

Dohrmann, Julia

Überzeugungen von Lehrkräften. Ihre Bedeutung für das pädagogische Handeln und die Lernergebnisse in den Fächern Englisch und Mathematik

Münster ; New York : Waxmann 2021, 240 S. - (Empirische Erziehungswissenschaft; 78) - (Dissertation, Universität Frankfurt, 2020)



Quellenangabe/ Reference:

Dohrmann, Julia: Überzeugungen von Lehrkräften. Ihre Bedeutung für das pädagogische Handeln und die Lernergebnisse in den Fächern Englisch und Mathematik. Münster ; New York : Waxmann 2021, 240 S. - (Empirische Erziehungswissenschaft; 78) - (Dissertation, Universität Frankfurt, 2020) - URN: urn:nbn:de:0111-pedocs-224986 - DOI: 10.25656/01:22498

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-224986>

<https://doi.org/10.25656/01:22498>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. der Inhalt darf nicht für kommerzielle Zwecke verwendet werden. Die neu entstandenen Werke bzw. Inhalte dürfen nur unter Verwendung von Lizenzbedingungen weitergegeben werden, die mit denen dieses Lizenzvertrages identisch oder vergleichbar sind.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en> - You may copy, distribute and transmit, adapt or exhibit the work in the public and alter, transform or change this work as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work. If you alter, transform, or change this work in any way, you may distribute the resulting work only under this or a comparable license.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der:


Leibniz-Gemeinschaft

Julia Dohrmann

Überzeugungen von Lehrkräften

**Ihre Bedeutung für das pädagogische Handeln
und die Lernergebnisse in den
Fächern Englisch und Mathematik**

Empirische Erziehungswissenschaft

herausgegeben von

Rolf Becker, Sigrid Blömeke, Wilfried Bos,
Hartmut Ditton, Cornelia Gräsel, Eckhard Klieme,
Kai Maaz, Thomas Rauschenbach, Hans-Günther Roßbach,
Knut Schwippert, Ludwig Stecher, Christian Tarnai,
Rudolf Tippelt, Rainer Watermann, Horst Weishaupt

Band 78

Julia Dohrmann

Überzeugungen von Lehrkräften

Ihre Bedeutung für das pädagogische Handeln
und die Lernergebnisse in den Fächern
Englisch und Mathematik



Waxmann 2021
Münster • New York

Der Fachbereich Erziehungswissenschaften der Johann Wolfgang Goethe-Universität zu Frankfurt am Main hat diese Arbeit unter dem Titel „Die Bedeutung von Überzeugungen für das pädagogische Handeln von Lehrkräften und für die Lernergebnisse von Schülerinnen und Schülern in den Fächern Englisch und Mathematik“ als Inaugural-Dissertation zur Erlangung der Doktorwürde im Fach Erziehungswissenschaften im Jahr 2020 angenommen. Die Disputation fand am 12.05.2020 statt. Begutachtet wurde die Arbeit von Prof. Dr. Tobias Feldhoff und Prof. Dr. Eckhard Klieme.

Die Publikation wurde gefördert durch den Open-Access-Publikationsfonds für Monografien der Leibniz-Gemeinschaft.

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.dnb.de> abrufbar.

D.30

Empirische Erziehungswissenschaft, Bd. 78

ISSN 1862-2127

Print-ISBN 978-3-8309-4417-1

E-Book-ISBN 978-3-8309-9417-6

DOI: 10.31244/9783830994176

Waxmann Verlag GmbH, 2021

www.waxmann.com

info@waxmann.com

Dieses Werk ist unter der Lizenz CC BY-NC-SA 4.0 veröffentlicht:

Namensnennung – Nicht-kommerziell –

Weitergabe unter gleichen Bedingungen 4.0 International (CC BY-NC-SA 4.0)



Inhalt

1	Einleitung.....	8
1.1	Einführung in das Forschungsthema.....	8
1.2	Aufbau der Arbeit.....	11
2	Überzeugungen von Lehrkräften	12
2.1	Überzeugungen als eine Komponente der professionellen Handlungskompetenz	12
2.2	Überzeugungen – ein unscharfes Konstrukt	15
2.3	Definition von Überzeugungen.....	17
2.4	Die Beziehung von Überzeugungen zu verwandten Begriffen.....	18
2.4.1	Überzeugungen und verwandte Begriffe im Überblick	18
2.4.2	Überzeugungen und Wissen	22
2.4.3	Überzeugungen und Werte	26
2.4.4	Überzeugungen und Einstellungen	29
2.4.5	Fazit zur Abgrenzung von Überzeugungen zu anderen Begriffen	30
3	Empirisch relevante Bereiche bei Überzeugungen von Lehrkräften.....	32
3.1	Epistemologische Überzeugungen.....	32
3.2	Subjektive Theorien.....	37
3.3	Pädagogische Überzeugungen	39
4	Überzeugungen und Handeln von Lehrkräften im Unterricht.....	44
4.1	Funktionen von Überzeugungen.....	44
4.2	Wirkbeziehungen von Überzeugungen auf das Handeln von Lehrkräften	48
4.2.1	Vier Sichtweisen der Beziehung zwischen den Überzeugungen und dem Handeln von Lehrkräften	49
4.2.2	Mögliche Ursachen für die fehlende Übereinstimmung von Überzeugungen und unterrichtlichem Handeln	51
4.3	Handeln von Lehrkräften im Unterricht	54
4.3.1	Unterstützendes Unterrichtsklima.....	55
4.3.2	Klassenführung.....	56
5	Fragestellungen und Hypothesen	58
6	Die Datenbasis der Drei-Länder-Studie und ihre Aufbereitung.....	65
6.1	Die Datenbasis der Drei-Länder-Studie.....	65
6.1.1	Entstehungsgeschichte	65
6.1.2	Stichprobenauswahl der Schulen	66

6.1.3	Die Auswahl der Klassen in den Schulen der Stichprobe.....	69
6.1.4	Die Erhebungsinstrumente.....	70
6.1.5	Das Erhebungsdesign.....	74
6.2	Die Aufbereitung der Daten.....	75
6.3	Die Neufassung der Fragebogenskalen.....	79
6.4	Testskalierung.....	106
6.4.1	IRT-Skalierung.....	106
6.4.2	Beschreibung der Tests.....	107
6.4.3	Vorbereitungen für die Skalierung.....	109
6.4.4	Curriculare Validität aus damaliger Sicht.....	110
6.4.5	Curriculare Validität aus heutiger Sicht.....	111
6.4.6	Die Auswahl der Testitems.....	112
7	Daten und Methoden der Sekundäranalysen.....	120
7.1	Die Stichprobe	120
7.2	Die Instrumente	124
7.2.1	Überzeugungen der Lehrkräfte	124
7.2.2	Unterrichtsmerkmale	129
7.2.3	Schulische Ergebnisse der Schülerinnen und Schüler	134
7.2.4	Soziobiografische Hintergrundmerkmale und Komposition der Unterrichtsgruppe	136
7.3	Auswertung.....	137
7.3.1	Mehrebenen-Modellierung	139
7.3.2	Alphafehler-Adjustierung	142
7.3.3	Zentrierung	142
7.3.4	Umgang mit fehlenden Werten.....	143
8	Empirische Analysen	147
8.1	Beschreibung der Datengrundlage.....	147
8.1.1	Deskriptive Verteilung der Überzeugungen, der Unterrichtsmerkmale und der schulischen Ergebnisse	147
8.1.2	Korrelationen der drei Überzeugungen der Lehrkräfte.....	151
8.1.3	Korrelationen der Überzeugungen mit Unterrichtsmerkmalen und schulischen Ergebnissen.....	152
8.2	Ergebnisse zur Beantwortung der Forschungsfragen.....	157
8.2.1	Zusammenhänge zwischen Überzeugungen und Unterricht.....	157
8.2.2	Zusammenhänge zwischen Überzeugungen und schulischen Ergebnissen	170
8.2.3	Direkte und indirekte Zusammenhänge zwischen Überzeugungen von Lehrkräften und schulischen Ergebnissen von Schülerinnen und Schülern	181
9	Zusammenfassung und Diskussion der Ergebnisse	197

9.1	Beantwortung der Fragestellungen und Prüfung der Hypothesen ..	197
9.2	Diskussion der Befunde	206
9.3	Fazit	210
Literatur		214
Abbildungsverzeichnis		230
Tabellenverzeichnis		233
Anhang		237

1 Einleitung

Die Fähigkeiten und Fertigkeiten von Lehrkräften sind in den Mittelpunkt der pädagogischen Lehr- und Unterrichtsforschung gerückt. Lehrkräfte werden als Expertinnen und Experten wahrgenommen, die durch ihr pädagogisches und didaktisches Handeln ihre Schülerinnen und Schüler dabei unterstützen, fachliche und überfachliche Bildungsziele zu erreichen. In der vorliegenden Dissertationsschrift wird mit den pädagogischen Überzeugungen von Lehrkräften ein zentraler Aspekt ihrer professionellen Kompetenz analysiert, der für die Qualität von Schule und Unterricht bedeutsam ist. Dabei wird den Fragen nachgegangen, mit welchen Unterrichtsmerkmalen allgemeine pädagogische Überzeugungen von Lehrkräften zusammenhängen und wie diese Überzeugungen, vermittelt über das Unterrichtshandeln, mit den Ergebnissen der Schülerinnen und Schüler zusammenhängen. Die allgemeinen Überzeugungen von Lehrkräften erlangen in ihrem Unterricht praktische Bedeutsamkeit und werden hier mit den Ergebnissen der Schülerinnen und Schüler im Englisch- und Mathematikunterricht verbunden. Dies geschieht durch eine Sekundäranalyse der Drei-Länder-Studie von Helmut Fend aus den Jahren 1978 und 1979.

1.1 Einführung in das Forschungsthema

Die empirische Bildungsforschung hat in den vergangenen Jahren verschiedene Aspekte des Wissens und Könnens als zentrale Komponenten des professionellen Handelns von Lehrkräften herausgearbeitet, die eine Bedeutung für den Unterricht haben. Baumert und Kunter (2006) zählen neben dem Professionswissen auch selbstregulative Fähigkeiten, motivationale Orientierungen sowie *Überzeugungen* und Werthaltungen zur professionellen Handlungskompetenz von Lehrenden.

Die zum letzteren Bereich vorliegenden empirischen Befunde aus Studien in Deutschland sind jedoch begrenzt (Oser & Blömeke, 2012). Es besteht Forschungsbedarf dahingehend, für welche Aspekte des Unterrichtshandelns die Überzeugungen der Lehrkräfte relevant sind und auf welche Schülerergebnisse sie sich auswirken (Baumert & Kunter, 2006; Reusser, Pauli & Elmer, 2011). Aktuell werden in empirischen Untersuchungen insbesondere epistemologische Überzeugungen betrachtet (siehe Kapitel 3.1). Diese beziehen sich konkret auf die Struktur und die Genese des Wissens in einem bestimmten Fach und wer-

den vor allem für den Mathematikunterricht, selten für andere Fächer untersucht. Inwieweit allgemeine pädagogische Überzeugungen hinsichtlich des Lehrens und Lernens bzw. lernförderlichen Unterrichts mit dem Unterrichtshandeln von Lehrpersonen und den Lernergebnissen der Schülerinnen und Schüler in Zusammenhang stehen, ist weitgehend unbekannt. Gleiches gilt für Überzeugungen in Bezug auf die Reform- bzw. Innovationsbereitschaft von Schulen, die vor allem in der Schulentwicklungsforschung im Kontext des Umgangs mit Schulreformen untersucht werden (siehe Kapitel 3.3).

Um zu klären, welche Zusammenhänge zwischen den pädagogischen Überzeugungen der Lehrkräfte, ihrem pädagogischen Handeln im Unterricht und den Lernergebnissen der Schülerinnen und Schüler bestehen, liegt mit der Drei-Länder-Studie (DLS) von Helmut Fend sowie Kolleginnen und Kollegen der Universität Konstanz aus den Jahren 1978 und 1979 (Fend, 1982) eine geeignete Datenbasis vor.

Der Rückgriff auf diese historischen Daten ist aus mehreren Gründen vielversprechend: Die Phase der gesellschaftlichen und bildungspolitischen Umbrüche der 1960er und 1970er Jahre ist schon zeithistorisch interessant, wie der Beitrag von Mattes und Reh (2019) zeigt. Die vorliegende Arbeit bearbeitet indes theoretische Grundfragen, für welche die Daten der Drei-Länder-Studie genauso einschlägig sind wie aktuelle Daten. Im Rückgriff auf die These einer „Grammar of Schooling“ (Tyack & Tobin, 1994) kann davon ausgegangen werden, dass Grundmuster der Relationierung von pädagogischen Orientierungen, Unterrichtshandeln und Lernen eine gewisse Stabilität besitzen. Außerdem lassen sich mit der Drei-Länder-Studie diejenigen normativen Aspekte von Überzeugungen untersuchen, die in der heutigen Professionalitätsdebatte vernachlässigt werden: allgemeine pädagogische Orientierungen, die bestimmte Überzeugungen über Schule, Lehren und Lernen abbilden. So lässt sich prüfen, ob auch solche Überzeugungen von Lehrkräften von Bedeutung sind, ob sie also mit ihrem Unterrichtshandeln und den Merkmalen von Schülerinnen und Schülern zusammenhängen. Die von Fend fachspezifisch erhobenen Skalen zur Unterrichtswahrnehmung der Schülerinnen und Schüler (z.B. Klassenführung, Lernunterstützung) haben die Schul- und Unterrichtsforschung nachhaltig beeinflusst (Klieme, 2016; Messner, 2016; Steffens, 2012; Steffens & Bargel, 2016; Terhart, 2000) und werden mit Adaptionen bis heute verwendet (Forschungsdatenzentrum Bildung, 2014b). Weniger beachtet wurden bislang diejenigen Items, die aus Sicht der aktuellen Forschung die Adaptivität des Unterrichts

erfassen. Sie werden in der vorliegenden Arbeit zu einer neuen Skala gebündelt. Schließlich eignet sich die Untersuchungsanlage der Drei-Länder-Studie für Analysen nach heutigen methodischen Standards: Im Design der Studie wurde die Mehrebenenstruktur der Schule mit Schul-, Klassen- und Individual-ebene so erfasst, dass entsprechende statistische Analysen durchgeführt werden können (Martínez, 2012; Opdenakker & Van Damme, 2000). Neben kognitiven wurden auch affektive Lernergebnisse der Schülerinnen und Schüler berücksichtigt.

Die Überzeugungen der Lehrkräfte spielten in Fends Studie „Die Pädagogik des Neokonservatismus“ (1984) eine zentrale Rolle. Seinerzeit wurde untersucht, wie Überzeugungen hinsichtlich der Notwendigkeit von Druck bei Lernprozessen mit dem Unterricht und den Lernergebnissen zusammenhängen. Allerdings wurde dieser Zusammenhang mit einem Extremgruppenvergleich von konservativen gegenüber liberalen Lehrkräften und nicht unter der gleichzeitigen Betrachtung von Unterricht und Lernergebnissen analysiert, wie es in der vorliegenden Arbeit geschieht.

Ziel dieser Dissertation ist es, die Beziehungen zwischen allgemeinpädagogischen Überzeugungen, Unterricht und Lernergebnissen unter Berücksichtigung des aktuellen theoretischen und methodischen Forschungsstandes zu analysieren.¹

Dabei stehen drei Überzeugungen von Lehrkräften im Mittelpunkt (ausführlich siehe Kapitel 3.3 Pädagogische Überzeugungen). Mit der Einstellung von Lehrkräften zur *Förderung schwacher Schülerinnen und Schüler* und der Einstellung von Lehrkräften zur *Notwendigkeit von Druck bei Lernprozessen* werden zwei Überzeugungen erfasst, die sich auf ein übergreifendes Verständnis lernförderlichen Unterrichts beziehen. Zu den allgemeinen pädagogischen Überzeugungen von Lehrkräften gehört ferner die *Reformbereitschaft von Lehrkräften*, die indiziert, wie beständig bzw. anpassungsfähig Schule auf gesellschaftliche Anforderungen reagieren sollte.

