.48
Doing jigsaws			1–8		6	-0.81	0.51
Building games			1–8		6	-0.91	-0.05
Doing crafts			1–8		7	-1.22	1.34
Role-playing			1–8		7	-1.56	2.11
Sports			1–8		6	-0.79	0.98
Music, singing or dancing			1–8		7	-1.29	1.64
Nature and gardening			1–8		6	-0.52	-0.83

Notes. *Md* = Median; *Sk* = Skewness; *Ku* = Kurtosis; ILA = Implicit learning activities.

Table 2 gives an overview of the correlations of the domain-unspecific ILA. In contrast to our expectations that they were positively correlated with one another, there were positive and negative interrelations between the domain-unspecific ILA variables. This casts into doubt the assumption that ILA is a latent reflective construct predicting the frequency of a child's activities.

Table 2: Correlation coefficients of the domain-unspecific ILA variables (HLE and ILE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Role-playing		.320*	.165*	.143*	.074*	-.151*	.196*	.265*	.121*
(2) Doing crafts	.203*		.221*	.299*	.220*	-.188*	.079*	.079*	.090*
(3) Comparing	.097*	.180*		.385*	.260*	.141*	.195*	.228*	.314*
(4) Counting	.040	.022	.145*		.441*	.086*	.068*	.166*	.226*
(5) Jigsaws	.069*	.230*	.191*	.066*		.073*	.002	.166*	.101*
(6) Building	-.042	-.076*	.161*	.246*	.014		.094*	-.003	.096*
(7) Sports	.038	-.016	.088*	.222*	-.072*	.155*		.234*	.275*
(8) Music	.205*	.217	.124*	.082*	.082*	-.009	.184*		.144*
(9) Nature	.087*	.144*	.164*	.085*	.085*	.123*	.168*	.157*	

Note. Correlation coefficient is Kendall's Tau. Coefficients for implicit learning activities at home are shown below the diagonal, while activities in the institution are above it.

* $p < .05$.

3.3.1 Modelling approach

The model was estimated using partial least squares structural equation modelling (PLS-SEM). All analyses were conducted with R (R Core Team, 2015), applying the package *plspm* by Sanchez, Trinchera, and Russolillo (2015).

PLS-SEM is a variance based-approach to structural equation modelling that estimates a measurement model for latent variables and its indicators, and a structural model, representing relations between the latent constructs (path model). It is an alternative approach to structural equation modelling that, according to Hair, Hult, Ringle, and Sarstedt (2014), is theoretically preferred over covariance-based SEM (CB-SEM) if the aim is more exploratory than confirmatory. As we focus on the question if HLE and ILE predict vocabulary independently from one another, we meet PLS-SEM's capacity: the aim of PLS-SEM is not to test how well a theoretical model fits the data but to maximize the explained variance in the dependent variable when the model is estimated (Hair et al., 2014).

In addition to these theoretical considerations, there were also statistical issues that spoke in favour of PLS-SEM in the current study: The NEPS data had different levels of measurement, and the HLE and ILE indices as well as the vocabulary

scores were non-normally distributed. In contrast to CB-SEM, PLS-SEM can handle these types of data as it makes no assumption on data distribution (Hair et al., 2014; Weiber & Mühlhaus, 2014). Moreover, it is suitable for constructs measured using a single-item approach and latently modelled variables within one model.

The pattern of correlations between ILA variables, as shown in Table 2, suggested that the construct might not be reflective as it is common in social sciences. It is reasonable to assume that implicit learning activities were not predicted by an underlying latent construct, but that the construct is formed by all activities taken together. If it is assumed that the construct is not homogenous but consists of different facets that together result in a scale, those constructs are formative (Jarvis, MacKenzie, & Podsakoff, 2003). As PLS-SEM can estimate formatively and reflectively measured constructs, we tested if a formative modelling of the domain-unspecific ILA had an advantage over the more common reflective approach.

To assess statistical significance, standard errors and confidence intervals were calculated via bootstrap validation on the basis of 500 samples.

3.3.2 The measurement model: operationalization and assessment

To specify the measurement model, we used a single-item approach with respect to the cultural capital, cultural practice, and the child's literacy-specific ILA, both for the HLE and ILE. The domain-unspecific ILA at home and in kindergarten was intended to be measured latently by nine indicators each. However, the weights calculated as covariance between the indicators of the latent variable and the construct loadings and communality did not accomplish the measurement model quality criteria (Henseler, Ringle, & Sinkovics, 2009). There were two learning activities – doing crafts and role-playing – that constituted the respective latent factors within the reflective model. They were relevant to variance in vocabulary with respect to both learning environments and met the statistical criteria to remain in the model ($ILA_{\text{home}}: \alpha = .38, DG. \rho = .73$; $ILA_{\text{institution}}: \alpha = .54, DG. \rho = .81$). In a next step, we used a formative approach to operationalize the ILA. The outer weights were calculated as partial least squares regression weights from the latent variable on its indicators (Henseler et al., 2009). As within the reflective measurement model, the remaining constructs were doing crafts and role-playing at home as well as in preschool. Therefore, the formative approach did not have an advantage over the reflective one and was not further investigated. Nevertheless, it stressed role-playing and doing crafts as important domain-unspecific ILA. These empirically derived domain-unspecific ILA can be described as creative activities.

Within the reflective model, domain-unspecific ILA was therefore modelled latently by the two items role-playing and doing crafts for HLE and ILE. Table 3 shows the weights, loadings, and communalities for the indicators of those constructs. The outer loadings for domain-unspecific ILA indicators were above .70 and, therefore, sufficient. To assess the quality of the reflective measurement mod-

el, we further checked convergent validity and discriminant validity (Hair et al., 2014; Sanchez, 2013). An indicator of convergent validity is the average variance extracted (AVE). The AVE was above the recommended cut off value of 0.50 for ILA within HLE and ILE. Moreover, there were no cross loadings indicating discriminant validity.

Table 3: Assessment of outer model (reflective)

Construct and indicators	Weight	Loading	Communality
ILA _{home}			
Doing crafts	0.645	0.799	0.638
Role-playing	0.621	0.781	0.610
ILA _{institution}			
Doing crafts	0.642	0.857	0.734
Role-playing	0.559	0.805	0.648

Note. ILA = Implicit learning activities.

3.3.3 The structural model: assessment

There are no global fit criteria available to assess the structural model, because PLS-SEM fits the model to the data, and not vice versa as in CB-SEM. To assess the structural model, recommendations are to evaluate the significance of the paths between the latent constructs, the explained variance of the endogenous constructs (R^2) and their effect sizes (f^2) (Hair et al., 2014). The effect sizes f^2 were interpreted as suggested by Cohen (1992): a value of .02 indicates a small, a value of .15 a medium, and a value of .35 a large effect.

To address our research questions, (1) to what extent family background and child characteristics predict vocabulary in preschool and primary school, (2) to what extent family background and child characteristics predict HLE and ILE, and (3) to what extent HLE and ILE predict vocabulary in preschool and primary school, we analysed the structural model's paths to get detailed information on the relation between the variables. Further, we tested the three main conditions to assume a compensatory effect within our integrative model.

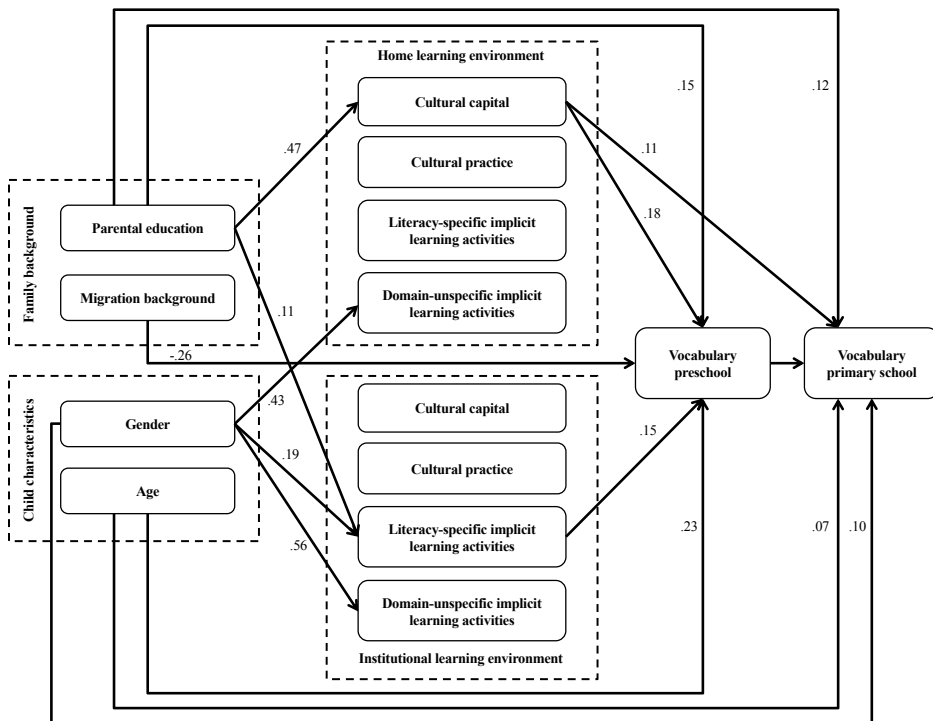
To test if the effect of family background on vocabulary is (at least partly) mediated by the child's learning environments (compensatory effect condition 1) we took the total effects of family background and child characteristics on vocabulary into account. Total effects are the sum of direct (family and child characteristics → vocabulary) and indirect effects (family and child characteristics → HLE/ ILE → vocabulary). If the total effects are higher than the direct effects, a mediation effect is indicated. With respect to compensatory effect condition 2, we analysed to what extent ILE is predicted by family background. Therefore, we summarized

family background and child characteristics respectively to assess their joint effects (f^2) on dimensions of the HLE and ILE in addition to analysing the path model. To assume a compensatory effect, the effect of family background on ILE should be small. To assess compensatory effect condition 3, that ILE and HLE have an independent effect on vocabulary, we summarized the dimensions of HLE and ILE and compared their effects (f^2) on vocabulary.

4. Results

Figure 2 shows the structural model. The path coefficients can be interpreted as standardized weights of an ordinary least squares regression and are comparable to each other. For reasons of simplicity, Figure 2 shows significant paths only. Based on this model our research questions will be answered.

Figure 2: Structural model to predict vocabulary acquisition (standardized path coefficients)



4.1 Research question 1: To what extent do family background and child characteristics predict vocabulary in preschool and primary school?

With regard to the family background, vocabulary in preschool was directly predicted positively by parental education ($B = .15$), while a migration background had a negative effect ($B = -.26$). Vocabulary in primary school was predicted positively by parental education ($B = .12$), whereas no significant direct effect of migration background was found. In terms of child characteristics, age had a positive effect on vocabulary in preschool ($B = .23$), while gender was not predictive. When it comes to vocabulary in primary school, both gender ($B = .10$) and age ($B = .07$) had a positive effect on vocabulary, indicating an advantage for girls and older children.

With respect to compensatory effect condition 1 (that the effect of family background on vocabulary is mediated by the child's learning environments), we can state that both, child characteristics and family background, predicted vocabulary in preschool indirectly, as their total effects on vocabulary in preschool and primary school were larger than their direct effects. Parental education had a significant total effect of .25 on vocabulary in preschool and a significant total effect of .30 on vocabulary in first grade. The migration background had a significant total effect of -.19 on vocabulary in first grade, although there was no significant direct effect. While gender had no significant total effect on vocabulary, the age of a child had a total significant effect on vocabulary in primary school of .20.

In sum, it can be stated that there were substantial direct effects from family background and child characteristics on vocabulary, while the larger total effects indicated indirect effects via HLE and ILE. One necessary condition for a compensatory effect was accomplished. This leads to our second research question.

4.2 Research question 2: To what extent do family background and child characteristics predict HLE and ILE?

With respect to the family background, Figure 2 reveals, that in terms of HLE parental education predicted cultural capital ($B = .47$), while migration background was not predictive towards any dimension of the HLE. In terms of the ILE, parental education predicted literacy-specific ILA ($B = .11$), but not cultural capital. A migration background was not predictive towards any dimension of the ILE.

In terms of child characteristics, we can state that age was not predictive for any dimension of the learning environments, while gender predicted the domain-unspecific ILA within both, the HLE ($B = .43$) and the ILE ($B = .56$). Moreover, it predicted literacy-specific ILA within the ILE ($B = .19$) but not within the HLE.

Focussing on compensatory effect condition 2, Table 4 summarizes the joint predictive power of family background and child characteristics towards the dimensions of HLE and ILE. Family background predicted three dimensions of HLE

with small (literacy-specific ILA, cultural practice) to medium (cultural capital) effects. Within the ILE, family background predicted literacy-specific ILA practically significant with a small effect, while it was neither relevant for domain-unspecific ILA, cultural practice nor for cultural capital. In turn, child characteristics predicted the domain-unspecific ILA as well as the literacy-specific ILA within the ILE with small to medium effects, while within the HLE the domain-unspecific ILA was predicted with a small effect.

Table 4: Effect sizes (f^2) of family background and child characteristics on HLE and ILE

	Home learning environment				Institutional learning environment			
	Cultural capital	Cultural practice	ILA		Cultural capital	Cultural practice	ILA	
			Specific	Unspecific			Specific	Unspecific
Family background	0.32	0.03	0.12	0.01	0.00	0.01	0.02	0.00
Child characteristics	0.00	0.00	0.00	0.23	0.00	0.00	0.04	0.45

Notes. ILA_{specific} = Literacy-specific implicit learning activities; $ILA_{\text{unspecific}}$ = Domain-unspecific implicit learning activities.

According to these results, compensatory effect condition 2, that ILE is not determined by family background, was met. Moreover, compared to the predictability of HLE by the family background, the effect of family background on ILE was rather small. Further, the results revealed that child characteristics seem to be more important within the ILE. Finally, the third research question investigated the effects of HLE and ILE on vocabulary.

4.3 Research question 3: To what extent do HLE and ILE predict vocabulary in preschool and primary school?

Within the HLE, the dimension cultural capital predicted vocabulary in preschool ($B = .18$) and in first grade ($B = .11$). No other dimension of the HLE was predictive towards vocabulary in preschool or primary school. Within the ILE, the dimension literacy-specific ILA predicted vocabulary in preschool ($B = .15$). No other dimension of the ILE was predictive towards vocabulary in preschool.

With respect to compensatory effect condition 3, that HLE and ILE have an independent effect on vocabulary, we focussed on the joint effects of the HLE and ILE dimensions on vocabulary. Taken all dimensions together, the HLE predicted vocabulary in preschool with a small effect ($f^2_{\text{HLE}} = 0.04$). The same applied to the ILE ($f^2_{\text{ILE}} = 0.06$). Further, HLE was predictive for vocabulary in primary school ($f^2 = 0.03$).

With regard to the conditions to assume a compensatory effect, this indicates that compensatory effect condition 3 was accomplished: HLE and ILE predicted vocabulary independently of one another.

In sum, family background, child characteristics, HLE, ILE, and vocabulary in preschool explained 46 % of variance in vocabulary in first grade, while 31 % of variance in vocabulary during preschool could be explained by child characteristics, family background, HLE, and ILE.

5. Discussion

In the present study, we investigated the predictive power of family background and child characteristics as well as characteristics of home and institutional learning environment on children's vocabulary in preschool and first grade within an extended theoretical framework based on Niklas (2015).³ The aim was to evaluate the conditions which are necessary to assume a compensatory effect of learning in kindergarten. Building upon prior theoretical considerations, we proposed a model integrating home and institutional learning environment in a parallelized way, aiming to make their effects on vocabulary and their predictability by family background and child characteristics distinguishable and comparable. Our study led to three main findings according to our research questions and with respect to a compensatory effect: Firstly, family background and child characteristics predicted vocabulary in preschool and primary school substantially; and this effect was partly mediated by both learning environments (meeting compensatory effect condition 1).

Secondly, parental education and a child's gender were substantially related to learning environment facets of the home and preschool. All in all, the connection between parental education and home-related learning facets was the strongest. Moreover, the predictive power of family background towards the preschool learning environment was not as strong as the predictive power child characteristics had towards this learning environment. Our study revealed that there were differences in the literacy-specific and domain-unspecific implicit learning activities between boys and girls within preschool, while at home these gender preferences occurred for the children's domain-unspecific activities only. This indicates that compensatory effect condition 2 was met.

Thirdly, both learning environments were predictive towards vocabulary. Within preschool, implicit learning activities were more relevant compared to cultural capital and practice. Based on our results, we can state that compensatory effect condition 3 was met: characteristics of the preschool learning environment were predictive towards vocabulary. In this context, the influence of gender on children's engagement with books is especially important. For boys, it could hinder a possible compensatory effect of learning in preschool.

3 For interpretation of the findings please note that no casual relations were investigated.

In general, our findings are in line with previous studies (e.g., Sylva et al., 2013; Weinert & Ebert, 2013) that found small to medium interrelations between children's early literacy skills and their home and preschool learning environment. Beyond that, our findings suggest that both environments represent distinguishable concepts with independent relations to endogenous family background variables and child characteristics as well as to exogenous outcome variables as vocabulary. While home characteristics (especially cultural capital) were more closely related to the family background, preschool facets (especially implicit learning activities) were more closely related to child characteristics. These findings have implications for theory building and practice.

In the present study, home-related cultural capital was measured by the amount of books available in the household. In other studies, this variable is used as an indicator for a family's socioeconomic status (e.g., Bos, Tarelli, Bremerich-Vos, & Schwippert, 2012). Therefore, it could be productive to launch a discussion about the theoretical localization of cultural capital. If cultural capital is seen as part of the learning environment, the home learning environment would be theoretically strongly linked to parental education. In Niklas' (2015) model, on which the present assumptions were built, cultural capital is treated as such a part. Following the sociologic perspective of Bourdieu (1983) one could argue, however, that cultural capital is a part of the family background which, in turn, influences how learning at home is supported (see also Bos et al., 2012). In future studies, consequently, it would be of interest to separate the effects of cultural capital from other home-related characteristics on a child's competence development (see also Linberg, 2017).

Our study can contribute to theory building as it revealed that a compensatory effect of learning in preschool cannot be assumed offhand. In line with several studies (e.g., Baumert & Maaz, 2006; McElvany, Becker, & Lüdtke, 2009; OECD, 2013) we replicated the link between family background and a child's academic competencies, while it can be assumed that part of this effect is mediated by characteristics of the home learning environment (e.g., Foster et al., 2005; Lehrl et al., 2012). With regard to the question if preschools compensate for less stimulating homes, we did find an independent effect of the preschool learning environment towards vocabulary, which is a necessary condition for a compensatory effect. However, in line with Becker and Biedinger (2016) and Högbe (2016) we also found that preschool characteristics were not fully independent of family background.

This finding has not only theoretical, but also practical implications. The gender bias found referred to the child's implicit learning activities and was medium for domain-unspecific activities and small, but substantial, for literacy-specific activities, which, in turn, was linked to vocabulary. As the domain-unspecific activities referred to creative activities and the literacy-specific activity referred to interaction with books, it is likely that educators might consider these activities as girl-typical. As a consequence, they may not encourage boys to get involved in these activities as much as they encourage girls. As the child's involvement with books in preschool was directly linked to vocabulary, this could prevent a compensatory effect

of the preschool learning environment for boys. Gender sensitive pedagogy encouraging boys to engage with literacy-related materials such as books could positively influence a boy's early literacy skills.

Several limitations of the present study should be considered when generalizing the results. First, even though we used a nationwide sample and, therefore, expanded the findings of former studies (e.g., Weinert & Ebert, 2013), the representativeness is questionable as the sample was selective due to a low percentage of children with migration background and a large percentage of children from privileged families. Moreover, the sample had a partly clustered structure which had to be neglected in the analyses and the treatment of missing values was suboptimal, as it may have led to minor distortions. This restricts the validity of the results. Further, there was little variance in home learning activities which might lead to an underestimation of their effects. Furthermore, we used a single-item approach to operationalize cultural capital, cultural practice, and literacy-specific implicit learning activities restricting the construct validity of the results. On the other hand, the single-item approach increased the validity of the comparison between the two learning environments. In future studies it should be investigated if the present findings can be replicated within a multi-item approach.

Notwithstanding that these limitations have to be considered when judging the results of our study, we can conclude that both the home and institutional learning environment have distinguishable effects on a child's early literacy skills and relate differently to family and child characteristics. With respect to a compensatory effect of the preschool learning environment, our study directs the focus on the children's implicit learning activities that can be influenced by the educators.

Acknowledgements

This paper uses data from the National Educational Panel Study (NEPS): Starting Cohort Kindergarten, *doi:10.5157/NEPS:SC2:3.0.0*. From 2008 to 2013, NEPS data was collected as part of the Framework Program for the Promotion of Empirical Educational Research funded by the German Federal Ministry of Education and Research (BMBF). As of 2014, NEPS is carried out by the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg in cooperation with a nationwide network.

References

- Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A., Rohwer, G., Rässler, S., & Blossfeld, H.-P. (2011). Sampling designs of the National Educational Panel Study: challenges and solutions. In H.-P. Blossfeld, H.-G. Roßbach, & J. von Maurice (Eds.), *Education as a Lifelong Process – The German National Educational Panel Study (NEPS)* [Special issue]. *Zeitschrift für Erziehungswissenschaft*, 14, 51–65. *doi:10.1007/s11618-011-0181-8*
- Autorengruppe Bildungsberichterstattung [Author Group Educational Reporting]. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit ei-*

- ner Analyse zu Bildung und Migration* [Education in Germany 2016. An indicator-based report with an analysis on education and migration]. Bertelsmann: Bielefeld. doi:10.3278/6001820ew
- Baumert, J., & Maaz, K. (2006). Das theoretische und methodische Konzept von PISA zur Erfassung sozialer und kultureller Ressourcen der Herkunftsfamilie: Internationale und nationale Rahmenkonzeption [The theoretical and methodological concept of PISA for measurement of social and cultural resources of the family of origin: international and national framework conception]. In J. Baumert, P. Stanat, & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000* [Disparities in educational system due to origin. Analyses in the context of PISA 2000] (pp. 11–29). Wiesbaden: VS. doi:10.1007/978-3-531-90082-7
- Becker, B. (2010). Wer profitiert mehr vom Kindergarten? Die Wirkung der Kindergartenbesuchsdauer und Ausstattungsqualität auf die Entwicklung des deutschen Wortschatzes bei deutschen und türkischen Kindern [Who profits more from preschool? Effects of duration and material resources on vocabulary development in German and Turkish children]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 139–163. doi:10.1007/s11577-010-0090-5
- Becker, B., & Biedinger N. (2016). Ethnische Ungleichheiten in der vorschulischen Bildung [Ethnic disparities in preschool education]. In C. Diehl, C. Hunkler, & C. Kristen (Eds.), *Ethnische Ungleichheiten im Bildungsverlauf* [Ethnic disparities in education] (pp. 433–474). Wiesbaden: VS. doi:10.1007/978-3-658-04322-3_10
- Bos, W., Schwippert, K., & Stubbe, T. C. (2007). Die Kopplung von sozialer Herkunft und Schülerleistung im internationalen Vergleich [The linkage of social background and students' achievements in international comparison]. In W. Bos, S. Hornberg, K. Arnold, G. Faust, L. Fried, E. Lankes, K. Schwippert, & R. Valtin (Eds.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* [IGLU 2006. Reading skills of elementary school students in Germany in international comparison] (pp. 225–248). Münster: Waxmann.
- Bos, W., Tarelli, I., Bremerich-Vos, A., & Schwippert, K. (2012). *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* [Reading competencies of primary school children in Germany from an international comparative perspective]. Münster: Waxmann.
- Blossfeld, H.-P., Roßbach, H.-G., & von Maurice, J. (Eds.). (2011). Education as a lifelong process – The German National Educationa Panel Study (NEPS) [Special issue]. *Zeitschrift für Erziehungswissenschaft*, 14. doi:10.1007/s11618-011-0198-z
- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital [Economic, cultural, and social capital]. In R. Kreckler (Ed.), *Soziale Ungleichheiten. Soziale Welt Sonderband 2* [Social disparities. social world special Issue 2] (pp. 183–198). Göttingen: Schwartz. doi:10.1007/978-3-531-18944-4_15
- Bronfenbrenner, U. (1979). Contexts of child rearing. Problems and prospects. *American Psychologist*, 34, 844–850. doi:10.1037/0003-066X.34.10.844
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human-development – Research perspectives. *Developmental Psychology*, 22, 723–742. doi:10.1037/0012-1649.22.6.723
- Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, 25, 140–165. doi:10.1016/j.ecresq.2009.11.001
- Burgess, S. R., Hecht, S. A., & Lonigan, C. J. (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading Research Quarterly*, 37, 408–426. doi:10.1598/RRQ.37.4.4

- Bynner, J., Londra, M., & Jones, G. (2004). *The impact of government policy on social exclusion among young people*. London: Social Exclusion Unit. doi:10.1080/13676260120028529
- Cloney, D., Cleveland, G., Hattie, J., & Tayler, C. (2016). Variations in the availability and quality of early childhood education and care by socioeconomic status of neighborhoods. *Early Education and Development, 27*, 384–401. doi:10.1080/10409289.2015.1076674
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.
- de Jong, P. F., & Leseman, P. M. (2001). Lasting effects on home literacy on reading achievement in school. *Journal of School Psychology, 39*, 389–414. doi:10.1016/S0022-4405(01)00080-2
- Dumais, S. (2006). Elementary school students' extracurricular activities: The effects of participation on achievement and teachers' evaluations. *Sociological Spectrum, 26*, 117–147.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody picture vocabulary test – Revised (PPVT-R)*. Circle Pines: American Guidance Service.
- Ebert, S., Lockl, K., Weinert, S., Anders, Y., Kluczniok, K., & Roßbach, H.-G. (2013). Internal and external influences on vocabulary development in preschool children. *School Effectiveness and School Improvement, 24*, 138–154. doi:10.1080/09243453.2012.749791
- ESSA – Every Student Succeeds Act. (2015). Every student succeeds act. 20 USC 6301 note. *Public Law 114-95, 114th Congress*, December 10. Retrieved from <https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf>
- Foster, M. A., Lambert, R., Abbott-Shim, M., McCarty, F., & Franze, S. (2005). A model of home learning environment and social risk factors in relation to children's emergent literacy and social outcomes. *Early Childhood Research Quarterly, 20*, 13–36. doi:10.1016/j.ecresq.2005.01.006
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: SAGE.
- Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*, 277–319. doi:10.1108/S1474-7979(2009)0000020014
- Hoff, E. (2010). Context effects on young children's language use: The influence of conversational setting and partner. *First Language, 30*, 461–472. doi:10.1177/0142723710370525
- Hogrebe, N. (2016). Segregation im Elementarbereich – Mobilität und Trägerschaft [Segregation in preschool education. Mobility and providers of settings]. *Zeitschrift für Grundschulforschung, 9*, 20–33.
- Jarvis, C., MacKenzie, S., & Podsakoff, P. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*, 199–218. doi:10.1086/376806
- Kluczniok, K., Lehr, S., Kuger, S., & Roßbach, H.-G. (2013). Quality of the home learning environment during preschool age – Domains and contextual conditions. *European Early Childhood Education Research Journal, 21(3)*, 420–438. <https://doi.org/10.1080/1350293X.2013.814356>
- Kluczniok, K., & Roßbach, H.-G. (2014). Conceptions of Educational Quality for Kindergartens. *Zeitschrift für Erziehungswissenschaft, 17*, 145–158. doi:10.1007/s11618-014-0578-2
- Kuger, S., & Kluczniok, K. (2008). Prozessqualität im Kindergarten. Konzept, Umsetzung und Befunde [Process quality in preschools. Concepts, implementations and findings] (Special issue). *Zeitschrift für Erziehungswissenschaft, 11*, 159–178. doi:10.1007/978-3-531-91452-7_11

- Lechert, Y., Schroedter, J., & Lüttinger, P. (2006). Die Umsetzung der Bildungsklassifikation CASMIN für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensus 1976–2004 [Implementation of the education classification CASMIN for the Census in 1970, the additional microcensus survey in 1971 and the Microcensuses between 1976 and 2004]. *ZUMA-Methodenbericht 2006*. Mannheim: ZUMA.
- LeFevre, J., Skwarchuk, S., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2009). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *41*, 55–66. doi:10.1037/a0014532
- Lehrl, S., Ebert, S., Roßbach, H., & Weinert, S. (2012). Die Bedeutung der familiären Lernumwelt für Vorläufer schriftsprachlicher Kompetenzen im Vorschulalter [Effects of the HLE on children's emerging literacy]. *Zeitschrift für Familienforschung*, *24*, 115–133.
- Lehrl, S., Kuger, S., & Anders, Y. (2014). Soziale Disparitäten beim Zugang zu Kindergartenqualität und differenzielle Konsequenzen für die vorschulische mathematische Entwicklung [Social disparities in access towards preschool quality and their consequences for early numeracy skills]. *Unterrichtswissenschaft*, *42*, 132–151. doi:10.3262/UW1402132
- LifBi – Leibniz-Institut für Bildungsverläufe [Leibnitz Institute for Educational Trajectories]. (2016). Starting Cohort 2: Kindergarten (SC2). Wave 1 Questionnaires (SUF Version 4.0.0). Retrieved from https://www.neps-data.de/Portals/o/NEPS/Datenzentrum/Forschungsdaten/SC2/4-0-0/SC2_4-0-0_W1_en.pdf. doi:10.5157/NEPS:SC2:4.0.0
- Linberg, T. (2017). Kind und Kontext. Häusliche Lernumwelt und soziale Ungleichheiten im vorschulischen Sprachstand [Child and context. Home learning environment and social disparities in early literacy development] Wiesbaden: VS.
- Lüttinger, P., & König, W. (1988). Die Entwicklung einer international vergleichbaren Klassifikation für Bildungssysteme [Development of an international comparable classification for educational systems]. *ZUMA Nachrichten*, *12*, 1–14.
- McElvany, N., Becker, M., & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten [The role of family variables in reading literacy, vocabulary, reading motivation, and reading behaviour]. *Zeitschrift Für Entwicklungspsychologie und Pädagogische Psychologie*, *41*, 121–131. doi:10.1026/0049-8637.41.3.121
- Melhuish, E. (2010). *Impact of the HLE on Child Cognitive Development. Secondary analysis of data from 'Growing up in Scotland'*. Edinburgh: Scottish Government.
- Melhuish, E. C., Phan, M. B., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2008). Effects of the home learning environment and preschool center experience upon literacy and numeracy development in early primary school. *Journal of Social Issues*, *64*, 95–114. doi:10.1111/j.1540-4560.2008.00550.x
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology*, *40*, 665–681. doi:10.1037/0012-1649.40.5.665
- Niklas, F. (2015). Die familiäre Lernumwelt und ihre Bedeutung für die kindliche Kompetenzentwicklung [The learning environment provided by the family and its impact on the development of child competencies]. *Psychologie in Erziehung und Unterricht*, *62*, 106–120. doi:10.2378/peu2015.art11d
- Niklas, F., Möllers, K., & Schneider, W. (2013). Die frühe familiäre Lernumwelt als Mediator zwischen strukturellen Herkunftsmerkmalen und der basalen Lesefähigkeit am Ende der ersten Klasse [The early learning environment provided by the family as a mediator between structural family background and basic read-