1 Darüber hinaus wird mit den Reanalysen der Daten aus der Drei-Länder-Studie eine Grundlage für Untersuchungen zum historischen Wandel von Schulen bereitgestellt. In einer Vorstudie (Hausen et al., 2019) wurde dargelegt, dass ein Großteil der damaligen Tests und Unterrichtsskalen auch heute noch zuverlässig einsetzbar ist, sodass mit einer Neuerhebung an denselben Schulen die Situation 40 Jahre danach beschrieben werden könnte.

1.2 Aufbau der Arbeit

Die vorliegende Arbeit gliedert sich in drei übergeordnete Teile. Im ersten Teil werden theoretische Zugänge zu Überzeugungen von Lehrkräften und deren Verbindung mit ihrem Handeln im Unterricht dargelegt. Im zweiten Teil folgen die Beschreibung der Datenbasis dieser Arbeit, der Drei-Länder-Studie, und ausführlich die Beschreibung der Daten und Methoden der vorliegenden Sekundäranalysen. Im dritten Teil werden die aufgestellten Forschungsfragen zu den Überzeugungen von Lehrkräften auf der Datenbasis der Drei-Länder-Studie untersucht. Anschließend werden die Ergebnisse diskutiert.

Im zweiten Kapitel werden die Überzeugungen von Lehrkräften theoretisch eingeordnet, definiert und zu benachbarten Begriffen in Beziehung gesetzt. Empirisch relevante Bereiche bei den Überzeugungen von Lehrkräften werden im dritten Kapitel beschrieben. Im vierten Kapitel wird die Verbindung von Überzeugungen mit dem Unterrichtshandeln von Lehrkräften und somit auch den Lernergebnissen von Schülerinnen und Schülern hergeleitet. Die Fragestellungen und Hypothesen dieser Arbeit werden im fünften Kapitel erläutert. Die Datenbasis dafür, die Drei-Länder-Studie, wird, einschließlich der umfangreichen Aufbereitungsarbeiten, im sechsten Kapitel vorgestellt. Im siebten Kapitel werden dann die Datenlage und die Methoden der hier vorgenommenen Sekundäranalysen erläutert. Die Ergebnisse der empirischen Analysen werden im achten Kapitel präsentiert, abschließend werden diese im neunten Kapitel diskutiert.

2 Überzeugungen von Lehrkräften

Überzeugungen sind eine der zentralen Komponenten der professionellen Handlungskompetenz von Lehrkräften (Bromme, 1997; Seifried, 2009, S. 42; Shulman, 1986). Dies wird im folgenden Abschnitt 2.1 näher dargestellt. Nachdem im Abschnitt 2.2 erklärt wird, warum das Konstrukt der Überzeugungen vielschichtig ist, wird im Abschnitt 2.3 eine Arbeitsdefinition dieses Konstrukts für die vorliegende Arbeit vorgestellt. Im abschließenden Abschnitt 2.4 werden die Überzeugungen zu verwandten Begriffen in Beziehung gesetzt, und der Bezug von Überzeugungen zu Wissen, Werten und Einstellungen wird aus theoretischer Sicht näher betrachtet.

Pajares (1992, S. 316) hat treffenderweise festgestellt, dass immer die Überzeugungen von Lehrkräften hinsichtlich Schule, Unterricht, Lernen sowie Schülerinnen und Schüler gemeint sind, wenn man allgemein von „Überzeugungen von Lehrkräften“ spricht. Das unterrichtliche Handeln von Lehrkräften wird auch von Überzeugungen beeinflusst, die sich nicht konkret auf das Lehren und Lernen beziehen, sondern beispielsweise auf das Demokratieverständnis der Lehrkräfte oder ihre Überzeugung bezüglich Geschlechterrollen. Wenn im Folgenden von „Überzeugungen“ gesprochen wird, sind daher immer diejenigen Überzeugungen der Lehrkräfte gemeint, die sich auf ihr Berufsumfeld, also auf die Schule, den Unterricht und das Lernen beziehen.

2.1 Überzeugungen als eine Komponente der professionellen Handlungskompetenz

Modelle, welche die professionelle Kompetenz von Lehrkräften zum Gegenstand haben, dienen dazu, empirische Befunde zu verschiedenen Aspekten der Lehrkräfte (ihrer Qualifikation, ihrer Kompetenz und ihrer Persönlichkeit) in ihrer Bedeutung für den Unterricht und das Lernen der Schülerinnen und Schüler zu ordnen und theoriebezogen zu diskutieren (Baumert & Kunter, 2006, S. 470). Deshalb werden im Folgenden die hier untersuchten Überzeugungen im Rahmen einiger dieser Modelle betrachtet. Nach König (2014, S. 94) werden in solchen Modellen Erkenntnisse der Forschung zur Expertise von Lehrkräften berücksichtigt, sie enthalten aber auch Annahmen zum Wissen und zum Können der Lehrkräfte. Schließlich wird der Kompetenzbegriff angewandt, wie er aktuell in der empirischen Bildungsforschung Verwendung findet. Es gibt

verschiedene Arten, Überzeugungen unter dem Gesichtspunkt der professionellen Kompetenz zu systematisieren. Beispiele finden sich unter anderem bei Baumert und Kunter (2006), Reusser et al. (2011) sowie Kunter und Pohlmann (2015).

Kunter und Pohlmann (2015, S. 268–271) stellen in ihrer Übersicht exemplarisch drei Überzeugungsbereiche vor, für die entsprechend gut abgesicherte Erkenntnisse zu den Wirkungen der verschiedenen Überzeugungen von Lehrkräften vorliegen. Dazu zählen die Selbstwirksamkeitsüberzeugungen der Lehrkräfte, ihre Erwartungen an die Schülerinnen und Schüler sowie ihre lerntheoretischen Überzeugungen (fachbezogen oder auf Lehren und Lernen allgemein bezogen). Allgemeine pädagogische Überzeugungen, wie sie hier im Fokus stehen, kommen lediglich beispielhaft am Rande als „Kulturelle Normen und Werte, die Bildung und Schule betreffen“ (Kunter & Pohlmann, 2015, S. 267) bzw. als „Allgemeine Wirksamkeitsüberzeugung bezogen auf die Lehrfähigkeit allgemein“ (Kunter & Pohlmann, 2015, S. 268) vor.

Reusser et al. (2011, S. 485–487) ordnen die Befunde zu den berufsbezogenen Überzeugungen von Lehrkräften in die folgenden drei Gegenstandsbereiche ein: epistemologische Überzeugungen zu Lerninhalten und Lernprozessen, personenbezogene Überzeugungen zu Lehrkräften und Lernenden sowie kontextbezogene Überzeugungen zu Schule und Gesellschaft. Die hier untersuchten allgemeinen pädagogischen Überzeugungen lassen sich teils dem zweiten Bereich der personenbezogenen Überzeugungen (die Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler und zur Notwendigkeit von Druck bei Lernprozessen) zuordnen, teils dem dritten Bereich der kontextbezogenen Überzeugungen (Reformbereitschaft).

Nach Baumert und Kunter (2006, S. 481) entsteht professionelle Handlungskompetenz aus dem Zusammenwirken von

„– spezifischem, erfahrungsgesättigten deklarativen und prozeduralen Wissen (Kompetenzen im engeren Sinne: Wissen und Können);

– professionellen Werten, Überzeugungen, subjektiven Theorien, normativen Präferenzen und Zielen;

– motivationalen Orientierungen sowie

– metakognitiven Fähigkeiten und Fähigkeiten professioneller Selbstregulation.“

Die erste Facette des Professionswissens wird dabei besonders genau ausdifferenziert. Innerhalb der zweiten Facette der Werthaltungen und Überzeugungen werden Wertbindungen, epistemologische Überzeugungen (siehe Kapitel 3.1), subjektive Theorien über Lehren und Lernen (siehe Kapitel 3.2) sowie Zielsysteme für Curriculum und Unterricht unterschieden. Auch in dieser Systematisierung können allgemeine pädagogische Orientierungen, wie sie in der vorliegenden Arbeit untersucht werden, nicht vollständig zugeordnet werden. Man könnte sie als Wertbindungen (siehe Abschnitt 2.4.3) verstehen, wobei Baumert und Kunter diese wie Oser (1998b) als Berufsmoral bzw. Professionsethik ansehen. Bei der Forschungsgruppe um Fritz Oser (1998b) liegt der Fokus darauf, wie pädagogische Entscheidungen zustande kommen, und nicht auf den Ergebnissen, welche diese Entscheidungen als Folge des unterrichtlichen Handelns von Lehrkräften bei den Schülerinnen und Schülern hervorrufen. Bei Oser macht sich die Berufsethik (und damit auch die Überzeugungen als ein Teil davon) nicht an konkreten Inhalten oder Handlungen fest, sondern an dem Prozess, wie unterrichtliche Handlungen von Lehrkräften entstehen (Scheunpflug, 2005, S. 61). Die professionelle Kompetenz von Lehrkräften, wie sie von Baumert und Kunter (2006) diskutiert wird, legt den Fokus auf die Lernergebnisse von Schülerinnen und Schülern, um an diesen die Berufsethik der Lehrkräfte im Sinne einer besonderen Qualität ihres Wissens und Könnens zu überprüfen. Hier umfasst die Berufsethik diejenigen Handlungen der Lehrkräfte, welche der Erfüllung ihrer professionellen Aufgaben entsprechen.

Die Bedeutung der professionellen Kompetenzen von Lehrkräften für das Lernen von Schülerinnen und Schülern lässt sich nicht nur in Modellen abbilden, sie lässt sich auch empirisch untermauern. So können laut Hattie (2003) 30 Prozent der Leistungsunterschiede bei den Schülerinnen und Schülern durch das Wissen, das Handeln und die Einstellungen der Lehrkräfte erklärt werden. Betrachtet man die Lernfortschritte der Schülerinnen und Schüler, dann lassen sich zwischen drei und 18 Prozent der Varianz des Lernfortschritts auf die Lehrperson zurückführen (American Educational Research Association, 2005). Eine Forschungsübersicht von Lipowsky (2006) zeigt, dass das Wissen und die Überzeugungen von Lehrkräften direkte und auch indirekte Effekte auf den Lernerfolg von Schülerinnen und Schülern in Mathematik und Naturwissenschaften haben können.

2.2 Überzeugungen – ein unscharfes Konstrukt

Für die Überzeugungen von Lehrkräften werden in den verschiedenen Teildisziplinen der Bildungswissenschaften, die sich mit Lehrkräften befassen, jeweils eigene Begriffe verwendet. Die Bezeichnungen, die sich für die Überzeugungen von Lehrkräften finden lassen, reichen von „Konzeptionen“² in der Hochschulforschung (Trautwein, 2013) und „pedagogical beliefs“ in der Unterrichtsforschung (König, 2012) bis hin zu „Sichtweisen“³ (Seifried, 2009) bzw. „Vorstellungen“ in der Didaktik (Hartinger, Kleickmann & Hawelka, 2006) oder „professionellen pädagogischen Haltungen“ in der Schulpädagogik (Schwer & Solzbacher, 2014). In Arbeiten aus der Lehrkräfteforschung finden sich Begriffe wie „subjektive Theorien“ (Groeben, Wahl, Schlee & Scheele, 1988) oder „educational beliefs“ (Pajares, 1992). „Teachers’ beliefs“ hingegen tauchen auch in der Lehrkräfteforschung auf (Fives & Gill, 2015; Turner, Christensen & Meyer, 2009) sowie in der Fachdidaktik (Philipp, 2007; Thompson, 1992) oder in der Schuleffektivitätsforschung (Muijs & Reynolds, 2002). Wie Seifried (2009, S. 103) festgestellt hat, werden in manchen Arbeiten einige dieser Begriffe auch synonym gebraucht, sicherlich häufig nicht aus inhaltlichen Überlegungen heraus, sondern zwecks der Vermeidung von sprachlichen Wiederholungen und der Verbesserung der Lesbarkeit von Texten.

An dieser Begriffsvielfalt wird deutlich, dass es im erziehungswissenschaftlichen Diskurs keinen einheitlichen Konzept für die Benennung, die Bestandteile und die Funktionsweise der Überzeugungen von Lehrkräften gibt. In seinem vielzitierten Artikel über „teachers’ beliefs“ als ein „messy construct“ schreibt Pajares (1992, S. 307): „The difficulty in studying teachers’ beliefs has been caused by definitional problems, poor conceptualizations, and differing understandings of beliefs and belief structures.“ Die Bezeichnung von Lehrerüberzeugungen als „messy construct“ wird auch in aktuelleren Arbeiten zu diesem Thema vielfach geteilt (siehe u.a. Fives & Buehl, 2012; Reusser et al., 2011).

2 Trautwein (2013) unterscheidet zwischen dem Begriff der „(Lehr-)Konzeptionen“, der vorwiegend im Hochschulkontext Verwendung findet, und dem Begriff „Überzeugungen zu Lehren und Lernen“, der im schulischen Kontext verwendet wird.

3 Seifried (2009, S. 36) verwendet den Begriff „Sichtweisen“ forschungslinienübergreifend, um Lehr-Lern-Vorstellungen, Überzeugungen, subjektive Theorien oder implizite/naive Theorien anzusprechen. Dabei lehnt er sich an den englischen Ausdruck „world views“ des Mathematikdidaktikers Schoenfeld (1985, 2002) an.

Der am häufigsten benutzte Ausdruck in der internationalen Literatur ist „teachers’ beliefs“, er findet sich in einschlägigen Forschungsübersichten (Calderhead, 1996; Kagan, 1992; Pajares, 1992; Philipp, 2007; Richardson, 1996; Woolfolk Hoy, Davis & Pape, 2006). Thompson (1992, S. 129) stellt jedoch fest, dass trotz der Popularität des Begriffs „beliefs“ in der englischsprachigen Bildungsforschung dieser häufig überhaupt nicht definiert wird, da er bei der Leserschaft als bekannt vorausgesetzt wird (siehe auch Furinghetti & Pehkonen, 2002). Im deutschen Sprachraum hat sich als Übersetzung dieser Bezeichnung der Begriff „Überzeugungen“ durchgesetzt (Reusser et al., 2011). Zur zunehmenden Verwendung des Begriffs „Überzeugungen“ beigetragen hat auch die Berücksichtigung von berufsbezogenen Überzeugungen in Modellen der professionellen Kompetenz von Lehrkräften (siehe Kapitel 2.1) und in der empirischen Lehrerbildungsforschung (Blömeke, Kaiser & Lehmann, 2008, 2010) sowie der Unterrichtsforschung (Kunter et al., 2011; OECD, 2009).

Wie an der Vielfalt der verwendeten Begriffe bzw. Bezeichnungen, die sich durch die unterschiedlichen Forschungsdisziplinen ziehen, deutlich wird, lässt sich dieser Forschungsgegenstand aus unterschiedlicher Perspektive betrachten. Für die Auseinandersetzung mit dem Forschungsthema der vorliegenden Arbeit, den allgemeinen pädagogischen Orientierungen (insbesondere zu Themen wie Förderung, Druck bei Lernprozessen und Reformbereitschaft; siehe Kapitel 1.1), eignet sich der Begriff „Überzeugungen“ am besten, da dieser breit und umfassend genug ist, um auch die normativ-wertenden Aspekte dieses Konstrukts zu berücksichtigen (siehe die Definition von Überzeugungen in Kapitel 2.3; auch Oser & Blömeke, 2012).