- ing abilities at the end of grade 1]. *Psychologie in Erziehung und Unterricht*, 60, 94–111. doi:10.2378/peu2013.art08d
- Niklas, F., & Schneider, W. (2017). Home learning environment and development of child competencies from kindergarten until the end of elementary school. *Contemporary Educational Psychology*, 49, 263–274. doi:10.1016/j.cedpsych.2017.03.006
- Niklas, F., & Tayler, C. (2018). Room quality and composition matters: Children’s verbal and numeracy abilities in Australian early childhood settings. *Learning and Instruction*, 54, 114–124. doi:10.1016/j.learninstruc.2017.08.006
- OECD – Organisation of Economic Co-operation and Development. (2013). *PISA 2012 Assessment and analytical framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris: OECD. doi:10.1787/9789264190511-en
- R Core Team (2015). *R: A language and environment for statistical computing*. R foundation for statistical computing. Vienna, Austria. Retrieved from <http://www.R-project.org/>
- Retelsdorf, J., & Möller, J. (2008). Familiäre Bedingungen und individuelle Prädiktoren der Lesekompetenz von Schülerinnen und Schülern [Family conditions and individual predictors for students’ reading comprehension]. *Psychologie in Erziehung und Unterricht*, 55, 227–237. Retrieved from <https://www.reinhardt-journals.de/index.php/peu/article/view/638>
- Roßbach, H.-G. (2005). Effekte qualitativ guter Betreuung, Bildung und Erziehung im frühen Kindesalter auf Kinder und ihre Familien [Effects of good quality of early child care, education and upbringing on children and their families]. In Sachverständigenkommission Zwölfter Kinder- und Jugendbericht (Eds.), *Materialien zum Zwölften Kinder- und Jugendbericht. Band 1. Bildung, Betreuung und Erziehung von Kindern unter sechs Jahren* [Materials for the twelfth children and youth report, Volume 1. Education, care and upbringing of children under six years] (pp. 55–174). München: DJI.
- Roßbach, H.-G., Kluczniok, K., & Kuger, S. (2008). Auswirkungen eines Kindergartenbesuchs auf den kognitiv-leistungsbezogenen Entwicklungsstand von Kindern. In H.-G. Roßbach, & H.-P. Blossfeld (Eds.), *Frühpädagogische Förderung in Institutionen* [Sonderheft]. *Zeitschrift für Erziehungswissenschaft*, 11, 139–158.
- Roßbach, H.-G., Tietze, W., & Weinert, S. (2005). *Peabody picture vocabulary test – revised*. Adapted German research version, original of Dunn, L. M., and L. M. Dunn. 1981. Universität Bamberg, FU Berlin.
- Sanchez, G. (2013). *PLS path modeling with R. Trowchez editions*. Berkeley, 2013. Retrieved from http://gastonsanchez.com/PLS_Path_Modeling_with_R.pdf
- Sanchez, G., Trinchera, L., & Russolillo, G. (2015). *PLSPM: Tools for partial least squares path modeling (PLS-PM)*. R Package Version 0.4.7. Retrieved from <http://CRAN.R-project.org/package=plspm>
- Schmerse, D., Anders, Y., Flöter, M., Wieduwilt, N., Roßbach, H.-G., & Tietze, W. (2018). Differential effects of home and preschool learning environments on early language development. *British Educational Research Journal*, 44(2), 338–357. <https://doi.org/10.1002/berj.3332>
- Siraj-Blatchford, I. (2010). Learning in the home and at school: How working class children ‘Succeed against the Odds’. *British Educational Research Journal*, 36, 463–482. doi:10.1080/01411920902989201
- Speece, D. L., Ritchey, K. D., & Cooper, D. H., Roth, F. P., & Schatschneider, C. (2004). Growth in early reading skills from kindergarten to third grade. *Contemporary Educational Psychology*, 29, 312–332. doi:10.1016/j.cedpsych.2003.07.001
- Sylva, K., Sammons, P., Chan, L. S., Melhuish, E., Siraj-Blatchford, I., & Taggart, B. (2013). The effects of early experiences at home and pre-school on gains in English and Mathematics in Primary School: a Multilevel Study in England. *Zeitschrift für Erziehungswissenschaft*, 16, 277–310. doi:10.1007/s11618-013-0364-6

- Tarelli, I., Valtin, R., Bos, W., Bremerich-Vos, A., & Schwippert, K. (2013). IGLU 2011. Wichtige Ergebnisse im Überblick [IGLU 2011. Overview of important findings]. In W. Bos, I. Tarelli, A. Bremerich-Vos, & K. Schwippert (Eds.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* [Reading skills of Primary School children in Germany in international comparison] (pp. 11–25). Münster: Waxmann.
- Tietze, W., Becker-Stoll, F., Bensel, J., Eckhardt, A. G., Haug-Schnabel, G., Kalicki, B., Keller, H., & Leyendecker, B. (Eds.). (2013). Nationale Untersuchung zur Bildung, Betreuung und Erziehung in der frühen Kindheit (NUBBEK). Weimar & Berlin: das netz.
- Weiber, R., & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS* [Structural equation modeling. An applied introduction into causal analysis with AMOS, SmartPLS, and SPSS] (2nd ed.). Heidelberg: VS. doi:10.1007/978-3-642-35012-2
- Weinert, S., & Ebert, S. (2013). Spracherwerb im Vorschulalter. Soziale Disparitäten und Einflussvariablen auf den Grammatikerwerb [Language acquisition in pre-school. Social disparities and predictors to the acquisition]. *Zeitschrift für Erziehungswissenschaft*, 16, 303–332.
- Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child Psychology: Cognition, perception, and language* (5th ed., pp. 523–573). New York, NY: Wiley & Sons.
- Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 69, 848–872. doi:10.1111/j.1467-8624.1998.tb06247.x

Appendix

Table A1: Correlations between all indicators of family background, child characteristics, home learning environment, institutional learning environment, and children's vocabulary

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Family Background</i>															
(1) Parental education	1	-.085	-.043	-.056	-.391	.077	.125	.066	.034	-.006	.006	.004	.071	-.004	-.040
(2) Migration background		1	.056	-.042	-.117	.040	-.129	-.017	.052	.001	.024	-.069	.038	-.031	-.074
<i>Child characteristics</i>															
(3) Gender			1	-.093	-.033	.076	.021	.011	-.028	.117	-.347	.288	-.088	.170	.013
(4) Age				1	-.026	.011	-.017	.007	.065	.022	.044	-.047	-.077	-.091	.003
<i>Home learning environment</i>															
(5) Cultural capital					1	.124	.258	.106	-.063	-.025	.003	.058	.092	-.004	-.030
(6) Cultural practice						1	.092	.071	.059	.040	.018	.028	.066	.070	.011
(7) Literacy-specific ILA							1	.244	.078	.183	.038	.214	.017	.154	.042
<i>Implicit learning activities</i>															
(8) Comparing/collecting things								1	.145	.191	.161	.180	.088	.124	.164
(9) Number games or counting									1	.066	.246	.022	.222	.082	.115
(10) Doing jigsaws										1	.014	.230	-.072	.082	.085
(11) Building games											1	-.076	.155	.009	.123
(12) Doing crafts												1	-.016	.217	.144
(13) Sports													1	.184	.168
(14) Music, singing or dancing														1	.157
(15) Nature and gardening															1
(16) Role-playing															1

Notes. Correlation coefficient is Kendall's Tau. Significant correlations are bold.

Table A1 continued

Table A1 continued

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<i>Institutional learning activities</i>															
(17) Cultural capital															
(18) Cultural practice															
(19) Literacy-specific ILA															
<i>Implicit learning activities</i>															
(20) Comparing/ collecting things															
(21) Number games or counting															
(22) Doing jigsaws															
(23) Building games															
(24) Doing crafts															
(25) Role-playing															
(26) Sports															
(27) Music, singing or dancing															
(28) Nature and gardening															
<i>Vocabulary</i>															
(29) Vocabulary t1															
(30) Vocabulary t2															

Table A1 continued

Table A1 continued

	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
<i>Family Background</i>															
(1) Parental education	.038	-0.005	.019	.068	.033	.049	-0.060	.010	-0.027	.011	-0.005	-0.026	.037	.161	.188
(2) Migration background	-0.018	-0.005	-0.001	-.100	-0.009	-0.009	-0.004	-0.059	.005	-0.001	-0.057	-.074	-.111	-.278	-.201
<i>Child characteristics</i>															
(3) Gender	.260	-0.024	.034	.166	.059	.127	.200	-.533	.436	.381	-0.039	.194	-0.043	-0.067	-0.016
(4) Age	-0.038	.031	-0.032	.001	.025	.040	.000	.061	-0.054	-0.047	.082	.015	.076	.207	.140
<i>Home learning environment</i>															
(5) Cultural capital	.068	.033	.040	.084	.054	.058	.023	.043	.019	-0.011	.006	.023	.090	.210	.213
(6) Cultural practice	.044	-0.007	.167	-0.027	-0.011	-0.018	.045	-0.014	-0.036	-0.041	.002	-0.026	-0.014	.054	.028
(7) Literacy-specific ILLA	.109	.022	.001	.041	-0.002	.026	.003	.010	.029	.011	.013	.070	.034	.099	.100
<i>Implicit learning activities</i>															
(8) Comparing/collecting things	.097	-0.011	-0.025	.017	.038	.061	.018	-0.022	.070	.012	.002	.015	-0.018	.012	.026
(9) Number games or counting	.040	-0.014	-0.035	.031	.053	.096	.088	.059	.021	-0.023	.040	.014	.031	-0.023	.004
(10) Doing jigsaws	.069	.003	-0.035	.016	-0.013	.018	.089	-.068	.060	.022	-0.013	.046	-0.034	-0.053	-0.040
(11) Building games	-0.042	.009	.009	-0.012	.009	-0.012	-0.044	.223	-.141	-.145	.060	-0.057	.015	.042	.035
(12) Doing crafts	.203	-0.009	.008	.081	.026	.030	.037	-.173	.233	.110	-0.004	.080	.014	-0.048	-0.021
(13) Sports	.038	-0.060	.010	-0.067	.013	-0.040	-0.052	.035	-.083	-.074	.072	-0.038	-0.016	.000	.014
(14) Music, singing or dancing	.205	-0.033	-0.001	.021	.008	.021	.037	-.121	.107	.084	-0.007	.094	-0.036	.000	-0.016
(15) Nature and gardening	.087	-0.023	-0.023	-0.006	.051	-0.012	-0.007	-0.013	.030	-0.009	.034	.050	-0.028	-0.006	-0.033
(16) Role-playing	1	.021	-0.022	.073	.061	.025	.002	-.162	.080	.147	-0.042	.072	.029	-0.022	-0.001

Table A1 continued

Table A1 continued

	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
<i>Institutional learning activities</i>															
(17) Cultural capital		1	-.038	.113	.041	.022	.030	.060	-.022	-.009	-.056	-.015	.088	.073	.078
(18) Cultural practice			1	.017	.051	-.075	-.058	-.029	.011	.026	.085	.089	.041	.037	-.022
(19) Literacy-specific ILA				1	-.402	.342	.260	.025	.249	.173	.083	.132	.181	.092	.061
<i>Implicit learning activities</i>															
(20) Comparing/collecting things					1	-.385	.260	.141	.221	.165	.195	.228	.314	.060	.036
(21) Number games or counting						1	.441	.086	.299	.143	.068	.166	.226	.047	.043
(22) Doing jigsaws							1	.073	.220	.074	.016	.166	.101	-.028	-.013
(23) Building games								1	-.188	-.151	.094	-.031	.096	.073	.036
(24) Doing crafts									1	.320	.079	.204	.090	-.049	-.035
(25) Role-playing										1	.196	.265	.121	-.027	-.017
(26) Sports											1	.234	.275	.056	-.011
(27) Music, singing or dancing												1	.144	-.001	.033
(28) Nature and gardening													1	.089	.058
<i>Vocabulary</i>															
(29) Vocabulary t1														1	.480
(30) Vocabulary t2															1

Notes. Correlation coefficient is Kendall's Tau. Significant correlations are bold.

Eva-Kristina Franz, Vera Heyl, Albrecht Wacker & Tobias Dörfler

Konstruktvalidierung eines Tests zur Erfassung von adaptiver Handlungskompetenz in heterogenen Gruppen

Zusammenfassung

Der kompetente Umgang mit Heterogenität stellt für Lehrkräfte vor dem Hintergrund aktueller Entwicklungen eine zunehmend bedeutsamere Kompetenz dar. Im Projekt Effektive Kompetenzanalyse in der Lehrerbildung (EKoL) wurde auf der Grundlage des kompetenztheoretischen Professionalisierungsansatzes hierzu ein Theoriemodell adaptiver Handlungskompetenz konzeptualisiert, darauf bezogen ein Vignettest entwickelt und dieser anhand der Daten einer Befragung von Lehramtsstudierenden (n = 706) empirisch geprüft. Der Beitrag berichtet Befunde aus dieser Querschnittsbefragung zur Güte des angenommenen Theoriemodells, das empirisch repliziert werden konnte. Die Ergebnisse geben Anhaltspunkte zur Spezifikation von Facetten der generisch-pädagogischen Kompetenz und zu möglichen sie bedingenden Einflussfaktoren.

Schlagworte

Professionalisierung; Lehrerkompetenzen; Umgang mit Heterogenität; adaptive Handlungskompetenz; Konstruktvalidität; Strukturgleichungsmodelle

Dr. Eva-Kristina Franz (corresponding author), Institut für Erziehungswissenschaft, Fakultät für Bildungswissenschaften, Pädagogische Hochschule Freiburg, Kunzenweg 21, 79117 Freiburg, Deutschland

E-Mail: eva.franz@ph-freiburg.de

Prof. Dr. Vera Heyl, Institut für Sonderpädagogik, Fakultät für Erziehungs- und Sozialwissenschaften, Pädagogische Hochschule Heidelberg, Keplerstraße 87, 69120 Heidelberg, Deutschland

E-Mail: hey@ph-heidelberg.de

Prof. Dr. Albrecht Wacker, Institut für Erziehungswissenschaft, Fakultät für Erziehungs- und Sozialwissenschaften, Pädagogische Hochschule Heidelberg, Keplerstraße 87, 69120 Heidelberg, Deutschland

E-Mail: wacker@ph-heidelberg.de

Prof. Dr. Tobias Dörfler, Institut für Psychologie, Fakultät für Erziehungs- und Sozialwissenschaften, Pädagogische Hochschule Heidelberg, Keplerstraße 87, 69120 Heidelberg, Deutschland

E-Mail: doerfler@ph-heidelberg.de

Construct validation of a test measuring adaptive teaching competency in heterogeneous groups

Abstract

Competencies in dealing with heterogeneity can be regarded as one of the most important skills for teachers. The Effective Competence Assessment in Teacher Education (EKoL) project conceptualises a theoretically derived model of adaptive teaching competency in order to develop a vignette test. The empirical examination of the theoretically derived model is based on a large sample of preservice teacher trainees (n = 706). Results show that the assumed model fits well with the data.

The reported results contribute to the specification of the facets and structure of competencies in dealing with heterogeneity. In addition, first hypotheses about determining individual difference variables are generated.

Keywords

Professionalization; Teacher competencies; Dealing with heterogeneity; Adaptive teacher competence; Construct validity; Structural equation modelling

1. Einleitung

Der kompetente Umgang mit heterogenen Lerngruppen ist zunehmend ein zentrales Thema der Lehrerbildung (Bohl, Budde & Rieger-Ladich, 2017). Zurückzuführen ist diese Entwicklung nach Budde (2017) auf vier weitreichende Transformationen im Schulsystem in der jüngeren Vergangenheit: auf die ungleichheitsverstärkenden Effekte der deutschen Schule (Baumert et al., 2001; Prenzel, Sälzer, Klieme & Köller, 2013), den Anspruch der Inklusion (United Nations, 2006), zahlreiche Strukturreformen in der Sekundarstufe (Wacker, 2017) und schließlich auch auf die daraus hervorgehenden Veränderungen der Unterrichtspraxis (z.B. Differenzierungsformen). Nach Terhart (2013) können die universitäre Phase der Lehrerbildung und auch der Vorbereitungsdienst hier Voraussetzungen schaffen, die bezogen auf diese Transformationen zu einem nachhaltigen Kompetenzaufbau der Lehrkräfte in den ersten Berufsjahren führen. An den Hochschulen wird in diesem Zusammenhang verstärkt der Ansatz einer didaktischen Adaptivität (Hardy et al., 2011; Hertel, 2014; Hertel, Fingerle & Rohlf, 2016; Warwas, Hertel & Labuhn, 2011) diskutiert. Unter der in diesem Kontext fokussierten adaptiven Lehrkompetenz (Beck et al., 2008; Brühwiler, 2014) wird einerseits die Fähigkeit zu diagnostizieren und darauf aufbauend Lernangebote adaptiv zu gestalten, andererseits die Fähigkeit, eine Gruppe (beziehungsweise Klasse) strukturiert zu führen und mit herausfordernden Verhaltensweisen der Lernenden adäquat umzugehen, gefasst (Brühwiler, 2017; Helmke & Weinert, 1997). Adaptive Lehrkompetenz bezeichnet „kognitive Prozesse [...], die als Basis für die Planung und Durchführung

adaptiven Unterrichts gelten“ (Brühwiler, 2017, S. 124). Dabei wird zwischen der präaktionalen Planungskompetenz und einer interaktionalen Handlungskompetenz unterschieden (Brühwiler, 2014, 2017).

Im vorliegenden Artikel wird die Entwicklung eines darauf bezogenen Tests zur Messung adaptiver Handlungskompetenz ausgeführt und Ergebnisse seiner Konstruktvalidierung referiert.

2. Zur adaptiven Handlungskompetenz im Umgang mit heterogenen Gruppen

Adaptivität stellt einen Sammelbegriff für Strategien und Verfahren der Differenzierung und Individualisierung von Unterricht dar (Hertel et al., 2016). In einem adaptiven Unterricht wird versucht, diesen in Planung und Durchführung inhaltlich und methodisch so an die Lernenden anzupassen, dass möglichst viele Lernende günstige Lernbedingungen vorfinden und von diesen profitieren können (Beck et al., 2008; Brühwiler, 2017). Adaptiver Unterricht gilt innerhalb der Lehr-Lernforschung als das „gegenwärtig wissenschaftlich fundierteste und didaktisch aussichtsreichste Konzept“ (Häcker, 2017, S. 280), um auf interindividuelle Unterschiede angemessen zu reagieren und zielt damit insbesondere auf eine bestmögliche Ausschöpfung des individuellen Entwicklungspotenzials der Lernenden (Arnold, 2008; Hardy et al., 2011). Hinsichtlich der schulpraktischen Umsetzung adaptiven Unterrichts lassen sich drei praktizierte Varianten herausarbeiten (Warwas et al., 2011): (a) Eine erste Variante stellen offene Unterrichtssettings dar. In ihnen werden beispielsweise organisatorische, methodische und soziale, zuweilen auch inhaltliche Öffnungen angeboten. Dieser Ansatz ist jedoch mit der Gefahr einer Überforderung leistungsschwächerer Schülerinnen und Schüler verbunden (Munser-Kiefer, 2014). (b) Als eine zweite Variante adaptiver Lernumgebungen kann daher eine gezielte Variation des Arbeitsmaterials beschrieben werden. Dieser Ansatz bietet über die Zuteilung differenzierter inhaltlicher Lernangebote unter Berücksichtigung der individuellen Lernvoraussetzungen der Schülerinnen und Schüler ein großes adaptives Potential (Hertel et al., 2016). Hierfür ist jedoch bedeutsam, als Lehrkraft über ausreichend diagnostische Informationen sowie über eine Expertise in Bezug auf die Angemessenheit von Lernaufgaben zu verfügen (Franz & Lange, 2014; Haag & Lohrmann, 2007; Hellmann, 2015; Schrader & Helmke, 1987; Schrader, 2011, 2013). (c) Eine dritte Variante adaptiven Unterrichtens fokussiert im Gegensatz zu den zwei skizzierten Ansätzen auf die Prozesse zwischen den Lernenden und Lehrkräften und sucht die interaktionalen Unterstützungsprozesse hinsichtlich der Strukturierung von Inhalten und Prozessen sowie die kognitive Aktivierung von einzelnen Lernenden oder Gruppen zu variieren. In wissenschaftlicher Hinsicht wird diese adaptive Unterstützung und Hilfestellung bei Schwierigkeiten vielfach im Begriff des *Scaffolding* gefasst

(Einsiedler & Hardy, 2010; Gibbons, 2015; Hertel et al., 2016; Wood, Bruner & Ross, 1976).

Einer adaptiven Lehrperson gelingt es nach Beck und Kollegen (2008, S. 37) „das Lehr-Lern-Geschehen unter bestmöglicher Berücksichtigung der inhaltlichen Anforderungen des Unterrichtsinhaltes (*Sachkompetenz*), der Vielfalt der Wissens- und Lernvoraussetzungen und der Lernverläufe der Schülerinnen und Schüler sowie der situativen Aspekte des Lernens (*diagnostische Kompetenz*), der Möglichkeiten und Chancen der didaktischen Gestaltung der Lernsituation (*didaktische Kompetenz*), der pädagogischen Maßnahmen zur Steuerung, Führung und Begleitung einer Schülergruppe oder Klasse (*Klassenmanagement*) erfolgreich zu orchestrieren.“ Brühwiler (2017, S. 124) fasst „das Zusammenspiel der an der *adaptiven Lehrkompetenz* beteiligten kognitiven Prozesse“ in folgendem Modell zusammen: Auf einer präaktionalen Ebene der Planungskompetenz werden auf Basis diagnostischer Überlegungen Unterrichtsverläufe antizipiert und geeignete Maßnahmen ausgewählt. Während der interaktionalen Phase des Unterrichts werden die gewählten didaktischen Maßnahmen permanent mit diagnostischen Informationen rückgekoppelt. Die im Unterricht gewonnenen diagnostischen Erkenntnisse bilden die Basis für die Planung weiterer Unterrichtseinheiten. Die bei Beck und Kollegen (2008) angelegten Dimensionen *Sachkompetenz* und *Klassenführung* schaffen dabei günstige Rahmenbedingungen für den *adaptiven Unterricht* (Brühwiler, 2017, S. 125).

Adaptivere Lehrpersonen erreichen gemäß Beck und Kollegen (2008, S. 126 f.) in zeitlich begrenzten Unterrichtsreihen einen signifikant größeren Leistungszuwachs bei Schülerinnen und Schülern im Vergleich zu weniger adaptiven Lehrkräften. Diese positive Wirkung adaptiven Unterrichts auf den Lernerfolg der Lernenden konnte auch für den angloamerikanischen Sprachraum (Van Lehn, Graesser, Jackson, Jordan, Olney & Rosé, 2007; Walberg & Paik, 2000; Waxman, Wang, Anderson & Walberg, 1985) und für den computergestützten adaptiven Unterricht in Deutschland (Leutner, 2006) festgestellt werden.

Die nachfolgenden Ausführungen rücken die theoretische Modellierung und empirische Überprüfung der adaptiven Handlungskompetenz in den Mittelpunkt. Darunter wird die Kompetenz von Lehrenden gefasst, im interaktiven Unterrichtsprozess situierte Informationen zu erfassen und davon ausgehend geplante Unterrichtsmaßnahmen mit dem Ziel anzupassen, eine möglichst optimale Unterstützung der individuellen Lernprozesse zu ermöglichen (Beck et al., 2008; Brühwiler, 2014, S. 88). Als Ausgangspunkt wurden die bereits inhaltlich konturierten Facetten (1) „diagnostizieren können“, (2) „differenzieren können“ und (3) „Klassenführung beherrschen“ ausgewiesen, deren Operationalisierung nachfolgend begründet wird.

Eine diagnostisch kompetente Lehrkraft kann nach Beck und Kollegen (2008, S. 41) „die Lernenden bezüglich ihrer Lernvoraussetzungen und -bedingungen (Vorwissen, Lernweisen, Lerntempo, Lernschwächen usw.) sowie ihrer Lernergebnisse zutreffend einschätzen“. Obwohl aktuell divergierende Definitionen von Diagnostik und daraus resultierend von diagnostischer

Kompetenz im Kontext der Lehrerbildung diskutiert werden, herrscht Konsens hinsichtlich der Grundthese, dass aus diagnostischen Prozessen stets Aussagen über daran anschließende Maßnahmen folgen sollten (Aufschnaiter et al., 2015). Ein wesentlicher Aspekt diagnostischer Kompetenz ist das personenbezogene Wissen einer Lehrperson (Weinert, Schrader & Helmke, 1990). Dieses beinhaltet einerseits allgemeines Wissen über Schülerinnen und Schüler in bestimmten Alters- und Schulstufen, über anzunehmende Leistungsstände, Vorwissen, Fehlvorstellungen und Schwierigkeiten (Beck et al., 2008, S. 42), andererseits umfasst es individuelles Wissen über Lernvoraussetzungen, Stärken und Schwächen einzelner Lernenden und ebenso über die eigene Klasse (Clark & Peterson, 1986). Diagnostische Kompetenz kommt im Kontext adaptiver Lehrkompetenz sowohl in der proaktiven Planungsphase als auch in der interaktiven Handlungsphase zum Tragen. Der im Projekt *Effektive Kompetenzanalyse in der Lehrerbildung* (EKoL) definierte Begriff der diagnostischen Kompetenz beinhaltet analog zu Beck und Kollegen (2008, S. 43) vor allem die Kenntnis und die Anwendung qualitativer Methoden in der Diagnostik, wie beispielsweise die Nutzung von Schülerprodukten zur Diagnose von Präkonzepten, Lese-Rechtschreib-Leistungen und motivationalen Situationen oder die Schülerbeobachtung im Unterrichtsgeschehen für die begründete Auswahl an geeigneten Unterrichts- und Fördermaßnahmen.

Eine didaktisch kompetente Lehrkraft verfügt über ein reichhaltiges didaktisches und methodisches Wissen, kennt die Vor- und Nachteile unterschiedlicher Vorgehensweisen und vermag Hypothesen zu bilden, unter welchen Bedingungen diese mit Erfolg eingesetzt werden können (Beck et al., 2008, S. 41). Den Begriff der didaktischen Kompetenz (Beck et al., 2008, S. 44f.) fassen wir darauf bezogen als Kompetenz zur inneren Differenzierung, mit der einhergeht, dass eine kompetente Lehrperson in der Lage ist, Differenzierungsmaßnahmen begründet auszuwählen und einzusetzen (Saalfrank, 2012, S. 73). Dazu benötigt sie neben den Kenntnissen der Lernvoraussetzungen, der Interessen, der Lernbereitschaft, der Lern- und Arbeitstempi und der Lernstile der Schülerinnen und Schüler auch ein großes Repertoire an Unterrichtsmethoden und -medien, die die Lehrperson in die Lage versetzen, unterschiedliche didaktisch-methodische Vorgehensweisen gegeneinander abzuwägen und gezielt einzusetzen.

Wie die bisherige Darstellung bereits zeigt, besteht eine gewisse Nähe der beiden Konzepte Diagnostik und Didaktik. Es verwundert daher nicht, dass aus den von Brühwiler (2017) analysierten Daten für sämtliche Dimensionen diagnostischer und didaktischer Planungs- und Handlungskompetenz positive Korrelationen hervorgehen, wobei sich der Zusammenhang zwischen diagnostischer und didaktischer Handlungskompetenz als am stärksten ausgeprägt erweist.

Damit das beschriebene Wechselspiel aus Diagnostik und Didaktik wirksam werden kann, ist eine Klassenführung in dem Sinne erforderlich, dass die Lehrkräfte den Unterricht so gestalten, dass die Schülerinnen und Schüler sich möglichst aktiv mit dem Lerngegenstand befassen können und Unterrichtsstörungen vermieden werden (Beck et al., 2008, S. 42). Nach Kounin (1976) gelingt dies insbesondere dann, wenn Schülerinnen und Schüler durchgängig im Blick der Lehrkraft

stehen, der Unterricht reibungslos verläuft, Übergänge gut geplant und gestaltet werden und der Umgang mit auftretenden Unterrichtsstörungen möglichst wenig den Unterrichtsfluss tangiert. Im Kontext adaptiver Lehrkompetenz kommt die Klassenführung besonders in der Phase des Unterrichtens zum Tragen und wird daher ausschließlich der Handlungskompetenz zugeordnet. Hierbei stellt sie nach Brühwiler (2017) eine günstige Rahmenbedingung adaptiver Interaktionen dar.

Die Sachkompetenz, welche ebenfalls als günstige Rahmenbedingung adaptiven Unterrichts aufgefasst werden kann, wurde bei der im Folgenden beschriebenen Testentwicklung nicht berücksichtigt, da der Fokus des Forschungsvorhabens auf dem Kern adaptiver Handlungskompetenz lag.

3. Entwicklung des Testinstruments und Fragestellungen

Um die Kompetenzen angehender Lehrkräfte empirisch einzuschätzen, sind zwei Vorgehensweisen verbreitet: (a) die direkte Beobachtung sowie (b) Verfahren der Selbsteinschätzung (Oser, Curcio & Düggeli, 2007, S. 17). Beide Verfahren sind mit Nachteilen behaftet: Während die Selbstauskunft einer Person zwar mittels Fragebogen ökonomisch eingeholt werden kann, lassen sich damit Kompetenzprofile – beispielsweise diagnostische oder didaktisch-differenzierende Kompetenzen – nicht objektiv vorhersagen (Schladitz, Groß Ophoff & Wirtz, 2015). Die direkte Beobachtung kann demgegenüber einen inhaltlich und ökologisch validen Einblick geben, evoziert aber eine hohe personelle und finanzielle Ressourcenbindung und leidet häufig unter schwer zu kontrollierenden Störvariablen. Praxisnahe Vignettentests stellen diesbezüglich nach Oser und Kollegen (2007, S. 17) eine probate Alternative dar.

Video- und Textvignetten sind kurze, dichte Beschreibungen von Unterrichtssituationen und aufgrund ihrer Nähe zur Praxis gut geeignet, sich empirisch der beruflichen Handlungskompetenz von Lehrkräften anzunähern: „Über die Beobachtung und Beurteilung des Unterrichts [einer] Drittlehrperson soll der Zugang zur Struktur des Kompetenzprofils der beobachtenden Person ermöglicht werden“ (Oser et al., 2007, S. 18). Aufgrund ihres in praktischen Situationen angesiedelten Ansatzes spiegeln Vignetten recht authentisch die hohe Komplexität des beruflichen Alltags von Lehrkräften wider (Brovelli, Bölsterli, Rehm & Wilhelm, 2013, 2014). Darüber hinaus erscheint das Wissen von Lehrpersonen häufig an konkrete Situationen geknüpft zu sein und kann deshalb erst durch eine entsprechende Kontextualisierung in einer vergleichbaren Situation abgerufen und angewendet werden (Borko, 2004; Meschede, 2014). Diese kontextualisierte Erfassung von Lehrerkompetenzen hat in den vergangenen Jahren stark zugenommen (König, 2015) und begründet sich im Bestreben, die innerhalb der Kompetenzforschung aufzufindende Schlagseite, einseitig auf Wissen zu rekurrieren, zu überwinden (König, 2016). Das vorliegende Projekt reiht sich hier in eine ganze Reihe pro-

minenter Forschungsvorhaben ein (TEDS-FU, u.a. Blömeke, König, Suhl, Hoth & Döhrmann, 2015; TEDS-M, u.a. Blömeke et al., 2014; Observer, u.a. Stürmer, Seidel & Kunina-Habenicht, 2015; ProWin, u.a. Tepner et al., 2012; FALKO, u.a. Krauss et al., 2017; EKoL, u.a. Rutsch, Rehm, Vogel, Seidenfuß & Dörfler, 2018).

Vignettentests erscheinen als möglicher Königsweg: Sie beinhalten ökologisch valide Situationen, die sorgfältig ausgewählte Stimuli enthalten und dadurch gut kontrollierbar sind. Über daran anschließende Testfragen erhofft man sich – wie auch im hier berichteten Projekt – Einblicke in die aus den Leistungsdispositionen resultierenden Kompetenzen der Probanden. Basierend auf dem theoretisch begründeten Arbeitsmodell zur adaptiven Handlungskompetenz wurden zu den angenommenen drei inhaltlich getrennten Dimensionen adaptiver Handlungskompetenz in heterogenen Gruppen zunächst je 15 Vignetten entwickelt, die in den Fächern Deutsch, Mathematik und in den natur- und sozialwissenschaftlichen Fächerverbänden kontextuiert sind. Diese bestanden aus der verdichteten Beschreibung einer Unterrichtssequenz, dem sogenannten Vignettenstamm, und dazugehörigen Items in Form von vorformulierten Handlungsalternativen, deren Angemessenheit durch die (angehenden) Lehrkräfte jeweils auf einer 6-stufigen Likert-Skala eingeschätzt wurde. Beispielitems für die drei Subfacetten der adaptiven Handlungskompetenz sind in Tabelle 1 dargestellt.

Die Beurteilung der Inhaltsvalidität erfolgte über eine zweistufige Befragung von Expertinnen und Experten ($n = 200$; vgl. dazu Franz, Wacker & Heyl, 2018). Als Expertinnen und Experten wurden Personen definiert, die an Hochschulen an Instituten für Erziehungswissenschaft oder Sonderpädagogik zum Umgang mit Heterogenität forschen und lehren. Darüber hinaus wurden Personen an Staatlichen Seminaren für Didaktik und Lehrerbildung, die Lehramtsanwärterinnen und -anwärter in der zweiten Phase begleiten, ausgewählte Lehrkräfte an Schulen, die für den Deutschen Schulpreis nominiert wurden sowie Kolleginnen und Kollegen an Schulen, welche Lehrerfortbildungen zum Themenschwerpunkt Umgang mit Heterogenität anbieten, als Expertinnen und Experten gefasst.

In einem ersten Schritt wurde zur Selektion der Vignetten die Zustimmung der Expertinnen und Experten zu den Kriterien der Relevanz und der Eindeutigkeit der Aufgabenstellung herangezogen (Franz et al., 2018). Darüber hinaus wurde die Urteileinigkeit der Expertinnen und Experten hinsichtlich der Güte der Handlungsalternativen (Items) als weiteres Selektionskriterium genutzt. Die fünf Vignetten mit der größten Übereinstimmung der Expertinnen und Experten pro Subfacette wurden für den finalen Test ausgewählt.

In einem zweiten Schritt wurden anhand der Urteile der Expertinnen und Experten die gewünschte Itemanzahl von vier Items pro Vignettenstamm ausgewählt. Dafür wurden – ähnlich dem Vorgehen bei Tepner und Dollny (2014) – bei einer Abweichung von mehr als 25 % der Experten vom Modalwert um $|1|$ die entsprechenden Items aus dem Itempool entfernt. Die verbleibenden Antwortalternativen wurden gemäß der Modi in vier Gruppen aufgeteilt: schlechte Antwortalternative (Modus 1); eher schlechte Antwortalternative (Modus 2 oder 3); eher gute Antwortalternative (Modus 4 oder 5); gute Antwortalternative (Modus

Tabelle 1: Vignettenbeispiele für die drei Facetten zur Messung der adaptiven Handlungskompetenz und die dazugehörigen Handlungsalternativen (Items) sowie die Einschätzung der Angemessenheit der Items durch die Expertinnen und Experten ($n = 140$) auf einer 6-stufigen Likert-Skala von unangemessen bis angemessen

Facette	Vignettenstamm	Handlungsalternativen (Items)	Modus der Expert*innen
diagnostizieren können	Leistungsrückgang in Mathematik Ein Schüler der Klasse 6 fällt dadurch auf, dass er immer schlechtere Leistungen in den Mathematiktests zeigt, wodurch seine Versetzung gefährdet werden könnte. In den anderen Fächern hingegen ist die Leistung konstant. Der Fachlehrer, Herr Schneider, möchte die Ursache für den Leistungsabfall in Mathematik herausfinden. Er spricht im Lehrerzimmer mit einer Kollegin: Wie zielführend ist aus pädagogisch-psychologischer Perspektive Herrn Schneiders Aussage?	a) „Dass der Max aber halt auch so faul ist, das bringt ihn nicht weiter.“ b) „Könntest Du mal auf die Klausur schauen, vielleicht fällt dir ja was auf.“ c) „Könntest Du Max mal in meinem Unterricht beobachten?“ d) „Können wir uns gemeinsam seine Klassenarbeiten und Hefte noch einmal systematisch anschauen?“	1 5 6 6
differenzieren können	Berufseinstieg Sie kommen als junge Lehrkraft an eine der neuen Gemeinschaftsschulen und unterrichten in Klassenstufe 7 und 8 jahrgangsübergreifend Mathematik. In den beiden Jahrgängen müssen Sie mindestens drei verschiedene Kompetenzniveaus bedienen. Wie würden Sie vorgehen?	e) Ich arbeite anhand von Lernspuren, d. h. ich halte Material für alle Kompetenzniveaus in meinem Klassenraum bereit. f) Ich setze Wochenpläne ein, in denen ich den Kindern unterschiedliche „Arbeitspäckchen“ zusammenstelle. g) Ich lasse immer einen Teil der Klasse still arbeiten und arbeite mit den anderen frontal. h) Ich führe das „Chefprinzip“ ein: Immer ein Schüler ist Chef einer Aufgabe, zu dem dann die Mitschüler bei Fragen kommen können.	6 5 1 5
Klassenführung beherrschen	Mein Traumberuf Ihre Anwärtlerin (Referendarin) thematisiert zum Ende des Schuljahres in einer Klasse 8 der Realschule das Thema „Mein Traumberuf“, um die Schüler(innen) schon einmal auf das im darauffolgenden Schuljahr stattfindende Praktikum vorzubereiten. Dazu lässt sie die Jugendlichen Steckbriefe des jeweiligen „Traumberufs“ erstellen. Die Jugendlichen haben vorbereitetes Material erhalten, waren im Berufsinformationszentrum und haben im Klassenraum Internetzugang. Sie arbeiten alleine oder in Gruppen. Wie sollte sich die Anwärtlerin aus pädagogisch-psychologischer Perspektive über den Arbeitsstand auf dem Laufenden halten?	i) Sie sollte immer im Klassenraum umhergehen und sehen, was die Jugendlichen arbeiten. j) Sie sollte jede Unterrichtsstunde mit einer Reflexionsrunde im Plenum beenden. k) Sie sollte jede Unterrichtsstunde mit einer Präsentation im Plenum beginnen und beenden. l) Sie sollte regelmäßig unterbrechen und nachfragen.	6 5 1 1

6) (Franz et al., 2018). Der Versuch, Antwortalternativen in den Aufgaben im Test zu belassen, welche zwar einerseits sehr einheitlich beurteilt wurden, andererseits aber möglichst divergierende Modi aufwiesen, gelang nicht durchgängig, weshalb letztendlich bei gleicher Güte der Items zusätzlich inhaltliche Kriterien bei der Auswahl zugrunde gelegt wurden.

Eine weitere Expertenbefragung (Rücklauf $n = 74$) an derselben Grundgesamtheit von Personen aus Hochschule, Seminar und Schule wurde zur Gewinnung eines Normwerts (Franz et al., 2018) für den finalen Vignettentest genutzt. Dieser dient als Grundlage für einen zu bildenden Summenscore, welcher als Kompetenzscore in die Analysen eingeht: Schätzt die Probandin, der Proband auf einer 6-stufigen Likert-Skala die Güte der Handlungsalternative (Testitem) exakt gleich ein wie die Expertinnen und Experten, so erhält er oder sie 2 Punkte. Unterscheidet sich die studentische Einschätzung um $|1|$, so wird noch ein Punkt gegeben. Größere Abweichungen wurden mit 0 Punkten gewertet.

3.1 Überlegungen zur Konstruktvalidierung

Im Hinblick auf bedeutsame Bedingungsfaktoren für einen kompetenten Umgang mit Heterogenität kann hinsichtlich einer konvergenten Validierung angenommen werden, dass die Einstellung zu Inklusion in einem positiven Zusammenhang mit der Performanz von Lehrkräften in heterogenen Gruppen steht (Heyl, Janz, Seifried & Trumpa, 2014; Miller, 2013; Seifried & Heyl, 2016). Die systemische Betrachtung von Organisationen zeigt, dass die Handlungskompetenz von Akteuren nicht nur vom individuell divergierenden spezifischen Können beeinflusst wird, sondern auch stark von Einstellungen, pädagogischen Haltungen und subjektiven Deutungen abhängt (König & Volmer, 2008). Heyl und Kollegen (2014, S. 200) benennen diesbezüglich explizit die Einstellung „der Akteure im Bildungsbereich“ als einen „wesentlichen Faktor für die Umsetzung von schulischer Inklusion“. Dies liegt darin begründet, dass Lehrkräfte mit einer positiven Einstellung zu Inklusion sich bereitwilliger fortbilden und sich auch im eigenen schulischen Alltag aktiver um wirksame pädagogische Rahmungen bemühen als Lehrkräfte mit einer negativeren Haltung. Für das vorliegende Projekt wird daher analog angenommen, dass Studierende mit einer positiveren Einstellung zu Inklusion auch höhere Werte im Kompetenztest zum Umgang mit Heterogenität aufweisen.

Das berufsbezogene Selbstkonzept lässt sich den motivationalen Orientierungen von Lehrpersonen subsumieren. Damit stellt es einen Teilbereich professioneller Kompetenz von Lehrkräften dar, was einen positiven Zusammenhang mit dem professionellen Handlungskompetenz erwarten lässt. Darüber hinaus ist das berufsbezogene Selbstkonzept eine in Bezug auf die Einstellung zu Inklusion bedeutsame Größe: Lehrkräfte haben dann eine positivere Einstellung zu Integration, wenn sie sich als diesbezüglich kompetent einschätzen (Sermier Dessemontet, Benoit & Bless, 2011). Die Frage, inwieweit das eigene Selbstkonzept mit der tatsächlichen Performanz im Unterrichtsgeschehen, der eigentlichen Handlungskompetenz,

übereinstimmt, ist bis dato noch nicht konsensuell geklärt. Brodhäcker (2014) stellt diesbezüglich signifikante Übereinstimmungen fest. Schladitz und Kollegen (2015) dagegen vermögen dies in ihrer Studie nicht zu replizieren.

Schulze-Stocker, Holzberger, Kunina-Habenicht, Terhart und Kunter (2016) können aufzeigen, dass sich das individuelle Selbstkonzept durch den gezielten Besuch von Lehrveranstaltungen an der Hochschule verändert. Im Detail zeigt sich, dass sich Studierende nach dem Besuch selbstgewählter Lehrveranstaltungen einen signifikant höheren Kenntnisstand attestieren (Schulze-Stocker et al., 2016), zu welchem auch in den Ergebnissen standardisiert gemessenem bildungswissenschaftlichen Wissens Evidenzen vorliegen. Dies macht die Frage nach der Teilnahme an ausgewählten Lehrveranstaltungen zum Thema Inklusion (Franz, 2014; Lüders, Eisenacher & Pleßmann, 2006; Schulze-Stocker et al., 2016) zu einer interessanten Kovariaten eines kompetenten Umgangs mit Heterogenität. Es kann darauf begründet angenommen werden, dass der Besuch von gezielten Lehrveranstaltungen zum Thema Inklusion mit dem Kompetenzscore positiv zusammenhängt.

Befunde zu Bedingungsfaktoren studentischer Kompetenzentwicklung in der Lehrerbildung identifizieren zudem die Studienwahlmotive (Pohlmann & Möller, 2010; Rauin, 2011): Während ein genuin pädagogisches Interesse als förderlich für den Kompetenzerwerb im Studium angenommen werden kann, stellen hedonistische Motive – beispielsweise der Wunsch, ein leichtes Studium zu absolvieren – negative Prädiktoren für den Kompetenzerwerb und die spätere berufliche Tätigkeit dar. Diesbezüglich ist davon auszugehen, dass ein pädagogisches Interesse als Studienwahlmotiv auch mit einer höheren adaptiven Handlungskompetenz in heterogenen Gruppen einhergeht.

Weiterhin stellt auch die Studiendauer ein die Kompetenzentwicklung beeinflussendes Moment dar. Studien verweisen hier insbesondere im Sinne eines Angebot-Nutzungs-Modells (Fend, 2008; Helmke, 2009; König, Wagner, Valtin & Schmude, 2011; Lüders et al., 2006) auf die aktive Nutzung sogenannter Lerngelegenheiten (Darge, Schreiber, König & Seifert, 2012; Hascher, 2006; Lüders et al., 2006). Dazu zählen (erziehungswissenschaftliche) Lehrveranstaltungen an der Hochschule einerseits und schulpraktische Studien andererseits. Studierende mit absolviertem Praktikum weisen bezüglich ihres pädagogischen Unterrichtswissens ebenfalls einen signifikanten Wissensvorsprung auf (Hascher, 2006; König, Tachtsoglou & Seifert, 2012). Es kann demnach vermutet werden, dass die Studiendauer mit der gemessenen Kompetenz positiv zusammenhängt.

Vignettentests basieren zu großen Teilen auf verschriftlichten Szenen der Unterrichtspraxis, sind aber auch beim Einsatz von Videovignetten leseintensiv und damit eventuell nicht unabhängig von konstruktfernen Variablen wie Leseleistung und sprachlichem schlussfolgerndem Denken respektive der verbalen Intelligenz. Um abzuschließen, dass der vorliegende Test nicht ausschließlich auf einer minder oder stärker ausgeprägten Fähigkeit des sprachlich schlussfolgernden Denkens beruht, wird dieser Faktor in der Validierungsstudie über eine Subskala

des WIT-2 (Kersting, Althoff & Jäger, 2008) miterhoben und der Zusammenhang diesbezüglich geprüft.

3.2 Forschungsfragen und Hypothesen

Ziel der vorliegenden Studie ist es, ein Testinstrument zur Erfassung einer adaptiven Handlungskompetenz zu entwickeln und zu prüfen. Untersucht werden die psychometrischen Eigenschaften der Skala. Fokussiert werden dabei verschiedene Aspekte der Konstruktvalidität. Ausgangspunkt ist die Frage nach der faktoriellen Validität. Darüber hinaus werden vor dem Hintergrund der dargestellten Forschungsbefunde Hypothesen zur konvergenten sowie zur diskriminanten Validität geprüft.

3.2.1 Faktorielle Validität

Lässt sich die angenommene Struktur einer die drei Subfacetten (1) „diagnostizieren können“, (2) „differenzieren können“ und (3) „Klassenführung beherrschen“ umfassenden *Adaptiven Handlungskompetenz* empirisch replizieren?

3.2.2 Konvergente Validität

Lassen sich signifikante Zusammenhänge zwischen den Werten im Testinstrument zur Erfassung der Adaptiven Handlungskompetenz und konstruktnahen Variablen nachweisen? Auf der Grundlage der dargestellten Forschungsbefunde werden zur Prüfung der konvergenten Validität folgende Hypothesen aufgestellt:

1. Eine positivere Einstellung zu Inklusion geht mit höheren Werten im Kompetenztest zum Umgang mit Heterogenität einher.
2. Die Einschätzung des berufsbezogenen Selbstkonzepts von Studierenden hängt mit der gemessenen Kompetenz positiv zusammen.
3. Der Besuch von gezielten Lehrveranstaltungen zum Thema Inklusion hängt mit dem Kompetenzscore positiv zusammen.
4. Eine stärkere Ausprägung des pädagogischen Interesses als Studienwahlmotiviv geht mit einer höheren adaptiven Handlungskompetenz einher.
5. Die Studiendauer hängt mit der gemessenen Kompetenz positiv zusammen.

3.2.3 Diskriminante Validität

Das Testinstrument zur Erfassung der Adaptiven Handlungskompetenz sollte weitgehend unabhängig sein von konstruktfernen Variablen wie der verbalen Intelligenz. Daher wird der Zusammenhang zwischen dem Kompetenzscore und

der Fähigkeit des sprachlich schlussfolgernden Denkens untersucht. Es wird davon ausgegangen, dass zwischen dem Kompetenzscore und der Fähigkeit des sprachlich schlussfolgernden Denkens lediglich ein sehr geringer positiver Zusammenhang besteht.

4. Methoden

Im Folgenden werden der neu entwickelte Vignettentest sowie die zur Konstruktvalidierung eingesetzten Verfahren beschrieben. Darüber hinaus wird die Stichprobe charakterisiert und abschließend das datenanalytische Vorgehen vorgestellt.

4.1 Instrumente

Zu Beginn des Fragebogens werden soziodemografische Variablen (Alter, Geschlecht, Fachsemester) und Fragen nach dem Besuch von Seminaren zum Themenfeld Inklusion sowie nach bereits gesammelter praktischer Lehrerfahrung erfasst. Daran schließt sich der Vignettentest zur adaptiven Handlungskompetenz an.

4.1.1 Vignettentest zur Erfassung einer adaptiven Handlungskompetenz

Der im Projekt entwickelte *Vignettentest zur Erfassung einer adaptiven Handlungskompetenz* umfasste ursprünglich insgesamt 15 Vignetten, 9 Text- und 6 Videovignetten. Für jede der drei angenommenen Subfacetten wurden 5 Vignetten entwickelt (siehe Tabelle 2).

Je nach Nähe des eigenen Urteils zur Expertennorm (siehe die Ausführungen zur Testentwicklung) erhielt der Proband bzw. die Probandin Punkte: Wählte der Proband, die Probandin exakt den Modus der Expertinnen und Experten, erhielt er zwei Punkte. Divergierte sein Urteil um $|1|$, wurde noch 1 Punkt vergeben. Bei einer größeren Differenz wurden 0 Punkte gewertet. Pro Vignette waren somit maximal 8 Punkte zu erzielen. Die Summe der Punkte ging als Kompetenzscore in alle weiteren Berechnungen ein. Zur Konstruktvalidierung des Vignettentests wurden folgende weitere Variablen erhoben:

Tabelle 2: Überblick über die Zuordnung der Vignetten pro Facette

Facette	Vignetten	Inhaltlicher Fokus	Format	Anzahl der Items	Verbleib im finalen Test
diagnostizieren können	Leistungsrückgang in Mathematik	Maßnahmen zur Ursachenanalyse eines Leistungsabfalls eines Schülers in Mathematik passend auswählen	Text	4	ja
	Lese-Rechtschreibschwäche (LRS)	Maßnahmen auf Basis einer LRS-Diagnose passend ergreifen	Text	4	ja
	Elwir	Maßnahmen auf Basis einer Schülerbeschreibung passend ergreifen	Text	4	ja
	Ahmeds Wortschatz	Maßnahmen auf Basis einer Schülerbeschreibung passend ergreifen	Text	4	ja
	Der Frosch	Schülerinnen in einer Gruppe wertneutral beobachten	Video	4	nein
differenzieren können	Berufseinstieg	Maßnahmen der inneren Differenzierung auf Basis einer beschriebenen Klassensituation ergreifen	Text	4	ja
	Der Arbeitsauftrag	Maßnahmen der inneren Differenzierung auf Basis einer beschriebenen Klassensituation ergreifen	Video	4	ja
	Marius	Maßnahmen der inneren Differenzierung auf Basis einer beschriebenen Klassensituation ergreifen	Video	4	nein
	Max	Maßnahmen der inneren Differenzierung auf Basis einer beschriebenen Klassensituation ergreifen	Text	4	nein
	Jan	Eine methodische Entscheidung kritisch reflektieren	Video	4	nein
Klassenführung beherrschen	Der coole Neue	Auf eine auftretende Unterrichtsstörung passend reagieren	Video	4	ja
	Buchpräsentation	Den Unterrichtsfluss aufrecht erhalten	Text	4	nein
	Kevin und die Zwiebelzelle	Auf eine auftretende Unterrichtsstörung passend reagieren	Video	4	nein
	Mein Traumberuf	Die Klasse im Blick behalten	Text	4	ja
	Timo	Kognitive Aktivierung gewährleisten	Text	4	ja

4.1.2 Einstellungen zu Inklusion

Die Einstellung zu Inklusion wurde mit dem Einstellungsfragebogen zu Inklusion für Lehrkräfte (EFI-L) von Seifried und Heyl (2016) erfasst. Der Fragebogen besteht aus drei Subskalen: Die erste Subskala *Fachliche Förderung im inklusiven Unterricht* ($\alpha_{\text{EKoL}} = .81$; $\alpha_{\text{Manual}} = .90$) umfasst sechs Items, z. B. „Die Qualität des Unterrichts wird besser, wenn Kinder mit besonderen Bedürfnissen dabei sind und miteinbezogen werden“. Die zweite Subskala *Persönliche Bereitschaft zu inklusivem Unterricht* ($\alpha_{\text{EKoL}} = .83$; $\alpha_{\text{Manual}} = .85$) umfasst fünf Items, z. B. „Aufgrund meiner bisherigen Ausbildung fühle ich mich qualifiziert, eine inklusive Klasse zu übernehmen“. Die dritte Subskala kann mit *Soziale Inklusion im Unterricht* ($\alpha_{\text{EKoL}} = .79$; $\alpha_{\text{Manual}} = .81$) überschrieben werden und beinhaltet vier Items, z. B. „Wenn Kinder mit besonderen Bedürfnissen in einer inklusiven Schulklasse sind, dann finden sie dort auch Freundinnen und Freunde“. EFI-L stellt ein Messinstrument dar, das sowohl kognitive Zuschreibungen und Überzeugungen als auch verhaltensbezogene und affektive Dispositionen im Hinblick auf inklusiven Unterricht valide misst: Die Faktoren- und Itemstruktur der schülerbezogenen, kognitiven Komponenten zeigt sich deckungsgleich mit den Ergebnissen von Stanley, Grimbeek, Bryer und Beamish (2003) sowie Kunz, Luder und Moretti (2010), behaviorale und affektive Items der selbstbezogenen Lehrereinstellung werden auf einem prägnanten Faktor vereint.

4.1.3 Berufsbezogene Selbstkonzepte

Das berufliche Selbstkonzept der angehenden Lehrkräfte wurde mit dem Fragebogen zur Erfassung berufsbezogener Selbstkonzepte (ERBSE-L; Retelsdorf, Bauer, Gebauer, Kauper & Möller, 2014, S. 98ff.) erhoben. Der Fragebogen enthält 20 Items, die den sechs Selbstkonzeptdimensionen *Fach* (3 Items, $\alpha_{\text{EKoL}} = .70$; $\alpha_{\text{Manual}} = .76$, z. B. „Ich denke, meine Studienfächer liegen mir besonders“), *Erziehung* (4 Items, $\alpha_{\text{EKoL}} = .61$; $\alpha_{\text{Manual}} = .77$, z. B. „Ich kann gut mit Kindern und Jugendlichen umgehen“), *Diagnostik* (4 Items, $\alpha_{\text{EKoL}} = .65$; $\alpha_{\text{Manual}} = .72$, z. B. „Es fällt mir leicht einzuschätzen, bei welchen Aufgaben andere Probleme haben“), *Beratung* (3 Items, $\alpha_{\text{EKoL}} = .67$; $\alpha_{\text{Manual}} = .79$, z. B. „Ich kann andere gut beraten“), *Innovation* (3 Items, $\alpha_{\text{EKoL}} = .61$; $\alpha_{\text{Manual}} = .72$, z. B. „Es liegt mir, neue Projekte zu planen und umzusetzen“) und *Medien* (3 Items, $\alpha_{\text{EKoL}} = .79$; $\alpha_{\text{Manual}} = .83$, z. B. „Der Umgang mit neuen Medien fällt mir leicht“) zugeordnet werden. Die aus den Faktorenanalysen resultierenden Skalen wurden auf plausible Gruppenunterschiede getestet und Korrelationen zu Außenkriterien aus dem Bereich Studienwahlmotivation und Studienleistung überprüft. Die Ergebnisse sprechen insgesamt für die Validität des Fragebogens zur Erfassung berufsbezogener Selbstkonzepte (ERBSE-L; Retelsdorf et al., 2014).

4.1.4 Studienwahlmotive

Zur Erhebung von sogenannten Studienwahlmotiven wurde der Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA) von Pohlmann und Möller (2010) eingesetzt. Der Test besteht aus 33 Items, die sich auf folgende sechs Subskalen verteilen: *Nützlichkeit für die Balance zwischen Beruf und Familie* ($\alpha_{\text{EKoL}} = .89$; $\alpha_{\text{Manual}} = .90$, 8 Items, z. B. „Ich habe das Lehramtsstudium gewählt, weil ich auch neben dem Beruf noch Zeit für Familie, Freunde und Hobbies haben will“), *Pädagogisches Interesse* ($\alpha_{\text{EKoL}} = .80$; $\alpha_{\text{Manual}} = .83$, 6 Items, z. B. „Ich habe das Lehramtsstudium gewählt, weil ich einen Beruf ausüben möchte, in dem ich die Erziehung von Jugendlichen mitgestalten kann“), *Fähigkeitsüberzeugung* ($\alpha_{\text{EKoL}} = .70$; $\alpha_{\text{Manual}} = .74$, 5 Items, z. B. „Ich habe das Lehramtsstudium gewählt, weil ich denke, dass ich eine gute Lehrerin/ein guter Lehrer sein werde“), *Soziale Einflüsse* ($\alpha_{\text{EKoL}} = .75$; $\alpha_{\text{Manual}} = .84$, 5 Items, z. B. „Ich habe das Lehramtsstudium gewählt, weil mir von Freunden und Bekannten zum Lehramtsstudium geraten wurde“), *Erwartete geringe Schwierigkeit des Lehramtsstudiums* ($\alpha_{\text{EKoL}} = .73$; $\alpha_{\text{Manual}} = .88$, 4 Items, z. B. „Ich habe das Lehramtsstudium gewählt, weil das Studium nicht so anstrengend ist“) sowie *Fachliches Interesse* ($\alpha_{\text{EKoL}} = .81$; $\alpha_{\text{Manual}} = .84$, 5 Items, z. B. „Ich habe das Lehramtsstudium gewählt, weil ich mich gern mit den Inhalten meiner Fächer beschäftige“). Die Konstruktvalidität des Instruments konnte in insgesamt vier Teilstudien mittels Faktorenanalysen sichergestellt und über Zusammenhangsanalysen mit konstruktnahen Variablen zusätzlich gestützt werden.

4.1.5 Sprachliches und schlussfolgerndes Denken

Zur Erfassung des sprachlichen und schlussfolgernden Denkens wurde die Subskala *Analogien* aus dem Wilde-Intelligenztest (WIT-2) eingesetzt (Kersting et al., 2008). Sie enthält insgesamt 20 Items ($\alpha_{\text{EKoL}} = .74$; $\alpha_{\text{Manual}} = .87$). Hierbei sind zwei Wörter auf der linken Seite eines Gleichheitszeichens angegeben, die in einer klaren Beziehung zueinander stehen. Auf der rechten Seite des Gleichheitszeichens steht ein Wort. Aus fünf Vorschlägen ist das Wort auszuwählen, das eine analoge Beziehung herstellt (Kersting et al., 2008).

4.2 Stichprobe

Das Instrument wurde im Sommersemester 2015 in der ersten Phase der Lehrerbildung erstmals breitflächig eingesetzt. Es konnten insgesamt 706 Lehramtsstudierende an vier Pädagogischen Hochschulen in Baden-Württemberg für die Teilnahme gewonnen werden (siehe Tabelle 3). Pädagogische Hochschulen existieren in Deutschland nur im Bundesland Baden-Württemberg und sind bil-

dungswissenschaftliche Universitäten mit allen akademischen Rechten. An Pädagogischen Hochschulen werden die Lehrämter Grundschule, Sekundarstufe I und Sonderpädagogik studiert, nicht jedoch das gymnasiale Lehramt.

Tabelle 3: Deskriptive Statistiken der Gesamt- und Teilstichproben

<i>N</i>	706	
Alter <i>M</i> (<i>SD</i>)	22.71 (2.77)	
Geschlecht	Absolut	Prozent
Weiblich	570	80.7 %
Männlich	135	19.1 %
Studiengänge	Absolut	Prozent
Lehramt an Grundschulen	303	43.0 %
Lehramt an Haupt-, Werkreal- und Realschulen	279	39.6 %
Lehramt an Sonderschulen	89	12.6 %
Europalehramt	34	4.8 %
Studiendauer <i>M</i> (<i>SD</i>)	4.82 (2.48)	
Fachsemester	Absolut	Prozent
1. Fachsemester	88	12.6 %
2. Fachsemester	107	15.3 %
3. Fachsemester	19	2.7 %
4. Fachsemester	105	14.9 %
5. Fachsemester	25	3.5 %
6. Fachsemester	173	24.5 %
7. Fachsemester	69	9.8 %
8. Fachsemester	99	14.0 %
9. Fachsemester und mehr	16	2.1 %

Die Studierenden verteilen sich auf das Lehramt an Grundschulen (43.0 %), Lehramt an Haupt-, Werkreal- und Realschulen (39.5 %) und das Lehramt an Sonderschulen (12.6 %) sowie das Europalehramt (4.8 %). Diese Verteilung ist repräsentativ für die Gesamtverteilung von Lehramtsstudierenden an den Pädagogischen Hochschulen in Baden-Württemberg. Auch die hohe Zahl von Probandinnen (80.7 %) entspricht der Quote der Grundgesamtheit. Die Studiendauer (in Anzahl studierter Fachsemester) liegt im Mittel bei 4.8 Semestern, die Verteilung der Fachsemester zeigt eine Ausgewogenheit bezüglich der Zugehörigkeit zu an den Hochschulen implementierten Studienstufen (Studienstufe 1: Semester 1 und 2; Studienstufe 2: Semester 3, 4 und 5; Studienstufe 3: Semester 6, 7 und 8 bzw. 9 für das Lehramt Sonderpädagogik).

4.3 Datenanalyse

Die Grundlage der Konstruktvalidierung des Vignettestes zum kompetenten Umgang mit Heterogenität bildete zunächst eine skalenanalytische Evaluation (Adams & Wu, 2002; Bühner, 2004). Die psychometrische Prüfung der Skalen wurde mit SPSS Statistics 22 und Amos 22 (IBM) durchgeführt. Vorab wurde festgelegt, Fälle mit einem Anteil fehlender Werte (Missings) von über 50 % aus den Analysen auszuschließen, um keine gravierenden Verzerrungen und Verfälschungen zu evozieren (Graham, 2012; Schafer & Graham, 2002; Wirtz, 2004). Dieses Kriterium traf jedoch auf keinen Datensatz zu. Einzelne fehlende Werte im Kompetenztest und in der verwendeten Subskala des Wilde-Intelligenz-Tests wurden als „nicht gelöst“ gewertet und gingen so in die Analysen ein, was ein übliches Verfahren in der Testkonstruktion darstellt (König, Rothland, Darge, Lünemann & Tachtsoglou, 2013). In den anderen eingesetzten Messverfahren wurden einzelne fehlende Werte mit dem Expectation-Maximization-Algorithmus (Dempster, Laird & Rubin, 1977) geschätzt.

In einem ersten Schritt wurde dann das theoretisch postulierte Arbeitsmodell als g-Faktormodell mit den 15 Vignetten des Tests in einer konfirmatorischen Faktorenanalyse (CFA) modelliert. Anschließend wurden alle Vignetten mit einer Faktorladung von $< .3$ aus dem Modell entfernt und ein modifiziertes g-Faktormodell mit neun Vignetten spezifiziert. Diesem wurde ein entsprechendes Dreifaktorenmodell gegenübergestellt, welches lediglich Kovarianzen zwischen den Faktoren zulässt (vgl. Tabelle 4). Alle Modelle beinhalten die Subfacetten (1) *diagnostizieren können*, (2) *differenzieren können* und (3) *Klassenführung beherrschen*. Die Modellgüte wird zunächst klassisch auf Grundlage der Signifikanz des χ^2 -Tests beurteilt. Da der χ^2 -Test bei großen Stichproben leicht signifikant wird, berücksichtigen wir folgende zusätzliche Fit-Indizes: den normierten Chi-Quadrat-Wert, den Root Mean Square Error of Approximation (RMSEA), den Tucker-Lewis-Index (TLI) und den Comparative-Fit-Index (CFI). Zusätzlich zu den oben angeführten Fit-Indizes werden die informationstheoretischen Maße Akaike Information Criterion (AIC) und Bayesian Information Criterion (BIC) berichtet. Der Vergleich aller Fit-Indizes miteinander soll schließlich das angemessenste Modell herausstellen. Nach dieser skalenanalytischen Evaluation werden die bivariaten Zusammenhänge der adaptiven Handlungskompetenz in heterogenen Gruppen (Kompetenzscore) mit den eingangs beschriebenen Kovariaten dargestellt.

5. Ergebnisse

5.1 Psychometrische Prüfung des Kompetenztests

Die konfirmatorische Faktorenanalyse mit Amos 22 (IBM) zeigte eine akzeptable Passung der Daten zum g-Faktormodell.

Tabelle 4: Übersicht der geprüften Modelle sowie der Modellfits der konfirmatorischen Faktorenanalysen ($N = 706$)

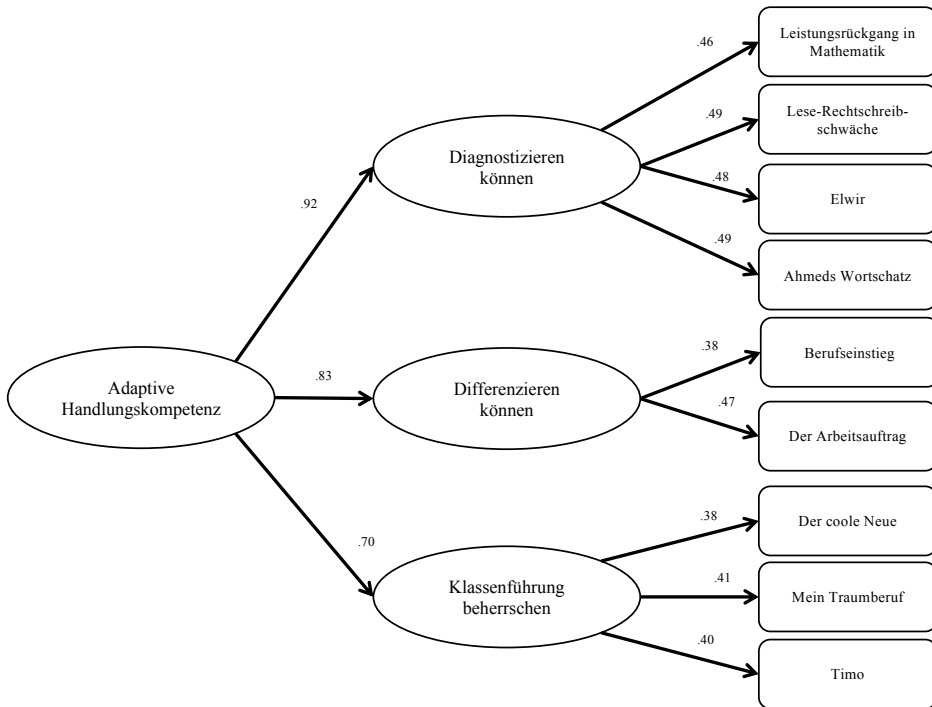
	g-Faktormodell	Modifiziertes g-Faktormodell	Dreifaktorenmodell	Vergleichsgrößen ^a
<i>Anzahl Vignetten</i>	15	9	9	
davon				
Textvignetten	9	7	7	
Videovignetten	6	2	2	
<i>Modellfit</i>				
<i>df</i>	51	25	24	
χ^2/df	2.263	1.417	1.385	< 3 gut < 2 sehr gut
RMSEA	.039	.024	.023	≤ .05
CFI	.821	.974	.977	≥ .97
TLI	.755	.963	.966	≥ .97
AIC ^b	293.164	75.430	75.231	
BIC ^b	414.261	166.622	170.983	

^a Schermelleh-Engel und Kollegen (2003). ^b Hier zeigen niedrigere Werte eine bessere Passung des Modells an die Daten an.

Durch die beschriebene Modifizierung, in Zuge derer Vignetten mit einer Ladung von $< .3$ auf den Faktor entfernt wurden (König, Buchholz & Dohmen, 2015; Meschede, 2014), konnte der Modell-Fit noch optimiert werden (vgl. Tabelle 4).

Durch die Modifikation ergab sich eine Reduktion um eine Vignette für den Faktor *diagnostizieren können*, eine Reduktion um 3 Vignetten für den Faktor *differenzieren können* und eine Reduktion um 2 Vignetten für den Faktor *Klassenführung beherrschen*.