Allerdings ist der theoretische Zugang zu diesem Feld, nicht nur aufgrund der mangelnden Einheitlichkeit der Begrifflichkeit, nicht ganz einfach. Der Begriff der Überzeugungen ist ein Abstraktum, bei dem man nur sehr vage erkennen kann, was es genau umfasst (Oser & Blömeke, 2012; dazu auch Abschnitt 2.4.1). Dieses Abstraktum wird benötigt, da seine Bestandteile je nach Kontext anders definiert werden („Containerbegriff“). Für den Rahmen der vorliegenden Arbeit wird der Begriff der Überzeugungen im folgenden Kapitel definiert (Kapitel 2.3), im Anschluss daran wird das Verhältnis von Überzeugungen und verwandten Begriffen diskutiert (Kapitel 2.4).

2.3 Definition von Überzeugungen

Überzeugungen im allgemeinen Verständnis sind in vielen wissenschaftlichen Disziplinen Untersuchungsgegenstand, und es besteht Übereinstimmung darüber, was Überzeugungen ausmacht: Richardson (1996, S. 103) beschreibt diesen Kern der verschiedenen Definitionen von Überzeugungen als „psychologically-held understandings, premises or propositions about the world that are felt to be true“. Eine Definition der Überzeugungen von Lehrkräften liefert Kagan (1990, S. 423), sie bezeichnet diese als „the highly personal ways in which a teacher understands classrooms, students, the nature of learning, the teacher’s role in the classroom, and the goals of education“. Philipp (2007, S. 259) versteht Überzeugungen von Lehrkräften als „dispositions toward action“, also als Neigungen in Bezug auf eine bestimmte Handlung. Nach Turner et al. (2009) spiegeln die Überzeugungen Lehrender nicht ein forschungsbasiertes Wissen über das Lernen, sondern ihre Einstellungen,⁴ ihren gesunden Menschenverstand („common sense“) und ihre Erfahrungserfahrungen wider. Im folgenden Kapitel 2.4 wird deutlich werden, dass für den theoretischen Zugang zu Überzeugungen ihr Bezug zu anderen Begriffen (wie Einstellungen und Wissen) zentral ist und häufig diskutiert wird.

Reusser, Pauli und Elmer (2011, S. 478) beziehen die Definition der Überzeugungen, wie sie beispielsweise bei Richardson zu finden ist (siehe oben), auf den Kontext von Lehrkräften. Berufsbezogene Überzeugungen von Lehrkräften sind demnach „Vorstellungen über das Wesen und die Natur von Lehr-Lernprozessen, Lerninhalten, [über] die Identität und Rolle von Lernenden und Lehrenden (sich selbst) sowie [über] den institutionellen und gesellschaftlichen Kontext von Bildung und Erziehung“. Diese Begriffsbestimmung fasst grundlegende Komponenten der verschiedenen Definitionen von Überzeugungen zusammen (siehe auch Pajares, 1992) und enthält durch die Vorstellungen über Lehr-Lernprozesse sowie den institutionellen Kontext auch allgemeine pädagogische Orientierungen. Überzeugungen sind, wie Reusser et al. (2011, S. 478) weiter darstellen, affektiv aufgeladen, enthalten eine Bewertungskomponente, werden von den Lehrpersonen „für wahr oder wertvoll gehalten (...) und [geben] ihrem berufsbezogenen Denken und Handeln Struktur, Halt, Sicherheit und Orientierung“. Sie sind somit äußerst bedeutsam für das unterrichtliche Handeln der Lehrkräfte.

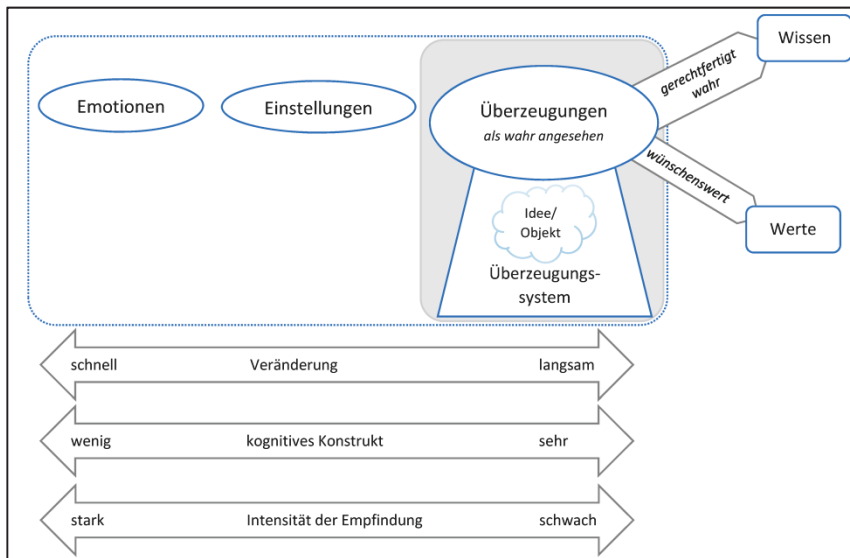
4 Zum Verhältnis von Überzeugungen und Einstellungen siehe Abschnitt 2.4.4.

Die folgenden Überlegungen orientieren sich an der Definition von Reusser et al. (2011). In dieser Definition finden sich bereits alle Aspekte, die den Begriff der „Überzeugungen von Lehrkräften“ ausmachen und auf die im Verlauf der theoretischen Ausarbeitung in dieser Arbeit näher eingegangen wird: 1. Überzeugungen werden von den Lehrkräften für wahr gehalten (auch wenn anderen Personen abweichende Überzeugungen zugestanden werden, siehe den folgenden Abschnitt 2.4.1), 2. durch die Bewertungskomponente enthalten Überzeugungen das Bekenntnis der Lehrkräfte zu bestimmten Werten (zum Verhältnis von Überzeugungen und Werten siehe Abschnitt 2.4.3), 3. durch den Bezug auf Vorstellungen wird deutlich, dass Überzeugungen bewusst, aber auch unbewusst sein können (siehe Kapitel 4.1), 4. Überzeugungen haben eine Beziehung zum Denken und Handeln der Lehrkräfte (siehe Kapitel 4.1) und geben diesen Struktur und Orientierung.

2.4 Die Beziehung von Überzeugungen zu verwandten Begriffen

2.4.1 Überzeugungen und verwandte Begriffe im Überblick

Da der Begriff der Überzeugungen von Lehrkräften ein „messy construct“ ist (siehe Kapitel 2.2), der keine eindeutige Definition und Operationalisierung mitliefert, ist es für die folgende Arbeit wichtig, das Verhältnis zu verwandten Begriffen wie Wissen, Werten und Einstellungen zu klären und ihn davon abzugrenzen. Zur Systematisierung des Verhältnisses von Überzeugungen zu den anderen Begriffen wird hier auf Schemata von Philipp (2007) zurückgegriffen, der sich wiederum auf Thompson (1992) bezieht. Wie in Abbildung 2.1 dargestellt, lassen sich Überzeugungen im Spannungsfeld von Emotionen, Einstellungen, Wissen und Werten betrachten. Welche Begriffe man wie zu dem der Überzeugungen ins Verhältnis setzt, hängt jedoch immer von der Betrachtungsweise ab, wie die folgende Systematisierung verdeutlicht.



Anmerkung: eigene Darstellung, nach Philipp (2007) und Thompson (1992)

Abbildung 2.1: Verhältnis von Überzeugungen zu benachbarten Begriffen

Emotionen, Einstellungen und Überzeugungen unterscheiden sich jeweils in ihrer Veränderbarkeit, darin, wie sehr sie ein kognitives Konstrukt sind, und in der Intensität ihrer Empfindung (Philipp, 2007, S. 259):

„Attitudes change more slowly than emotions, but they change more quickly than beliefs. [...] Attitudes are more cognitive than emotion but less cognitive than beliefs. [...] Attitudes, like emotions, may involve positive or negative feelings, and they are felt with less intensity than emotions. [...] Beliefs are [...] felt less intensely [...] than attitudes.“

Abbildung 2.1 zeigt diese graduelle Abfolge von Emotionen – Einstellungen – Überzeugungen für die Veränderbarkeit, die Kognition und die empfundene Intensität. Dabei soll die Kategorisierung dieser graduellen Abfolge als beschreibende Annäherung an die hier vorgestellten Konstrukte verstanden werden. So konstatieren beispielweise Furinghetti und Pehkonen (2002, S. 40), dass man die Verortung von Überzeugungen auf der Dimension „Affektivität – Kognitivität“ durchaus unterschiedlich betrachten kann: Wollte man den Zusammenhang von Überzeugungen und Wissen betonen, dann würde man Überzeugungen eher bei der kognitiven Struktur von Individuen sehen (wie in Abbildung 2.1). Werden Überzeugungen jedoch als eine Form der Reaktion auf eine bestimmte Situation angesehen, dann würde man Überzeugungen stärker als ein affektives denn als ein kognitives Konstrukt betrachten.

Auch das Wissen liegt, wie bereits erwähnt, im Umfeld von Überzeugungen. Wissen und Überzeugungen weisen eine enge Verbindung auf und sind deshalb nicht immer leicht voneinander zu unterscheiden (siehe auch Furinghetti & Pehkonen, 2002). Für Pajares (1992) und Richardson (1996) ist die Unterscheidung zwischen Überzeugungen und Wissen ein äußerst komplexes Thema, das hinsichtlich der Unschärfe in der Begriffsbestimmung von Überzeugungen zentral ist. Da in der Selbstwahrnehmung von Lehrkräften Überzeugungen auch als Wissen gesehen werden können, wird von Thompson (1992, S. 129) vorgeschlagen, dass die Forschung sich nicht mit der Suche nach Unterschieden zwischen Überzeugungen und Wissen beschäftigen sollte, sondern lieber damit, ob und wie die Überzeugungen von Lehrkräften (oder was sie als ihr Wissen betrachten) ihre Erfahrungen beeinflussen. Die Frage nach einem Zusammenhang zwischen den Überzeugungen von Lehrkräften und ihrem Handeln wird in dieser Arbeit empirisch beantwortet, doch zuvor werden in Abschnitt 2.4.2 Differenzierungen zwischen Überzeugungen und Wissen dargestellt. Die vermeintliche Deckungsgleichheit von Überzeugungen und Wissen von Lehrkräften wird dahingehend aufgelöst, dass Wissen als eine gerechtfertigt wahre Überzeugung („justified true belief“, Philipp, 2007, S. 259; siehe Abbildung 2.1) definiert wird, wohingegen Überzeugungen lediglich von der betreffenden Person als wahr angesehen („thought to be true“, Philipp, 2007, S. 259; siehe Abbildung 2.1) werden.

Eine Überzeugung, die man in sich trägt, sehr hochhält und gewissermaßen verehrt, kann als Wert bezeichnet werden (Philipp, 2007, S. 259). Werte unterscheiden sich vor allem dadurch von Überzeugungen, dass sie als desiderabel angesehen werden. Auch wenn Überzeugungen und Werte einander ähneln, werden Werte eher mit der Dichotomie wünschenswert/nicht wünschenswert und Überzeugungen mit der Dichotomie wahr/falsch in Zusammenhang gebracht. Werte werden, ähnlich wie Überzeugungen, als sehr stabile, hochkognitive Konstrukte angesehen, die aber im Gegensatz zu Überzeugungen stark affektiv sind (Philipp, 2007, S. 301). In Abschnitt 2.4.3 wird das Verhältnis von Überzeugungen zu Werten genauer beleuchtet.

Einstellungen sind Handlungs-, Gefühls- oder Denkweisen, welche die eigene Disposition oder Meinung anzeigen (Philipp, 2007, S. 259; siehe Abschnitt 2.4.4). Emotionen sind schnell wechselnde Gefühlszustände, die zumeist in einen Kontext eingebettet sind (Philipp, 2007, S. 301). Ein Zusammenhang zwischen Emotionen und Einstellungen liegt darin, dass wiederholte emotionale

le Reaktionen auf eine Erfahrung dazu führen können, dass diese Emotion in eine Einstellung zu dieser Erfahrung automatisiert wird (Philipp, 2007, S. 261). Das Verhältnis von Emotionen zu Überzeugungen und die Rolle von Emotionen bei der Veränderung von Überzeugungen wird von Gill und Hardin (2015) ausführlich dargestellt. Zur Wirkung der Emotionen von Lehrkräften auf ihr Unterrichtshandeln bieten Frenzel, Götz und Pekrun (2008) einen ersten Überblick.

Überzeugungen gruppieren sich um bestimmte Ideen oder Objekte, sie bilden also Überzeugungssysteme aus (Green, 1971; Pajares, 1992; Philipp, 2007, S. 260; Rokeach, 1960; Thompson, 1992, S. 130). Da Überzeugungen nie isoliert vorkommen, ist eine Betrachtung solcher Überzeugungssysteme⁵ sinnvoll. Überzeugungssysteme sind dynamische Systeme, die sich umstrukturieren und verändern können, wenn Individuen ihre Überzeugungen im Hinblick auf ihre Erfahrungen neu beurteilen (Thompson, 1992). Innerhalb dieser Systeme lassen sich drei Aspekte bezüglich der Organisation von Überzeugungen ausmachen (Green, 1971; Philipp, 2007; Thompson, 1992): Überzeugungen sind 1. primär oder abgeleitet, 2. zentral oder peripher und 3. sie existieren nicht isoliert, sondern immer in Bündeln.

Die *primären Überzeugungen* bilden eine quasi logische Grundlage für die abgeleiteten Überzeugungen. Ein Beispiel für eine primäre Überzeugung wäre: Die Schule als Institution soll sich dem Wandel der Gesellschaft anpassen. Eine davon abgeleitete Überzeugung wäre es, wenn eine Lehrkraft es für wichtig erachtet, ihre Unterrichtsmethoden zu verändern oder neue zu erproben, um beispielsweise stärker auf die heterogenen Voraussetzungen der Schülerinnen und Schüler eingehen zu können. *Zentrale Überzeugungen* sind von großer Bedeutung für eine Person, während periphere Überzeugungen dies weniger sind und sich deshalb leichter verändern lassen. Dabei muss eine primäre Überzeugung nicht notwendigerweise zugleich eine zentrale sein. Im eben genannten Beispiel könnte es sein, dass die Lehrkraft großen Wert auf die Verwendung moderner Unterrichtsmethoden legt, dass also diese abgeleitete Überzeugung für sie besonders zentral ist und die eher allgemeine Anpassung der Schule an eine sich wandelnde Gesellschaft hingegen einen peripheren Stellenwert einnimmt. Der umgekehrte Fall wäre natürlich ebenso möglich. Die Ausgestaltung des Überzeugungssystems ist schließlich bei jeder Person anders. Die

5 Pajares (1992, S. 318) sieht Einstellungen und Werte bei Rokeach (1969) als Substrukturen von Überzeugungen und somit als Teile von Überzeugungssystemen an.

Tatsache, dass Überzeugungen *in Bündeln* (englisch „cluster“) vorkommen, die relativ isoliert voneinander bestehen können, kann dazu führen, dass die Struktur der verschiedenen Überzeugungen einer Person innerhalb eines Überzeugungssystems sich von außen betrachtet als inkonsistent oder als widersprüchlich darstellt. Es gibt nur wenig Kontakt und Beeinflussung zwischen den Überzeugungssystemen, und Überzeugungen, die inkompatibel sind, können in verschiedenen Bündeln aufrechterhalten werden. Solange inkompatible Überzeugungen niemals nebeneinandergestellt und auf Konsistenz hin geprüft werden, kann die Inkompatibilität bestehen bleiben (Green, 1971; siehe auch Richardson, 1996). So wäre es beispielsweise denkbar, dass eine Lehrkraft zwar grundsätzlich der Meinung ist, die Schule als Institution solle sich dem Wandel der Gesellschaft anpassen, sie aber keinen Wert darauf legt, ihre eigenen Unterrichtsmethoden anzupassen, wenn sich die Schülerschaft verändert. Hier läge dann ein Überzeugungsbündel zum Thema „Schule als Institution“ und ein Bündel „eigene Unterrichtsmethoden“ vor, die nicht fest miteinander verknüpft wären, sondern nebeneinander bestünden.