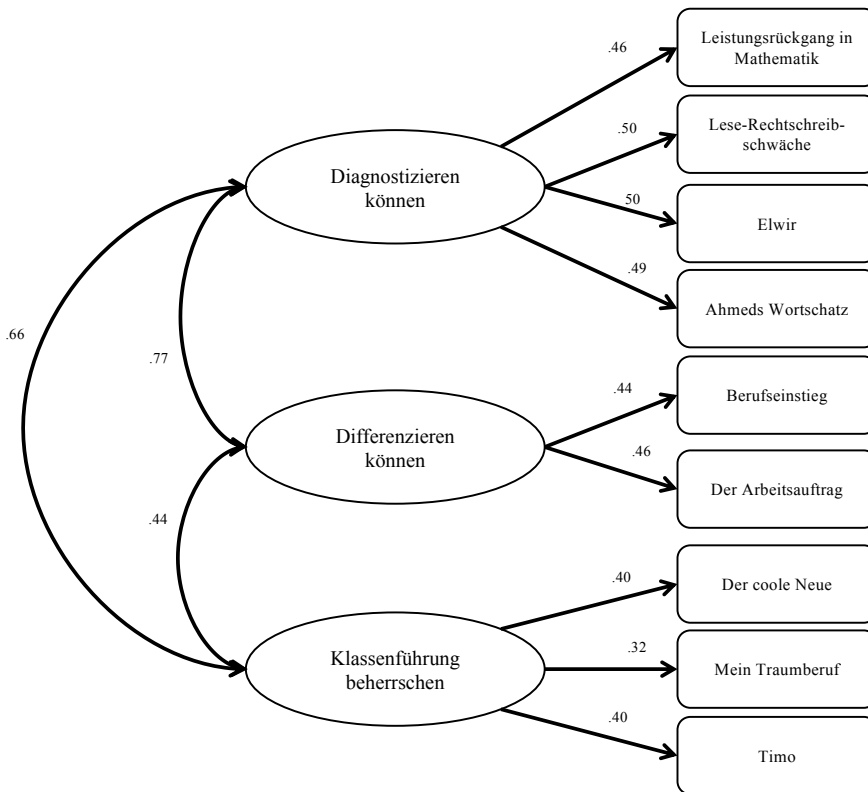
Abbildung 1: Modifiziertes g-Faktormodell: Ausschluss aller Vignetten mit Faktorladung < 3



Anmerkungen. $df = 25$, $\chi^2/df = 1.417$, RMSEA = .024, CFI = .974, TLI = .963, AIC = 75.430, BIC = 166.622. Aus Gründen der Übersichtlichkeit wurde auf die Darstellung von Fehlertermen verzichtet.

Das modifizierte g-Faktormodell (siehe Abbildung 1) weist einen hervorragenden Modellfit auf. Der normierte χ^2 -Wert liegt mit 1.417 im sehr guten Bereich (Weiber & Mühlhaus, 2014), der RMSEA ist mit .024 ebenfalls als hervorragend zu bezeichnen, dasselbe gilt für den TLI von .963 sowie den CFI von .974 (Hair, Black, Babin & Anderson, 2010; Schermelleh-Engel, Moosbrugger & Müller, 2003). Im Hinblick auf die Informationskriterien BIC und AIC zeigt sich darüber hinaus eine Verbesserung im Vergleich zum ursprünglichen g-Faktormodell. Cronbachs Alpha kann mit einem Wert von .613 für einen Kompetenztest als grenzwertig zufriedenstellend angesehen werden (König, Blömeke & Kaiser, 2015; Meschede, 2014). Darüber hinaus wurde ein Dreifaktorenmodell (siehe Abbildung 2) geprüft, welches ebenfalls die nach der Modifizierung verbleibenden neun Vignetten enthält und lediglich Kovarianzen zwischen den Subfacetten annimmt. Dieses weist eine dem modifizierten g-Faktormodell vergleichbare Passung auf.

Abbildung 2: Dreifaktorenmodell ohne g-Faktor



Anmerkungen. $df = 24$, $\chi^2/df = 1.385$, RMSEA = .023, CFI = .977, TLI = .966, AIC = 75.231, BIC = 170.983. Aus Gründen der Übersichtlichkeit wurde auf die Darstellung von Fehlertermen verzichtet.

Aufgrund der Gleichwertigkeit der Modellfits sowie der größeren Sparsamkeit des modifizierten g-Faktormodells wird dieses als endgültiges Modell angenommen. In beiden Modellen zeigt sich, dass die Dimensionen (1) *diagnostizieren können*, (2) *differenzieren können* und (3) *Klassenführung beherrschen* positiv korrelieren, wobei die Korrelation zwischen der Dimension (1) *diagnostizieren können* und (2) *differenzieren können* höher ausgeprägt ist als zwischen den beiden benannten und der Dimension (3) *Klassenführung beherrschen*.

Für die im Folgenden reportierten Zusammenhänge zwischen gemessener Kompetenz und ausgewählten Variablen wird daher der auf dem modifizierten g-Faktormodell basierende Kompetenzscore verwendet. Dieser setzt sich aus der Summe der neun Vignettenscores zusammen. Im Mittel erreichen die 706 Probandinnen und Probanden 40.22 von 72 möglichen Punkten ($r = .27$). Im Minimum wurden 18, im Maximum 62 Punkte erreicht. Die Daten weisen eine leichte Linksschiefe von $-.263$ ($r = .09$) und eine positive Kurtosis von $.119$ ($r = .18$) auf. Für die im Folgenden beschriebenen Zusammenhänge mit Außenkriterien

können demnach Decken- oder Bodeneffekte als Varianzeinschränkungen ausgeschlossen werden.

5.2 Zusammenhänge mit Außenkriterien

Konstruktvalidität setzt voraus, „dass das von einem Messinstrument erfasste Konstrukt mit möglichst vielen anderen Variablen in theoretisch begründbaren Zusammenhängen steht“ (Diekmann, 2010, S. 224). Aus diesem Grund wurden für eine erste Orientierung die Zusammenhänge zwischen Testergebnis (Kompetenzscore) und den für die Konstruktvalidierung als bedeutsam angenommenen Variablen betrachtet.

Tabelle 5: Übersicht über die verwendeten Kovariaten

Skala	N	Itemanzahl	Mittelwert	Standardabweichung
<i>EFI-L (Seifried & Heyl, 2016)</i>	690	15	57.31	10.66
<i>ERBSE-L (Retelsdorf et al., 2014)</i>				
Fach	706	3	3.86	0.61
Erziehung	706	4	4.24	0.46
Diagnostik	706	4	3.92	0.49
Beratung	706	3	3.66	0.59
Innovation	706	3	3.94	0.70
Medien	706	3	4.07	0.53
<i>FEMOLA (Pohlmann & Möller, 2010)</i>				
Nützlichkeit für die Balance zwischen Beruf und Familie	706	8	2.93	0.87
Pädagogisches Interesse	706	6	4.56	0.45
Fähigkeitsüberzeugung	706	5	4.01	0.06
Soziale Einflüsse	706	5	2.55	0.96
Fachliches Interesse	706	5	1.79	0.80
<i>WIT-2 (Kersting et al., 2008)</i>				
Analogien	699	20	12.36	3.84

Tabelle 6: Bivariate Korrelationen zur Konstruktvalidierung ($N = 706$)

Kompetenzscore	
Studiendauer	.24**
<i>FEMOLA</i>	
Balance zwischen Familie und Beruf	-.17**
Pädagogisches Interesse	.19**
Fähigkeitsüberzeugung	.09*
Soziale Einflüsse	-.14**
Geringe Schwierigkeit des Lehramtsstudiums	-.23**
Fachliches Interesse	.03
<i>ERBSE-L</i>	
Selbstkonzeptdimension Fach	.07
Selbstkonzeptdimension Erziehung	.17**
Selbstkonzeptdimension Diagnostik	.11**
Selbstkonzeptdimension Beratung	.10**
Selbstkonzeptdimension Innovation	.05
Selbstkonzeptdimension Medien	-.02
Einstellung zu Inklusion	.10**
Sprachliches und schlussfolgerndes Denken	.03

* $p \leq .05$, ** $p \leq .01$.

Wie erwartet erzielen Studierende mit einer positiveren Einstellung zu Inklusion auch höhere Werte im Kompetenztest zum Umgang mit Heterogenität, die Korrelation zur Einstellung zu Inklusion ist bei kleiner Effektstärke signifikant. Auch die Einschätzung des berufsbezogenen Selbstkonzepts von Studierenden hängt hypothesenkonform mit der gemessenen Kompetenz positiv zusammen: Die Korrelationen zwischen Kompetenzscore und den Selbstkonzeptdimensionen Diagnostik, Beratung und Erziehung sind bei geringen Effektstärken statistisch signifikant. Wie erwartet hängt auch der Besuch von gezielten Lehrveranstaltungen zum Thema Inklusion mit dem Kompetenzscore positiv zusammen.

Geringe Zusammenhänge werden bezüglich der Studienwahlmotive wie dem Pädagogischen Interesse oder der erwarteten geringen Schwierigkeit des Lehramtsstudiums deutlich, dennoch ergeben die Daten Hinweise darauf, dass ein pädagogisches Interesse als Studienwahlmotiv mit einer höheren adaptiven Handlungskompetenz einhergeht. Auch der Zusammenhang mit der Studiendauer stellt sich erwartungskonform positiv dar: Studierende in oberen Fachsemestern erreichen höhere Werte im Test im Vergleich zu Studienanfängerinnen und -anfängern. Es zeigt sich kein signifikanter Zusammenhang zwischen der Kompetenz im Umgang mit Heterogenität und dem sprachlichen und schlussfolgernden Denken.

6. Zusammenfassung und Diskussion

6.1 Zusammenfassung der vorliegenden Studie

Ziel dieses Beitrags war es, einen neu entwickelten Vignettest zur Messung des kompetenten Umgangs von Lehrkräften mit Heterogenität empirisch zu prüfen. Die angenommene Adaptive Handlungskompetenz baut auf einem theoretisch hergeleiteten Arbeitsmodell auf (Beck et al., 2008; Franz, Wacker & Heyl, 2016). Zur Überprüfung des Modells wurden Daten aus einer querschnittlichen Untersuchung von Lehramtsstudierenden in unterschiedlichen Fachsemestern herangezogen ($n = 706$). Die gewonnenen Daten wurden hierzu einerseits im Hinblick auf das postulierte Arbeitsmodell konfirmatorisch überprüft und bezüglich ihrer Güte evaluiert, andererseits wurden die Zusammenhänge zwischen dem Testergebnis und ausgewählten Variablen analysiert. Insgesamt liegt die Stichprobengröße mit über 700 Personen in einem Bereich, für den „verzerrte Ergebnisse“ bei Faktorenanalysen weitgehend ausgeschlossen werden können (Eid, Gollwitzer & Schmitt, 2010, S. 918).

Es bleibt festzuhalten, dass das angenommene Modell auf der Grundlage der erhobenen Querschnittsdaten empirisch bestätigt werden kann. Die Datenstruktur ergibt eine übergeordnete Adaptive Handlungskompetenz, welche durch die Faktoren erster Ordnung (1) *diagnostizieren können*, (2) *differenzieren können* und (3) *Klassenführung beherrschen* konstituiert wird. Die Modellierung weist hervorragende Gütekriterien auf. Der Zusammenhang zwischen den Dimensionen (1) und (2) ist augenscheinlich höher als der Zusammenhang zwischen der Dimension (1) bzw. (2) und der Dimension (3). Auch bei Brühwiler (2017) zeigte sich der Zusammenhang zwischen diagnostischer und didaktischer Handlungskompetenz als am stärksten ausgeprägt. Die Bestätigung der postulierten Faktorenstruktur sowie die nachfolgend diskutierten Ergebnisse zur konvergenten und diskriminanten Validität zeigen, dass es möglich ist, die angenommene Adaptive Handlungskompetenz valide zu erfassen.

6.2 Adaptive Handlungskompetenz im Kontext potenzieller Zusammenhangsvariablen

Die Ergebnisse der Korrelationsanalysen zur konvergenten Konstruktvalidierung zeigen erwartungskonform, dass die Adaptive Handlungskompetenz mit der Studiendauer positiv korreliert, was im Sinne einer effektiven Nutzung der angebotenen Lerngelegenheiten auf eine Kompetenzentwicklung über den Studienverlauf hinweg verweisen könnte. Dies stützt die Annahme, dass die adaptive Handlungskompetenz grundsätzlich lehr- und lernbar ist (Beck et al., 2008) und dass im Rahmen der Lehramtsausbildung an Baden-Württembergischen Hochschulen diesbezüglich Lerngelegenheiten vorhanden sind sowie von Studierenden

entsprechend genutzt werden. Als solche förderlichen Lerngelegenheiten können Lehrveranstaltungen ebenso angenommen werden wie die Teilnahme an schulpraktischen Studien (Darge et al., 2012; Hascher, 2006; König et al., 2012).

Darüber hinaus scheinen ebenfalls theoriekonform personenbezogene Merkmale für die Kompetenz im Umgang mit Heterogenität von Bedeutung zu sein. Die gefundenen signifikanten Zusammenhänge mit ausgewählten Studienwahlmotiven stellen lediglich kleine Effekte dar, entsprechen in ihrer Richtung aber den theoretischen Annahmen: Während das Pädagogische Interesse als Motiv für die Wahl des Studiums positiv mit der Kompetenz im Umgang mit Heterogenität zusammenhängt, korreliert die Motivation, mit dem Lehramtsstudium ein leichtes Studium zu absolvieren, signifikant negativ mit dem Kompetenzscore. Auch das Motiv, dass Lehrkräfte Familie und Beruf leicht in Balance bringen können, sowie eine von sozialen Einflüssen wie der Meinung der Eltern geprägte Motivation, ein Lehramtsstudium zu beginnen, hängen mit der Adaptiven Handlungskompetenz negativ zusammen.

Der Zusammenhang zwischen der Adaptiven Handlungskompetenz und der Einstellung zu Inklusion zeigt sich deutlich geringer als angenommen: Theoretisch lässt sich ableiten, dass Lehrkräfte, die eine positivere Einstellung zu inklusiven Bildungsprozessen haben, diese auch kompetenter umsetzen. Diese Annahme lässt sich wenngleich mit geringer Effektstärke mit den vorliegenden Befunden auch empirisch untermauern. Die deskriptive Analyse der Daten zeigt jedoch, dass die befragten Studierenden bezüglich ihrer Einstellung zu Inklusion eine starke Tendenz zur Mitte der Skala aufweisen, was sich einerseits als ein Unentschieden hinsichtlich der eigenen Haltung deuten ließe, andererseits könnte sich darin auch eine Inkompatibilität der Skala mit der Situation von Studierenden verbergen. Im Gegensatz zu Lehrkräften der dritten Phase sind Studierende noch nicht mit praktischen Fragen von Inklusion konfrontiert und verfügen deshalb vielleicht noch nicht über Inszenierungsmuster, auf die die zur Hypothesenbildung genutzten Studien von Heyl und Kollegen (2014) bzw. Miller (2013) rekurrieren.

Der zur diskriminanten Validierung herangezogene Zusammenhang zwischen der Fähigkeit sprachlogisch zu denken einerseits und der Performanz im Vignettentest andererseits weist keine statistische Signifikanz auf. Dies kann als Hinweis interpretiert werden, dass die Bearbeitung des Tests nicht abhängig ist von einer minder oder stärker ausgeprägten Fähigkeit des sprachlich schlussfolgernden Denkens, sondern es anderer Kompetenzen bedarf, die Vignetten erfolgreich zu bearbeiten.

6.3 Einschränkungen der Untersuchung und Forschungsausblick

Als eine erste Limitation der vorliegenden Untersuchung ist auszuweisen, dass zum jetzigen Zeitpunkt noch keine Längsschnittdaten vorliegen, die es ermöglichen, über die vorliegenden Querschnittdaten und ihre Befunde hinaus auch

Veränderungshypothesen und kausale Zusammenhänge zu prüfen. So kann beispielsweise die Korrelation mit der Studiendauer nur als erster Hinweis interpretiert werden, der einer längsschnittlichen Überprüfung bedarf. In der zweiten Projektphase wird im Rahmen einer längsschnittlichen Erhebung im Referendariat (Vorbereitungsdienst) auch ein Nachweis der prognostischen Validität des Verfahrens angestrebt, indem zusätzlich zu einer Prä- und Post-Testung der Lehramtsanwärterinnen und -anwärter die Einschätzung der Unterrichtsqualität durch die Schülerinnen und Schüler erhoben wird.

Darüber hinaus erweisen sich die Ergebnisse der Korrelationsanalysen zur konstruktvalidierenden Überprüfung der Zusammenhänge der Testscores mit theoretisch bedeutsamen Variablen insgesamt affin zu den getroffenen theoretischen Annahmen. Aufgrund der geringen Effektstärken ($r < |.3|$) sind hier dennoch vorsichtige Interpretationen angebracht, die in weiteren Erhebungen der Überprüfung bedürfen.

Ferner kann aus den vorliegenden Befunden nicht eindeutig geschlussfolgert werden, dass die Vignetten des vorliegenden Kompetenztests ausschließlich adaptive Handlungskompetenz erfassen. Diagnostische und didaktische Kompetenz als Kern adaptiver Handlungskompetenz sind ebenso auch im Bereich der Planungskompetenz konstituierend. In wie weit die Vignetten tatsächlich einfordern, Maßnahmen in der Interaktion mit Schülerinnen und Schülern zu ergreifen, kann nicht sicher geklärt werden. Eventuell schwingen hier auch der interaktiven Phase des Unterrichts vorgelagerte Handlungsentscheide mit.

Im Gesamtfazit bleibt festzuhalten, dass der vorliegende Vignettentest die zunächst theoretisch konzeptualisierte Kompetenz zum Umgang mit Heterogenität auch empirisch zu replizieren vermag. Darüber hinaus können aus den Befunden erste Hypothesen über mögliche bedingende institutionelle und personale Einflussfaktoren spezifiziert werden. Folgt man der Annahme, dass es adaptiv kompetenteren Lehrerinnen und Lehrern besser gelingt, einen individuellen Leistungszuwachs bei Schülerinnen und Schülern zu fördern als weniger kompetenten Lehrpersonen (Beck et al., 2008), so trägt die Untersuchung der Kompetenzgenese und -entwicklung insbesondere über die Identifikation positiv wirkender Einflussfaktoren in der universitären Lehrerbildung zur Verbesserung der Unterrichtsqualität an Schulen bei.

Forschungsmethodisch stellen kontextualisierte Messverfahren wie der in EKoL entwickelte Vignettentest zur adaptiven Handlungskompetenz einen viel diskutierten Weg zur „Erhellung des Theorie-Praxis-Problems in der Lehrerbildung“ dar (Neuweg, 2015). Adaptive Lehrkompetenz bedeutet das Vorhandensein kognitiver Prozesse zum situationsadäquaten Handeln. Auf einer präaktionalen Ebene der Planungskompetenz werden auf Basis diagnostischer Überlegungen Unterrichtsverläufe antizipiert und geeignete Maßnahmen ausgewählt. Während der interaktionalen Phase des Unterrichts werden die gewählten didaktischen Maßnahmen permanent mit diagnostischen Informationen rückgekoppelt. Die im Unterricht gewonnenen diagnostischen Erkenntnisse bilden die Basis für die Planung weiterer Unterrichtseinheiten. Will man Kompetenzen erfassen,

so bedarf es einer Operationalisierung in Wissenskomponenten. Nicht nur die Testentwicklung stellt für die meisten Domänen Neuland dar, auch die Definitionen der dahinterliegenden Kompetenzprofile sind vielfach als Forschungsdesiderat anzusehen (Neuweg, 2015). Auch wenn nicht abschließend geklärt werden kann, ob die erfasste Kompetenz tatsächlich ausschließlich die interaktive Handlungsseite abbildet oder ob sich Planungs- und Handlungskompetenz in den Vignetten vermischen, kann das vorliegende Projekt über die entwickelten Vignetten einen Beitrag zur Operationalisierung adaptiver Lehrkompetenz leisten. Es wird ein Testinstrument für weitere Forschungsvorhaben bereitgestellt, das mit geschlossenen Antwortalternativen operiert und das angenommene Arbeitsmodell valide abbilden kann. Insbesondere im Ausblick auf die aktuell mit dem Instrument durchgeführte längsschnittliche Erhebung im Vorbereitungsdienst (Referendariat) lassen sich interessante Impulse für die Lehrerbildung erhoffen.

Literatur

- Adams, R. J. & Wu, M. L. (2002). *PISA 2000 Technical Report*. Paris: OECD.
- Arnold, K.-H. (2008). Chancengleichheit herstellen – neue Aufgaben für die Forschung, neue Perspektiven für die Praxis. In J. Ramseger & M. Wagener (Hrsg.), *Chancenungleichheit in der Grundschule. Ursachen und Wege aus der Krise* (S. 65–74). Wiesbaden: VS.
- Aufschnaiter, C. v., Cappell, J., Dübbelde, G., Ennemoser, M., Mayer, J., Stiensmeier-Pelster, J., Sträßer, R. & Wolgast, A. (2015). Diagnostische Kompetenz. Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Zeitschrift für Pädagogik*, 61(5), 738–758.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schielefe, U., Schneider, W., Stanat, P., Tillmann, K.-J & Weiß, M. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Beck, E., Baer, M., Guldiman, T., Bischoff, S., Brühwiler, C., Müller, P., Niedermann, R., Rogalla, M. & Vogt, F. (Hrsg.). (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.
- Blömeke, S., König, J., Busse, A., Suhl, U., Benthien, J., Döhrmann, M. & Kaiser, G. (2014). Von der Lehrerbildung in den Beruf – Fachbezogenes Wissen als Voraussetzung für Wahrnehmung, Interpretation und Handeln im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 17(3), 509–542.
- Blömeke, S., König, J., Suhl, U., Hoth, J. & Döhrmann, M. (2015). Wie situationsbezogen ist die Kompetenz von Lehrkräften? Zur Generalisierbarkeit der Ergebnisse von videobasierten Performanztests. *Zeitschrift für Pädagogik*, 61(3), 310–327.
- Bohl, T., Budde, J. & Rieger-Ladich, M. (Hrsg.). (2017). *Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen*. Bad Heilbrunn: Klinkhardt.
- Borko, H. (2004). Professional development and teacher learning. Mapping the terrain. *Educational Researcher*, 33(8), 3–15.
- Brodhacker, S. (2014). *Unterrichtsplanningkompetenz im Praktikum. Einflussfaktoren auf die Veränderung der wahrgenommenen Kompetenz von Studierenden*. Münster: Waxmann.

- Brovelli, D., Bölsterli, K., Rehm, M. & Wilhelm, M. (2013). Erfassen professioneller Kompetenzen für den naturwissenschaftlichen Unterricht – ein Vignettestest mit authentisch komplexen Unterrichtssituationen und offenem Antwortformat. *Unterrichtswissenschaft*, 41(4), 306–329.
- Brovelli, D., Bölsterli, K., Rehm, M. & Wilhelm, M. (2014). Using vignette testing to measure student science teachers' professional competencies. *American Journal of Educational Research*, 2(7), 555–558.
- Brühwiler, C. (2014). *Adaptive Lehrkompetenz und schulisches Lernen. Effekte, handlungssteuernde Kognitionen von Lehrpersonen auf Unterrichtsprozesse und Lernergebnisse der Schülerinnen und Schüler*. Münster: Waxmann.
- Brühwiler, C. (2017). Diagnostische und didaktische Kompetenz als Kern adaptiver Lehrkompetenz. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften. Theoretische und methodische Weiterentwicklungen* (S. 123–134). Münster: Waxmann.
- Budde, J. (2017). Heterogenität: Entstehung, Begriff, Abgrenzung. In T. Bohl, J. Budde & M. Rieger-Ladich (Hrsg.), *Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen* (S. 13–26). Bad Heilbrunn: Klinkhardt.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Clark, C. M. & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (3. Aufl., S. 255–296). New York, NY: Macmillan.
- Darge, K., Schreiber, M., König, J. & Seifert, A. (2012). Lerngelegenheiten im erziehungswissenschaftlichen Studium. In J. König & A. Seifert (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung* (S. 87–118). Münster: Waxmann.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1), 1–38.
- Diekmann, A. (2010). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen* (21. Aufl.). Reinbek bei Hamburg: Rowohlt.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden*. Weinheim: Beltz.
- Einsiedler, W. & Hardy, I. (2010). Kognitive Strukturierung im Unterricht. Einführung und Begriffsklärungen. *Unterrichtswissenschaft*, 38(3), 194–209.
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS.
- Franz, E.-K. (2014). Entwicklungsaufgaben der Lehrerprofessionalisierung im Kontext von Inklusion (ProfI). In S. Trumpa, S. Seifried, E.-K. Franz & T. Klauß (Hrsg.), *Inklusive Bildung. Erkenntnisse und Konzepte aus Fachdidaktik und Sonderpädagogik* (S. 122–136). Weinheim: Beltz.
- Franz, E.-K. & Lange, B. (2014). Lernaufgaben im Rahmen inklusiver Didaktik. In E.-K. Franz, S. Trumpa & I. Esslinger-Hinz (Hrsg.), *Inklusion: Eine Herausforderung für die Grundschulpädagogik* (S. 97–107). Baltmannsweiler: Schneider.
- Franz, E.-K., Wacker, A. & Heyl, V. (2016). Lehrerprofessionalität im Spannungsfeld inklusiver Bildung: Theoretische und empirische Modellierung eines erweiterten Kompetenzmodells als Grundlage für didaktisches Handeln von Lehrkräften. In G.-B. v. Carlsburg (Hrsg.), *Strategien der Lehrerbildung. Zur Steigerung von Lehrkompetenzen und Unterrichtsqualität* (S. 307–320). Frankfurt a. M.: Peter Lang.
- Franz, E.-K., Wacker, A. & Heyl, V. (2018). Entwicklung von Testitems zur Erfassung Pädagogisch-psychologischer Handlungskompetenz. In J. Rutsch, M. Rehm, M.

- Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 47–73). Wiesbaden: VS.
- Gibbons, P. (2015). *Scaffolding language, scaffolding learning. Teaching English language learners in the mainstream classroom* (2. Aufl.). Portsmouth, NH: Heinemann.
- Graham, J. W. (2012). *Missing data*. New York, NY: VS.
- Haag, L. & Lohrmann, K. (2007). Diagnostische (In-)Kompetenz von Lehrern. In B. Fuchs & C. Schönherr (Hrsg.), *Urteilkraft und Pädagogik. Beiträge zu einer pädagogischen Handlungstheorie. Lutz Koch zum 65. Geburtstag* (S. 239–249). Würzburg: Königshausen & Neumann.
- Häcker, T. (2017). Individualisierter Unterricht. In T. Bohl, J. Budde & M. Rieger-Ladich (Hrsg.), *Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen* (S. 275–290). Bad Heilbrunn: Klinkhardt.
- Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R. E. (2010). *Multivariate data analysis* (7. Aufl.). Upper Saddle River, NJ: Prentice Hall.
- Hardy, I., Hertel, S., Kunter, M., Klieme, E., Warwas, J., Büttner, G. & Lühken, A. (2011). Adaptive Lerngelegenheiten in der Grundschule. Merkmale, methodisch-didaktische Schwerpunktsetzungen und erforderliche Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 57(6), 819–833.
- Hascher, T. (2006). Veränderungen im Praktikum – Veränderungen durch das Praktikum: Eine empirische Untersuchung zur Wirkung von schulpraktischen Studien in der Lehrerbildung. In C. Allemann-Ghionda (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern. Ausbildung und Beruf* (S. 130–148). Weinheim: Beltz.
- Hellmann, K. A. (2015). *Diagnostische Kompetenz als Perspektivenübernahmefähigkeit der Einfluss von Lehrerfahrung, Vertrautheit mit dem Urteilsobjekt und Aufgabendesign auf die Urteilsgüte*. Dissertation. Freiburg im Breisgau.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 71–176). Göttingen: Hogrefe.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett.
- Hertel, S. (2014). Adaptive Lerngelegenheiten in der Grundschule. Merkmale, methodisch-didaktische Schwerpunktsetzungen und erforderliche Lehrerkompetenzen. In B. Kopp, S. Martschinke, M. Munser-Kiefer, M. Haider, E.-M. Kirschhock, G. Ranger & G. Renner (Hrsg.), *Jahrbuch Grundschulforschung: Individuelle Förderung und Lernen in der Gemeinschaft* (S. 19–34). Wiesbaden: VS.
- Hertel, S., Fingerle, M. & Rohlfs, C. (2016). Gestaltung adaptiver Lerngelegenheiten in der Schule. In K. Rabenstein & B. Wischer (Hrsg.), *Individualisierung schulischen Lernens. Mythos oder Königsweg?* (S. 64–75). Seelze: Klett.
- Heyl, V., Janz, F., Seifried S. & Trumpa, S. (2014). Einstellungen zu Inklusion von Lehrkräften und Eltern an der Grundschule. In E.-K. Franz, S. Trumpa & I. Esslinger-Hinz (Hrsg.), *Inklusion: Eine Herausforderung für die Grundschulpädagogik* (S. 200–212). Baltmannsweiler: Schneider.
- Kersting, M., Althoff, K. & Jäger A. O. (2008). *WIT-2. Wilde-Intelligenz-Test 2*. Göttingen: Hogrefe.
- König, E. & Volmer, G. (2008). *Handbuch systemische Organisationsberatung*. Weinheim: Beltz.
- König, J. (2015). Kontextualisierte Erfassung von Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 61(3), 305–309.
- König, J. (2016). Lehrerexpertise und Lehrerkompetenz. In M. Rothland (Hrsg.), *Beruf Lehrer/Lehrerin. Ein Studienbuch* (S. 127–148). Münster: Waxmann.

- König, J., Blömeke, S. & Kaiser, G. (2015). Early career mathematics teachers' general pedagogical knowledge and skills. Do teacher education, teaching experience, and working conditions make a difference? *International Journal of Science and Mathematics Education*, 13(2), 331–350.
- König, J., Buchholtz, C. & Dohmen, D. (2015). Analyse von schriftlichen Unterrichtsplanungen: Empirische Befunde zur didaktischen Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte. *Zeitschrift für Erziehungswissenschaft*, 18(2), 375–404.
- König, J., Rothland, M., Darge, K., Lünemann, M. & Tachtsoglou, S. (2013). Erfassung und Struktur berufswahlrelevanter Faktoren für die Lehrerausbildung und den Lehrerberuf in Deutschland, Österreich und der Schweiz. *Zeitschrift für Erziehungswissenschaft*, 16(3), 553–577.
- König, J., Tachtsoglou, S. & Seifert, A. (2012). Individuelle Voraussetzungen, Lerngelegenheiten und der Erwerb von pädagogischem Professionswissen. In J. König & A. Seifert (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung* (S. 235–284). Münster: Waxmann.
- König, J., Wagner, C., Valtin, R. & Schmude, C. (2011). *Jugend, Schule, Zukunft. Psychosoziale Bedingungen der Persönlichkeitsentwicklung; Ergebnisse der Längsschnittstudie AIDA*. Münster: Waxmann.
- Kounin, J. S. (1976). *Techniken der Klassenführung*. Bern: Huber.
- Krauss, S., Lindl, A., Schilcher, A., Fricke, M., Göhring, A., Hofmann, B., Kirchhoff, P., Mulder, R. H. & Baumert, J. (Hrsg.). (2017). *FALKO: Fachspezifische Lehrerkompetenzen. Konzeption von Professionswissenstests in den Fächern Deutsch, Englisch, Latein, Physik, Musik, Evangelische Religion und Pädagogik mit neuen Daten aus der COACTIV-Studie*. Münster: Waxmann.
- Kunz, A., Lüder, R. & Moretti, M. (2010). Die Messung von Einstellungen zur Integration (EZI). *Empirische Sonderpädagogik*, 2(3), 83–94.
- Leutner, D. (2006). Programmierter und Computerunterstützter Unterricht. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 595–602). Weinheim: Beltz.
- Lüders, M., Eisenacher, S. & Pleßmann, S. (2006). Der Umgang mit Studienzeit. Eine empirische Untersuchung bei Studierenden der Lehramter und im Diplomstudiengang Erziehungswissenschaft. In C. Allemann-Ghionda (Hrsg.), *Zeitschrift für Pädagogik: Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern. Ausbildung und Beruf* (S. 116–129). Weinheim: Beltz.
- Meschede, N. (2014). *Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. Theoretische Beschreibung und empirische Erfassung*. Berlin: Logos.
- Miller, S. (2013). Die Sicht der Lehrkräfte auf Heterogenität. Ergebnisse einer quantitativen Erhebung in NRW. In E. Jürgens & S. Miller (Hrsg.), *Ungleichheit in der Gesellschaft und Ungleichheit in der Schule. Eine interdisziplinäre Sicht auf Inklusions- und Exklusionsprozesse* (S. 235–251). Weinheim: Beltz.
- Munser-Kiefer, M. (2014). Formen und Qualitätsmerkmale offenen Unterrichts. In W. Einsiedler, M. Götz, H. Hacker, J. Kahlert, R. W. Keck & U. Sandfuchs (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (4. Aufl., S. 351–355). Bad Heilbrunn: Klinkhardt.
- Neuweg, G. H. (2015). Kontextualisierte Kompetenzmessung. Eine Bilanz zu aktuellen Konzeptionen und forschungsmethodischen Zugängen. *Zeitschrift für Pädagogik*, 61, 377–383.
- Oser, F., Curcio, G.-P. & Düggeli, A. (2007). Kompetenzmessung in der Lehrerbildung als Notwendigkeit Fragen und Zugänge. *Beiträge zur Lehrerbildung*, 25, 14–25.

- Pohlmann, B. & Möller, J. (2010). Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). *Zeitschrift für Pädagogische Psychologie*, 24(1), 73–84.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (Hrsg.). (2013). *Fortschritte und Herausforderungen in Deutschland: PISA 2012: Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Rauin, U. (2011). Forschungen zur Lehrerbildung aus nationaler und internationaler Perspektive. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 441–450). Münster: Waxmann.
- Retelsdorf, J., Bauer, J., Gebauer, S. K., Kauper, T. & Möller, J. (2014). Erfassung berufsbezogener Selbstkonzepte von angehenden Lehrkräften (ERBSE-L). *Diagnostica*, 60(2), 98–110.
- Rutsch, J., Rehm, M., Vogel, M., Seidenfuß, M. & Dörfler, T. (Hrsg.). (2018). *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen*. Wiesbaden: VS.
- Saalfrank, W.-T. (2012). Differenzierung. In E. Kiel (Hrsg.), *Unterricht sehen, analysieren, gestalten* (2. Aufl., S. 65–98). Bad Heilbrunn: Klinkhardt.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Schladitz, S., Groß Ophoff, J. & Wirtz, M. (2015). Konstruktvalidierung eines Tests zur Messung bildungswissenschaftlicher Forschungskompetenz. In S. Blömeke & O. Zlatkin-Troitschanskaia (Hrsg.), *Zeitschrift für Pädagogik, Beiheft: Kompetenzen von Studierenden* (S. 167–184). Weinheim: Beltz.
- Schrader, F.-W. (2011). Lehrer als Diagnostiker. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 683–698). Münster: Waxmann.
- Schrader, F.-W. (2013). Diagnostische Kompetenz von Lehrpersonen. *Beiträge zur Lehrerbildung*, 31(2), 154–165.
- Schrader, F.-W. & Helmke, A. (1987). Diagnostische Kompetenz von Lehrern. Komponenten und Wirkungen. *Empirische Pädagogik*, 1(1), 27–52.
- Schulze-Stocker, F., Holzberger, D., Kunina-Habenicht, O., Terhart, E. & Kunter, M. (2016). Spielen Studienschwerpunkte wirklich eine Rolle? Zum Zusammenhang von bildungswissenschaftlichen Studienschwerpunkten, selbst eingeschätzten Kenntnissen und gemessenem Wissen am Ende eines Lehramtsstudiums. *Zeitschrift für Erziehungswissenschaft*, 19(3), 599–623.
- Seifried, S. & Heyl, V. (2016). Entwicklung und Validierung eines Einstellungsfragebogens zu Inklusion für Lehrkräfte (EFI-L). *Empirische Sonderpädagogik*, 1(1), 22–35.
- Sermier Dessemontet, R., Benoit, V. & Bless, G. (2011). Schulische Integration von Kindern mit einer geistigen Behinderung. Untersuchung der Entwicklung der Schulleistungen und der adaptiven Fähigkeiten, der Wirkung auf die Lernentwicklung der Mitschüler sowie der Lehrereinstellungen zur Integration. *Empirische Sonderpädagogik*, 3(4), 291–307.
- Stanley, A., Grimbeek, P., Bryer, F. & Beamish, W. (2003). Comparing parents' versus teachers' attitudes to inclusion: When PATI meets TATI. In B. Bartlett, F. Bryer, & D. Roebuck (Hrsg.), *Reimagining practice: Researching change* (S. 62–69). Nathan, Qld: Griffith University.
- Stürmer, K., Seidel, T. & Kunina-Habenicht, O. (2015). Unterricht wissensbasiert beobachten. Unterschiede und erklärende Faktoren bei Referendaren zum Berufseinstieg. *Zeitschrift für Pädagogik*, 61(3), 345–360.

- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., Leutner, D., Neuhaus, B. J., Sandmann, A., Sumfleth, E., Thillmann, H. & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7–28.
- Tepner, O. & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 311–323). Berlin: VS.
- Terhart, E. (2013). Universität und Lehrerbildung: Perspektiven einer Partnerschaft. In E. Terhart (Hrsg.), *Erziehungswissenschaft und Lehrerbildung* (S. 184–202). Münster: Waxmann.
- United Nations (2006). *United nations convention on the rights of persons with disabilities*.
- Van Lehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A. & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science*, 31, 3–62.
- Wacker, A. (2017). Schulstruktur und Zweigliedrigkeit: Umbau des Bildungssystems. In T. Bohl, J. Budde & M. Rieger-Ladich (Hrsg.), *Umgang mit Heterogenität in Schule und Unterricht. Grundlagentheoretische Beiträge, empirische Befunde und didaktische Reflexionen* (S. 191–206). Bad Heilbrunn: Klinkhardt.
- Walberg, H. J. & Paik, S. J. (2000). *Effective educational practices*. Brussels: International Academy of Education & International Bureau of Education.
- Warwas, J., Hertel, S. & Labuhn, A. S. (2011). Bedingungsfaktoren des Einsatzes von adaptiven Unterrichtsformen im Grundschulunterricht. *Zeitschrift für Pädagogik*, 57(6), 854–867.
- Waxman, H. C., Wang, M. C., Anderson, K. A. & Walberg, H. J. (1985). Synthesis of research on the effects of adaptive education. *Educational Leadership*, 43(1), 26–29.
- Weiber, R. & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung in die Kausalanalyse mit Hilfe von AMOS, SmartPLS und SPSS* (2. Aufl.). Berlin: VS.
- Weinert, F. E., Schrader, F.-W. & Helmke, A. (1990). Unterrichtsexpertise – Ein Konzept zur Verringerung der Kluft zwischen zwei theoretischen Paradigmen. In L.-M. Alisch, J. Baumert & K. Beck (Hrsg.), *Professionswissen und Professionalisierung* (Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft, Bd. 28; S. 173–206). Braunschweig: Copy-Center Colmsee.
- Wirtz, M. (2004). Über das Problem fehlender Werte. Wie der Einfluss fehlender Informationen auf Analyseergebnisse entdeckt und reduziert werden kann. *Die Rehabilitation*, 43(2), 109–115.
- Wood, D. J., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychiatry and Psychology*, 17(2), 89–100.

Magdalena Muckenthaler¹, Teresa Tillmann¹, Sabine Weiß,
Andreas Hillert & Ewald Kiel

Belastet Kooperation Lehrerinnen und Lehrer? Ein Blick auf unterschiedliche Kooperationsgruppen und deren Belastungserleben

Zusammenfassung

Die vorliegende Studie identifiziert Gruppen von Lehrkräften hinsichtlich Kooperation, basierend auf verschiedenen Formen der Zusammenarbeit, und vergleicht diese nach an der Schule bestehenden Bedingungen für Kooperation und nach Belastungserleben. Bestehende Befunde dokumentieren positive Effekte von Zusammenarbeit für Schule, Schülerinnen und Schüler sowie die Lehrpersonen selbst, gleichzeitig sind vor allem komplexe Formen, wie kokonstruktives Arbeiten, in nur geringem Maße realisiert. Die Identifizierung der Gruppen erfolgte mittels hierarchischer Clusteranalyse, die Clusterlösung wurde folgend durch eine iterativ-partitionierende Clusterzentrenanalyse (K-Means) optimiert. Es zeigen sich drei differierende Cluster. Lehrende im ersten Cluster nutzen alle Formen von Zusammenarbeit am intensivsten, äußern die günstigsten Rahmenbedingungen hinsichtlich Zeit für und Zwang zur Kooperation und beschreiben sich als am wenigsten belastet. Die dritte Gruppe verhält sich gegensätzlich, nutzt alle Formen am wenigsten, die Rahmenbedingungen erfahren die ungünstigste Einschätzung, das Belastungserleben ist am höchsten. Das zweite Cluster pflegt fachlichen und schülerbezogenen Austausch intensiver, nutzt aber arbeitsteilige Formen deutlich weniger; das Belastungserleben nimmt eine Mittelposition ein. Lehrkräfte, die am stärksten arbeitsteilig kooperieren, zeigen sich als am wenigsten belastet. Zudem

1 Geteilte Erstautorenschaft

Magdalena Muckenthaler, M.A. (corresponding author) · Teresa Tillmann, M.Sc. · PD Dr. Sabine Weiß · Prof. Dr. Ewald Kiel, Lehrstuhl für Schulpädagogik, Ludwig-Maximilians-Universität München, Leopoldstr. 3, 80802, München, Deutschland
E-Mail: magdalena.muckenthaler@edu.lmu.de
teresa.tillmann@edu.lmu.de
sabine.weiss@edu.lmu.de
kiel@lmu.de

Prof. Dr. Dr. Andreas Hillert, Schön Klinik Roseneck, Am Roseneck 6, 83209 Prien am Chiemsee, Deutschland
E-Mail: ahillert@schoen-kliniken.de

führt Zwang zur Zusammenarbeit „von oben“ bzw. ein als solcher wahrgenommener nicht zu (intensiverer bzw. gelingender) Kooperation. Die Befunde werden im Kontext der Divergenz von Bedeutungszuschreibung und tatsächlicher Umsetzung von Zusammenarbeit sowie von Implikationen hinsichtlich Kooperationsverhalten und struktureller Bedingungen diskutiert.

Schlagworte

Arbeitsteilung; Kooperation; Belastungserleben; Lehrkräfte

Are teachers burdened by collaboration?

A view on different collaboration groups and their perceived stress

Abstract

The presented study identified groups of teachers based on different forms of collaboration teachers use. The groups were compared with regard to school-related conditions of collaboration and teachers' perceived stress. Existing research findings show positive effects of collaboration on the individual school, students and teaching staff. Concurrently, especially complex forms, such as construction, are realised less in teachers' everyday work. Teacher groups were identified using a hierarchical cluster analysis, and the revealed clusters were optimised by k-means clustering. Three different clusters on collaboration were extracted. The first cluster used all forms of collaboration most intensive, stressed the most positive view on school-related conditions (time and compulsion to collaborate) and stated the lowest stress experience. The third group is contrary to the first one: teachers implemented all form of collaboration to the lowest extent and estimated both conditions and stress experience most unfavourable. The second cluster used forms of exchange, but more complex forms basing on common work organisation were less implemented; stress experience was in the middle range. Teachers implementing complex forms to the greatest extent showed the lowest stress. In addition, compulsion to collaborate did not result in (more intensive or successful) collaboration. The findings were discussed with regard to the divergence of importance and actual implementation of collaboration as well implications on collaborative work and structural conditions.

Keywords

Collaboration; Collective work organisation; Stress experience; Teachers

1. Einleitung

Lehrerkooperation ist ein bedeutsamer Gegenstand im internationalen Diskurs von Erziehungswissenschaft und verwandter Disziplinen (z. B. Gräsel, Fussangel & Pröbstel, 2006; Ronfeldt, Owens Farmer, McQueen & Grissom, 2015; Vangrieken,

Grosemans, Dochy & Kyndt, 2017), der durch die steigende Heterogenität der Schülerschaft durch Inklusion und Migration eine zunehmende Bedeutungszuschreibung erfährt (Da Fonte & Barton-Arwood, 2017; McElvany, Gebauer, Bos & Holtappels, 2013). Diese kontrastiert allerdings damit, dass Kooperation nach wie vor für viele Lehrkräfte ungewohntes Neuland bedeutet und oftmals mit Skepsis bis hin zu Ablehnung behaftet ist (Fabel-Lamla, 2018; Fussangel, 2018). Vor allem intensive, arbeitsteilige und kokonstruktive Formen sind in der schulischen Praxis selten realisiert (Cheng & Ko, 2009; Gräsel et al., 2006; Vangrieken, Dochy, Raes & Kyndt, 2015). Lehrende selbst benennen Kooperation immer wieder als zusätzlich zu leistende Aufgabe, die sie mit Mehraufwand, der Erwartung übermäßigen Engagements und Belastung assoziieren (vgl. Bovbjerg, 2006; Klusmann, Kunter, Trautwein, Lüdtke & Baumert, 2008). Wirkungszuschreibung und Praxis divergieren, es bestehen offene Fragen hinsichtlich der meist positiv konnotierten Effekte von Zusammenarbeit, insbesondere zu Kooperation und Belastungserleben.

Die vorliegende Studie verfolgt das Ziel, ausgehend von dem Modell von Gräsel et al. (2006) verschiedene Gruppen von Lehrkräften hinsichtlich Kooperation zu ermitteln, die auf den verschiedenen Formen von Zusammenarbeit basieren. Die Gruppen werden in einem zweiten Schritt, hinsichtlich an der Schule wahrgenommener Bedingungen von Zusammenarbeit und Dimensionen des Belastungserlebens, verglichen.

2. Kooperation und Belastung im Lehrerberuf: ein Spannungsfeld

2.1 Lehrerkoooperation: ein Feld der Divergenz von Bedeutungszuschreibung und tatsächlicher Realisierung

Kooperationsmodelle zeigen, welche Formen der innerschulischen Zusammenarbeit möglich und zielführend sind sowie auch tatsächlich realisiert werden. Das im deutschsprachigen Raum populärste Modell von Gräsel et al. (2006) unterscheidet, basierend auf dem Kooperationschema von Little (1990) und organisationspsychologischen Ansätzen, drei Formen der Zusammenarbeit:

- Beim Austausch informieren sich Lehrende gegenseitig über berufliche Inhalte und tauschen Material aus. Sie arbeiten überwiegend unabhängig voneinander. Da keine Zielinterdependenz erforderlich ist, müssen sich Lehrkräfte nicht über (divergierende) Standpunkte einigen und es sind kaum negative Konsequenzen für den Einzelnen, wie Konflikte, zu erwarten. Daher spricht man von einer „low cost-Form“ der Kooperation.
- Arbeitsteilige Kooperation setzt eine gemeinsame Verständigung über Ziele voraus, die durch Erledigung von Teilaufgaben erreicht wird. Die einzelne Lehrperson arbeitet weitgehend autonom, trägt aber einen bestimmten Part zur gemeinsamen Gesamtaufgabe bei. Dafür ist ein höheres Maß an Vertrauen und

Absprache erforderlich, da man sich darauf verlassen muss, dass jeder seinen Teil erledigt.

- Kokonstruktion bedeutet, Wissen so aufeinander zu beziehen, dass gemeinsames Wissen und/oder gemeinschaftliche Problemlösungen entstehen. Diese intensive Form der Zusammenarbeit schränkt die Autonomie der Beteiligten ein, fordert eine Abstimmung der individuellen Ziele auf eine gemeinsame Perspektive und birgt in den nötigen Aushandlungsprozessen Potenzial für Konflikte. Daher wird von einer „high-cost“-Form der Zusammenarbeit gesprochen.

Betrachtet man die Literatur- bzw. Befundlage zu Kooperation im Lehrerberuf, ist diese mit Blick die Bedeutungszuschreibung und die tatsächliche Umsetzung durch eine Divergenz geprägt. Auf der einen Seite wird die Zusammenarbeit von Lehrkräften als in hohem Maße erstrebenswert und nötig beschrieben und fast schon zwingend gefordert. Sie gilt etwa als gewinnbringend für die Qualität und Entwicklung der Einzelschule (z.B. Clement & Vandenberghe, 2000; Steinert et al., 2006; Vangrieken et al., 2015), für die Schülerinnen und Schüler (z.B. Goddard, Goddard & Tschannen-Moran, 2007; Lomos, Hofman & Bosker, 2011; Ronfeldt et al., 2015) und nicht zuletzt für die Lehrerinnen und Lehrer selbst, beispielsweise für deren Reflexionsfähigkeit, den Aufbau neuen Wissens oder die Anwendung innovativer Unterrichtsstrategien (Bertrand, Roberts & Buchanan, 2006; Vangrieken et al., 2015; siehe auch die Professionellen Lerngemeinschaften, Bonsen & Rolff, 2006). Insbesondere das Desiderat nach der „höchsten“ Form der Zusammenarbeit, der Kokonstruktion, kennzeichnet die Literatur, da diese positiv mit Schul- und Unterrichtsqualität sowie der professionellen Weiterentwicklung von Lehrpersonen assoziiert ist (Doppenberg, Bakx & den Brok, 2012). Allerdings, so Fussangel und Gräsel (2012), ist nicht per se davon auszugehen, dass eine intensivere Form der Kooperation eine stärkere Wirkung hat, sondern dies muss situativ beurteilt werden.

Dem stehen auf der anderen Seite Studien gegenüber, die wiederholt replizieren, dass intensivere Kooperationsformen im schulischen Alltag kaum realisiert werden und die Zusammenarbeit von Lehrenden vor allem durch ein Minimum an Austausch von Informationen und Materialien erfolgt (Cheng & Ko, 2009; Gräsel et al., 2006; Vangrieken et al., 2015; Visscher & Witziers, 2004). Als eine Erklärung hierfür dient das Autonomie-Paritäts-Muster (Lortie, 1975). Lehrkräfte wollen im Alltag ihrer Arbeit nachgehen, ohne dass jemand in ihr Handeln oder ihren Unterricht eingreift (Autonomie), und wollen dabei gleichberechtigt agieren (Parität) (Vangrieken et al., 2017). Dem folgend werden intensivere Formen der Kooperation nicht zuletzt deshalb selten umgesetzt, da sie die Autonomie der Lehrenden stärker einschränken als einfachere Formen (z.B. Gräsel et al., 2006). Lehrkräfte fürchten beispielsweise, ihren freien Handlungsspielraum zu beschränken oder im Schulalltag stärker fremdbestimmt zu sein.

Ergänzend verweisen Untersuchungen auf die Rolle von schulischen Rahmenbedingungen (Gräsel et al., 2006; Scruggs, Mastropieri & McDuffie, 2007). Kooperation ist kontextuell gebunden; soll sie gelingen, bedarf sie

- institutioneller Bedingungen, z.B. geteilte Ziele, ausreichend Zeit, geeignete Räume, kooperationsbefürwortende Schulleitung,
- personeller Bedingungen, z.B. positive Einstellung, Bereitschaft zur Zusammenarbeit, Rollen-Reflexion, kommunikative Kompetenz und
- organisationskultureller Bedingungen, z.B. Klima der Offenheit und des Vertrauens, gemeinsame Werte, respektvoller Umgang (Huber, Ahlgrimm & Hader-Popp, 2012, S. 354f.).

Diese Bedingungen tragen zur Umsetzung intensiver und qualitativer Kooperationsaktivitäten im Kollegium bei (Gräsel et al., 2006; Huber et al., 2012; Pröbstel & Soltau, 2012) und können diese befördern wie auch einschränken und verhindern.

2.2 Belastung von Lehrenden: begriffliche Klärung

Die internationale Belastungsforschung, mit Schwerpunkten auf personenbezogenen Merkmalen (z.B. Cramer & Binder, 2015) und sozialer Unterstützung (für einen Überblick siehe Prilleltensky, Neff & Bessell, 2016), ist dadurch charakterisiert, dass der eher allgemeine Begriff „Belastung“ nicht einheitlich geklärt und operationalisiert ist. So wird Belastung in der Literatur sowohl für die Beschreibung des Befindens als auch für bestimmte Auslöser hinsichtlich negativen Befindens genutzt (Prilleltensky et al., 2016). Um einer objektiven Beschreibung der wahrgenommenen Belastung im Lehrerberuf nachzukommen, legt die vorliegende Studie Definitionen der Arbeitspsychologie bzw. nach DIN ISO 10075 (Joiko, Schmauder & Wolff, 2010) zugrunde. Psychische Belastung wird demnach als „die Gesamtheit aller erfassbarer Einflüsse, die von außen auf den Menschen zukommen und psychisch auf ihn einwirken“ (Joiko et al., 2010, S. 9) definiert. Diese Einflüsse schließen den Arbeitsplatz mit Arbeitsaufgaben, -bedingungen, -umgebungen und die Arbeitsorganisation mit ein. Böhm-Kasper (2004) beschreibt, dem folgend, in seinem Modell Belastung als Gesamtheit der äußeren Einflüsse auf die Lehrkraft (z.B. Rahmenbedingungen, Kooperation, Lehrdeputat oder individuelle Faktoren). Diese mündet schließlich in Beanspruchung. Unter negativer Beanspruchung werden beispielsweise psychosomatische Erkrankungen oder emotionale Erschöpfung verstanden.

Um die individuelle Wirkung von Belastung erfassen zu können, bietet sich eine Operationalisierung durch psychische und psychiatrische Variablen wie Depressivität, der Empfehlung von Schonfeld (1992) folgend, und Angst an. Beide Konstrukte sind international wissenschaftlich etabliert, können durch validierte Instrumente erhoben werden (vgl. Lovibond & Lovibond, 1995) und sind in internationalen Klassifikationssystemen psychischer Störungen, der ICD-10 (WHO; Dilling, Mombour & Schmidt, 2014) und dem DSM-5 (American Psychiatric Association, 2013), zu finden. Depressivität wird hier unter anderem über eine depressive Stimmungslage und weitere Symptome beschrieben. Angst wird als

„ein mit bestimmten Situationsveränderungen *intraindividuell* (innerhalb eines Individuums) variierender affektiver Zustand (state) des Organismus verstanden“ (Krohne, 2010, S. 15).

2.3 Kooperation und Belastungserleben von Lehrenden: Zuschreibung und Evidenz

Die empirische Lehrgesundheitsforschung ordnet Kooperation häufig in den Kontext der sozialen Unterstützung ein, weniger im Sinne konkreter Kooperationsaktivitäten, sondern als Eingebundenheit und Miteinander mit Kolleginnen und Kollegen. Für die soziale Eingebundenheit im Kollegium sind positive Effekte für das Wohlbefinden und die psychische Gesundheit von Lehrkräften dokumentiert (Egodawatte, McDougall & Stoilescu, 2011; Johnson, 2003; im Überblick Rothland, 2013). Nur wenige Studien erfassen Kooperation bzw. deren Formen differenziert. Ein Beispiel dafür sind die Befunde von Böhm-Kasper und Kolleginnen (vgl. Böhm-Kasper, Dizinger & Gausling, 2016; Fussangel, Dizinger, Böhm-Kasper & Gräsel, 2010; Dizinger, Fussangel & Böhm-Kasper, 2011), die zeigen, dass die kokonstruktive Form der Zusammenarbeit in einem negativen, wenn auch geringen, Zusammenhang mit Belastung und Beanspruchung von Lehrenden steht; für die anderen Kooperationsformen finden sich keine Zusammenhänge. Die Autorinnen und Autoren schlussfolgern daraus, dass vor allem die kokonstruktive Zusammenarbeit zwar zunächst zusätzlichen persönlichen und zeitlichen Einsatz von den Lehrkräften verlangt, aber schließlich entlastend wirken kann (Böhm-Kasper, Dizinger & Heitmann, 2013). Die Studien sind allerdings in der Ganztagschule entstanden und daher nicht auf alle Kooperationskontexte übertragbar. Die Befunde weiterer Studien weisen darauf hin, dass Zusammenarbeit protektiv bezüglich Arbeitsunzufriedenheit und Burnout wirkt, emotional entlastet und sich positiv auf die Erholungs- und Leistungsfähigkeit auswirken kann (Egodawatte et al., 2011; Forte & Flores, 2014; Reeves, Pun & Chung, 2017). Im Widerspruch dazu stehen Studien, wonach Kooperation von den Lehrenden als Belastung und in Zusammenhang mit Mehraufwand, Stress und emotionaler Erschöpfung wahrgenommen wird (Bovbjerg, 2006; Johnson, 2003; Klusmann et al., 2008).

Die Verallgemeinerbarkeit und Vergleichbarkeit der inkonsistenten Befundlage ist, wie für den Belastungsbegriff bereits skizziert, durch unterschiedliche Operationalisierungen von Kooperation erschwert. Diese wird vielfach eindimensional betrachtet, über unterschiedliche Messinstrumente (Fussangel & Gräsel, 2014; Reeves et al., 2017; Vangrieken et al., 2017) oder verwandte Konstrukte wie soziale Unterstützung erhoben und in ihrem Zusammenhang mit Belastung analysiert (z. B. Morgenroth & Buchwald, 2015).

3. Ziel und Forschungsfragen

Bestehende Befunde dokumentieren positive Effekte der Zusammenarbeit von Lehrenden für Schule, Schülerinnen und Schüler sowie die Lehrpersonen selbst (u. a. Lomos et al., 2011; Ronfeldt et al., 2015; Vangrieken et al., 2015). Gleichzeitig sind aber vor allem komplexe Formen der Kooperation in nur geringem Maße realisiert. Die vorliegende Studie nimmt das Modell von Gräsel et al. (2006) als Ausgangspunkt. Es ist, den Desideraten folgend, hinsichtlich Kooperation die Mehrdimensionalität des Konstrukts zu berücksichtigen und das Zusammenwirken von Kooperationsformen, Kooperationsbedingungen und Belastungserleben in den Blick zu nehmen.

Folgende Forschungsfragen und Annahmen werden verfolgt:

- 1) Lassen sich Gruppen von Lehrkräften identifizieren, die sich hinsichtlich der genutzten Kooperationsformen unterscheiden?
Studien verweisen darauf, dass Lehrende unterschiedlich intensiv kooperieren bzw. unterschiedliche Formen der Kooperation auf unterschiedlichen Niveaus nutzen (Cheng & Ko, 2009; Gräsel et al., 2006; Vangrieken et al., 2015; Visscher & Witziers, 2004). Daher ist davon auszugehen, dass sich verschiedene Gruppen hinsichtlich der Umsetzung von Kooperation identifizieren lassen.
- 2) Sind Lehrende verschiedener Schularten in den ermittelten Clustern unterschiedlich verteilt?
Befunde zeigen, dass Lehrkräfte an Grundschulen intensiver kooperieren als die an weiterführenden Schulen (Klieme, Baumert & Schwippert, 2000; Richter & Pant, 2016; Steinert et al., 2006). Daher ist zu erwarten, dass sich Grundschullehrkräfte vermehrt in Kooperationsclustern finden, in denen die Zusammenarbeit intensiver ausgeprägt ist.
- 3) Bestehen Unterschiede zwischen den Gruppen hinsichtlich der wahrgenommenen Bedingungen von Kooperation an der Schule (Zeitstrukturen, Zwang zur Zusammenarbeit)?
Es ist anzunehmen, dass günstiger beurteilte zeitliche Strukturen zu intensiverer Kooperation führen (z. B. Fussangel, 2008; Richter & Pant, 2016). Hinsichtlich des Zwangs gibt es kaum eine Befundlage; bei Fussangel (2008) zeigt dieser wenig Einfluss.
- 4) Bestehen Unterschiede zwischen den Gruppen von Lehrenden hinsichtlich der wahrgenommenen Belastung (operationalisiert durch Depressivität und Angst)?
Die Forschungslage zu Kooperation und Belastung ist inkonsistent und eine Annahme daher erschwert, denn Befunde berichten einen Konnex zwischen Kooperation und Belastung (z. B. Bovbjerg, 2006; Klusmann et al., 2008), andere widerlegen dies.

4. Methode

4.1 Projektkontext

Die vorliegende Studie ist Teil des Projekts *Risiko-Check für das Lehramt* an der Ludwig-Maximilians-Universität München (LMU). Dieses wurde im Rahmen der gemeinsamen Qualitätsoffensive Lehrerbildung von Bund und Ländern durch das Bundesministerium für Bildung und Forschung für einen Zeitraum von drei Jahren (2015–2018) gefördert, danach fest etabliert. Ziel von Projekt und Studie ist es, domänenübergreifende Kriterien und Gelingensbedingungen für eine erfolgreiche Berufsausübung von Lehrkräften zu identifizieren.

4.2 Datenerhebungsprozess und Stichprobenbeschreibung

Die Stichprobe wurde im Zeitraum Spätherbst und Winter 2017/18 mit Hilfe einer universitätsinternen Plattform generiert. Ein Zugang zu dieser Plattform erfolgte ohne Barrieren und entsprechend Datenschutzrichtlinien komplett anonym, d. h. ohne Klarnamen, ohne Generierung von Benutzernamen oder Passwörtern sowie ohne vorherige Anmeldung bei der Plattform. Die Nutzung der Plattform bzw. der Aufruf zur Beteiligung an der Befragung wurden über den Bayerischen Lehrer- und Lehrerinnenverband sowie den Bayerischen Beamtenbund beworben, d. h. interessierte Lehrerinnen und Lehrer wurden angesprochen, an der Online-Befragung teilzunehmen. Zusätzlich standen hier Kontaktdaten der Ansprechpartnerin in der LMU zur Verfügung, über die sich Lehrkräfte bei Fragen melden konnten.

Vor der Teilnahme wurden alle Lehrkräfte über datenschutzrechtliche Aspekte sowie das Ziel des Projekts informiert. Die Plattform bzw. der Fragebogen war so programmiert, dass ein Überspringen von Items nicht möglich war. Daher musste keine Bereinigung der Daten aufgrund fehlender Werte erfolgen. Der komplette Online-Fragebogen umfasste insgesamt sechs Messinstrumente zu verschiedenen Bereichen der Tätigkeit von Lehrpersonen mit je nach Messinstrument 6-24 Items, die aber nicht Teil der vorliegenden Untersuchung sind, neben Kooperation und Belastungserleben beispielsweise zu personalen Merkmalen und zu den Anforderungen des Lehrerberufs. Die Bearbeitungszeit wurde elektronisch erfasst und liegt je nach Lehrkraft bei 15-20 Minuten.

Die Gesamtstichprobe umfasst $N = 175$ Lehrerinnen und Lehrer an allgemeinbildenden Schulen im Bundesland Bayern. Alle Lehrkräfte, die eine Berufserfahrung von weniger als drei Jahren hatten (und sich somit noch im bzw. kurz im Anschluss an das Referendariat befinden), wurden zuvor aus dem Datensatz herausgenommen, da diese aufgrund der geringen beruflichen Erfahrung auch über wenig Erfahrung in der Kooperation verfügen. Alle erfassten Lehrkräfte sind derzeit im Schuldienst tätig. Das Durchschnittsalter liegt bei 43,9 Jahren ($SD = 10,8$ Jahre), die durchschnittliche Berufserfahrung bei 17,5 Jahren ($SD = 10,2$ Jahre).

63.4 % der Lehrenden sind weiblich. 38.9 % unterrichten an einer Grundschule, 29.4 % arbeiteten an einer Mittelschule, etwa jede fünfte Lehrkraft (21.7 %) an einem Gymnasium. Realschullehrkräfte waren 10.0 % aller Befragten. In der Stichprobe sind Lehrkräfte aller Fächer und Fachkombinationen vertreten. Bei Grundschullehrkräften geben, der Schulart entsprechend, fast alle Lehrerinnen und Lehrer an, Deutsch und Mathematik zu unterrichten, auch Fächer wie Englisch, Sport und Kunst werden häufig genannt. In den weiterführenden Schulen sind die angeführten Fächer divergenter; neben auch hier häufig angeführten Fächern Deutsch, Mathematik und Englisch werden eine Vielzahl weiterer Fächer wie Erdkunde, Biologie, Wirtschaft etc. unterrichtet.

4.3 Verwendete Messinstrumente

Kooperation: Zur Erfassung der Kooperation wird das Messinstrument von Fussangel (2008) genutzt. Aus dem Instrument gehen die vier Skalen zur Kooperationsausübung und zwei zu den Kooperationsbedingungen ein. Alle Items sind auf einer Ratingskala von 1 = „trifft überhaupt nicht zu“ bis 4 = „trifft völlig zu“ zu beantworten. Mit Ausnahme der Kokonstruktion/Reflexion weisen alle Skalen zufriedenstellende bis gute Reliabilitäten auf. Tabelle 1 stellt die Skalen mit Itemanzahl, Beispielitem und Cronbachs Alpha dar.

Tabelle 1: Überblick über die Skalen zur Kooperation mit Anzahl der Items, Beispielitems sowie Reliabilitäten (Cronbachs α)

Skala	Subskala	Itemanzahl	Beispielitem	Reliabilität
Formen von Kooperation	Fachlicher Austausch	4	Wichtige berufsbezogene Informationen teile ich meinen Kolleg/innen mit.	$\alpha = .82$
	Schülerbezogener Austausch	3	Mit meinen Kolleg/innen bespreche ich mich, wenn ich Schwierigkeiten mit einzelnen Schüler/innen habe.	$\alpha = .79$
	Gemeinsame Arbeitsorganisation	3	Mit meinen Kolleg/innen erstelle ich gemeinsam Arbeitsblätter.	$\alpha = .89$
	Kokonstruktion/ Reflexion	4	Mit meinen Kolleg/innen unterrichte ich gemeinsam eine Klasse.	$\alpha = .64$
Bedingungen von Kooperation	Zeitstrukturen	4	Für Kooperation planen wir bewusst Zeit ein.	$\alpha = .82$
	Zwang zur Kooperation	3	Die Kooperation wird „von oben“ angeordnet.	$\alpha = .79$

Belastung: Zur Erfassung der Belastung, operationalisiert durch die wissenschaftlich etablierten und valide messbaren Konstrukte Depressivität (vgl. Schonfeld, 1992) und Angst (siehe Dilling et al., 2014), wurden die Depressions-, Angst- und Stress-Skalen (DASS, Nilges & Essau, 2015; im Original von Lovibond & Lovibond, 1995) eingesetzt. Teilnehmende bewerteten auf einer 4-stufigen Likert-Skala

(0 = „gar nicht“ bis 3 = „sehr stark“), wie die dort beschriebenen Symptome auf sie zutreffen. Als zeitlicher Bezugsrahmen fungiert die vorherige Woche. Tabelle 2 stellt eine Beschreibung der beiden Skalen dar. Die Reliabilitäten sind gut bis sehr gut.