2.4.2 Überzeugungen und Wissen

Wie im einleitenden Teil (2.4.1) dieses Kapitels erwähnt, weisen Wissen und Überzeugungen eine enge Verbindung untereinander auf. Die Ähnlichkeit von Wissen und Überzeugungen trägt zur Ungenauigkeit des Begriffs „Überzeugungen“ bei, die in Kapitel 2.2 diskutiert wurde. Wie Pajares (1992, S. 315) feststellt, enthält auch im Fall von Überzeugungen und Wissen die Definition dieser Begriffe keine unumstrittene Wahrheit, sondern ist eine Unterscheidung, die zu dem Zweck getroffen wurde, mit Hilfe einheitlicher Begriffe kommunizieren zu können. Der Unterschied zwischen Überzeugungen und Wissen stellt somit eine paradigmatische Annahme dar, und es ist wichtig, die vereinbarten Definitionen in der Forschung konsistent, exakt und angemessen zu verwenden. Im Folgenden werden Überlegungen vorgestellt, die den Begriff „Wissen“ in ein Verhältnis zum Begriff „Überzeugungen“ stellen, und die Begriffe voneinander abgrenzen.

Um das Verhältnis von Überzeugungen und Wissen darzustellen, liegen unterschiedliche theoretische Ansätze vor: Überzeugungen werden als Teil des Wissens, gleichsam als eine bestimmte Art des Wissen (Furinghetti & Pehkonen, 2002; Turner et al., 2009), oder als ein für sich stehender Teilbereich von Kompetenz, also getrennt vom Wissen (Baumert & Kunter, 2006), betrachtet.

Kagan (1990, S. 421) verwendet die Begriffe „Überzeugungen“ und „Wissen“ synonym, da nach ihrer Ansicht ein Großteil des professionellen Wissens von Lehrkräften subjektiv („appears to be defined in highly subjective terms“) und daher den Überzeugungen ähnlich ist.

Werden Überzeugungen als eine Art von Wissen betrachtet, dann gelten sie zumeist als subjektives Wissen im Gegensatz zu objektivem Wissen (Furinghetti & Pehkonen, 2002, S. 43; Turner et al., 2009, S. 361): Als objektiv wird ein Wissen bezeichnet, das von einer Gemeinschaft akzeptiert wird (z.B. offizielle Fachkenntnisse). Persönliches Wissen, das nicht unbedingt der Bewertung durch einen Außenstehenden unterliegt, wird als subjektives Wissen bezeichnet. Überzeugungen repräsentieren hier das subjektive Wissen eines Einzelnen. Eine genauere Beschreibung von Überzeugungen als subjektivem Wissen⁶ liefern Furinghetti und Pehkonen (2002, S. 39–40):

„Individuals continuously receive signals from the world around them. According to their perceptions and experiences based on these messages, they draw conclusions about different phenomena and their nature. Individuals' subjective knowledge, i.e., their beliefs (including affective factors), is a compound of these conclusions.“

Überzeugungen repräsentieren subjektives Wissen, da sie nicht wie objektives Wissen von einer ganzen Gemeinschaft geteilt und auch nicht von außerhalb evaluiert werden (Philipp, 2007; Turner et al., 2009). Alba Thompson (1992, S. 130) erläutert diese Eigenschaft von Wissen aus epistemologischer Perspektive: Es gibt eine generelle Übereinkunft darüber, wie man die Validität von Wissen evaluiert und beurteilt. Dagegen ist einem Einzelnen bewusst, dass andere Menschen nicht immer die eigenen Überzeugungen teilen und ihre Standpunkte nicht unbedingt widerlegt werden können. Sie merkt auch an, dass sich die Evidenz, anhand derer ein Anspruch auf Wissen bewertet wird, im Laufe der Zeit ändern kann und Wissen dabei zu Überzeugungen wird (z.B. „die Erde ist eine Scheibe“).

Anstatt Überzeugungen als eine Art von Wissen anzusehen, wie es Furinghetti und Pehkonen (2002) tun, fassen Philipp (2007, S. 266) und Thompson (1992)

6 Dass Furinghetti und Pehkonen (2002) Überzeugungen als „subjektives Wissen“ bezeichnen, wird von Douglas und Susan McLeod (2002, S. 119) im selben Band als eine informelle Definition (von ihnen auch als „Daumenregel“-Definition bezeichnet) eingeordnet, die besonders gut dazu geeignet sei, das Konstrukt einem allgemeinen Publikum näherzubringen, weniger jedoch einem Fachpublikum.

Wissen als eine Form von Überzeugungen auf, die eine Person gerechtfertigterweise für wahr hält (siehe Abbildung 2.1).⁷ Reine Überzeugungen sind lediglich gefühlt wahr (Richardson, 1996). Dass der Begriff der „Wahrheit“ sowohl bei der Definition von Wissen als auch bei Überzeugungen eine wichtige Komponente darstellt, macht die Unterscheidung zwischen beiden Konstrukten nicht einfacher und sorgt für Ungenauigkeiten bei der Verwendung dieser Begriffe.

Überzeugungen sind im Gegensatz zu objektivem Wissen individuell, sie sind affektiv aufgeladen und haben eine episodische Struktur (Calderhead, 1996; Turner et al., 2009). Informationen in Wissenssystemen sind vor allem in semantischen Netzwerken gespeichert, während sich Überzeugungssysteme aus episodischem Material zusammensetzen, das aus persönlichen Erfahrungen oder kulturellen bzw. institutionellen Quellen gewonnen wird. Dass die episodische Struktur Überzeugungen klar von Wissen unterscheidet, ist umstritten (Nespor, 1987, S. 320), da es auch episodische Wissensmodelle gibt. Des Weiteren gibt es verschiedene Grade von Überzeugungen, beispielsweise starke zentrale Überzeugungen und schwache periphere. Diese Graduierung weist objektives Wissen dagegen nicht auf (Philipp, 2007; Thompson, 1992; Turner et al., 2009): Man kann von einer Überzeugung sehr stark überzeugt sein, eine Tatsache dagegen kann man nicht sehr stark wissen. Eine Zusammenfassung der Unterschiede zwischen (objektivem) Wissen und Überzeugungen gibt Tabelle 2.1.

7 Baumert und Kunter (2006, S. 496) teilen die Ansicht, dass „Wissen [...] gerechtfertigter wahrer Glaube [sei]“ auch wenn dies angesichts der getrennten Kompetenzfacetten von Wissen und Überzeugungen in ihrem Modell nicht deutlich wird (siehe Kapitel 2.1). Infolge der unterschiedlichen Rechtfertigungsansprüche von Wissen und Überzeugungen wird die kategoriale Differenz dieser beiden Begriffe an ihrem unterschiedlichen epistemologischen Status festgemacht. Allerdings seien die „Übergänge fließend“ (Baumert & Kunter, 2006, S. 496).

Tabelle 2.1: Unterschiede zwischen Wissen und Überzeugungen

	Wissen	Überzeugungen
Einvernehmen („consensuality“)	ja	nein
Individuell	nein	ja
Affektive Elemente	nein	ja
Episodische Struktur	(nein ⁸)	ja
Verschiedene Grade der Überzeugtheit („conviction“)	nein	ja

Anmerkung: eigene Darstellung nach Thompson (1992, S. 129–130) bzw. Philipp (2007, S. 259–260) sowie Turner et al. (2009)

Trotz der bisher beschriebenen Unterschiede werden Überzeugungen und Wissen in manchen Forschungsarbeiten sehr ähnlich behandelt, da dort beide als kognitive Konstrukte angesehen werden. Da die Beziehung zwischen Überzeugungen und Wissen unscharf bleibt, sollte es für die empirische Bildungsforschung, wenn sie sich mit der Professionalisierung von Lehrkräften beschäftigt, wichtig sein, ob die untersuchten Lehrkräfte selbst ihre Aussagen als Überzeugung oder als Wissen ansehen (Philipp, 2007, S. 266–268).

Auch der allgemeine Kompetenzbegriff (nicht nur der für die professionelle Kompetenz von Lehrkräften) stellt eine Beziehung zwischen Überzeugungen und Wissen her. In der Handlung kommen die drei Dimensionen Wissen, Können und Haltungen (und somit auch Überzeugungen)⁹ zusammen und bilden das „eigentliche Kompetenzbündel“ der Handlungskompetenz (Wildfeuer, 2009, S. 1796). Die Dimension des Wissens wird in diesem Abschnitt behandelt und von Überzeugungen abgegrenzt, die Dimension des Könnens allerdings kann im Rahmen dieser Arbeit nicht behandelt werden. Beim Modell der professionellen Handlungskompetenz von Lehrkräften von Baumert und Kunter (2006) sind Wissen und Überzeugungen kategorial getrennte Kompetenzfacetten, auch wenn sie die Übergänge zwischen den beiden als fließend ansehen (siehe dazu auch Fußnote 7).

8 Mit episodischer Struktur ist in Bezug auf Wissen nicht ein episodisches Wissen im Sinne eines prozeduralen Wissens bzw. Gedächtnisses (im Gegensatz zu deklarativem Wissen bzw. Gedächtnis) gemeint, sondern die Struktur eines Wissensnetzwerks im Vergleich zu einem Überzeugungsnetzwerk. Tiefergehende Beschreibungen zur episodischen Struktur von Überzeugungen im Gegensatz zu Wissen finden sich bei Nespor (1987) und Abelson (1979).

9 Wildfeuer (2009) verwendet sowohl den Begriff der Haltung als auch den der Werte zur Benennung dieses Teils der Handlungskompetenz. Beide Begriffe gehen sowohl ineinander als auch in den Begriff der Überzeugung über, wie in Abschnitt 2.4.3 deutlich wird.

Wenn man versucht, das Verhältnis von Überzeugungen und Wissen zu beleuchten, stößt man in der Forschungsliteratur, die sich mit den Überzeugungen von Lehrenden befasst, auch immer wieder auf den Begriff der „conceptions“, was häufig als „Vorstellungen“ übersetzt wird. Dieser Begriff hat sich vor allem in der Hochschulforschung im Zusammenhang mit Lehr- und Lernvorstellungen verbreitet (Reusser et al., 2011; als Beispiel siehe Trautwein, 2013). Ähnlich wie bei Wissen und Überzeugungen ist auch hier das Verhältnis der beiden Begriffe zueinander nicht eindeutig; bei Alba Thompson (1992) beispielsweise sind die „conceptions“ den Überzeugungen übergeordnet. Trotzdem stellt Philipp (2007, S. 259) fest, dass sie selbst die Begriffe „conceptions“ und Überzeugung (*belief*) teilweise auch synonym verwendet. Allerdings bemerkt Thompson (1992, S. 130): Die Unterscheidung zwischen *beliefs* und *conceptions* „may not be a terribly important one“. Sie wird deshalb im Rahmen dieser Arbeit auch nicht weiter behandelt.

Nach der bisherigen Diskussion der Unterscheidung zwischen Wissen und Überzeugungen kann Folgendes festgehalten werden: Wissen lässt sich als gerechtfertigt für wahr gehaltene Überzeugungen ansehen, Überzeugungen dagegen sind lediglich gefühlt wahr. In dieser Unterscheidung spiegelt sich der Unterschied zwischen der Einvernehmlichkeit des Wissens und der Individualität von Überzeugungen wider.

2.4.3 Überzeugungen und Werte

Wie bereits in Abschnitt 2.4.1 erläutert, lässt sich eine Überzeugung, die man tief in sich trägt und hochhält, als Wert bezeichnen (Philipp, 2007, S. 259). Insofern sind Werte nach dieser Betrachtungsweise eine bestimmte Art von Überzeugungen, sie unterscheiden sich jedoch von diesen. Es gibt weitere Aspekte, nach denen sich Werte von Überzeugungen unterscheiden: Werte werden eher mit der Dichotomie wünschenswert/nicht wünschenswert, Überzeugungen hingegen mit der Dichotomie wahr/falsch verbunden. Somit sind Überzeugungen auch kontextabhängiger als Werte, da ein Urteil über die Richtigkeit bzw. die Falschheit immer ein Objekt benötigt. Da Werte weniger kontextabhängig und stärker internalisiert sind als Überzeugungen, sind sie schwieriger zu verändern als Überzeugungen (Bishop, Seah & Chin, 2003). Beiden ist allerdings gemein, dass sie in Systemen organisiert sind (Rokeach, 1969). Rokeach (1969, S. 124) schreibt zum Verhältnis der Größenordnung von Überzeugungen und Werten, dass Personen „hundreds of thousands of beliefs, [...]“

but only dozens of values“ besitzen. Ähnlich wie bei den Überzeugungen in einem Überzeugungssystem (siehe Abschnitt 2.4.1) kann es auch Konstellationen von Werten geben, nach denen eine Person handelt, die untereinander inkompatibel sind und miteinander konfliktieren (Bishop et al., 2003; Philipp, 2007; Rokeach, 1969).

„A belief *that* is about beliefs, but a belief *in* is about values. [Hervorhebung im Original]“; so beschreibt Philipp (2007, S. 265) in einem Satz den Unterschied zwischen Überzeugungen und Werten. Ein Wert stellt für ihn somit ein Bekenntnis („creed“) dar, an dem Menschen beispielsweise auch in Krisenzeiten festhalten. Eine Überzeugung hat für ihn im Gegensatz dazu einen größeren Alltagsbezug. Angewandt auf den Schulalltag lässt sich der Unterschied folgendermaßen beschreiben: Wenn eine Lehrkraft denkt, dass die Schülerinnen und Schüler zum Lernen Druck benötigen und ohne Druck von außen keine Lernfortschritte erzielen werden, dann handelt es sich um eine Überzeugung. Wenn eine Lehrkraft meint, Autorität sollte generell der wichtigste Bestandteil der Erziehung von Kindern sein, handelt es sich um einen Wert.

Demnach kann ein Wert also als wünschenswerte Überzeugung bezeichnet werden (siehe Abbildung 2.1). Es gibt jedoch noch andere Ansätze, die beschreiben, wie Überzeugungen und Werte zueinander in Bezug stehen können. Werte werden manchmal auch als Teil von Überzeugungen betrachtet: so Rokeach (1969), der Werte als dauerhafte Überzeugungen beschreibt, oder bei Clarkson und Bishop (2000; siehe Bishop et al., 2003), wo Werte als „beliefs in action“, also Handlungsüberzeugungen, verstanden werden. Bei Raths, Harmin und Simon (1987) sind Werte Überzeugungen, die nach einer Reflexion frei gewählt werden und die man schätzt und bekräftigt und nach denen man seine Handlungen ausrichtet (siehe auch Rokeach, 1969). Überzeugungen können Werte auch hervorbringen. So beschreibt Pajares (1992, S. 319) Überzeugungskuster (siehe Abschnitt 2.4.1), die um ein bestimmtes Objekt oder eine bestimmte Situation herum Einstellungen bilden, die wiederum zu Handlungsagenden werden. Die Überzeugungen stehen miteinander und mit anderen Überzeugungen in anderen Einstellungen in Verbindung.¹⁰ Pajares (1992) beschreibt als Beispiel die Einstellung einer Lehrkraft zu einem bestimmten Bildungsproblem, die solche Überzeugungen beinhalten kann, die mit Einstellungen hinsichtlich der Natur der Gesellschaft, der Gemeinschaft und sogar der

10 Weitere Ausführungen zum Verhältnis von Überzeugungen und Einstellungen siehe Abschnitt 2.4.4.

Familie verbunden sind. Diese eben genannten Verbindungen schaffen die Werte, die das Leben leiten, andere Einstellungen entstehen lassen und aufrechterhalten sowie Informationen interpretieren und das Verhalten bestimmen.