Tabelle 2: Überblick über die verwendeten Skalen zur Belastung mit Anzahl der Items, Beispielitems sowie Reliabilitäten (Cronbachs α)

Subskala	Item-anzahl	Beispielitem	Reliabilität
Depressivität	7	Ich konnte überhaupt keine positiven Gefühle mehr erleben.	$\alpha = .91$
Angst	7	Ich zitterte (z. B. an den Händen).	$\alpha = .87$

4.4 Statistische Analysen

Die Identifikation der Clusterzahl auf Grundlage der Kooperationsformen erfolgte durch eine hierarchische Clusteranalyse nach dem Ward-Verfahren (quadrierte euklidische Distanz). Dieses Vorgehen legt das Varianzkriterium als Heterogenitätsmaß zugrunde (Backhaus, Erichson, Plinke & Weißer, 2016). Das Ziel ist eine Zuordnung von Objekten zu Gruppen, sodass sich das Heterogenitätsmaß am wenigsten vergrößert. Gleichzeitig sollen die Cluster untereinander aber gut voneinander trennbar sein (Bortz, 2010; Härdle & Simar, 2015). Zur Identifizierung und Bestimmung der optimalen Clusteranzahl wurden das Elbow-Kriterium und das Dendrogramm genutzt (Backhaus et al., 2016). Die Clusterlösung wurde anschließend durch eine interaktiv-partitionierende Clusterzentrenanalyse (K-Means-Verfahren) optimiert.

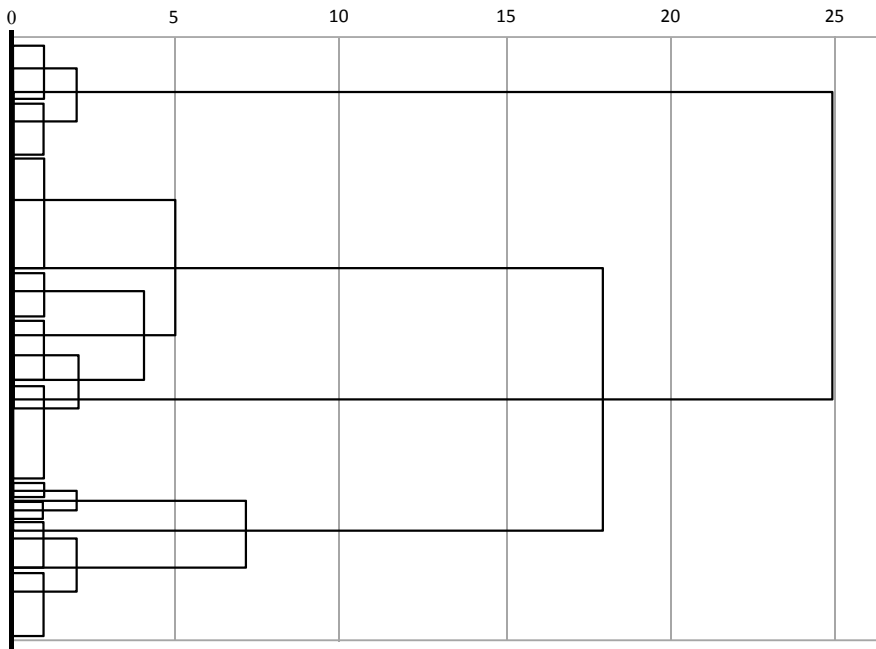
Die Überprüfung von Mittelwertsunterschieden erfolgte mit univariaten Varianzanalysen. Hier wurde überprüft, ob sich die identifizierten Cluster hinsichtlich der Bedingungen von Kooperation an der Schule sowie Depressivität und Angst als Indikatoren psychischer Belastung unterscheiden. Zur Berechnung von Zusammenhängen zwischen den kategorialen Variablen Clusterzugehörigkeit und Schulart wurde ein Chi²-Test genutzt (Schäfer, 2016).

5. Ergebnisse

5.1 Analyse und Überprüfung von Gruppen zu Kooperation mittels Clusteranalyse

Bestimmung der Clusterzahl durch hierarchische Clusteranalyse: Die Bestimmung der Clusterzahl erfolgte durch eine Clusteranalyse nach dem Ward-Verfahren mit dem Ähnlichkeitsmaß quadrierte euklidische Distanz. Abbildung 1 stellt das Dendrogramm zu den Stufen der Agglomeration dar.

Abbildung 1: Dendrogramm der Ward-Clusteranalyse



Das Dendrogramm legt eine drei-Cluster-Lösung bzw. eine Beendigung des Verschmelzungsprozesses an dieser Stelle nahe. Dem folgend umfasst das erste Cluster 33 (18.9 %), das zweite 96 (54.9 %) und das dritte 46 Lehrkräfte (26.2 %). Tabelle 3 stellt die Ergebnisse der ANOVA zum Vergleich der Cluster, einschließlich der Effektstärke Cohens f (Cohen, 1988), dar. Zwischen den drei Clustern zeigen sich signifikante Unterschiede mit durchgängig hohen Effektstärken.

Tabelle 3: Deskriptive und varianzanalytische Statistik der Clusterzentren für die drei Cluster

Subskala	Deskriptive Statistiken $M(SD)$			F (2,172)	f	p
	Cluster 1 ($n = 33$)	Cluster 2 ($n = 96$)	Cluster 3 ($n = 46$)			
Fachlicher Austausch	3.83 (0.21)	3.39 (0.45)	2.14 (0.46)	194.33	1.60	< .001
Schülerbezogener Austausch	3.69 (0.42)	3.17 (0.50)	2.10 (0.48)	119.64	1.20	< .001
Gemeinsame Arbeitsorganisation	3.57 (0.46)	1.93 (0.56)	1.64 (0.80)	109.39	1.11	< .001
Konstruktion/Reflexion	2.45 (0.50)	1.74 (0.54)	1.49 (0.53)	35.69	0.63	< .001

Optimierung der Clusterlösung durch interaktiv-partitionierende Clusterung (K-Means): Die Optimierung der Clusterlösung erfolgte durch eine iterativ-partitionierende Clusterzentrenanalyse nach dem K-Means-Algorithmus. Ausgehend von den drei identifizierten Clusterzentren werden die Daten neu ihrem am nächsten liegenden Clustermittelpunkt zugeordnet. Dies führte zu einer Neuuzuordnung von 17 Personen (9.7 %). Die Konvergenz wurde nach fünf Iterationen erreicht. Die Distanzen zwischen den Clusterzentren der finalen Clusterlösung zeigen, dass die euklidische Distanz zwischen den Gruppen 1 und 3 (Distanzwert 2.96) größer ist als die Distanz zwischen den Gruppe 1 und 2 (Distanzwert 1.88) sowie 2 und 3 (Distanzwert 1.69). Die aus dem Iterationsprozess resultierenden Zentrenänderungen bzw. finalen Clusterzentren sind, einschließlich der ANOVA zum Vergleich der Gruppen, in Tabelle 4 dargestellt.

Tabelle 4: Deskriptive und varianzanalytische Statistik der Clusterzentren für die drei finalen Cluster

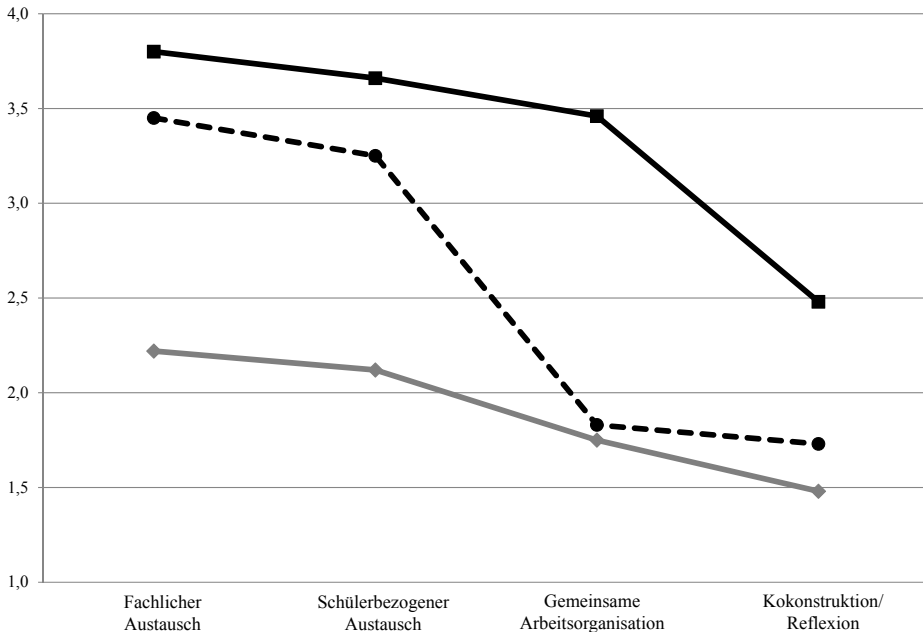
Subskala	Deskriptive Statistiken			F (2,172)	f	p
	M (SD)					
	Cluster 1 (n = 37)	Cluster 2 (n = 88)	Cluster 3 (n = 50)			
Fachlicher Austausch	3.80 (0.28)	3.45 (0.39)	2.22 (0.49)	209.49	1.58	< .001
Schülerbezogener Austausch	3.66 (0.44)	3.25 (0.42)	2.12 (0.44)	171.76	1.40	< .001
Gemeinsame Arbeitsorganisation	3.46 (0.51)	1.83 (0.55)	1.75 (0.80)	109.01	1.09	< .001
Kokonstruktion/Reflexion	2.48 (0.49)	1.73 (0.52)	1.48 (0.49)	46.00	0.73	< .001

Zwischen den Clustern bestehen hinsichtlich der Formen der Kooperation signifikante Unterschiede mit durchgängig hohen Effektstärken. Die Cluster werden folgend nach ihrem Kooperationsverhalten charakterisiert. Da sich mittels Chi²-Test ein Zusammenhang zwischen Cluster und Schulart feststellen ließ ($\chi^2(6) = 15.23$, $p < .05$), der die Lehrenden an Grundschulen betrifft, berücksichtigt die folgende Clusterbeschreibung dies. Abbildung 2 stellt die Clusterzentren grafisch dar.

Cluster 1: Lehrende in Cluster 1 (21.1 %) schreiben es sich zu, übergreifend alle Formen der Kooperation am intensivsten zu nutzen ($ps < .001$ in allen Dimensionen). Fachlicher und schülerbezogener Austausch sowie gemeinsame Arbeitsorganisation zeigen die höchsten Mittelwerte (alle > 3); dies gilt auch für die Kokonstruktion/Reflexion, die ebenfalls die höchste Einschätzung erfährt, in ihrer Ausprägung aber hinter den anderen Formen der Zusammenarbeit zurückbleibt. In diesem Cluster befinden sich viele Grundschullehrkräfte.

Cluster 2: Lehrkräfte in Cluster 2 (50.3 %) sind in ihrem Kooperationsverhalten durch eine Divergenz charakterisiert. Verschiedene Formen von Zusammenarbeit erfahren eine sehr unterschiedliche Einschätzung. Kooperationsformen, die auf Austausch beruhen (fachlich, schülerbezogen), fallen zwar gegenüber dem ersten

Abbildung 2: Grafische Darstellung der Clusterzentren der drei Cluster auf den verschiedenen Kooperationsdimensionen



Cluster ab ($ps > .001$), zeigen aber eine deutlich günstigere Ausprägung gegenüber dem dritten Cluster ($ps > .001$). Die Mittelwerte liegen über 3. Betrachtet man allerdings komplexere, arbeitsteilige Formen der Zusammenarbeit, erweisen sich diese als in geringerem Maße realisiert und ähnlich dem dritten Cluster ($ps > .001$ geg. Cluster 1; Kokonstruktion/Reflexion: $ps > .05$ geg. Cluster 3).

Cluster 3: Das dritte Cluster verhält sich konträr zu Cluster 1: Lehrkräfte (28.6 %) kooperieren insgesamt am wenigsten (für alle Dimensionen $ps < .001$, außer Gemeinsame Arbeitsorganisation geg. Cluster 2). Besonders hinsichtlich des fachlichen und schülerbezogenen Austauschs sind die Unterschiede zu den Clustern 1 und 2 sehr ausgeprägt und bei den komplexeren Formen der Zusammenarbeit sind die Mittelwerte mit < 2 sehr niedrig. In diesem Cluster befindet sich eine besonders geringe Anzahl von Grundschullehrkräften.

5.2 Vergleich der Cluster nach den Kooperationsbedingungen

Zum Vergleich der Cluster hinsichtlich wahrgenommener Zeitstrukturen für und Zwang zur Kooperation wurde eine ANOVA mit Post-hoc Vergleichen nach Bonferroni durchgeführt. Tabelle 5 fasst die Ergebnisse zusammen.

Tabelle 5: Deskriptive und varianzanalytische (F-)Statistik sowie Effektstärke der Clusterzentren zu den Bedingungen von Kooperation an der Schule

Subskala	Deskriptive Statistiken <i>M (SD)</i>			<i>F</i> (2,172)	<i>f</i>	<i>p</i>
	Cluster 1 (<i>n</i> = 37)	Cluster 2 (<i>n</i> = 88)	Cluster 3 (<i>n</i> = 50)			
Zeitstrukturen	2.80 (0.82)	2.11 (0.80)	1.70 (0.68)	21.91	0.51	< .001
Zwang zur Kooperation	1.58 (0.63)	1.87 (0.74)	2.43 (0.93)	14.08	0.41	< .001

Lehrende in Cluster 1, die übergreifend alle Formen der Kooperation am intensivsten nutzen, zeichnen hinsichtlich der Bedingungen, unter denen Kooperation an der Schule stattfindet, das positivste Bild. Sie beschreiben die für Zusammenarbeit vorhandenen Zeitstrukturen als am günstigsten ($ps < .001$) und fühlen sich am wenigsten zur Zusammenarbeit gezwungen ($p < .001$ geg. Cluster 3, $p < .05$ geg. Cluster 2).

Lehrkräfte in Cluster 2, die ihr Verhalten bezüglich Zusammenarbeit divergent einschätzen, nehmen mit Blick auf Kooperationsbedingungen an der Schule eine mittlere Position ein. Die zeitlichen Strukturen der Zusammenarbeit sind ungünstiger als in Cluster 1 ($p > .001$), aber günstiger als in Cluster 3 beschrieben ($p > .01$). Ein Zwang zur Zusammenarbeit wird stärker als in Cluster 1, aber geringer als in Cluster 3 wahrgenommen ($ps > .001$).

Das dritte Cluster mit Lehrkräften, die in allen Dimensionen am wenigsten zusammenarbeiten, ist auch bezüglich der Kooperationsbedingungen durch eine ungünstigere Einschätzung charakterisiert: Lehrende äußern die negativste Einschätzung zu vorhandenen Zeitstrukturen für Zusammenarbeit ($p < .001$ geg. 1; $p > .01$ geg. 2) und berichten vom stärksten Zwang zur Kooperation ($ps < .001$).

5.3 Vergleich der Cluster hinsichtlich der wahrgenommenen Belastung

Der Vergleich der Cluster hinsichtlich der Belastungsdimensionen Depressivität und Angst erfolgte ebenfalls mittels einer ANOVA mit Post-hoc Vergleichen (siehe Tabelle 6).

Tabelle 6: Deskriptive und varianzanalytische (F-)Statistik sowie Effektstärke der finalen Clusterzentren zu den Variablen des Belastungserlebens (Depressivität und Angst)

Subskala	Deskriptive Statistiken <i>M (SD)</i>			<i>F</i> (2,172)	<i>f</i>	<i>p</i>
	Cluster 1 (<i>n</i> = 37)	Cluster 2 (<i>n</i> = 88)	Cluster 3 (<i>n</i> = 50)			
Depressivität	1.19 (0.25)	1.51 (0.56)	1.78 (0.84)	10.0	0.35	< .001
Angst	1.14 (0.21)	1.27 (0.38)	1.51 (0.59)	7.61	0.33	< .001

Lehrende, die am intensivsten kooperieren, wie dies im ersten Cluster der Fall ist, berichten das geringste Belastungserleben (Depressivität: $p > .001$ geg. Cluster 3, $p > .05$ geg. Cluster 2; Angst: $p > .001$ geg. Cluster 3). In der zweiten Gruppe, die vor allem auf Austausch basierende und weniger komplexe Formen der Zusammenarbeit nutzt, erfahren die Indikatoren des Belastungserlebens eine mittlere Einschätzung, höher als in der ersten und niedriger als in der dritten Gruppe (alle $ps > .05$, außer bei Angst geg. Cluster 1). Lehrende, die durchweg am wenigsten zusammenarbeiten (Cluster 3), berichten das höchste Belastungserleben (für Depressivität und Angst $ps < .001$ geg. 1; $ps > .05$ geg. 2).

6. Diskussion

6.1 Kooperation und Belastung

In der dargestellten Studie wurden Gruppen von Lehrkräften hinsichtlich Kooperation identifiziert und folgend nach ihrem Belastungserleben verglichen. Mit Blick auf das Kooperationsverhalten repliziert die Studie zunächst bestehende Befunde, wonach Lehrkräfte eher Kooperationsformen nutzen, die auf Austausch über fachliche Inhalte, Material und die Anliegen von Schülerinnen und Schülern basieren und mit weniger Einschränkungen der eigenen Autonomie verbunden sind (z. B. Gräsel et al., 2006; Vangrieken et al., 2015). Komplexere Formen, die auf arbeitsteilig zu erreichenden Zielen, Kokonstruktion und gegenseitigem Feedback beruhen, werden weniger realisiert.

Trotz dieses übergreifenden Musters bestehen Unterschiede in der Ausübung bzw. Intensität von Zusammenarbeit, wie die Clusterbildung zeigt. Lehrkräfte des ersten Clusters nutzen alle Formen von Zusammenarbeit intensiver. Sie berichten auch von vergleichsweise günstigeren zeitlichen Strukturen für Kooperation. Das dritte Cluster verhält sich gegensätzlich: Alle Kooperationsformen werden am wenigsten genutzt, die Rahmenbedingungen erfahren eine ungünstigere Einschätzung, es wird vor allem ein stärkerer Zwang zur Zusammenarbeit geäußert. Gelingende Kooperation ist mit günstigen Effekten für die Qualität und Entwicklung einer Schule (z. B. Steinert et al., 2006; Vangrieken et al., 2015), für die Schülerinnen

und Schüler (z.B. Goddard et al., 2007; Ronfeldt et al., 2015) sowie für die Lehrkräfte mit Blick auf Unterricht (Bonsen & Rolff, 2006; Egodawatte et al., 2011; Vangrieken et al., 2015) und Berufszufriedenheit (Johnson, 2003; Reeves et al., 2017) assoziiert. Doch wird Zusammenarbeit von Lehrerseite immer wieder als eine zusätzliche, mit Mehraufwand einhergehende Aufgabe benannt, die Belastung verursacht. In der vorliegenden Studie sind es gerade die Lehrkräfte, die intensiver kooperieren und dabei aufwändigere, arbeitsteilige Formen nutzen, die ein geringeres Belastungserleben äußern. Das zeigt sich mit Blick auf das erste Cluster, das nicht nur Austausch nutzt, sondern auch aufwändige Formen realisiert – und die geringste Belastung angibt. Umgedreht äußern diejenigen, die weniger intensiv und aufwändig kooperieren, eine höhere Belastung (vgl. z.B. Böhm-Kasper et al., 2016; Dizinger et al., 2011; Fussangel et al., 2010). Dieser Befund lässt sich durch das zweite Cluster weiter schärfen: Im Vergleich zu den beiden kontrastiven Gruppen befindet sich Cluster 2 zwischen den beschriebenen Extremen. Die Lehrenden pflegen durchaus fachlichen und schülerbezogenen Austausch, nutzen arbeitsteilige Formen aber deutlich weniger. Sie äußern zwar mehr Belastungserleben als das erste intensiver kooperierende Cluster, sie zeigen sich aber weniger belastet als die dritte Lehrergruppe.

Dieser Befund lässt sich zum einen dahingehend diskutieren – nimmt man Cluster 3 als Vergleichsgruppe – dass Lehrkräfte auch bei Nutzung einfacher, nicht arbeitsteilig angelegter Formen der Zusammenarbeit wie Austausch weniger Belastung angeben als die, die auch hier (deutlich) abfallen. Die Bedeutung von Austausch als Form der Zusammenarbeit sollte daher nicht pauschal negiert werden. Allerdings: Nimmt man Cluster 1 als Vergleichsgruppe, Lehrende, die auch arbeitsteilige Formen der Zusammenarbeit stärker umsetzen, so ist deren wahrgenommene Belastung gegenüber dem zweiten Cluster geringer. Dies stellt der Wert besonders von arbeitsteiligen Formen der Kooperation heraus.

Zudem zeigt sich, dass sich der Zwang zur Zusammenarbeit bzw. ein als solcher wahrgenommener nicht in (intensiverer bzw. gelingender) Kooperation niederschlägt. Lehrende kooperieren nicht intensiver, wenn man dies „von oben anordnet“. Oder anders formuliert, ein struktureller Rahmen an der Schule, in dem die Schulleitung die Lehrkräfte dazu zwingt bzw. zwingen möchte, (mehr) zusammenarbeiten, erweist sich zumindest in der vorliegenden Studie als nicht erfolgsbringend. Welche Gründe es dafür gibt, ob beispielsweise Zwang einfach ignoriert wird oder möglicherweise zu Reaktanz führt (vgl. Dickenberger, Gniech & Grabitz, 2001), lässt sich im Kontext der hier präsentierten Studie nicht feststellen. Gerade dieses Thema erscheint aber für weiterführende Forschung interessant, ebenso mit Blick auf Implikationen.

Ähnlich bestehender Befunde kooperieren Lehrkräfte an Grundschulen intensiver als die an weiterführenden Schulen (Harazd & Drossel, 2011; Klieme et al., 2000; Richter & Pant, 2016; Soltau & Mienert, 2009; Steinert et al., 2006). Als Gründe sind als günstiger wahrgenommene Rahmenbedingung beschrieben, beispielsweise hinsichtlich Klassenlehrerprinzip und zeitlicher Strukturen (Richter & Pant, 2016); zudem werden Grundschullehrkräfte als sehr teamfähig charak-

terisiert (Soltau & Mienert, 2009). Allerdings lassen sich auf Basis der Ungleichverteilung innerhalb der Stichprobe diesbezüglich lediglich Tendenzen ableiten, sodass an dieser Stelle keine vertieften Rückschlüsse auf einen Zusammenhang von Schulform und Kooperation ziehen lassen.

6.2 Implikationen

Aus dem Befund, dass intensiver kooperierende Lehrerinnen und Lehrer ein geringeres Belastungserleben kommunizieren, lassen sich Schlussfolgerungen hinsichtlich Implikationen ableiten. Mit Blick auf die individuelle Ebene, die der Lehrkräfte, ließe sich nun erneut auf die seit vielen Jahren immer wieder geäußerte Forderung verweisen, Lehrkräfte sollten mehr und vor allem auf höheren Niveaustufen kooperieren. Zur Umsetzung dieser Forderung wurden vielfältige Maßnahmen zu vor allem (schulinternen) Fortbildungsangeboten initiiert. Dennoch erreicht man mit dieser Forderung viele Lehrende nicht bzw. trägt nicht zu einer Veränderung der Situation bei. Implikationen müssen daher auch thematisieren, warum dies so ist. Die Forschung argumentiert hier mit Lorties Autonomie-Paritäts-Muster, wonach Lehrkräfte nicht in ihrer Autonomie eingeschränkt werden wollen (Vangrieken et al., 2017) und daher vor allem „low-cost“-Varianten der Zusammenarbeit bevorzugen (siehe Gräsel et al., 2006). Andere Untersuchungen im Kontext der Lehrerforschung thematisieren auch Rahmenbedingungen als Hemmnis und Barriere von Zusammenarbeit – und nehmen somit nicht nur das Verhalten der Lehrenden, sondern auch den Arbeitskontext in den Blick.

Nimmt man dies als Ausgangspunkt, sind strukturelle Maßnahmen anzudenken, schulische Rahmenbedingungen für Kooperation, soweit dies möglich ist, günstiger zu gestalten. Möglicherweise lassen sich bestehende Zeitstrukturen an Schulen effektiver gestalten; gerade hierzu ließen sich die angeführten Fortbildungsangebote modifizieren. Allerdings ist hier aber die Ressourcenfrage zu stellen, denn Kooperation erfordert Zeit, besonders dann, wenn die Anliegen und das Unterrichten einer immer heterogeneren Schülerschaft Prozesse intensiver Absprachen und Formen der Zusammenarbeit nötig machen (Da Fonte & Barton-Arwood, 2017; Scruggs et al., 2007; Weiß, Markowitz & Kiel, 2017).

Auf struktureller Ebene ist auch „angeordnete Kooperation“, der Zwang zur Zusammenarbeit, in den Blick zu nehmen. Lehrkräfte, die von mehr Zwang berichten, kooperieren weder mehr noch beschreiben sie sich als weniger belastet. Dieser Aspekt adressiert die Leitungsebene. Schulleitungen haben eine entscheidende Funktion für die Initiierung von Veränderungen und die Gestaltung von Neuerungsprozessen (DeMatthews, 2014; Huber, 2012), wie (der Ausbau von) Kooperation einer sein kann. Eine der Grundlagen in der Schulentwicklungsforschung ist die Einsicht, dass Entwicklung nicht verordnet und Neuerung nicht „gegen“ ein Kollegium eingeführt werden kann, sondern die Perspektiven der in der Organisation Handelnden einbezogen werden müssen (Dalin, 1999). Die Aufgabe der Schulleitung mit Blick auf Teambuildingprozesse und arbeitsteilig zu

erreichende Ziele (siehe Berkemeyer, Berkemeyer & Schwikal, 2015) liegt darin, diese mit einem Grad an Partizipation zu etablieren, der von den Lehrenden nicht als Zwang oder Anordnung verstanden wird und zu Widerstand führt. Kooperation sollte, soweit möglich, gemeinschaftlich mit den Lehrkräften gestaltet werden, ohne das Kollegium dazu „von oben“ zu verpflichten (Huber, 2012; Wissinger, 2014). Studien wie die von Harazd und Drossel (2011) zeigen, dass sich die Niveaustufen der Kooperation proportional positiv zu unterstützenden Strukturen verhalten und die Schulleitung dabei eine zentrale Rolle einnimmt. Es lohnt sich daher, nicht das Defizit nichtgelingender Kooperation in den Blick zu nehmen, sondern ressourcenorientiert diese unterstützenden Strukturen zu forcieren.

6.3 Einschränkungen und Desiderate

Abschließend sind einige methodische Aspekte bei der Interpretation der Ergebnisse zu berücksichtigen. Anzusprechen ist die niedrige Reliabilität der Skala Kokonstruktion/Reflexion. Hier ist darauf hinzuweisen, dass diese auch in der Originalpublikation des Messinstruments gegenüber den anderen Kooperationskalen niedriger ausfällt (Fussangel, 2008). Da die Formulierung und Komplexität der Items nicht von den Items der anderen Skalen abweicht, ließe sich auf inhaltlicher Ebene spekulieren, ob Lehrkräften kokonstruktives Arbeiten möglicherweise so wenig bekannt ist (die Skala hat die geringste Ausprägung), dass hier Verständnisschwierigkeiten auftreten. Ein anderer zu diskutierender Punkt liegt darin, dass Clusteranalysen ein exploratives Verfahren mit dem Ziel einer Zuordnung von Objekten zu Gruppen sind, sodass sich das Heterogenitätsmaß am wenigsten vergrößert; gleichzeitig sollen die Cluster gut voneinander trennbar sein (Bortz, 2010; Härdle & Simar, 2015). Hier ist anzumerken, dass die Clusterbildung sowohl von der Wahl des Proximitätsmaßes abhängt, als auch einer Festlegung der optimalen Clusterzahl bzw. einer Entscheidung darüber bedarf, an welcher Stelle der Verschmelzungsprozess abgebrochen werden soll. Hierzu gibt es kein festgelegtes Abbruchkriterium. In der vorliegenden Studie wurden für die Bestimmung der Clusterzahl das Elbow-Kriterium und das Dendrogramm herangezogen sowie die Clusterbildung anschließend über eine K-Means-Clusterung optimiert. Allerdings wären auch andere Möglichkeiten hinsichtlich der Heterogenitätsmaße denkbar. Zudem muss sich die Clusteranalyse die Kritik gefallen lassen, von der Stichprobe abhängig zu sein. Ein an die vorliegende Studie anschließendes Forschungsdesiderat kann daher eine Weiterarbeit an den ermittelten Clustern sein.

Weitere Einschränkungen betreffen die limitierte Größe sowie die Eigenschaften der Stichprobe. Da eine Teilnahme anonym erfolgte, ist nicht nachvollziehbar, an welchen Schulen die teilnehmenden Lehrkräfte genau unterrichteten, um detaillierter und mit Hilfe komplexerer statistischer Modelle Ähnlichkeiten bezüglich der individuellen schulischen Bedingungen berücksichtigen zu können. Zudem folgen die Schularten in der Stichprobe keiner gleichmäßigen Verteilung und die

Stichprobe ist insgesamt klein, was die Verallgemeinerbarkeit der Ergebnisse einschränkt. Ein Forschungsdesiderat stellt daher der Einbezug von personalen Variablen dar; beispielsweise ist Selbstwirksamkeit als Einflussfaktor insbesondere auf das Belastungserleben von Lehrenden beschrieben (Dizinger et al., 2011; Kunnari, Ilomäki & Toom, 2018). Zudem erlauben die Analysen keine kausalen Schlüsse zum Verhältnis von Kooperation und Belastungserleben. Auch hier ließen sich weiterführende Forschungsarbeiten ansetzen. Die erfassten Dimensionen von Kooperation und Belastung ließen sich beispielsweise in Strukturgleichungsmodellen überprüfen, um mögliche Zusammenhänge herstellen zu können, aber auch den Einfluss möglicher moderierende/mediierender Variablen überprüfen zu können.

Literatur

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders. DSM-5* (5. Aufl.). Washington, DC: American Psychiatric Association.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2016). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (14., überarb. u. akt. Aufl.). Berlin: VS.
- Berkemeyer, J., Berkemeyer, N. & Schwikal, A. (2015). Lernen als Leitbild. Internationale Erfahrungen zum Schulleitungshandeln im Kontext von Professionalisierungsprozessen von Lehrkräften. In J. Berkemeyer, N. Berkemeyer & F. Meetz (Hrsg.), *Professionalisierung und Schulleitungshandeln* (S. 12–32). Weinheim: Beltz.
- Bertrand, L., Roberts, R. A. & Buchanan, R. (2006). Striving for success: Teacher perspectives of a vertical team initiative. *National Forum of Teacher Education Journal*, 16, 1–10.
- Böhm-Kasper, O. (2004). *Schulische Belastung und Beanspruchung*. Münster: Waxmann.
- Böhm-Kasper, O., Dizinger, V. & Gausling, P. (2016). Multiprofessional collaboration between teachers and other educational staff at German all-day schools as a characteristic of today's professionalism. *International Journal for Research on Extended Education*, 4(1), 29–51.
- Böhm-Kasper, O., Dizinger, V. & Heitmann, V. (2013). Interprofessionelle Kooperation an offenen und gebundenen Ganztagsgrundschulen. *Zeitschrift für Grundschulforschung*, 6, 53–68.
- Bonsen, M. & Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern. *Zeitschrift für Pädagogik*, 52(2), 167–184.
- Bortz, J. (2010). *Statistik für Human- und Sozialwissenschaftler* (7., überarb. und akt. Aufl.). Heidelberg: VS.
- Bovbjerg, K. M. (2006). Teams and collegiality in educational culture. *European Educational Research Journal*, 5, 244–253.
- Cheng, L. P. & Ko, H. (2009). Teacher-team development in a school-based professional development program. *The Mathematics Educator*, 19, 8–17.
- Clement, M. & Vandenberghe, R. (2000). Teachers' professional development: a solitary or collegial (ad)venture? *Teaching and Teacher Education*, 1(1), 81–101.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Taylor & Francis.

- Cramer, C. & Binder, K. (2015). Zusammenhänge von Persönlichkeitsmerkmalen und Beanspruchungserleben im Lehramt. Ein internationales systematisches Review. *Zeitschrift für Erziehungswissenschaft*, 18(1), 101–123.
- Da Fonte, M. A. & Barton-Arwood, S. M. (2017). Collaboration of general and special education teachers: perspectives and strategies. *Intervention in School and Clinic*, 53(2), 99–106.
- Dalin, P. (1999). *Theorie und Praxis der Schulentwicklung*. Neuwied: Luchterhand.
- DeMatthews, D. E. (2014). Principal and teacher collaboration: An exploration of distributed leadership in professional learning communities. *International Journal of Educational Leadership and Management*, 2(2), 176–206.
- Dickenberger, D., Gniech, G. & Grabitz, H.-J. (2001). Die Theorie der psychologischen Reaktanz. In M. Irle & D. Frey (Hrsg.), *Theorien der Sozialpsychologie. Band I. Kognitive Theorien* (2., vollst. überarb. u. erw. Aufl., S. 243–275). Bern: Huber.
- Dilling, H., Mombour, W. & Schmidt, M. H. (2014). *Internationale Klassifikation psychischer Störungen ICD-10 Kapitel V (F); klinisch-diagnostische Leitlinien* (9. Aufl. unter Berücksichtigung der Änderungen entsprechend ICD-10-GM 2014). Bern: WHO.
- Dizinger, V., Fussangel, K. & Böhm-Kasper, O. (2011). Lehrer/in sein an der Ganztagschule: Neue Kooperationsanforderungen – neue Belastungen? *Zeitschrift für Erziehungswissenschaft*, 14(3), 43–61.
- Doppenberg, J. J., Bakx, A. W. E. A. & den Brok, P. J. (2012). Collaborative teacher learning in different primary school settings. *Teachers and Teaching: Theory and Practice*, 18, 547–566.
- Egodawatte, G., McDougall, D. & Stoilescu, D. (2011). The effects of teacher collaboration in Grade 9 applied mathematics. *Educational Research for Policy and Practice*, 10, 189–209.
- Fabel-Lamla, M. (2018). Vertrauen in der interprofessionellen Zusammenarbeit an Schulen. *Journal für Schulentwicklung*, 22(1), 25–30.
- Forte, A. M. & Flores, M. A. (2014). Teacher collaboration and professional development in the workplace: A study of Portuguese teachers. *European Journal of Teacher Education*, 37(1), 91–105.
- Fussangel, K. (2008). *Subjektive Theorien von Lehrkräften zur Kooperation. Eine Analyse der Zusammenarbeit von Lehrerinnen und Lehrern in Lerngemeinschaften*. Unveröffentlichte Dissertation, Universität Wuppertal. Zugriff am 02.07.18 unter <http://elpub.bib.uni-wuppertal.de/servlets/DerivateServlet/Derivate-1129/dgo802.pdf>
- Fussangel, K. (2018). Lehrerkooperation – ein Thema nur für die Praxis oder bereits für die Lehrerbildung? *Seminar*, 24(2), 5–15.
- Fussangel, K., Dizinger, V., Böhm-Kasper, O. & Gräsel, C. (2010). Kooperation, Belastung und Beanspruchung von Lehrkräften an Halb- und Ganztagschulen. *Unterrichtswissenschaft*, 38(1), 51–67.
- Fussangel, K. & Gräsel, C. (2012). Lehrerkooperation aus Sicht der Bildungsforschung. In T. S. Idel, H. Ullrich & E. Baum (Hrsg.), *Kollegialität und Kooperation in der Schule. Theoretische Konzepte und empirische Befunde* (S. 29–40). Wiesbaden: VS.
- Fussangel, K. & Gräsel, C. (2014). Forschung zu Kooperation im Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. u. erw. Aufl., S. 846–864). Münster: Waxmann.
- Goddard, Y. L., Goddard, R. D. & Tschannen-Moran, M. (2007). A theoretical and empirical investigation of teacher collaboration for school improvement and student achievement in public elementary schools. *Teachers College Record*, 109(4), 877–896.
- Gräsel, C., Fussangel, K. & Pröbstel, C. (2006). Lehrkräfte zur Kooperation anregen – eine Aufgabe für Sisyphos? *Zeitschrift für Pädagogik*, 52(2), 205–219.

- Harazd, B. & Drossel, K. (2011). Formen der Lehrerkooperation und ihre schulischen Bedingungen. Empirische Untersuchung zur kollegialen Zusammenarbeit und Schulleitungshandeln. *Empirische Pädagogik*, 25(2), 145–160.
- Härdle, W. K. & Simar, L. (2015). *Applied multivariate statistical analysis* (4. Aufl.). Berlin: VS.
- Huber, S. G. (2012). Zwölf Thesen zur guten Schulleitung. *SchulTrends. Sonderausgabe der Schulverwaltung*, 1, 8–11.
- Huber, S. G., Ahlgrimm, F. & Hader-Popp, S. (2012). Kooperation in und zwischen Schulen sowie mit anderen Bildungseinrichtungen: Aktuelle Diskussionsstränge, Wirkungen und Gelingensbedingungen. In S. G. Huber & F. Ahlgrimm (Hrsg.), *Kooperation. Aktuelle Forschung zur Kooperation in und zwischen Schulen sowie mit anderen Partnern* (S. 323–372). Münster: Waxmann.
- Johnson, B. (2003). Teacher collaboration: Good for some, not so good for others. *Educational Studies*, 29(4), 337–350.
- Joiko, K., Schmauder, M. & Wolff, G. (2010). *Psychische Belastung und Beanspruchung im Berufsleben: Erkennen – Gestalten*. Zugriff am 02.07.18 unter https://www.baua.de/DE/Angebote/Publikationen/Praxis/A45.pdf?__blob=publicationFile
- Klieme, E., Baumert, J. & Schwippert, K. (2000). Schulbezogene Evaluation und Schulleistungsvergleiche. Eine Studie im Anschluss an TIMSS. In H.-G. Rolff, W. Bos, K. Klemm, H. Pfeiffer & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung* (Bd. 11, S. 387–420). Weinheim: Beltz.
- Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O. & Baumert, J. (2008). Engagement and emotional exhaustion in teachers: Does the school context make a difference? *Applied Psychology: An International Review*, 57, 127–151.
- Krohne, H. W. (2010). *Psychologie der Angst. Ein Lehrbuch*. Stuttgart: Kohlhammer.
- Kunnari, I., Ilomäki, L. & Toom, A. (2018). Successful teacher teams in change: The role of collective efficacy and resilience. *International Journal of Teaching and Learning in Higher Education*, 30(1), 111–126.
- Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91(4), 509–536.
- Lomos, C., Hofman, R. H. & Bosker, R. J. (2011). Professional communities and student achievement – a meta-analysis. *School Effectiveness and School Improvement*, 22(2), 121–148.
- Lortie, D. C. (1975). *Schoolteacher – a sociological study*. Chicago: The University of Chicago Press.
- Lovibond, P. F. & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335–343.
- McElvany, N., Gebauer, M. M., Bos, W. & Holtappels, H. G. (Hrsg.). (2013). *Sprachliche, kulturelle und soziale Heterogenität in der Schule als Herausforderung und Chance der Schulentwicklung*. Weinheim: Beltz.
- Morgenroth, S. & Buchwald, P. (2015). Burnout und Ressourcenerhaltung bei Lehrkräften. *Unterrichtswissenschaft*, 43(2), 136–149.
- Nilges, P. & Essau, C. (2015). Die Depressions-Angst-Stress-Skalen. *Der Schmerz*, 6, 649–657.
- Prilleltensky, I., Neff, M. & Bessell, A. (2016). Teacher stress: What it is, why it's important, how it can be alleviated. *Theory Into Practice*, 55(2), 104–111.
- Pröbstel, C. H. & Soltau, A. (2012). Wieso Lehrkräfte (nicht) kooperieren – Die Bedeutung „personaler Faktoren“ in der Zusammenarbeit am Arbeitsplatz Schule. In T. S. Idel, H. Ullrich & E. Baum (Hrsg.), *Kollegialität und Kooperation in der Schule. Theoretische Konzepte und empirische Befunde* (S. 55–76). Wiesbaden: VS.

- Reeves, P. M., Pun, W. H. & Chung, K. S. (2017). Influence of teacher collaboration on job satisfaction and student achievement. *Teaching and Teacher Education*, 67, 227–236.
- Richter, D. & Pant, H.-A. (2016). *Lehrerkooperation in Deutschland. Eine Studie zu kooperativen Arbeitsbeziehungen bei Lehrkräften der Sekundarstufe I*. Zugriff am 02.07.2018 unter <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/lehrerkooperation-in-deutschland/>
- Ronfeldt, M., Owens Farmer, S., McQueen, K. & Grissom, J. A. (2015). Teacher collaboration in instructional teams and student achievement. *American Educational Research Journal*, 52, 475–514.
- Rothland, M. (2013). Soziale Unterstützung. Bedeutung und Bedingungen im Lehrerberuf. In M. Rothland (Hrsg.), *Belastung und Beanspruchung im Lehrerberuf. Modelle, Befunde, Interventionen* (2., vollst. überarb. Aufl., S. 232–250). Wiesbaden: VS.
- Schäfer, T. (2016). *Methodenlehre und Statistik. Einführung in Datenerhebung, deskriptive Statistik und Inferenzstatistik*. Wiesbaden: VS.
- Schonfeld, I. S. (1992). Assessing stress in teachers: Depressive symptoms scales and neutral self-reports of the work environment. In J. C. Quick, L. R. Murphy, J. Joseph & J. R. Hurrell (Hrsg.), *Stress & well-being at work: Assessments and interventions for occupational mental health* (S. 270–285). Washington, DC: American Psychological Association.
- Scruggs, T. E., Mastropieri, M. A. & McDuffie, K. A. (2007). Co-teaching in inclusive classrooms: A metasynthesis of qualitative research. *Exceptional Children*, 73(4), 392–416.
- Soltau, A. & Mienert, M. (2009). Teamorientierung und Einstellungen zu Formen der Lehrerkooperation bei Lehrkräften. *Psychologie in Erziehung und Unterricht*, 56(3), 213–223.
- Steinert, B., Klieme, E., Maag Merki, K., Döbrich, P., Halbheer, U. & Kunz, A. (2006). Lehrerkooperation in der Schule: Konzeption, Erfassung, Ergebnisse. *Zeitschrift für Pädagogik*, 52(2), 185–204.
- Vangrieken, K., Dochy, F., Raes, E. & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, 15, 17–40.
- Vangrieken, K., Grosemans, I., Dochy, F. & Kyndt, E. (2017). Teacher autonomy and collaboration: A paradox? Conceptualising and measuring teachers' autonomy and collaborative attitude. *Teaching and Teacher Education*, 67, 302–315.
- Visscher, A. J. & Witziers, B. (2004). Subject departments as professional communities? *British Educational Research Journal*, 30, 785–800.
- Weiß, S., Markowetz, R. & Kiel, E. (2017). Multiprofessionelle Kooperation inner- und außerschulischer Akteure im Förderschwerpunkt Geistige Entwicklung – Die Sicht von Lehrkräften. *Zeitschrift für Heilpädagogik*, 68(7), 316–328.
- Wissinger, J. (2014). Schulleitung und Schulleitungshandeln. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. u. erw. Aufl., S. 144–176). Münster: Waxmann.

Ilonca Hardy, Jasmin Decristan & Eckhard Klieme

Adaptive teaching in research on learning and instruction

Abstract

In education, adaptive teaching is commonly viewed as adjusting instruction to students' individual differences in abilities, motivation, and linguistic background. According to Corno (2008), adaptive teaching involves the adaptation of instruction on a macro level and a micro level, using methods of differentiating instruction. Despite of these broad defining properties, there is no consensus on how to assess the fit to individual learners' needs and the effects of successful adaptive teaching, as they largely depend on the specific theoretical and methodological considerations involved. In this paper, we systematize different approaches to adaptive teaching on a conceptual level. Further, we summarize the main approaches to investigating the effects of adaptive teaching and discuss the respective results. We distinguish between studies on overall effects of adaptive teaching, on ATI effects and differential effects, and studies on effects of within-class variation on student achievement. Exemplary results are highlighted to illustrate the different methodological approaches. The paper ends with implications for theoretical clarification and empirical investigation.

Keywords

Adaptive teaching; Scaffolding; Classroom process quality; Teacher expertise

Prof. Dr. Ilonca Hardy (corresponding author), Faculty of Educational Sciences, Goethe-University, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main, Germany
e-mail: hardy@em.uni-frankfurt.de

Prof. Dr. Jasmin Decristan, School of Education, Bergische University Wuppertal, Gaußstraße 20, 42119 Wuppertal, Germany
e-mail: decristan@uni-wuppertal.de

Prof. Dr. Eckhard Klieme, German Institute for International Educational Research (DIPF), Department of Educational Quality and Evaluation, Rostocker Straße 6, 60323 Frankfurt am Main, Germany & Faculty of Educational Sciences, Goethe-University, Theodor-W.-Adorno-Platz 6, 60629 Frankfurt am Main, Germany
e-mail: klieme@dipf.de

Adaptiver Unterricht in der Lehr-Lernforschung

Zusammenfassung

Im erziehungswissenschaftlichen Kontext wird adaptiver Unterricht als Anpassung von Lehrkräftenhandeln an die individuellen kognitiven, motivationalen und sprachlichen Voraussetzungen von Schülerinnen und Schülern verstanden. Nach Corno (2008) beinhaltet dies Anpassungen an die Voraussetzungen der Lernenden auf der Makroebene und der Mikroebene des Unterrichts, so dass der Einsatz von differenzierenden methodischen Arrangements erforderlich ist. Jenseits dieser breiten Unterscheidungen gibt es jedoch keinen Konsens hinsichtlich der Erfassung einer gelungenen Passung und deren Wirkungen, da diese auf entsprechenden theoretischen und methodologischen Annahmen basieren. In diesem Beitrag fassen wir die unterschiedlichen Konzeptualisierungen des adaptiven Unterrichts sowie die wesentlichen methodischen Herangehensweisen bei der Untersuchung von Adaptivität zusammen. Wir unterscheiden dabei zwischen Studien zu Gesamteffekten, ATI-Studien bzw. Studien mit einer differenziellen Perspektive und Studien zur Variabilität von Schülerleistungen. Beispielhaft werden Ergebnisse vorgestellt, welche die unterschiedlichen methodischen Herangehensweisen verdeutlichen. Der Beitrag endet mit Implikationen für die Konzeptualisierung und empirische Erfassung von Adaptivität im Unterricht.

Schlagworte

Adaptivität; Scaffolding; Unterrichtsqualität; Lehrerexpertise

1. Introduction

Educational contexts typically involve students from a variety of social and linguistic background with differing cognitive, motivational, and self-regulatory resources. Recently, the consideration of individual student needs has reemerged as a major issue in theoretical, empirical, and practice-oriented work, especially in contexts with increased student heterogeneity such as inclusive educational settings. In educational research, teachers' adjustments to students' individual developmental states have been repeatedly considered a core element of effective teaching (for an overview see Parsons et al., 2018). As Corno (2008) stated, these adaptations may refer either to the macro level of instruction, as in planned programs for groups of similar students, differentiated learning material and tasks, or they may refer to the micro level of instruction, as in contingent support, on-going diagnosis and didactical moves. Interestingly, although intuitively appealing, the theoretical considerations as well as the methodological approaches regarding the construct of adaptive teaching are by no means unitary. For example, Parsons (2008) refers to adaptive teaching as "teacher action that (a) is non-routine, proactive, thoughtful, and improvisational; (b) includes a change in professional knowledge or practice;

and (c) is done to meet the needs of a student or an instructional situation” (p. 20). According to this definition, beyond a teacher’s acknowledgement of individual student needs, teachers’ professional knowledge, beliefs, and skills are considered prerequisites of adaptive teaching. Furthermore, the notion of adaptive teaching as non-routine and flexible is in line with Glaser’s (1972) early view of adaptive teaching as “a wide range and variety of instructional methods and opportunities for success” (p. 6). The definition of adaptive teaching put forward by Corno (2008), in contrast, refers to pedagogical choices in terms of the degree of instructional support given with respect to different levels of student ability, thus focusing on teachers’ considerations in instructional design and its implementation (macro-adaptivity) and in teacher-student interaction (micro-adaptivity). Weinert and Helmke (1997) conceptualize adaptive teaching as a meta-category of effective teaching and instructional quality, referring to active and proactive teacher behavior in terms of meeting individual student needs. Similarly, Walberg and Paik (2000) list “adaptive education” as one of ten categories of effective educational practices.

Besides these definitions of adaptive teaching, in a recent review of empirical literature published between 1975 and 2014 Parsons et al. (2018) identified different terminologies and theoretical foundations of adaptive teaching. They showed that in the early literature, teachers’ decision making was focused on, while literature between 1995 and 2004 often referred to scaffolding and teacher reflection. The term of adaptive teaching was not used until 2008. When applying these conceptualizations in research and practice, major questions arise: How may we assess adaptive teaching on different levels of analysis? How may we judge the degree of success of adaptive teaching for individual students? What are the relations between adaptive teaching and teachers’ professional competence? Which methodological approaches are useful for investigating outcomes of adaptive teaching? In this paper, we will review the core literature on adaptive teaching, point to aspects in need of theoretical clarification, and highlight promising analytical approaches.

2. Adaptive teaching: A priori considerations

According to Parsons et al. (2018), adaptive teaching is considered to be “socially constructed as teachers metacognitively reflect on students’ needs before, during, and after instruction” (p. 209). Adaptive teaching thus may be regarded a type of social practice of reflective teachers in classroom settings that enables adaptations to students’ individual differences and learning needs. In their review, they identified teacher factors relevant for adaptive teaching such as beliefs, experiences, knowledge, and thinking. They also identified affordances for adaptive teaching in educational settings such as the instructional approaches used and the assessment practices involved. Socio-constructivist views presume an intricate relation between the affordances for co-construction of knowledge in social or cultural contexts and individual cognitive development (Brown, Collins, & Duguid, 1989; Greeno, 1998;

Lave & Wenger, 1991; Rogoff, Baker-Sennett, Lacasa, & Goldsmith, 1995; Vygotsky, 1986). In educational settings, affordances include the physical and social resources within a learning environment, thus leaving teachers with a variety of options in instructional design to provide learning opportunities, and leaving learners with a variety of options for their enactment.

Against this theoretical background and based on the categories identified by Parsons et al. (2018), we put forward some preliminary considerations on adaptive teaching before reviewing the literature and correspondent methodological approaches for assessing adaptive teaching. We suggest to differentiate between *intended* adaptive teaching and *implemented* adaptive teaching, presuming that teachers vary in the degree to which their instructional design and planned learning activities allow for adaptations to student needs in the first place. Intended adaptive teaching thus refers to teachers' pedagogical intentions in instructional design and is primarily based on their use of diagnostic tools and their a priori considerations of students' learning prerequisites. Implemented adaptive teaching refers to adaptive teaching episodes during classroom instruction in which teacher-intended activities are actually taken up by the students; in this sense, there is an alignment of intention and situational enactment. Successful adaptive episodes will in the long run support self-regulated learning of students within educational settings (Corno, 2008). In-situ use of diagnostic tools and strategies of questioning and feedback support teachers in their implementation of adaptive teaching on a micro-level. These considerations of adaptive teaching also encompass a causal impact of adaptive teaching episodes on related student outcomes.

Consideration 1: Teacher decisions in instructional design, including the assembly of tasks, materials, and instructional methods, are conceptualized as *intended* adaptive teaching if they refer to teacher actions or instructional choices that are based on (formally or informally) diagnosed individual student needs and learning states (cf. Parsons, 2008). These teacher actions may involve decisions with respect to individualized student support with respective material and verbal supports, methods of differentiated instruction for groups of students, or long-term supports within a classroom community. Along these lines, Randi and Corno (1997) describe adaptive teaching as efficient for group instruction and, at the same time, respectful to individual learning profiles and patterns. As students are to achieve increasing autonomy by taking on responsibility of their learning, intended adaptive teaching will also incorporate goals of self-regulated learning (Corno, 2008). This implies that teachers teach students strategies for self-regulated learning and choose tasks in a way to enable increasing student autonomy, for example by fading scaffolds over time. Intended adaptive teaching also means that teachers are able to verbalize and explain their instructional decisions either in lesson planning or in post-hoc reflections, linking decisions to diagnostic information and student learning progressions.

Consideration 2: As students are active learners, features of teacher-designed learning environments may be perceived by students and taken up as intended, but they may also be transformed and re-interpreted by students. This corresponds

to the notion of individual learners' perception of social and physical affordances within learning environments and their respective acting on those. *Implemented* adaptive teaching means that an alignment of the intended adaptive design, respective in-situ didactical moves on the part of the teacher and an uptake of this learning environment on the part of students are observed. With respect to the level of analysis, implemented adaptive teaching may be segmented into adaptive teaching episodes of different grain sizes. Adaptive teaching episodes will involve evidence of teachers' on-going diagnosis and respective moves such as questioning, prompting, explaining or giving feedback. Similarly, Parsons, Dodman, and Burrowbridge (2013) propose that for differentiating instruction to be implemented successfully, student individual learning paths on a micro level need to be considered.

Consideration 3: If students' interpretations and actions match with a teacher's proposed learning needs on a macro level and a micro level of instruction, this alignment should eventually be evident in successful student learning trajectories, i.e., individual learning outcomes, developmental patterns, or in-situ evidence of learning.

Consideration 4: Given these constraints, we may then ask for facets of teacher professional competence associated with the frequency of adaptive teaching episodes such as teacher knowledge of individual differences, diagnostic competencies, frequency of reflection and meta-cognition, self-efficacy, and epistemological beliefs (Beck et al., 2008; Parsons et al., 2018).

In the following sections, we elaborate on elements of adaptive teaching that refer to the four considerations listed above. We consider the diagnosis of students' individual learning prerequisites and developmental states as a prerequisite of adaptive teaching under consideration 1. Similarly, we consider research approaches on adaptive teaching on a macro level and a micro level of instruction as relevant to the distinction between intended and implemented adaptive teaching (cf. considerations 1 and 2), although this distinction is rarely pursued explicitly. After summarizing research on teacher professional competence related to adaptive teaching (cf. consideration 4), we suggest methodological approaches to assessing successfully implemented adaptive teaching episodes (cf. consideration 3).

3. Review of research on aspects of adaptive teaching

3.1 Taking individual differences into account

A prerequisite for adaptive teaching episodes is teachers' consideration of relevant individual student characteristics prior to and/or during instruction. To describe interindividual learner variation, different categorizations of learner characteristics have been suggested. For instance, Heinzl (2008) defines five dimensions of student heterogeneity: socio-economic status, ethnicity or cultural background, abilities, gender, and generation/age – all of which have been related to students'

long-term academic success in empirical research. In addition, Hamre and Pianta (2005) distinguish between two central risk categories: demographic risks and functional risks of school failure. Demographic risks involve dimensions related to family background (i.e., socio-economic status, ethnicity or cultural background). Functional risks refer to individual student characteristics such as executive functioning, cognitive, behavioral, or domain-specific disorders, as well as students' varying ability to perform the adaptations required for social learning processes in the classroom (see also Hasselhorn, Andresen, Becker, Betz, Leuzinger-Bohleber, & Schmid, 2015). Beyond these categories frequently related to academic careers in large-scale studies and comparative group designs, instructional design is based on learner characteristics on an affective, motivational, cognitive, and social level. Along these lines, Tomlinson et al. (2003) list student readiness, interest and motivation, and learning profiles (including intelligence, gender, and culture) as relevant student characteristics for differentiating instruction.

To what extent do teachers perceive of and act on the heterogeneous student population within their classes? In order to enable the type of intended adaptive teaching described in the last section, teachers need to apply methods of diagnosing students' current understanding and preconditions for further learning. These methods include diagnostic instruments such as standardized achievement tests, language proficiency tests, and screenings as well as the ongoing diagnosis of student understanding using formative assessment tools. Formative assessment refers to the repeated use of assessment-based information to recognize and respond to student learning for the purpose of fostering development (Bell & Cowie, 2001; Black & William, 1998; Kingston & Nash, 2011), thus aligning assessment, instruction, and learning goals (Wilson & Sloane, 2000). It has been successfully used in secondary science and math education (Furtak, Morrison, & Kroog, 2014; Rakoczy, Harks, Klieme, Blum, & Hochweber, 2013) and recently in elementary science education to further students' conceptual understanding (Decristan, Hondrich et al., 2015). According to van de Pol, Volman, and Beishuizen (2012), the diagnosis of students' understanding is a prerequisite for tailored support – the type of support that has also been described within the context of formative assessment practices. Formative assessment has in fact been conceived of as a cycle starting with diagnostic strategies for assessing students' current understanding, followed by the validation of this estimate, and the use of this information for instructional task design and intervention (Ruiz-Primo & Furtak, 2007). Beyond assessment practices to gauge students' individual preconditions and current levels of understanding, teachers' overall judgment of within-class variability in student achievement may affect the way in which teachers respond to individual learners' needs in instruction. It is in this sense that adaptive teaching has been characterized as “deeply psychological” (Corno, 2008, p. 163). Corno suggests that teachers vary in the degree to which they view learner variation as “obstacles to be overcome” rather than as “opportunities for learning”. Similarly, Prengel (2006) argues that within educational settings, differences between students are to be acknowledged without accompanying value judgments, evident in patterns of supportive teacher-student

interactions. Corno proposes that teachers who view learner variation as obstacles tend to deal with students' achievement heterogeneity by using homogenous grouping or individualized instruction. Teachers who view learner variations as opportunities for learning tend to provide challenging and supportive learning environments that enable all students to profit from each other. Thus, while adaptive teaching is regarded as a way to recognize and (proactively) act on student individual differences in instruction, it needs to be kept in mind that instruction is a dynamic process, and adaptation is by no means a fixed parameter referring to single students. In this sense, the long-term goals of instructional adaptations include the creation of learning spaces for individual learners within a "community of learners", capitalizing on the strengths of heterogeneous groups, and thereby supporting the self-regulated learning of all students (Behrensen, Gläser, & Solzbacher, 2015; Corno, 2008).

3.2 Adaptive teaching on a macro level

Instructional adaptations include practices on a macro level and on a micro level of adaptation. While micro adaptations are those decisions that teachers make on a moment-to-moment basis in classroom instruction, Corno and Snow (1986) describe macro adaptations as rather large-scale adjustments in instruction, informed by formal assessments. Similarly, Corno (2008) and Klieme and Warwas (2011) refer to macro adaptations as structured programs that have been designed for students with similar individual capabilities such as in gifted education or in instruction for students with limited language proficiency, pursuing the goal of individual enhancement and prevention of further risks for educational careers. There is a large body of empirical evidence on the effectiveness of early intervention programs such as instruction of literacy and numeracy if these programs are implemented with high quality over an extended amount of time (e.g., Hasselhorn & Kuger, 2014; Weiland & Yoshikawa, 2013). Likewise, practices of pull-out or partial pull-out programs are found in many countries as a means to meet individual learner needs within homogeneous grouping approaches (for a comparison of program effects for students with cognitive disabilities see Marston, 1996). Beyond these programs, teachers' instructional decision-making concerns formats of differentiating instruction within classroom communities. According to Tomlinson et al. (2003), differentiating instruction includes teaching in a proactive way, where teachers modify curricula, teaching methods, or resources to create optimal learning environments for all students. Overall, three aspects of differentiating instruction may be distinguished. On an organizational level, students may be assigned to subgroups according to different methods, media, content, or forms of social interaction. On a didactical level, students may be provided with different learning material and tasks according to individual interest, motivation, level of competency, or cognitive ability. On a level of instructional design, students may take part in

long-term activities of self-regulated learning such as individualized schedules or methods of cooperative learning (overview in Haag & Streber, 2014). For example, task differentiation may involve the assignment of groups of students with different prior knowledge and/or self-regulatory capabilities to tasks with different levels of content-related learning goals (difference of learning goals) and a more or less structured sequence of solution steps toward the same learning goal (indifference of learning goals). Yet, teachers report that they employ formats of differentiating instruction only rarely, both in elementary and in secondary education (see Tomlinson et al, 2003). In Germany, differentiation via additional time rather than via variation in task difficulty seems to be most prominent (Bos, Hornberg, Bonsen, & Buddeberg, 2008).

Following Corno (2008), teacher practices of differentiation are successful if they capitalize on the variability of student learning prerequisites (cf. Parsons et al., 2018). That is, instructional contexts that take into account learners as part of a community of learners, intending to move them towards a common core (or “the center”) rather than to compensate for individual weaknesses will eventually serve a pedagogical goal of adaptivity. Therefore, one challenge to teachers is the coordination of individual and social activities that effectively consider individual students’ needs. According to the socio-constructivist view, the co-construction of knowledge in discourse is pivotal for individual cognitive development. Arrangements such as structured cooperative learning activities may then serve a two-fold goal of individual participation and productive use of heterogeneity in student achievement. For example, peer tutoring has proven to be an effective method especially with heterogeneous ability grouping of students, leading to persistent effects in social, motivational, and achievement measures (Ginsburg-Block, Rohrbeck, & Fantuzzo, 2006; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003). It has been defined as “a class of practices and strategies that employ peers as one-to-one teachers to provide individualized instruction, practice, repetition, and clarification of concepts” (Utley & Mortweet, 1997, p. 9). Presumably, the processes of co-construction of knowledge, the fulfilment of basic needs of autonomy, competency, and social inclusion, and the corresponding social attractiveness of peers lead to positive effects (e.g., Fuchs, Fuchs, Mathes, & Martinez, 2002; Jordan & Métais, 1997). This line of research also points to the necessity to regard adaptive teaching as a long-term process in which students are slowly enculturated into practices of increasing autonomy before positive effects on student outcomes emerge. For example, Adl-Amini (2018) found that only after an extended phase of implementation of peer tutoring in elementary science education, student conceptual understanding increased. Furthermore, only in classes with high levels of cognitive ability was peer learning implemented with a high degree of adherence to the program goals. Apparently, teachers of students with low cognitive ability preferred controlled environments with low degrees of student autonomy so that they were not inclined to implement peer-learning activities. Along these lines, Corno (2008) proposes two continua, the support continuum and the learners’ abilities continuum, which are co-considered when deciding on appropriate methods of in-

struction. Yet, the degree to which methods of differentiation employed on a macro level actually corresponds to students' individual preconditions is rarely investigated in detail; rather it is the outcomes of students' assignment to grouped formats of differentiation that is of interest.

3.3 Adaptive teaching on a micro level

The micro level focuses on teachers' processes of adaptive support on a moment-to-moment basis. One may distinguish between a contingency perspective, which is based on the construct of scaffolding, and a perspective of instructional quality, emphasizing relevant instructional processes on a classroom level. Reiser (2004) points to the difficulty of achieving an optimal level of instructional support for all learners. To realize these supports, teachers need to track individual students' content understanding and proactively structure learning activities in instructional task design. Here, learning may be conceptualized in terms of "multiple zones of proximal development" (cf. Palincsar, 1998), including supports by verbal interaction with the teacher and peers, by models, artifacts, and other symbolic or computerized support systems.

3.3.1 Scaffolding and contingent support

The construct of scaffolding was introduced by Wood, Bruner, and Ross (1976) to describe those forms of adaptive support enabling learners to successfully work through tasks which they otherwise would not be able to perform. Whereas early work on scaffolding concentrated on tutorial situations with a tutor – usually the teacher – and a student, the construct has also been transferred to classroom teaching in theoretical considerations (e.g., Clark & Graves 2005; Hogan & Pressley, 1997; Stone 1998a, 1998b) and recent empirical work (e.g., Nathan & Kim, 2009; Smit, van Eerde, & Bakker, 2013; van de Pol, Volman, Oort, & Beishuizen, 2015). The main characteristics of scaffolding in tutorial situations have been put forward by Puntambekar and Hübscher (2005) and more recently by van de Pol, Volman, and Beishuizen (2010). They include the characteristics of contingent support, transfer of responsibility (fading), and use of diagnostic strategies. In addition, Puntambekar and Hübscher (2005) refer to the construction of shared understanding between tutor and tutee as a prerequisite for contingent support. Different means of scaffolding are distinguished. Among these are: feeding back, explaining, modeling, questioning (van de Pol, Volman, & Beishuizen, 2011), cueing specific or general elements, cueing specific strategies (Beed, Hawkins, & Roller 1991), verifying and clarifying understanding (Roehler & Cantlon, 1997), framing the goal, refocusing the discussion, attending to conflicts and differences, prompts for refinement of language (Hogan & Pressley, 1997), and reacting to errors as opportunities for learning (Wischgoll, Pauli, & Reusser, 2015). Parsons

et al. (2018) refer to questioning, encouraging, managing, giving feedback, making connections, assessing, modeling, explaining, and challenging as the most frequent types of teacher adaptations in the research they reviewed. On a higher-order level of classification, Pea (2004) employs the functions of focusing and modeling. Focusing points to teachers' structuring of content with the intention of focusing learners' attention to essential aspects of a task by prompts or by reducing degrees of freedom during task solution. Modeling means teachers' intentions to gradually enculturate learners into relevant domain-specific practices, including strategies such as the use of representational means, explication of solutions, and content-directed prompts. Reiser (2004) refers to structuring content and problematizing (of concepts, preconditions, models) as functions of scaffolding, whereas Krammer (2010) distinguishes between scaffolding strategies on a level of emotions (initiate and keep up student motivation to solve a task and to manage frustration), of procedures (structuring and managing student solution processes), and of content (providing prompts for relevant task features and solution models). These functions emphasize that scaffolding employs adaptations in a multidimensional way, including cognition, affect, and self-regulation, with the goal of continuous task-engagement and transfer of responsibility (cf. van de Pol et al., 2010).