Doch auch der Begriff der „values“ in der englischsprachigen Bildungsforschung ist (ähnlich wie „beliefs“, siehe Kapitel 2.2) nicht immer scharf umrissen. So sieht Campbell (2003, S. 17) den Begriff der „values“ als einen begrifflichen Ausgleicher über verschiedene andere Begriffe (dazu zählt er auch Überzeugungen) hinweg, als „the great conceptual equalizer of all preference, opinion, belief, and attitude“. Das Verhältnis von Werten zu Überzeugungen wird in der deutschsprachigen Bildungsforschung nicht im gleichen Rahmen thematisiert wie im englischsprachigen Raum. Dort wird, wie bisher beschrieben, zwischen Werten („values“) und Überzeugungen („beliefs“) unterschieden. Im Deutschen wird in der Bildungsforschung eher der Begriff der „Werthaltungen“ (von Lehrkräften) (Baumert & Kunter, 2006) als der Begriff der „Werte“ verwendet,¹¹ wobei sich dann durch die Verbindung mit dem Wort „Haltungen“¹² (Schwer & Solzbacher, 2014) sogleich eine erneute begriffliche Vermischung ergibt: Die Unterscheidung von Werthaltungen, Haltungen und Überzeugungen von Lehrkräften ist, zumindest bei den Paarungen Werthaltungen – Haltungen und Haltungen – Überzeugungen fließend.

Werte von Lehrkräften werden thematisiert, wenn die Lehrkraft als einer Profession angehörig betrachtet wird. So sind Werte ein Bestandteil des „professionellen Selbst“ von Lehrkräften (Bauer, Kopka & Brindt, 1996, S. 234): „Das professionelle Selbst ist die auswählende, ordnende, entscheidende und wertorientiert handelnde Instanz, die den Zusammenhang zwischen beruflicher Erfahrung, Diagnosekompetenz, Handlungsrepertoire und pädagogischen Werten und Zielen herstellt.“ Bauer et al. (1996) wählten den Begriff des professionellen Selbst, um den Unterschied zwischen der Persönlichkeit einer Lehrkraft als Ganzes und dem für ihr berufliches Handeln relevanten Teil der Person zu verdeutlichen.

Beim Modell der professionellen Handlungskompetenz von Baumert und Kunter (2006) werden Werthaltungen gemeinsam mit Überzeugungen betrach-

11 Zur synonymen Verwendung der Begriffe „Werte“ und „Werthaltungen“ siehe Harder (2014).

12 Die Haltung des Lehrers oder der Lehrerin ist ein Begriff, der heute wieder vermehrt verwendet wird und in der Geschichte der Pädagogik eine lange Tradition hat, wenn es um „gute“ Lehrkräfte und Lehrerpersönlichkeiten geht (Fiebert & Solzbacher, 2014, S. 17).

tet und in vier Bereiche ausdifferenziert: Wertbindungen, epistemologische Überzeugungen (siehe Kapitel 3.1), subjektive Theorien über Lehren und Lernen (siehe Kapitel 3.2) und Zielsysteme für Curriculum und Unterricht (diese werden bei Baumert und Kunter zusammen mit den subjektiven Theorien abgehandelt). Die berufsbezogenen Wertbindungen von Lehrkräften werden auch häufig als „Berufsethos“ oder „Berufsmoral“ bezeichnet (Oser, 1998b). Die zentralen Dimensionen eines beruflichen Ethos von Lehrkräften sind nach Oser (1998a, S. 37–38) Gerechtigkeit, Fürsorge, Wahrhaftigkeit, Verantwortung und Engagement. Baumert und Kunter (2006, S. 498) sehen im Bereich des Berufsethos von Lehrkräften eine Forschungslücke hinsichtlich seiner Handlungsrelevanz:

„Es ist bis heute unklar, welche Auswirkungen spezifische Wertpräferenzen für das professionelle Handeln von Lehrkräften tatsächlich haben [...]. Es ist anzunehmen, dass die Ausprägungen der Berufsmoral sowohl für den Umgang mit Heterogenität als auch für die Unterstützungsqualität von Lernumgebungen [...] bedeutsam sind.“

Beim Berufsethos ist die Frage zentral, wie im Berufsalltag eine situationsadäquate Umsetzung gelingen kann, da etwa von Blömeke (2010, S. 219) festgestellt wurde, dass es Lehrkräften in konkreten Unterrichtssituationen häufig nicht zufriedenstellend gelingt, ihr Berufsethos in entsprechende Handlungen umzusetzen (siehe dazu Kapitel 4.1). Der Beitrag der Theorie des beruflichen Ethos von Lehrkräften liegt für diese Arbeit in der Betonung der Verantwortung der Lehrkräfte für das Wohl von Schülerinnen und Schülern als zentrale Dimension des Lehrerberufs.

Zusammenfassend lässt sich sagen, dass sich Werte vor allem dadurch von Überzeugungen unterscheiden, dass sie als wünschenswert angesehen werden, Überzeugungen hingegen werden mit der Dichotomie wahr/falsch verknüpft. Außerdem sind Überzeugungen kontextabhängiger als Werte. Gerade im deutschsprachigen Raum ist die Nähe von Überzeugungen zu Wertvorstellungen jedoch groß, sodass sich der theoretische Horizont in diesem Forschungsfeld nicht auf Arbeiten zu einem dieser Begriffe beschränken sollte.

2.4.4 Überzeugungen und Einstellungen

Auch das Verhältnis von Überzeugungen zu Einstellungen ist zu klären, da die in der vorliegenden Arbeit untersuchten Überzeugungen von Lehrkräften (Re-

formbereitschaft, Einstellung zur Notwendigkeit von Druck bei Lernprozessen und Einstellung zur Förderung schwacher Schülerinnen und Schüler) teilweise sogar in ihrer Bezeichnung die „Einstellungen“ im Namen tragen. Wie in Abschnitt 2.4.1 bereits erläutert wurde, sind Einstellungen Handlungs-, Gefühls- oder Denkweisen, welche die eigene Disposition oder Meinung anzeigen (Philipp, 2007, S. 259).

Es gibt unterschiedliche Auffassungen hinsichtlich des Unterschieds von Überzeugungen und Einstellungen. Bei Rokeach (1969, S. 132) sind Einstellungen dauerhaft miteinander verbundene Überzeugungen: Eine Einstellung („attitude“) ist „a relatively enduring organization of interrelated beliefs that describe, evaluate, and advocate action with respect to an object or situation“. Ähnlich werden Einstellungen von Fishbein und Ajzen (1975, S. 6) als „learned predispositions to respond in a consistently favourable or unfavourable manner with respect to a given object“ beschrieben, allerdings hier ohne Bezug zu Überzeugungen.¹³ Einstellungen weisen demnach einen gewissen Grad von Konsistenz auf, sie prädisponieren Handlungen und sie sind erlernt. Einstellungen und Überzeugungen unterscheiden sich jeweils hinsichtlich ihrer Veränderbarkeit, darin, wie sehr sie ein kognitives Konstrukt sind, und schließlich in der Intensität ihrer Empfindung (Philipp, 2007; siehe Abschnitt 2.4.1). Einen Überblick über das genauere Verhältnis von Einstellungen und Überzeugungen bietet Richardson (1996). Sie beschreibt darüber hinaus, wie sich in der Bildungsforschung der Fokus von der Betrachtung von Einstellungen zu der von Überzeugungen verschoben hat. Pajares (1992) stellt fest, dass sich die beiden Begriffe „Einstellungen“ und „Überzeugungen“ vor allem hinsichtlich des Kontextes ihrer Verwendung unterscheiden und dass sich im Schulkontext für die Einstellungen von Lehrerinnen und Lehrern der Begriff „Überzeugungen“ etabliert hat. Die vorliegende Arbeit folgt diesem Hinweis und verwendet den Begriff „Überzeugungen“.

2.4.5 Fazit zur Abgrenzung von Überzeugungen zu anderen Begriffen

In den vorangegangenen Abschnitten wurde das Verhältnis von Überzeugungen zu verwandten Begriffen wie Wissen, Werten und Einstellungen beleuchtet und Abgrenzungen wurden vorgenommen. Es wurde deutlich, dass es auch von der

13 Die Arbeit von Fishbein und Ajzen (1975) stellt aus sozialpsychologischer Sicht Überzeugungen, Einstellungen, Intentionen und Verhalten zueinander in Beziehung, allerdings mit einer starken Schwerpunktsetzung auf die Einstellungen.

Betrachtungsweise abhängt, wie streng man diese anderen Begriffe von Überzeugungen unterscheidet oder ob man die ihnen gemeinsamen Aspekte sucht. Als Fazit der Abgrenzung schlägt Trautwein (2013) vor, das Verhältnis von Überzeugungen zu solchen ihnen verwandten Begriffen wie Wissen oder Einstellungen nicht als ein kategorial Verschiedenes, sondern als ein graduell Anderes zu betrachten: Es gibt kognitive Überzeugungen, die explizit und veränderbar sind und deren Unterscheidung gegenüber Wissen sich schwierig gestaltet. Affektive Überzeugungen hingegen sind implizit, schwer veränderbar und die Unterscheidung zwischen ihnen und Einstellungen ist schwierig.

Pajares (1992, S. 329) findet schließlich das Konstrukt der Überzeugungen von Lehrkräften „less messy, far cleaner, and conceptually clearer“, als es scheint. Voraussetzung dafür ist, dass Überzeugungen klar begrifflich herausgearbeitet, ihre Grundannahmen untersucht, ihre genauen Bedeutungen konsequent verstanden und eingehalten sowie spezifische Überzeugungskonstrukte angemessen erfasst und untersucht werden. Diese Erläuterung des Gegenstands, auf den sich die Untersuchung bezieht, wird in der vorliegenden Arbeit aus theoretischer Perspektive in den Kapiteln 3.3 und 4.1 dargestellt. Im weiteren Verlauf der Arbeit werden die hier untersuchten Überzeugungskonstrukte empirisch gestützt vorgestellt (Abschnitt 7.2.1) und untersucht (Kapitel 8).

In der vorliegenden Arbeit werden konkret die Einstellungen von Lehrkräften gegenüber der Förderung schwacher Schülerinnen und Schüler und zur Notwendigkeit von Druck bei Lernprozessen sowie ihre Reformbereitschaft untersucht. Dabei wird der Schwerpunkt darauf gelegt, wie diese Einstellungen und die Reformbereitschaft mit ihrem Unterrichtshandeln und schließlich mit den Lernergebnissen der Schülerinnen und Schüler zusammenhängen. Will man dies untersuchen, dann liegt die größte Anschlussfähigkeit dieser drei Überzeugungen zusammengefasst unter dem Begriff der allgemeinen pädagogischen Überzeugungen an die Unterrichtsforschung und an die professionelle Kompetenz von Lehrkräften.

3 Empirisch relevante Bereiche bei Überzeugungen von Lehrkräften

Nachdem im vorangegangenen Kapitel 2 die Überzeugungen von Lehrkräften aus verschiedenen theoretischen Ansätzen erläutert und zu anderen Begriffen in Bezug gesetzt wurden, werden in den folgenden Kapiteln 3.1 und 3.2 epistemologische Überzeugungen und subjektive Theorien die Überzeugungen von Lehrkräften beschreiben. Die in dieser Arbeit untersuchten pädagogischen Überzeugungen werden, da sie nicht vollständig in einem der beiden genannten Bereiche aufgehen können, im Kapitel 3.3 beschrieben.

3.1 Epistemologische Überzeugungen

Theoretische Begriffsbestimmungen zu „Überzeugungen“ (siehe Kapitel 2.2 und 2.3) werden meistens sehr weit gefasst. In den empirischen Untersuchungen zu Überzeugungen von Lehrkräften dominiert jedoch eine sehr konkrete, unterrichtsnahe und häufig fachspezifische Operationalisierung (Läge & McCombie, 2015). Zurzeit werden in empirischen Untersuchungen vor allem epistemologische Überzeugungen betrachtet (Biedermann, Brühwiler & Krattenmacher, 2012; Blömeke, Müller, Felbrich & Kaiser, 2008; Hartinger et al., 2006; Staub & Stern, 2002; Voss, Kleickmann, Kunter & Hachfeld, 2011). Dabei handelt es sich um Überzeugungen, die sich konkret auf die Struktur und Genese des Wissens in einem bestimmten Fach beziehen. Sie werden vor allem für den Mathematikunterricht untersucht. Auch Überblicksarbeiten zu den Überzeugungen von Lehrkräften beziehen sich bei ihren theoretischen Überlegungen auf die zentrale Rolle, welche die Überzeugungen beim Lehren und Lernen von Mathematik spielen (Goldin, Rösken & Törner, 2009; Philipp, 2007; Thompson, 1992). In den empirischen Arbeiten, die im Folgenden vorgestellt werden, werden neben epistemologischen Überzeugungen im Fach Mathematik, also Überzeugungen hinsichtlich des mathematischen Wissens, auch Überzeugungen hinsichtlich des Lehrens und Lernens von Mathematik behandelt. Diese letzteren Überzeugungen sind strenggenommen nicht den epistemologischen Überzeugungen zuzuordnen, werden aber in diesem Kapitel gemeinsam damit abgehandelt. Eine strenge Unterscheidung zwischen diesen beiden Fach-Überzeugungen ist für lerntheoretische Arbeiten von Bedeutung, bei empirischen Untersuchungen hat sich eine integrative Betrachtung von epistemologischen Überzeugungen und Überzeugungen hinsichtlich des Lehrens und

Lernens als sinnvoll erwiesen, da sich beide entweder an konstruktivistischen oder transmissiven Lerntheorien orientieren (Voss et al., 2011, S. 238).

Epistemologische Überzeugungen sind subjektive Überzeugungssysteme, die sich auf die Struktur des Wissens oder die Struktur des Lernens (Wissenserzeugung) beziehen (Helmke, 2015, S. 115). Bei epistemologischen Überzeugungen von Lehrkräften wird, wie bereits erwähnt, häufig zwischen konstruktivistischen und transmissiven Überzeugungen unterschieden. Diese beruhen auf konstruktivistischen und transmissiven Lerntheorien, die Annahmen über den Erkenntnis- bzw. Wissensaufbau von Schülerinnen und Schülern enthalten, aber keine Empfehlung für eine bestimmte Unterrichtsform implizieren (Lipowsky, 2006, S. 59). Konstruktivistische Überzeugungen betonen, dass Schülerinnen und Schüler sich aktiv mit dem Lerngegenstand auseinandersetzen sollten, da sie diesem mit bestimmten Vorstellungen und Voraussetzungen entgegentreten. Wissen wird hier als prozesshaft und als Ergebnis subjektiver Konstruktionsprozesse verstanden. Hinsichtlich transmissiver Überzeugungen wird Lernen hauptsächlich als die Weitergabe von Wissen an die Schülerinnen und Schüler verstanden. Wissen wird hier als objektiv feststehend, und die Lernenden werden als passive Rezipienten angesehen (Voss et al., 2011, S. 238–239). Fives, Lacatena und Gerard (2015, S. 261) merken allerdings auch an, dass konstruktivistische Überzeugungen eine große Bandbreite von Überzeugungen über das Lernen umfassen und somit auch eine Vielzahl von Unterrichtshandlungen damit verbunden sein können. Transmissive Überzeugungen hingegen beruhen auf einem eindimensionalen, lehrkraftzentrierten Unterrichtsmodell.¹⁴ Epistemologische Überzeugungen von Lehrkräften wurden auch im Rahmen international-vergleichender empirischer Studien zur Lehrerbildung (MT-21, TEDS-M, TALIS, COACTIV) untersucht. Die wichtigsten Ergebnisse dieser Studien zu den Überzeugungen von Lehrkräften werden im Folgenden kurz vorgestellt.