When principles of scaffolding are transferred from individualized tutoring to classrooms, additional aspects are relevant. In this regard, the "2 sigma problem" raised by Bloom (1984) points to the empirically greater efficacy of one-to-one tutoring in comparison to learning situations in a classroom context – an issue rendering questionable the direct transfer of tutorial support principles to the classroom. It suggests that teachers will have to employ different methods in larger groups than in one-to-one tutoring in order to achieve positive outcomes for all students. Accordingly, when transferring scaffolding to the classroom, Hogan and Pressley (1997) refer to the consideration of multiple zones of development with large groups of students, their diverse communication styles, curriculum and time constraints, the need for student assessment, the ownership of ideas, and the uncertainty of endpoints in classroom discussions. Smit et al. (2013) add that the layered, distributed, and cumulative nature of scaffolded discourse in the classroom constitute the most distinctive elements. A prominent application of scaffolding to classroom contexts has been pursued by Hammond and Gibbons (2005) who use strategies of scaffolding to support students of English as a second language. In their approach, micro scaffolding refers to means of enriching classroom discourse to meet individual second language learners' needs. Macro scaffolding, in contrast, is concerned with teachers' planning of classroom activities based on the diagnosis of individual language competence and the formulation of language and content learning goals.

Apart from the means and functions of scaffolding (van de Pol et al., 2010), an important consideration concerns finding the appropriate level of support. According to Wood, Wood, and Middleton (1978), the *Contingent Shift Principle* shall determine whether a specific teacher action is contingent on student behavior: If a student does not solve a task, an increase of (teacher) control follows, with

control defined as the degree of explicit information and instruction provided. If a student does solve a task successfully, a decrease of (teacher) control follows (van de Pol et al., 2012). In the *Optimal Scaffolding Distance Parameter* (van Geert & Steenbeek, 2005) contingency is defined as support in relation to a student's current level of understanding. The optimal distance of support may be computed according to a rule assuming that learning effects tend toward zero if support is given at a distance too close or too distant from a student's current level of understanding. Beyond these theoretical distinctions, in a number of studies teachers' uses of scaffolding in instructional settings were analyzed (van de Pol et al., 2010). In regular classrooms, teachers' use of diagnostic strategies in discourse is generally low, therefore an adequate basis for the provision of contingent support may be lacking (Lockhorst, Wubbels, & van Oers, 2010; van de Pol et al., 2012). In other studies, teachers' adaptations in instructional classroom discourse and accompanying parameters of contingent support were investigated in detail. For example, Nathan and Kim (2009) found that teachers related their prompts and questions to prior assessment of students' mathematical understanding during class discussions. Van de Pol et al. (2015) found that in group work, teacher contingent support was most effective in instructional settings with a high degree of student independent work if the variable of student task effort was considered in the analyses. Van de Pol, Mercer, and Volman (2018) used a mixed methods design to investigate the extent to which students pick up teacher support in small group work and their effects on learning outcomes. Using a large sample of 35 lessons, their study employed both mediation analyses and exemplary in-depth description to pinpoint the role of contingency and fading of instructional support, revealing that uptake of students occurred especially with timely fading of support in contingent interactions. This study points to the necessity of further empirical evidence for the theoretically proposed mechanisms of a link between contingent support and learning within the scaffolding literature.

3.3.2 Processes of classroom instructional quality

Besides the tradition of scaffolding research, adaptive teaching may be considered a meta-category (Weinert & Helmke, 1997) or a specific component of effective, high quality instructional practice (Walberg & Paik, 2000). Thus, there is a partial convergence of adaptive teaching with constructs of teaching quality such as instructional support or supportive climate. Instructional support typically refers to the degree to which teachers aim to promote students' conceptual understanding, e.g., by employing teacher prompts to explore students' prior knowledge, using higher-order thinking tasks, and stimulating cognitive conflicts (Baumert et al., 2010; Lipowsky et al., 2009). When comparing classrooms with different degrees of instructional support using high inference ratings, one finds that these types of support are predictive of student cognitive learning outcomes (e.g., Fauth, Decristan, Rieser, Klieme, & Büttner, 2014; Kunter et al., 2013; Pianta & Hamre,

2009). These findings hold across different subjects (mainly mathematics, science) and age groups (secondary school, elementary school). Likewise, the construct of emotional support describes teacher behavior that takes student concerns seriously, provides constructive feedback, and takes student errors and misconceptions as learning opportunities, among others (e.g., Brophy, 2000; Hamre & Pianta, 2005; Klieme, Pauli, & Reusser, 2009). Empirical findings show an impact of emotional support on students' feelings of relatedness and their peer relationship, which in turn affects their interest in academic activities, engagement, and academic achievement (e.g., DeRosier, Kupersmith, & Patterson, 1994; Wentzel, Battle, Russell, & Looney, 2010). In contrast to definitions of scaffolding and the conditions for adaptive teaching episodes outlined above, the diagnosis of student preconditions in teacher action is taken up only implicitly in instruments of instructional quality. Accordingly, indicators of instructional quality aim at judging teachers' successful consideration of students' conceptual, linguistic, and motivational preconditions in classroom activities by high-inference ratings (cf. indicators used in the CLASS system; Hamre & Pianta, 2005).

The strong empirical findings within this perspective and the conceptual overlap with teacher scaffolding such as prompting, questioning, and explaining allude to a potential benefit of merging these process-oriented perspectives. For example, approaches to investigating adaptive teaching at the macro level such as forms of differentiating instruction or cooperative learning may be specified further by relating their implementation to a micro level of cognitive and emotional support. Decristan, Klieme, et al. (2015) showed an interaction between the use of formative assessment practices at the macro level and dimensions of classroom process quality with regard to student achievement. Students benefited most from teaching that combined formative assessment strategies with high levels of classroom process quality. A further combination of approaches is proposed by Howe (2013) in a framework for investigating peer learning activities with an emphasis on scaffolding on a micro-level, focusing on its potential to resolve cognitive conflicts between peers as a relevant condition for conceptual change. These types of analyses thus may provide insights into the quality of processes on a micro level necessary to successfully implement formats of adaptive teaching on a macro level.

4. Relation of adaptive teaching and teacher professional competence

As outlined above, the degree to which adaptive teaching is implemented in classrooms is likely associated with teacher professional competence. At least two lines of research are relevant: One prominent theoretical approach is based on Shulman's distinction of different aspects of professional knowledge associated with successful teaching. With regard to adaptive teaching, diagnostic knowledge and knowledge of assessment routines, pedagogical knowledge, epistemological be-

liefs of learning and teaching, beliefs on differentiating instruction, and beliefs of learner characteristics are likely relevant teacher characteristics (e.g., Beck et al., 2008; Dubberke, Kunter, McElvany, Brunner, & Baumert, 2008; Jordan, Glenn, & McGhie-Richmond, 2010; Jussim & Harber, 2005). There is empirical evidence of systematic relations between teacher beliefs, instructional arrangements employed, and student motivational and cognitive outcomes, accompanied by theoretical models of moderator variables on these relationships (Fives & Gill, 2014).

Another prominent approach conceives of teaching competence as flexible teaching (see also Beck et al., 2008; Glaser, 1972; Parson, 2008). Here, it is the degree to which teachers are able to deal with unplanned situations in the classroom that is conceptualized as a variable distinguishing between teachers, related to effectiveness of instruction (Fairbanks et al., 2010; Parsons, 2008, 2012). For example, Darling-Hammond and Bransford (2005) view adaptive expertise as teachers' ability to effectively establish routines and procedures in the classroom while also being able to deal with the complexity of classroom instruction. Parsons (2012) investigated teachers' adaptations in literacy instruction and compared two teachers' profiles selected from a larger sample of teachers. Based on observations, teachers' lesson plans, and interviews after each observed lesson, he found adaptations on the following levels of instruction: Modification of lesson objective, means by which objectives are met, inventing an example or analogy, inserting a mini-lesson, suggesting different solution procedures, omitting planned activities or assignments, inserting unplanned activities or assignments, changing the planned order of instruction. However, Parsons, Davis, Scales, Williams, and Kear (2010) showed that teachers' adaptations might not be as thoughtful as researchers have previously proposed.

5. Investigating the outcomes of adaptive teaching

5.1 Comparison of overall learning gains

An obvious approach to assessing adaptive teaching is to estimate its contribution to successful student learning outcomes. In fact, Oh (2005) suggests that with the construct of scaffolding, the success of the employed scaffold is already required *per definitionem*. It is especially in studies with experimental designs that these causal relations are investigated, ranging from laboratory experiments to quasi-experimental studies realized in actual school contexts. For example, Murphy and Messer (2000) investigated the effects of scaffolding activities with a balance beam by with five to seven-year-olds in a controlled laboratory setting. Individual pre- and posttests assessed children's development of conceptual knowledge in two conditions, varying in degree of explanations and prompts given in comparison to a control condition with individual work. They showed that scaffolding was superior in all of the investigated student outcomes except for an abstract nonverbal

task. In a teaching experiment with six intervened classrooms and three baseline classrooms, Hardy, Jonen, Möller, and Stern (2006) investigated the effects of instructional support (high/low) on third-graders' conceptual understanding. It was shown that instructional support in terms of structured instructional discourse and a sequenced curriculum was associated with long-term gains in conceptual knowledge. Similarly, in a sequence of experimental and quasi-experimental studies, Rakoczy et al. (2013) examined teachers' adoptions of different forms of formative assessment practices in secondary mathematics instruction.

In the recent cluster-randomized IGEL (Individual Support and Adaptive Learning Environments in Primary School) study with a total of 54 teachers and 1,070 third-grade students from 39 German primary schools three instructional conditions varying adaptive teaching methods were investigated. After intensive teacher professional development, the participating teachers employed the methods of scaffolding instructional discourse (SID), peer tutoring (PT), or formative assessment (FA) in their classes. In the SID classes, teachers guided students' learning by eliciting their preconceptions, by drawing attention to essential concepts, and by providing prompts within instructional discourse. In the PT classes, a reciprocal student tutoring approach was used, thus students switched the roles of the tutor (who explained the learning content) and tutee (who was assumed to learn from the tutor) according to given rules. This interaction was supported by material that included change-of-role-signs on the worksheets (for details see Adl-Amini, Decristan, Hondrich, & Hardy, 2014). In the FA groups, teachers used students' answers to diagnostic tasks to provide each student with an individual, written, and informative feedback on his or her current level of understanding and on his or her subsequent learning task on an appropriate level of difficulty (for details see Hondrich, Hertel, Adl-Amini, & Klieme, 2016). Results of this study showed that teachers in the FA classes were most successful in providing support for their students' conceptual development (Decristan, Hondrich, et al., 2015). However, it remains to be shown in detail how the intended adaptive teaching in the instructional design was actually implemented by the participating teachers to match students' learning needs. Among others, video data, transcribed verbal interactions, as well as statistical models with mediating and moderating variables may be employed in multi-method approaches to shed light on the proposed causal mechanisms implicit in the experimental design.

5.2 Methodological approaches within ATI research

Outcomes of adaptive teaching may be well investigated by methodological innovations within the aptitude-treatment-interaction (ATI) paradigm (see Cronbach & Snow, 1977). Within this paradigm it is assumed that learner characteristics and learning environments are differentially related to student learning outcomes. For example, Corno (1979) investigated the interaction of student cognitive and affective predispositions and teacher classroom behavior with respect to achieve-

ment. In retrospect, the early studies on ATI showed that only few interactions held across several contexts. Among the persisting results is the one that teaching approaches with high levels of structure and explicit instruction are beneficial for students with low abilities and from families with low socio-economic status (see Weinert & Helmke, 1997). Methodologically, these studies have several shortcomings. Most importantly, they did not consider appropriately that students are nested within classes. Furthermore, the ecological validity of many studies is low since treatments (teaching conditions) included highly artificial laboratory conditions. Recently, researchers renewed this strand of research. For example, in the IGEL study, Decristan, Hondrich, et al. (2015) analyzed the differential effects for students at risk within specific treatments: Using hierarchical linear modelling and the specification of cross-level-interactions (individual level: students' risk variable; classroom level: treatment), students with low language proficiency were shown to particularly benefit from teacher-scaffolded discourse (SID) as well as from FA. Treatments focusing on the diagnosis of student prior knowledge, on individualized feedback, and on structuring content knowledge thus provided extra support for conceptual growth of students at risk. Studies investigating differential instructional effects consistently show that at-risk students perform better in classrooms of high-quality instructional support than in classrooms of low instructional quality (e.g., Curby, Rimm-Kaufman, & Ponitz, 2009; Hamre & Pianta, 2005). Also, Decristan, Kunter, Fauth, Büttner, Hardy, and Hertel (2016) examined whether students with demographic risks (i.e., immigrant background) and functional risks (i.e., low cognitive abilities) benefit from high extents of process quality (i.e., classroom management and classroom climate/emotional support). The authors showed that in classes with high instructional quality, students from immigrant families scored higher than their peers. In addition, students with low cognitive abilities particularly benefited from an effective classroom management.

5.3 Estimating effects on variability in student characteristics

Instructions with individualized support presumably provide equal learning opportunities for students of different abilities. Yet, it is both a normative and an empirical question if adaptive teaching will go along with decreased variability of student achievement within classes. If individual students' learning needs are met, adaptive teaching might also result in increased intra-class variability in student achievement. Technically, heterogeneity refers to the extent to which students within a given group (e.g., classes, schools) differ with respect to a certain dimension (e.g., abilities). Hence, heterogeneity is considered a group-level variable rather than an individual characteristic. Kluczniok, Große, and Roßbach (2011) differentiate between three approaches: Heterogeneity as a measure describing an entire sample, heterogeneity describing variance within groups, and heterogeneity describing variance between groups. For example, in addition to class-level mean abilities, within-class variation in abilities has been used as an indicator of class composition

when studying differences within and between classes. Studies using within-class variation in student ability as a predictor typically failed to show systematic relationships to student learning (e.g., Gröhlich, Scharenberg, & Bos, 2009; Küsting, Post, Greb, Faust, & Lipowsky, 2010). However, recently, Decristan et al. (2017) examined the interplay between dimensions of instructional quality and intra-class variability in students' abilities. They showed that students in classes of heterogeneous ability particularly benefitted from high cognitive activation and a supportive climate. In addition, research on optimal learning outcomes, labeled "optimal classes" (for a summary see Schwippert & Walker, 2003), points to the role of instructional quality for learning outcomes. In optimal classes, two instructional goals are reached simultaneously: raising achievement for all students and increasing alignment of achievement between high-achieving and low-achieving students. Thus, above-average improvement in mean achievement is combined with a reduction of intra-class variability in student achievement. Helmke (1988) showed that the profiles of optimal classes display efficient classroom management, a focus on the learning content, higher-level tasks, diagnostic competence, a positive teacher-student relationship, and adaptive teaching. Using univariate analysis of variance it was shown that teachers in optimal classes showed a significantly higher focus on the learning content and on adaptive teaching than those of comparative classes (Helmke, 1988, p. 62f).

6. Discussion: Adaptive teaching as a messy construct?

In this paper, we focused on the construct of adaptive teaching on a macro level and a micro level of instruction and related it to methodological approaches for investigating outcomes of adaptive teaching. The research on adaptive teaching presented here shows that greater precision and alignment in conceptual and methodological approaches are needed: For example, while methodologically, research with a focus on outcomes of adaptive teaching and its differential effects on students seems to be promising, the respective theoretical models specifying macro- and micro-adaptations are still vague. Also, empirically validated models of adaptive teaching as an aspect of teacher professional competence are lacking (cf. Beck et al., 2008; Kunter et al., 2013). This may be due in large part to the lack of instruments allowing the reliable and valid assessment of adaptive teaching behavior, such as teachers' reflection on action and reflection in action (see Parsons et al., 2018). Here, analyses employed within the studies of adaptive expertise by Parsons (2012) need to be highlighted. He investigated teachers' choices and adaptations by comparing them to preplanned lessons within an innovative methodological approach. Small-scale observations of classroom instruction may therefore be combined with teachers' reflection on their action to arrive at an empirical indicator of adaptive teaching. Taking indicators of teacher expertise as reflection on action may also allow researchers to relate teacher competence to instructional be-

havior on a micro level and a macro level empirically, thereby considering adaptive teaching as an individual difference variable. Finally, the consideration of student perceptions of adaptive teaching can be employed in studies of triangulation, combining student-perceived teaching quality, observer ratings of adaptive teaching, and measures of teacher competence in multimethods designs. If one recurs to the four considerations on adaptive teaching outlined in section 1, it is especially the measures of intended adaptive teaching within units of contingent teacher-student behavior that present challenges. Choices on the level and detail of analysis (as in codings of utterances, time-based codings, qualitative analyses of episodes, or high inference ratings) typically are embedded within different theoretical models of contingency. For example, Wood, Bruner, & Ross (1976) and van de Pol et al. (2010) propose to take student understanding as the basis for subsequent contingent teacher reaction. In our model, in contrast, teacher-intended adaptive design (and the respective deviations from it) would have to be considered as well.

Conceptually, the relation between adaptive teaching and ongoing diagnosis needs to be addressed in more detail. Teachers' use of appropriate formative, summative, and standardized diagnostic instruments to assess students' preconditions and content comprehension during instruction have been put forward in models of formative assessment practices (Ruiz-Primo, 2011), scaffolding (van de Pol et al., 2010), and teacher diagnostic competence (Schrader, 2013). To meet students' needs, various strands of assessment practices typically are interwoven with instructional decisions on a macro level and a micro level. While the quality of teachers' formative assessment practices shows an impact on student cognitive and motivational outcomes (Briggs, Ruiz-Primo, Furtak, Shepard, & Yin, 2012; Hondrich et al., 2016), useful connections between formats of differentiating instruction and teachers' associated formative assessment practices still need to be pinpointed. Most of the proposed lines of research imply that a combination of methodological approaches is needed especially in the stages of research where analytical categories are still being refined. In sum, the concept of adaptive teaching is still, or again, a promising field of research that needs to be addressed in more detail with regard to relations between teacher competence, teacher expertise, diagnostics, and instructional behavior.

References

- Adl-Amini, K. (2018). *Tutorielleres Lernen im naturwissenschaftlichen Sachunterricht der Grundschule. Umsetzung und Wirkung*. Münster: Waxmann.
- Adl-Amini, K., Decristan, J., Hondrich, A. L., & Hardy, I. (2014). Umsetzung von peer-gestütztem Lernen durch Lehrkräfte im naturwissenschaftlichen Sachunterricht der Grundschule. *Zeitschrift für Grundschulforschung*, 7(2), 74–87.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Tsai, Y., Neubrand, M., & Krauss, S. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180.

- Beck, E., Baer, M., Guldimann, T., Bischoff, S., Brühwiler, C., Müller, P., Niedermann, R., Rogalla, M., & Vogt, F. (2008). *Adaptive Lehrkompetenz*. Analyse und Struktur, Veränderung und Wirkung handlungssteuernden Lehrerwissens. Münster: Waxmann.
- Beed, P. L., Hawkins, E. M., & Roller, C. M. (1991). Moving learners toward independence: The power of scaffolded instruction. *The Reading Teacher*, 44(9), 648–655.
- Behrensen, B., Gläser, E., & Solzbacher, C. (2015). Individuelle Förderung in der Grundschule: Eine bedeutsame Aufgabe aller Fachdidaktiken. In B. Behrensen, E. Gläser & C. Solzbacher (Eds.), *Fachdidaktik und individuelle Förderung in der Grundschule: Perspektiven auf Unterricht in heterogenen Lerngruppen* (pp. 1–10). Hohengehren: Schneider.
- Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85(5), 536–553.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16.
- Bos, W., Hornberg, S., Bonsen, M., & Buddeberg, I. (2008). Herausforderungen und Perspektiven für die Grundschule im Kontext von IGLU, KESS und LAU. In P. Hanke (Ed.), *Grundschule in Entwicklung* (pp. 17–41). Münster: Waxmann.
- Briggs, D., Ruiz-Primo, A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, 31(4), 13–17.
- Brophy, J. (2000). *Teaching*. Brussels, Belgium: International Academy of Education.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42.
- Clark, K. F., & Graves, M. F. (2005). Scaffolding students' comprehension of text. *The Reading Teacher*, 58(6), 570–580.
- Corno, L. (1979). A hierarchical analysis of selected naturally occurring aptitude-treatment interactions in the third grade. *American Educational Research Journal*, 16(4), 391–409.
- Corno, L. (2008). On teaching adaptively. *Educational Psychologist*, 43(3), 161–173.
- Corno, L., & Snow, R. E. (1986). Adapting teaching to individual differences in learners. In M. C. Wittrock (Ed.), *Third handbook of research on teaching* (pp. 605–629). Washington, DC: American Educational Research Association.
- Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York, NY: Irvington.
- Curby, T. W., Rimm-Kaufman, S. E., & Ponitz, C. C. (2009). Teacher-child interactions and children's achievement trajectories across kindergarten and first grade. *Journal of Educational Psychology*, 101(4), 912–925.
- Darling-Hammond, L., & Bransford, J. (Eds.). (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.
- Decristan, J., Fauth, B., Kunter, M., Büttner, G., & Klieme, E. (2017). The interplay between class heterogeneity and teaching quality in primary school. *International Journal of Educational Research*, 86, 109–121.
- Decristan, J., Hondrich, A. L., Büttner, G., Hertel, S., Klieme, E., Kunter, M., Lühken, A., Adl-Amiri, K., Djakovic, S.-K., Mannel, S., Naumann, A., & Hardy, I. (2015). Impact of additional guidance in science education on primary students' conceptual understanding. *The Journal of Educational Research*, 108(5), 358–370.
- Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., Hondrich, A. L., Rieser, S., Hertel, S., & Hardy, I. (2015). Embedded formative assessment and classroom process quality: How do they interact in promoting students' science understanding? *American Educational Research Journal*, 52(6), 1133–1159.

- Decristan, J., Kunter, M., Fauth, B., Büttner, G., Hardy, I., & Hertel, S. (2016). What role does instructional quality play for elementary school children's science competence? A focus on students at risk. *Journal for Educational Research Online*, 8(1), 66–89.
- DeRosier, M. E., Kupersmith, J. B., & Patterson, C. J. (1994). Children's academic and behavioral adjustment as a function of the chronicity and proximity of peer rejection. *Child Development*, 65(6), 1799–1813.
- Dubberke, T., Kunter, M., McElvany, N., Brunner, M., & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften: Einflüsse auf die Unterrichtsgestaltung und den Lernerfolg von Schülerinnen und Schülern. *Zeitschrift für Pädagogische Psychologie*, 22(34), 193–206.
- Fairbanks, C. M., Duffy, G. G., Faircloth, B. S., He, Y., Levin, B. B., Rohr, J., & Stein, C. (2010). Beyond knowledge: Exploring why some teachers are more thoughtfully adaptive than others. *Journal of Teacher Education*, 61(1–2), 161–171.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1–9.
- Fives, H., & Gill, M. G. (2014). *International handbook of research on teachers' beliefs*. New York, NY: Routledge.
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Martinez, E. A. (2002). Preliminary evidence on the social standing of students with learning disabilities in PALS and no-PALS classrooms. *Learning Disabilities Research & Practice*, 17(4), 205–215.
- Furtak, E., Morrison, D., & Kroog, H. (2014). Investigating the link between learning progressions and classroom assessment. *Science Education*, 98(4), 640–673.
- Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, 98(4), 732–749.
- Glaser, R. (1972). Individuals and learning: The new aptitudes. *Educational Researcher*, 1(6), 5–13.
- Greeno, J. (1998). The situativity of knowing, learning, and research. *American Psychologist*, 53(1), 5–26.
- Gröhlich, C., Scharenberg, K., & Bos, W. (2009). Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus? *Journal for Educational Research Online*, 1(1), 86–105.
- Haag, L., & Streber, D. (2014). *Individuelle Förderung*. Weinheim: Beltz.
- Hammond, J., & Gibbons, P. (2005). Putting scaffolding to work: The contribution of scaffolding in articulation ESL education. *Prospect*, 20(1), 6–30.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development*, 76(5), 949–967.
- Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology*, 98(2), 307–326.
- Hasselhorn, M., Andresen, S., Becker, B., Betz, T., Leuzinger-Bohleber, M., & Schmid, J. (2015). Children at risk of poor educational outcomes: In search of a transdisciplinary theoretical framework. *Child Indicators Research*, 8(2), 425–438.
- Hasselhorn, M., & Kuger, S. (2014). Wirksamkeit schulrelevanter Förderung in Kindertagesstätten. *Zeitschrift für Erziehungswissenschaft*, 17(2), 299–314.
- Heinzel, F. (2008). Umgang mit Heterogenität in der Grundschule. In J. Ramseger & M. Wagener (Eds.), *Chancenungleichheit in der Grundschule. Ursachen und Wege aus der Krise* (pp. 133–138). Wiesbaden: VS.

- Helmke, A. (1988). Leistungssteigerung und Ausgleich von Leistungsunterschieden in Schulklassen: unvereinbare Ziele? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 20(1), 45–76.
- Hogan, K. E., & Pressley, M. E. (1997). *Scaffolding student learning: Instructional approaches and issues*. Cambridge, MA: Brookline.
- Hondrich, A. L., Hertel, S., Adl-Amini, K., & Klieme, E. (2016). Implementing curriculum-embedded formative assessment in primary school science classrooms. *Assessment in Education: Principles, Policy & Practice*, 23(3), 353–376.
- Howe, C. (2013). Scaffolding in context: Peer interaction and abstract learning. *Learning, Culture and Social Interaction*, 2(1), 3–10.
- Jordan, A., Glenn, C., & McGhie-Richmond, D. (2010). The Supporting Effective Teaching (SET) project: The relationship of inclusive teaching practices to teachers' beliefs about disability and ability, and about their roles as teachers. *Teaching and Teacher Education*, 26(2), 259–266.
- Jordan, D. W., & Métais, J. L. (1997). Social skilling through cooperative learning. *Educational Research*, 39(1), 3–21.
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9(2), 131–155.
- Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28–37.
- Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Münster: Waxmann.
- Klieme, E., & Warwas, J. (2011). Konzepte der individuellen Förderung. *Zeitschrift für Pädagogik*, 57(6), 805–818.
- Kluczniok, K., Große, C., & Roßbach, H.-G. (2011). Heterogene Lerngruppen in der Grundschule. In W. Einsiedler, M. Götz, A. Hartinger, F. Heinzel, J. Kahlert & U. Sandfuchs (Eds.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (pp. 180–186). Bad Heilbrunn: Klinkhardt.
- Krammer, K. (2010). Individuelle Unterstützung im Unterricht mit 4- bis 8-jährigen Kindern. In M. Leuchter (Eds.), *Didaktik für die ersten Bildungsjahre: Unterricht mit 4- bis 8-jährigen Kindern* (S. 112–127). Zug: Klett.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on quality and student development. *Journal of Educational Psychology*, 105(3), 805–820.
- Künsting, J., Post, S., Greb, K., Faust, G., & Lipowsky, F. (2010). Leistungsheterogenität im mathematischen Anfangsunterricht – ein Risiko für die Leistungsentwicklung. *Zeitschrift für Grundschulforschung*, 3(1), 46–64.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean theorem. *Learning and Instruction*, 19(6), 527–537.
- Lockhorst, D., Wubbels, T., & van Oers, B. (2010). Educational dialogues and the fostering of pupils' independence: the practices of two teachers. *Journal of curriculum studies*, 42(1), 99–121.
- Marston, D. (1996). A comparison of inclusion only, pull-out only, and combined service models for students with mild disabilities. *The Journal of Special Education*, 30(2), 121–132.
- Murphy, N., & Messer, D. (2000). Differential benefits from scaffolding and children working alone. *Educational Psychology*, 20(1), 17–31.

- Nathan, M., & Kim, S. (2009). Regulation of Teacher Elicitations in the Mathematics Classroom. *Cognition and Instruction*, 27(2), 91–120.
- Oh, P. (2005). Discursive roles of the teacher during class sessions for students presenting their science investigations. *International Journal of Science Education*, 27(15), 1825–1851.
- Palincsar, A. S. (1998). Keeping the metaphor of scaffolding fresh – A response to C. Addison Stone’s “The metaphor of scaffolding: Its utility for the field of learning disabilities”. *Journal of Learning Disabilities*, 31(4), 370–373.
- Parsons, S. A. (2008). *Case studies of four teachers: The openness of the tasks they implement, the adaptations they make, and the rationales they offer for adapting*. Unpublished Doctoral Dissertation. University of North Carolina.
- Parsons, S. A. (2012). Adaptive teaching in literacy instruction case studies of two teachers. *Journal of Literacy Research*, 44(2), 149–170.
- Parsons, S. A., Davis, S. G., Scales, R. Q., Williams, J. B., & Kear, K. (2010). How and why teachers adapt their literacy instruction. In S. Szabo, M. B. Sampson, M. M. Foote & F. Falk-Ross (Eds.), *Mentoring Literacy Professionals: Continuing the Spirit of CRA/ALER after 50 Years* (pp. 221–236). Texas: A&M-University-Commerce.
- Parsons, S. A., Dodman, S. L., & Burrowbridge, S. C. (2013). Broadening the view of differentiated instruction. *Phi Delta Kappan*, 95(1), 38–42.
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., & Allen, M. (2018). Teachers’ Instructional Adaptations: A Research Synthesis. *Review of Educational Research*, 88(2), 205–242.
- Pea, R. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*, 13(3), 423–451.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38(2), 109–119.
- Prenzel, A. (2006). *Pädagogik der Vielfalt. Verschiedenheit und Gleichberechtigung in Interkultureller, Feministischer und Integrativer Pädagogik* (3. Aufl.). Wiesbaden: VS.
- Puntambekar, S., & Hübscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40(1), 1–12.
- Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students’ perception, moderated by goal orientation. *Learning and Instruction*, 27, 63–73.
- Randi, J., & Corno, L. (1997). Teachers as innovators. In B. J. Biddle, T. L. Good & I. F. Goodson (Eds.), *International handbook of teachers and teaching, Vol. I* (pp. 1163–1122). Dordrecht, NL: Kluwer.
- Reiser, B. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 13(3), 273–304.
- Roehler, L. R., & Cantlon, D. J. (1997). Scaffolding: A powerful tool in social constructivist classrooms. In K. E. Hogan & M. E. Pressley (Eds.), *Scaffolding student learning: Instructional approaches and issues* (pp. 6–42). Cambridge, MA: Brooklin.
- Rogoff, B., Baker-Sennett, J., Lacasa, P., & Goldsmith, D. (1995). Development through participation in sociocultural activity. *New Directions for Child and Adolescent Development*, 1995(67), 45–65.
- Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240–257.

- Ruiz-Primo, M. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, 37(1), 15–24.
- Ruiz-Primo, M., & Furtak, E. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44(1), 57–84.
- Schrader, F. (2013). Diagnostische Kompetenz von Lehrpersonen. *Beiträge zur Lehrerbildung*, 31(2), 154–165.
- Schwippert, K., & Walker, M. (2003). Homogenous and high performing classes: The case of optimal classes. *Studies in Educational Evaluation*, 29(2), 109–128.
- Smit, J., van Eerde, H., & Bakker, A. (2013). A conceptualisation of whole-class scaffolding. *British Educational Research Journal*, 39(5), 817–834.
- Stone, C. (1998a). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of Learning Disabilities*, 31(4), 344–364.
- Stone, C. (1998b). Should we salvage the scaffolding metaphor? *Journal of Learning Disabilities*, 31(4), 409–413.
- Tomlinson, C. A., Brighton, C., Hertberg, H., Callahan, C. M., Moon, T. R., Brimijoin, K., Conover, L. A., & Reynolds, T. (2003). Differentiating instruction in response to student readiness, interest, and learning profile in academically diverse classrooms: A review of the literature. *Journal for the Education of the Gifted*, 27(2–3), 119–145.
- Utley, C. A., & Mortweet, S. L. (1997). Peer-mediated instruction and interventions. *Focus on Exceptional Children*, 29(5), 1–23.
- van de Pol, J., Mercer, N., & Volman, M. (2018). Scaffolding student understanding in small-group work: Students' uptake of teacher support in subsequent small-group interaction. *Journal of the Learning Sciences*, 00, 1–34.
- van de Pol, J., Volman, M., Oort, F., & Beishuizen, J. (2015). The effects of scaffolding in the classroom: support contingency and student independent working time in relation to student achievement, task effort and appreciation of support. *Instructional Science*, 43(5), 615–641.
- van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22(3), 271–297.
- van de Pol, J., Volman, M., & Beishuizen, J. (2011). Patterns of contingent teaching in teacher-student interaction. *Learning and Instruction*, 21(1), 46–57.
- van de Pol, J., Volman, M., & Beishuizen, J. (2012). Promoting teacher scaffolding in small-group work: A contingency perspective. *Teaching and Teacher Education*, 28(2), 193–205.
- van Geert, P., & Steenbeek, H. (2005). The dynamics of scaffolding. *New Ideas in Psychology*, 23(3), 115–128.
- Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT.
- Walberg H. J., & Paik, S. (2000). *Effective educational practices*. Brussels: International Academy of Education.
- Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child Development*, 84(6), 2012–2130.
- Weinert, F. E., & Helmke, A. (1997). *Entwicklung im Grundschulalter*. Weinheim: Beltz.
- Wentzel, K. R., Battle, A., Russell, S. L., & Looney, L. B. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, 35(3), 193–202.
- Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13(2), 181–208.

- Wischgoll, A., Pauli, C., & Reusser, K. (2015). Scaffolding – How can contingency lead to successful learning when dealing with errors? *ZDM Mathematics Education*, *47*(7), 1147–1159.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100.
- Wood, D., Wood, H., & Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. *International Journal of Behavioral Development*, *1*(2), 131–147.