Die Studien MT-21¹⁵ (Blömeke, Müller et al., 2008) und die darauf aufbauende Studie TEDS-M¹⁶ (Felbrich, Schmotz & Kaiser, 2010; Schmotz, Felbrich &

14 „As such, the term ‘constructivism’ could mean any range of beliefs about *learning* and in turn could suggest a variety of *teaching* practices that reflect those beliefs. In contrast, beliefs that teaching is ‘transmission’ seem to reflect a one-dimensional model of instruction rooted in a behaviorist conception of learning and a teacher-centered belief about teaching“ (Fives, Lacatena & Gerard, 2015, S. 261). [Hervorhebungen im Original]

15 MT-21: Mathematics Teaching in the 21st Century

16 TEDS-M: Teacher education and development study: Learning to teach mathematics

Kaiser, 2010) haben sich intensiv mit epistemologischen Überzeugungen von angehenden Mathematiklehrkräften befasst. Allerdings waren im Untersuchungsdesign dieser Studien keine Verbindungen zum Unterrichtshandeln der Lehrkräfte oder zu den Lernergebnissen der Schülerinnen und Schülern vorgesehen, da sie sich auf Lehramtsstudierende und Lehrkräfte im Referendariat beschränkt hatten. Müller, Felbrich und Blömeke (2008) nehmen im Rahmen von MT-21 auch die handlungsnäheren Überzeugungen zum Lehren und Lernen von angehenden Mathematiklehrkräften in den Blick: Als unterrichtsbezogene Überzeugungen betrachten sie die Vorstellungen angehender Lehrkräfte bezüglich der Lernziele des Mathematikunterrichts, der unterrichtsmethodischen Präferenzen und der Überzeugungen zum „Classroom Management“. Sowohl für die epistemologischen als auch für die unterrichtsnäheren Überzeugungen konnte bei MT-21 herausgefunden werden, dass diese innerhalb ihrer Dimensionen und auch zwischen den Subdimensionen übereinstimmend mit ihren Erwartungen konsistent zusammenhängen (Blömeke, Müller et al., 2008; Müller et al., 2008). So hängen die begabungs- und erkenntnistheoretischen Subdimensionen wie erwartet zusammen (Blömeke, Müller et al., 2008, S. 235):

„Die Vorstellungen von mathematischen Fähigkeiten als anthropologisch konstant, determiniert durch Herkunftsmerkmale und im Lebenslauf nicht veränderbar und die auf Veränderbarkeit ausgerichtete Vorstellung eines Conceptual Change korrelieren ebenso latent negativ wie die erkenntnistheoretisch ausgerichteten Transmissions- und Konstruktionsvorstellungen zum Erwerb mathematischer Kompetenz [...].“

Blömeke und ihre Kolleginnen (2008, S. 240) sprechen von einem „umfassenden Überzeugungs-Syndrom“. In der TEDS-M-Studie wurden im internationalen Vergleich Überzeugungen angehender Lehrkräfte der Primarstufe (Felbrich et al., 2010) und der Sekundarstufe I (Schmotz et al., 2010) zum Wesen der Mathematik und zum Lehren und Lernen von Mathematik untersucht. Die Zuordnung von Überzeugungsprofilen zu kulturell geprägten Ländergruppen in dieser Studie deutet darauf hin, dass es eine Passung von grundlegenden gesellschaftlichen Überzeugungen und Überzeugungen zum Erwerb sowie zur Struktur mathematischen Wissens gibt (Felbrich et al., 2010; Schmotz et al., 2010).

Die beiden Studien TALIS¹⁷ (OECD, 2009) und COACTIV¹⁸ (Voss et al., 2011) haben neben den epistemologischen Überzeugungen von Mathematik-

17 TALIS: Teaching and Learning International Survey

18 COACTIV: Cognitive Activation in the Classroom

Lehrkräften auch ihr Unterrichtshandeln und die Ebene der Schülerinnen und Schüler einbezogen. In der TALIS-Studie wurden Überzeugungen, Praktiken und Einstellungen untersucht, die sich in früheren Untersuchungen als relevant für die Verbesserung und die Wirksamkeit von Schulen erwiesen hatten (OECD, 2009, S. 89). TALIS ist eine repräsentative internationale Large-Scale-Studie, die in 23 Ländern (allerdings nicht in Deutschland) durchgeführt wurde. Als Überzeugungen zum Lehren und Lernen wurden transmissive und konstruktivistische Überzeugungen erhoben. In den meisten der untersuchten Länder waren konstruktivistische Überzeugungen stärker vertreten als transmissive. Das bedeutet, dass die Lehrkräfte die Schülerinnen und Schüler eher als aktive Teilnehmende am Prozess des Wissenserwerbs betrachten, als die Hauptrolle der Lehrkraft in der Informationsvermittlung und dem Aufzeigen richtiger Lösungen zu sehen (OECD, 2009, S. 94). Daneben wurden auch Einstellungen zu bestimmten Unterrichtspraktiken wie Strukturiertheit und Schülerorientierung erfasst. In allen Ländern zeigte sich das Muster, dass die Lehrkräfte im Unterricht mehr Wert auf Strukturierung legen als auf schülerorientierte Aktivitäten, die den Schülerinnen und Schülern größere Autonomie geben (OECD, 2009, S. 97–99). Allerdings konnte die TALIS-Studie kein klares Bild zum Zusammenhang zwischen den Überzeugungen der Lehrkräfte zu bestimmten Unterrichtsmethoden und ihrem beruflichen Handeln zeichnen. In vielen Ländern wurden positive Zusammenhänge zwischen transmissiven Überzeugungen und strukturierenden Unterrichtspraktiken gefunden. In einigen Ländern trat dieser Zusammenhang auch für konstruktivistische Überzeugungen auf. Für die Schülerorientierung im Unterricht gab es in vielen Ländern einen positiven Zusammenhang mit konstruktivistischen Überzeugungen und häufig keinen Einfluss von transmissiven Überzeugungen (OECD, 2009, S. 131).

Bei COACTIV handelt es sich um eine repräsentative Längsschnittstudie, in der sowohl Lehrkräfte als auch Schülerinnen und Schüler untersucht wurden. Dort konnte gezeigt werden, dass die Überzeugungen von Mathematiklehrkräften für das Unterrichtshandeln und den Lernerfolg in Mathematik bedeutsam sind (Voss et al., 2011). Transmissive Überzeugungen stehen in einem negativen, konstruktivistische Überzeugungen in einem positiven Zusammenhang mit Unterrichtsqualität und Lernerfolg. Als Mediator konnte das Potential zur kognitiven Aktivierung ermittelt werden. Beim Mediationseffekt für konstruktive Unterstützung zeigte sich aufgrund von Suppressionseffekten ein weniger klares Bild (Dubberke, Kunter, McElvany, Brunner & Baumert, 2008). Ein Ein-

fluss der transmissiven und konstruktivistischen Überzeugungen der Lehrkräfte auf effiziente Klassenführung konnte nicht festgestellt werden (Voss et al., 2011, S. 250).

Im Folgenden werden weitere Arbeiten aufgeführt, welche die epistemologischen Überzeugungen der Lehrkräfte mit dem Unterrichtshandeln oder den Lernergebnissen der Schülerinnen und Schüler in Verbindung gebracht haben. Hartinger et al. (2006) fanden Hinweise auf Zusammenhänge zwischen einem konstruktivistisch orientierten Lehr-Lernverständnis von Grundschullehrkräften und einer Öffnung des Unterrichts im Hinblick auf die Entscheidungsmöglichkeiten von Schülerinnen und Schülern. Außerdem empfinden sich die Schülerinnen und Schüler als selbstbestimmter und schätzen den Unterricht als interessanter ein, wenn ihre Lehrkraft überwiegend konstruktivistische Vorstellungen von Lehren und Lernen mitbringt. Pauli, Reusser und Grob (2007) konnten einen Zusammenhang zwischen einer konstruktivistischen Orientierung von Mathematiklehrkräften und einem Unterricht, der aus Sicht der Lehrkräfte den Schülerinnen und Schülern Möglichkeiten für Problemlösen und selbstreguliertes Lernen bietet, nachweisen. Staub und Stern (2002) untersuchten die Beziehung zwischen den epistemologischen Überzeugungen von Grundschullehrkräften und dem Leistungszuwachs der Schülerinnen und Schüler im Fach Mathematik. Bedeutsame Anteile der Unterschiede bei den Leistungszuwächsen der Schülerinnen und Schüler konnten durch Unterschiede bei den konstruktivistischen Überzeugungen der Lehrkräfte erklärt werden.

In der Übersicht von Handal (2003) wird die Beziehung zwischen den epistemologischen Überzeugungen und der Unterrichtspraxis von Mathematiklehrkräften als dialektisch beschrieben, die durch viele widersprüchliche Faktoren vermittelt wird. Die Überzeugungen der Lehrkräfte sind zwar für ihre Unterrichtspraxis bedeutsam, jedoch können aufgrund der Überlagerung durch Eventualitäten, die in die Schul- und Klassenkultur eingebettet sind, keine Kausalzusammenhänge festgestellt werden. Solche vermittelnde Faktoren sind unter anderem der Prüfungsdruck, administrative Anforderungen, die Erwartungen von Seiten der Lernenden und ihren Eltern, ein Ressourcenmangel oder das Verhalten der Schülerinnen und Schüler. Weitere Zusammenstellungen empirischer Ergebnisse zu epistemologischen Überzeugungen im Mathematikunterricht aus dem englischen Sprachraum finden sich in den Handbuchartikeln von Thompson (1992) und Philipp (2007). Einen aktuelleren Überblick über Arbei-

ten zu den Überzeugungen von Lehrkräften, die das Fach Mathematik betreffen, bieten Cross Francis, Rapacki und Eker (2015).

In diesem Kapitel konnte aufgezeigt werden, dass es Hinweise auf die Bedeutung der Überzeugungen von Lehrkräften, insbesondere in Form epistemologischer Überzeugungen, auf ihr Unterrichtshandeln und auch auf die Lernergebnisse von Schülerinnen und Schülern gibt. Bisher liegen hauptsächlich Ergebnisse für Mathematik oder Naturwissenschaften vor (Baumert & Kunter, 2006, S. 498), für das Fach Englisch konnten keine einschlägigen empirischen Arbeiten gefunden werden. Epistemologischen Überzeugungen wird zwar eine erhebliche Bedeutung für die Unterrichtswahrnehmung und für das berufliche Handeln von Lehrkräften zugeschrieben, die empirische Forschungslage dazu sei aber immer noch unbefriedigend (Baumert & Kunter, 2006, S. 499). Helmke (2015, S. 115) folgert trotzdem, dass „die Annahme [...] plausibel [ist], dass solche intuitiven Überzeugungen für die Gestaltung des Unterrichts und für die Erklärung schulischer Erfolge oder Misserfolge der Schüler folgenreich sind.“

Epistemologische Überzeugungen sind von ihrer Struktur her verwandt mit subjektiven Theorien (Helmke, 2015, S. 115), die im folgenden Abschnitt beschrieben werden. Im Modell professioneller Handlungskompetenz von Baumert und Kunter (2006) bilden subjektive Theorien über Lehren und Lernen zusammen mit Wertbindungen, epistemologischen Überzeugungen und Zielsystemen für Curriculum und Unterricht den Komplex „Werthaltungen und Überzeugungen“ (siehe Kapitel 2.1).

3.2 Subjektive Theorien

Im Schulkontext wurde die Forschung zu den Überzeugungen von Lehrkräften lange vom Konzept der subjektiven Theorien dominiert (Trautwein, 2013). Eine Übersicht über empirische Arbeiten zu subjektiven Theorien in der Forschung zu Lehrkräften findet sich bei Seifried (2009, S. 82). Subjektive Theorien wurden vor allem hinsichtlich schwieriger Unterrichtssituationen, beispielsweise aufgrund aggressiver Schülerinnen und Schüler, erforscht sowie bezüglich der Gestaltung von Gruppenarbeit (Helmke, 2015, S. 115). Mittlerweile findet der Begriff der Überzeugungen breitere Verwendung (Seifried, 2009), da dieser im Vergleich mit internationaler Forschung und auch hinsichtlich des Begriffs „beliefs“ anschlussfähig ist (Reusser et al., 2011).

Das Verständnis von subjektiven Theorien hat im deutschsprachigen Raum eine starke Prägung durch das gleichnamige Forschungsprogramm (Groeben et al., 1988) erfahren (siehe Reusser et al., 2011; Seifried, 2009); dieses ist in der psychologischen Handlungstheorie zu verorten (Wahl, 2002). Demnach sind subjektive Theorien komplexe Aggregate von Kognitionen der Selbst- und Weltsicht mit impliziter Argumentationsstruktur (Groeben et al., 1988). Subjektive Theorien weisen Ähnlichkeiten mit den Überzeugungen, wie sie hier definiert werden, auf, jedoch auch Unterschiede. Subjektive Theorien besitzen die gleichen Strukturen und Funktionen wie wissenschaftliche Theorien. Sie dienen der Beschreibung, Erklärung und Vorhersage von Begebenheiten und kommen darüber hinaus bei der Herstellung und Gestaltung der Lebenswelt zum Einsatz (Schlee, 2017, S. 215), während Überzeugungen von Lehrkräften als Vorstellungen über Lehren, Lernprozesse bzw. Lerninhalte sowie Lehr- und Lernkontexte angesehen werden. Gemeinsam ist den Begriffen der subjektiven Theorien und der Überzeugungen, dass sie sowohl die Sicht auf das Selbst als auch die Sicht auf die Umwelt beschreiben. Ein Vorteil der subjektiven Theorien besteht darin, dass es dabei ein Erklärungsmodell zum theoretischen Zusammenhang zwischen Wissen (und subjektiven Theorien als Teil von Wissen) und dem Handeln von Lehrkräften gibt. Wahl (2006, S. 16–28) beschäftigt sich mit der Frage, warum die Umsetzung von (trägem) Wissen in (kompetentes) Handeln so schwierig ist und benennt drei kritische Punkte, die er im Bereich der Lehrerbildung untersucht hat: Erstens umfasst Handeln wesentlich mehr als Tun oder Verhalten. Handeln ist zielgerichtet und bewusst, die Akteurinnen und Akteure greifen dabei auf soziale wie auch auf individuelle Wissensstrukturen zurück. Allerdings ist einschränkend anzumerken, dass Handeln zwar ein bewusster Vorgang ist, aber nicht alle internen Prozesse beim Handeln „bewusstseinspflichtig“ sind (Wahl, 2002, S. 229). Beim Handeln findet die sinnhafte Integration von Denken, Fühlen und Agieren statt. Dabei springt die Aufmerksamkeit zwischen verschiedenen Ebenen der Handlungsregulation hin und her.

Zweitens kann der Handlungsablauf durch zwei Phasen konzeptualisiert werden: die Situationsorientierung und die Aktionsplanung. Bei beiden Handlungsschritten greifen die handelnden Personen auf subjektive Theorien zurück. Bei der Situationsorientierung versuchen die Akteurinnen und Akteure, die Situation einzuschätzen und greifen dabei auf biografisch erworbenes Wissen, gefasst als subjektive Theorien, zurück. Bei der Aktionsplanung werden vor dem Hintergrund der Situationsorientierung Ziele aufgestellt, wobei subjektive Theorien

hier das Wissen über mögliche Handlungen und ihre Wirkungen liefern. Ein dritter kritischer Punkt, warum verändertes Wissen sich nicht immer in veränderte Handlung übersetzen lässt, ist die Tatsache, dass es bei den Akteurinnen und Akteuren Wissensbestände gibt, die das Handeln leiten, und solche, die dies nicht tun. Übertragen auf subjektive Theorien wird zwischen subjektiven Theorien größerer und kürzerer Reichweite unterschieden (Wahl, 2002, 2006). Subjektive Theorien größerer Reichweite stellen semantische Netzwerke dar, die nicht direkt mit der Steuerung des Handelns verbunden sind. Subjektive Theorien kürzerer Reichweite hingegen sind handlungsnah, stabil und dem Bewusstsein leicht zugänglich. Sie werden vor allem dann eingesetzt, wenn schnelle Handlungen unter Druck erforderlich sind.

3.3 Pädagogische Überzeugungen

Im Angebots-Nutzungs-Modell zur Wirkungsweise des Unterrichts von Helmke (2015) werden bei den Lehrpersonen fachliche, didaktische, diagnostische und Klassenführungscompetenz, Erwartungen und Ziele sowie Engagement, Humor und Geduld und zusätzlich dazu pädagogische Orientierungen aufgeführt. Dieser Komplex wirkt sich nach diesem Modell auf den Unterricht (das Angebot) aus und hat über die Lernaktivitäten der Schülerinnen und Schüler (die Nutzung) Auswirkungen auf ihre fachlichen und überfachlichen Kompetenzen. Interessant ist, dass in den weiteren Ausführungen zu den unterrichtsrelevanten Merkmalen und Orientierungen von Lehrpersonen nur auf subjektive Theorien und epistemologische Überzeugungen näher eingegangen wird. Allgemeinere pädagogische Orientierungen oder eben Überzeugungen (siehe zu den Begriffen Kapitel 2.2) finden keine Erwähnung. Doch nicht nur in theoretischen Rahmenmodellen zum Unterricht oder zur professionellen Kompetenz von Lehrkräften (siehe Kapitel 2.1) tauchen pädagogische Überzeugungen eher am Rande auf, auch die empirische Forschungslage zum Bereich der pädagogischen Überzeugungen ist, gerade in Deutschland, begrenzt. Insbesondere besteht Forschungsbedarf in Bezug darauf, welche Überzeugungen einen Einfluss auf den Unterricht haben, für welche Aspekte des Unterrichtshandelns die Überzeugungen der Lehrkräfte relevant sind und auf welche Ergebnisse der Schülerinnen und Schüler sich diese auswirken (Baumert & Kunter, 2006; Reusser et al., 2011; Voss et al., 2011).

Ertmer, Ottenbreit-Leftwich und Tondeur (2015, S. 403–404) definieren pädagogische Überzeugungen als jede Aussage, die sich speziell auf das Lehren und

Lernen bezieht und mit dem Satz „Ich glaube, dass ...“ beginnt.¹⁹ In der vorliegenden Arbeit werden pädagogische Überzeugungen etwas enger gefasst und als Überzeugungen von Lehrkräften über das Lehren und Lernen im Allgemeinen definiert. Damit wird eine Perspektive eingenommen, die unabhängig vom fachdidaktischen Kontext und für alle Unterrichtsfächer relevant ist. Solche allgemeinen pädagogischen Überzeugungen lassen sich daher über die Fächergrenzen hinweg vergleichen: Im Rahmen dieser Arbeit werden allgemeine pädagogische Überzeugungen von Englisch- und Mathematiklehrkräften betrachtet. Außerdem können solche allgemeinen pädagogischen Überzeugungen bei Lehrkräften mit unterschiedlicher Lehrerfahrung erfasst werden, sei es eine Lehrkraft kurz nach dem Referendariat oder mit langjähriger Berufserfahrung, allgemeine pädagogische Überzeugungen sind bei ihnen allen zu finden. Fives et al. (2015, S. 249–250) führen aus, dass bei angehenden Lehrkräften die Überzeugungen zu Lehren und Lernen noch nicht so ausdifferenziert sind wie bei erfahrenen Lehrkräften. Allgemeine pädagogische Überzeugungen sind jedoch so umfassend formuliert, dass diese auch bei angehenden Lehrkräften erhoben werden können (siehe auch Richardson, 1996).

Pädagogische Überzeugungen streifen aufgrund ihrer Normativität auch das „propositional knowledge“ (Aussagenwissen), das nach Shulman (1986, S. 10–11) eine der drei Formen darstellt, die das Wissen von Lehrkräften repräsentieren (neben Fallwissen und strategischem Wissen). Eine Art von Aussagenwissen kann das normative Wissen sein, das Normen, Werte und ideologische sowie philosophische Verpflichtungen zu Gerechtigkeit und Fairness beinhaltet. Dieses Wissen ist weder theoretisch noch praktisch, sondern normativ und steht Shulman (1986) zufolge im Mittelpunkt dessen, was wir mit „teacher knowledge“ meinen. Pädagogische Überzeugungen sind nicht das Gleiche wie dieses normative Wissen (zum Unterschied von Überzeugungen und Wissen siehe Abschnitt 2.4.2), aber in den nun folgenden Ausführungen zu pädagogischen Überzeugungen, insbesondere solchen über Lernende und ihre Förderung, wird deutlich werden, dass diese auch, ähnlich wie der hier vorgestellte Wissensbestandteil bei Shulman (1986), einen normativen Bezug haben. Pädagogische Überzeugungen von Lehrkräften sind ein weites Feld: Muijs und Reynolds (2002) beschreiben das Universum der Überzeugungen von Lehrkräften als größer als das der Verhaltensweisen in einer Klasse – weswegen jede Studie sich darauf beschränken muss, nur ein begrenztes Überzeugungssystem

19 „[...] we define pedagogical beliefs as any proposition, specifically related to teaching and learning, which begins with the phrase ‘I believe that ...’“.

zu betrachten. Im Folgenden wird der Fokus auf bestimmte Überzeugungen von Lehrkräften über die Lernenden und ihre Förderung sowie die Reformbereitschaft von Lehrkräften gelegt.

Überzeugungen dahingehend, wie Schülerinnen und Schüler unterrichtet und wie ihnen dabei begegnet werden sollte, sind in empirischen Untersuchungen am ehesten bei den schülerzentrierten Überzeugungen von konstruktivistischen Lerntheorien zu finden (Fives et al., 2015; siehe auch Kapitel 3.1). Mit einem speziellen Fokus auf bestimmte Gruppen von Schülerinnen und Schülern werden häufig die Überzeugungen von Lehrkräften hinsichtlich Inklusion (Kielblock, 2018; Kiely, Brownell, Lauterbach & Benedict, 2015) oder sprachlicher und kultureller Heterogenität in der Unterrichtsgruppe untersucht (Fischer & Ehmke, 2019; Gay, 2015; Woolfolk Hoy et al., 2006). Beim Thema Inklusion werden beispielsweise die Einstellungen von Lehrkräften zur Inklusion und ihre Selbstwirksamkeitsüberzeugungen in Bezug auf die Inklusion betrachtet (Knigge & Rotter, 2015). Die Überzeugung der Lehrkräfte zur expliziten Förderung schwacher Schülerinnen und Schüler im Unterricht steht kaum im Zentrum von empirischen Arbeiten zu den Überzeugungen von Lehrkräften.

Die Studie von Davis (2006) hat gezeigt, dass einige Lehrkräfte Verantwortung für das Lernen von leistungsschwachen Schülerinnen und Schülern übernahmen und geeignete Lernangebote für sie planten. Andere Lehrkräfte waren der Überzeugung, es sei nicht ihre Aufgabe, angemessene Unterrichtsprogramme für Kinder mit Lernschwierigkeiten zu entwerfen, sie verwiesen dafür auf die Zuständigkeit externer Einrichtungen (siehe auch Rubie-Davies, 2015, S. 268). Die Unterrichtsentscheidungen hingen also davon ab, ob die Lehrkräfte Verantwortung für die Unterstützung oder Verbesserung leistungsschwächerer Lernender übernahmen. Zur Überzeugung von Lehrkräften hinsichtlich der Förderung schwacher Schülerinnen und Schüler gehört auch der Aspekt der generellen Förderbarkeit von Fähigkeiten. Überzeugungen von Lehrkräften hinsichtlich der Stabilität oder Veränderbarkeit der Fähigkeiten der Schülerinnen und Schüler haben einen Einfluss auf die Lernergebnisse der Schülerinnen und Schüler und können diese negativ beeinflussen (Turner et al., 2009, S. 365), wenn die Lehrkräfte Schülerinnen und Schüler nachteilig stereotypisieren.

Auch eine andere allgemeine pädagogische Überzeugung von Lehrkräften steht selten im Fokus empirischer Untersuchungen, nämlich, ob und wie sehr die

Lehrkräfte Druck beim Lernen der Schülerinnen und Schüler für notwendig halten. Woolfolk Hoy et al. (2006, S. 723–725) geben beispielsweise einen Überblick über Studien, die sich mit Überzeugungen hinsichtlich der Merkmale von Schülerinnen und Schülern beschäftigen. Diese Untersuchungen decken jedoch eher den Bereich der impliziten Theorien über die intellektuellen Fähigkeiten, die einzelnen Schülerinnen oder Schülern zugeschrieben werden, und über problematisches Verhalten ab. Weitere Studien beschäftigen sich mit den Überzeugungen von Lehrkräften hinsichtlich Gender und der sexuellen Orientierung der Lernenden. Die Frage, welche Überzeugungen Lehrkräfte hinsichtlich der Notwendigkeit von Druck bei Lernprozessen aufweisen, ist bedeutsam, da dies auf die Rolle der Motivation von Schülerinnen und Schülern beim Lernen verweist. Es konnte belegt werden, dass selbstbestimmt motiviertes Lernen mit positiven Konsequenzen für kognitive und emotionale Prozesse einhergeht (Rakoczy, 2006). Lernende haben ein Bedürfnis nach Autonomie. Nach der Selbstbestimmungstheorie von Deci und Ryan (1993) wird darunter das Bestreben verstanden, sich selbst als Verursachende der eigenen Handlungen zu erleben und aus eigenem Interesse heraus zu handeln. Solch eine Handlung kann zwar durchaus von außen, also von der Lehrkraft, initiiert werden. Wenn aber die Lehrkräfte der Überzeugung sind, dass sich ohne Zwang und Druck die Schülerinnen und Schüler nicht zum Lernen motivieren lassen, entspricht das weder den Prinzipien der Autonomie noch denen der Freiwilligkeit (externe Regulation, siehe Deci & Ryan, 1993, S. 227). Wie sehr die Lernmotivation einer Schülerin oder eines Schülers überhaupt von der Lehrkraft beeinflusst werden kann, zeigt auch die allgemeine Wirksamkeitsüberzeugung von Lehrkräften auf (Kunter & Pohlmann, 2015, S. 268–269). Eine allgemeine (im Gegensatz zur persönlichen) Wirksamkeitsüberzeugung beinhaltet die Einschätzung der Lehrkraft darüber, wie gut es allgemein (ohne Bezug auf die Fähigkeit der Lehrkraft selbst) gelingen kann, das Lernen und Verhalten – auch von schwächeren – Schülerinnen und Schülern zu fördern.

Im Unterschied zu den fach- und unterrichtsbezogenen Überzeugungen richten sich allgemeinere Überzeugungen von Lehrkräften auch häufig auf die Reform- bzw. Innovationsbereitschaft von Schulen sowie den Umgang von Schulen und Lehrkräften mit Schulreformen (Böse, Neumann, Becker, Maaz & Baumert, 2013; Feldhoff, 2011; Holtappels, 1997; Pant, Vock, Pöhlmann & Köller, 2008; Woolfolk Hoy et al., 2006). Mit dem Modell der „stages of concern“ wird seit vielen Jahren, ausgehend von Hall und Hord (1987), untersucht, wie Lehrkräfte auf Veränderungen und Reformbemühungen im Schulkontext reagieren. Dabei

gibt es sieben Entwicklungsstufen, die in unterschiedlichen Ausprägungen darstellen, wie die Lehrkräfte mit der Implementation von Innovationen umgehen. Die allgemeine Reformbereitschaft von Lehrkräften im Sinne einer grundsätzlichen Anpassungsbereitschaft an Veränderungen auf institutioneller Ebene wird in keiner der hier aufgeführten Studien mit quantitativen Methoden untersucht. Eine Studie, welche die generelle Einstellung zu Reformen mit qualitativen Methoden im Längsschnitt untersucht hat, ist die von Blossing und Ekholm (2005). Dort wurden Lehrkräfte in Interviews unter anderem dazu befragt, wie offen oder ablehnend sie Reforminitiativen bzw. schulischen Veränderungen gegenüber eingestellt sind. Sie finden an den Schulen unterschiedliche Verbesserungskulturen („improvement cultures“) vor: An Schulen, die bei der Implementation von Reformen systematisch vorgehen, haben die Lehrkräfte zumeist eine offene Einstellung gegenüber den Forderungen nach Veränderung – zumindest solange, wie die Forderungen mit der vorherrschenden Lehrerkultur übereinstimmen. An Schulen, die mit einer passiven Strategie auf die Implementation von Reformen reagierten, standen die Lehrkräfte den staatlichen Vorschlägen zur Umstrukturierung der Schulen verschlossen gegenüber.

Im folgenden Kapitel 4.1 wird auf die Funktion von Überzeugungen für das Handeln von Lehrkräften im Unterricht eingegangen und es wird ein besonderer Fokus auf die Funktionsweise pädagogischer Überzeugungen gelegt.

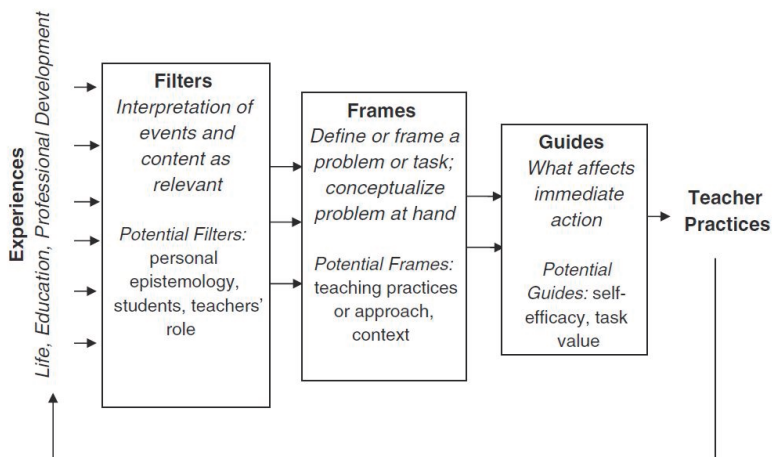
4 Überzeugungen und Handeln von Lehrkräften im Unterricht

In diesem Kapitel wird nun erklärt, wie sich Überzeugungen auf das Handeln von Lehrkräften im Unterricht auswirken können (Kapitel 4.1), wie diese Wirkbeziehungen aussehen können und welche Gründe es haben kann, wenn keine Zusammenhänge zwischen Überzeugungen und Unterrichtshandeln gefunden werden (Kapitel 4.2). Danach wird auf das Handeln von Lehrkräften im Unterricht eingegangen, das in der vorliegenden Arbeit als Vermittler von Überzeugungen der Lehrkräfte auf die Lernergebnisse der Schülerinnen und Schüler untersucht wird (Kapitel 4.3).

4.1 Funktionen von Überzeugungen

Um Überzeugungen zu beschreiben, bedarf es neben der Verortung von Überzeugungen hinsichtlich verwandter Begriffe (z.B. Wissen) und der Beschreibung der Charakteristika von Überzeugungen (siehe Kapitel 2.4) auch der Beschreibung der Funktionen von Überzeugungen (Buehl & Beck, 2015; Fives & Buehl, 2012). Pajares (1992, S. 307) betrachtet die Überzeugungen von Lehrkräften als Indikatoren für deren Entscheidungen, ja sogar als „best indicators of the decisions individuals make throughout their lives“. Im Folgenden wird nun aufgezeigt, wie sich diese Entscheidungen durch Überzeugungen beeinflussen lassen.

Für diese Arbeit wird eine Systematisierung der Funktionen von Überzeugungen gewählt, die auf ihrer Nähe zur Handlung fokussiert (Fives & Buehl, 2012; Läge & McCombie, 2015), da sich so die spezifischen Mechanismen und Prozesse der Wirkung von Überzeugungen erklären lassen. Nach Fives und Buehl (2012, S. 478–480) lassen sich drei verschiedene Funktionen von Überzeugungen in Bezug auf Handlungen identifizieren (siehe Abbildung 4.1): als (1) Filter für die Interpretation, (2) als Rahmen für die Definition von Problemen und (3) als handlungslenkend.



Quelle: Fives & Buehl, 2012, S. 478

Abbildung 4.1: Die drei Funktionen von Überzeugungen

(1) Überzeugungen wirken als *Filter*, welche die Wahrnehmung und die Interpretation von Informationen und Erfahrungen beeinflussen (Pajares, 1992; Seifried, 2009, S. 39). Diese Wirkung von Überzeugungen formt das nachfolgende Denken und die Informationsverarbeitung neu, kann sie aber auch verzerren (Pajares, 1992, S. 325): „[...] the filtering effect of belief structures ultimately screens, redefines, distorts, or reshapes subsequent thinking and information processing.“ Beispielsweise können sich Überzeugungen über angemessene Unterrichtspraktiken darauf auswirken, worauf Lehrkräfte bei der kognitiven Verarbeitung und Evaluierung von Informationen achten und wie sie diese Informationen interpretieren. Sembill und Seifried (2009, S. 346) formulieren es so: „Sichtweisen steuern mentale Vorgänge, wirken wie ein Filter, sind bei der Informationsaufnahme und beim Wissenserwerb aktiv und damit sui generis internal wie external handlungsleitend.“ Eine ähnliches Bild dieser Funktionsweise von Überzeugungen wählt Philipp (2007), der von Linsen spricht, durch die man bei der Interpretation der Welt blickt und die so die Interaktion von Menschen mit ihrer Welt beeinflussen.

(2) Überzeugungen wirken als *Rahmen* („frame“) für Probleme und Aufgaben. Nespor (1987, S. 323) schreibt dazu: „Task environments are defined in terms of belief systems.“ Je nach Überzeugung über das Lehren und Lernen, wie in konstruktivistischen oder transmissiven Lehrmodellen, werden Lehrkräfte ein

aktuelles Problem unterschiedlich rahmen und einordnen (Nespor, 1987) sowie den Unterricht unterschiedlich planen und reflektieren.

(3) Wenn ein Problem definiert ist, können Überzeugungen die Planung und die Handlungen von Lehrkräften *lenken* („guide“). Hierzu führen Fives und Buehl (2012) motivationale Konstrukte an, zum Beispiel Selbstwirksamkeitsüberzeugungen, welche die Ziele von Lehrkräften, ihren Aufwand zum Erreichen dieser Ziele, ihre Beharrlichkeit und ihr Engagement lenken bzw. beeinflussen (Bandura, 1997). Dieses Verhalten (also Entscheidungen, Anstrengung und Ausdauer) beeinflusst schließlich die Unterrichtsqualität („quality of teachers’ practices“ Fives & Buehl, 2012). Andere motivationale Konstrukte, die dem Unterrichtshandeln vorausgehen, sind beispielsweise die erwarteten Konsequenzen einer Handlung („outcome expectancy“) und der subjektive Wert einer Aufgabe („task value“). Solche Überzeugungen können die Handlung von Lehrkräften unmittelbar steuern und sind somit vergleichsweise verhaltensnah.

Im Folgenden werden die beiden Funktionen Filter und Rahmen auf die drei Überzeugungen aus der Drei-Länder-Studie (siehe Kapitel 7.2.2) angewendet, da solche allgemeinen Überzeugungen die Wahrnehmung und Situationseinstellung von Lehrkräften beeinflussen. Zuerst beleuchten wir die Überzeugungen hinsichtlich des allgemeinen Verständnisses von Unterricht und des gelingenden Lernens. Wenn Lehrkräfte eine positive Einstellung zur Förderung schwacher Schülerinnen und Schüler haben, kann dies dazu führen, dass die individuellen Stärken und Schwächen der einzelnen Schülerinnen und Schüler eher wahrgenommen werden (Filter), da die Lehrkraft jede und jeden einzelnen mit ihrem Unterricht erreichen möchte. Eine Lehrkraft hingegen, die von der Notwendigkeit von Druck im Unterricht überzeugt ist und Zwang, Kontrolle und Durchgreifen bei Schülerinnen und Schülern präferiert, nimmt die Schülerschaft eher als Ganzes und weniger die Unterschiede innerhalb einer Klasse wahr.

Wenn leistungsschwächere Schülerinnen und Schüler im Unterricht bei einer bestimmten Aufgabe nicht gut mitkommen, wird eine Lehrkraft mit positiver Einstellung zur Förderung schwacher Schülerinnen und Schüler die momentane Situation anders einordnen (Rahmen) als eine Lehrkraft mit negativer Einstellung zur Förderung. Das Eingehen auf individuelle Schwächen einzelner Schülerinnen und Schüler erscheint der ersten Lehrkraft als selbstverständlich, die andere Lehrkraft wird diese Situation eher als störend und als den Unterricht

behindernd empfinden. Ebenso wird eine Lehrkraft, die bei der Einstellung zur Notwendigkeit von Druck bei Lernprozessen hohe Ausprägungen aufweist, sich eher an der Durchschnittsschülerin bzw. am Durchschnittsschüler orientieren und deshalb leistungsschwachen Schülerinnen und Schülern, die vom durchschnittlichen Lerntempo der Klasse abweichen, weniger Aufmerksamkeit schenken (Standop, 2017, S. 268; Trautmann & Wischer, 2008).

Auch Überzeugungen zur Schule, zum Beispiel die Reformbereitschaft der Lehrkräfte, können als Filter und Rahmen wirken. Bei der Reformbereitschaft von Lehrkräften liegt eine Anpassungsbereitschaft an Veränderungen auf institutioneller Ebene vor. Dies können zum Beispiel unterschiedliche Ausgangslagen in den Klassen und der Umgang mit einer heterogenen Schülerschaft sein. Es kann angenommen werden, dass reformbereite Lehrkräfte einen schülerorientierten Unterricht bevorzugen und auf individuelle Unterschiede der Schülerinnen und Schüler eingehen (Trautmann & Wischer, 2008).

Gerade im Zusammenhang mit der Funktion von Überzeugungen als Filter wird oft auf die Stabilität und Beständigkeit bestimmter Überzeugungen von Lehrkräften hingewiesen. Zur Stabilität von Überzeugungen gibt es verschiedene Ansichten; häufig vertreten wird die Ansicht, dass diese robust und veränderungsresistent (Kane, Sandretto & Heath, 2002) bzw. zwar nicht unveränderlich, aber sehr stabil sind (Pajares, 1992). Am plausibelsten erscheint die Annahme, dass es je nach Art einer Überzeugung verschiedene Grade der Stabilität gibt. Besonders tiefsitzende Überzeugungen, die durch die Lebensgeschichte der Lehrkräfte geformt wurden, können laut Reusser et al. (2011, S. 481) in ihrer Filterfunktion auch als „Barrieren“ wirken, die sich nur gegen Widerstände, Druck und Krisen verändern lassen“. Überzeugungen lassen sich umso schwerer verändern, je zentraler ihre Stellung im persönlichen Überzeugungssystem ist (Pajares, 1992; siehe dazu Abschnitt 2.4.1). Prozesse der Veränderung von Überzeugungen werden unter anderem in der Conceptual-Change Forschung untersucht (siehe Pajares, 1992; Woolfolk Hoy et al., 2006) oder bei der Untersuchung subjektiver Theorien betrachtet (siehe Kapitel 3.2). Kane et al. (2002, S. 180) fassen die Erkenntnisse von Pajares (1992), Richardson (1996), Kagan (1992) und Nespor (1987) zur Veränderbarkeit der Überzeugungen von Lehrkräften wie folgt zusammen:

„Students enter teacher education programs with preexisting beliefs based on their experience as students in schools [...].

These beliefs are robust and resistant to change [...].

The beliefs act as filters allowing in or filtering out new knowledge that is deemed compatible or incompatible with current beliefs [...].”

Lehrkräfte besitzen also bereits zu Beginn ihrer Ausbildung bestimmte pädagogische Überzeugungen, die auf ihren eigenen Erfahrungen als Schülerinnen und Schüler basieren. Diese Überzeugungen sind robust gegenüber Veränderungen und fungieren als Filter bei der Aufnahme neuer Informationen.

In diesem Kapitel wurden die verschiedenen theoretisch hergeleiteten Funktionen von Überzeugungen dargestellt. Nun erhebt sich die Frage, wie Überzeugungen konkret auf das Handeln der Lehrkräfte im Unterricht wirken können. Dies ist Gegenstand des nun folgenden Kapitels 4.2.

4.2 Wirkbeziehungen von Überzeugungen auf das Handeln von Lehrkräften

Es gibt viele Studien, die zeigen, dass unterrichtsbezogene Überzeugungen von Lehrkräften mit ihrem Unterrichtshandeln und letztlich auch mit den Lernergebnissen der Schülerinnen und Schüler zusammenhängen (Buehl & Beck, 2015; Fives & Buehl, 2012; Pajares, 1992; Woolfolk Hoy et al., 2006). Eine Auswahl dieser Ergebnisse zu lerntheoretischen Überzeugungen, subjektiven Theorien und pädagogischen Überzeugungen wurde im Kapitel 3 geboten. Im vorigen Kapitel 4.1 konnten drei Funktionen von Überzeugungen in Bezug auf Handlungen identifizieren werden. Nun werden zuerst die Wirkbeziehungen von Überzeugungen auf das Handeln von Lehrkräften näher betrachtet (Abschnitt 4.2.1), dann wird auf mögliche Ursachen für eine fehlende Übereinstimmung von Überzeugungen und Handeln eingegangen (Abschnitt 4.2.2). Zuvor werden noch einige generelle Aspekte der Wirkung von Überzeugungen auf das Handeln dargestellt.

Oser und Blömeke (2012, S. 415) erklären in ihrer Einführung in den Themen teil zu „Überzeugungen“ in der Zeitschrift für Pädagogik den Begriff der Überzeugungen, indem sie diese als „nicht-wissenschaftliche Vorstellungen darüber, wie etwas beschaffen ist oder wie etwas funktioniert, mit dem Anspruch der Geltung für das Handeln“ bezeichnen. Durch den Zusatz „Anspruch der Geltung für das Handeln“ werden die Überzeugungen per definitionem als hand-

lungsrelevant bezeichnet, was aus theoretischer Sicht aber nicht immer der Fall ist (siehe Abschnitt 4.2.2 zu den möglichen Ursachen).

Überzeugungen von Lehrkräften sind auch deshalb für ihr Handeln im Unterricht von Bedeutung, weil dieses Handeln davon bestimmt ist, dass es vielen Anforderungen gleichzeitig genügt und deshalb häufig ein Handeln unter Druck darstellt. In solchen Situationen greifen Menschen eher auf Überzeugungen als auf Wissen zurück, um die komplexe Situation zu vereinfachen und Entscheidungen treffen zu können (Pajares, 1992).²⁰ Wegen dieser Wirkung werden Überzeugungen häufig mit Filtern (Pajares, 1992), Brillen (Reusser et al., 2011) oder Linsen (Kane et al., 2002) verglichen, die zwischen der Wahrnehmung von Ereignissen und der Reaktion der Lehrkräfte stehen (siehe Kapitel 4.1). Bei Fishbein und Ajzen (1975, S. 15) wird die Wirkbeziehung zwischen Überzeugungen und Handeln mittels Einstellungen dargestellt (siehe dazu auch Abschnitt 2.4.4). Die Überzeugungen einer Person haben bestimmten Einstellungen zu Folge, die wiederum zu Handlungsabsichten und somit zu möglichen Handlungen führen.

Die Überzeugungen von Lehrkräften gelten somit als relevant für die Qualität ihres professionellen Handelns (Reusser et al., 2011): Ihre Überzeugungen haben Auswirkungen auf die Auswahl von Zielen und Handlungsplänen einer Lehrkraft sowie auf ihre Wahrnehmung und Deutung von Situationen im Unterricht. Sie steuern ihr didaktisches und kommunikatives Handeln und Problemlösen (Calderhead, 1996; Goldin et al., 2009; Pajares, 1992; Philipp, 2007; Thompson, 1992; Woolfolk Hoy et al., 2006).

4.2.1 Vier Sichtweisen der Beziehung zwischen den Überzeugungen und dem Handeln von Lehrkräften

Auch wenn Überzeugungen häufig als eine Grundlage von Handlungen angesehen werden (z.B. bei Pajares, 1992), werden im Literatur-Review von Buehl

20 Dass Überzeugungen zur Entscheidungsfindung in komplexen Situationen zu Tragen kommen, findet sich bereits vor über 100 Jahren in pädagogischer Fachliteratur. Scholz (1909, S. 102) schreibt im Encyklopädischen Handbuch der Pädagogik (Hrsg. Wilhelm Rein) zu dem von Herbart geprägten Stichwort „pädagogischer Takt“: „Pädagogischer Takt ist „(...) die auf *pädagogischen Überzeugungen* und Erfahrungen beruhende Grundstimmung des Gemütes (...), welche den Erzieher befähigt, auf dem Gebiete der Erziehung und des Unterrichtes in jedem Einzelfall die zweckmäßigste Entscheidung rasch und sicher zu treffen“ [Hervorhebung der Verfasserin].

und Beck (2015, S. 68–73) verschiedene Perspektiven aufgezeigt, wie Überzeugungen und Handeln zueinander in Beziehung stehen: (1) Überzeugungen beeinflussen das Handeln, (2) das Handeln beeinflusst Überzeugungen, (3) die Überzeugungen von Lehrkräften sind von ihrem Handeln abgekoppelt und (4) es gibt eine reziproke, aber komplexe Beziehung zwischen den Überzeugungen und dem Handeln von Lehrkräften.

(1) Wenn Überzeugungen von Lehrkräften mit ihren Handlungen korreliert sind bzw. eine Passung vorliegt oder wenn sich die Überzeugungen im Handeln widerspiegeln, wird daraus häufig gefolgert, dass Überzeugungen das Handeln beeinflussen (Buehl & Beck, 2015, S. 68). Alternativ kann man Überzeugungen als von Handlungen geformt ansehen (2). Dieser Zusammenhang wird vor allem von solchen Studien gestützt, welche die Effekte der Ausbildung und der beruflichen Weiterbildung von (angehenden) Lehrkräften auf Überzeugungen betrachten (Taibi, 2012). Hierbei werden zumeist die Selbstwirksamkeit von Lehrkräften oder andere motivationale Konstrukte untersucht (Rothland, 2012). Um diese zu stärken, benötigt man Handlungserfahrungen, die Erfolgsergebnisse ermöglichen (Bandura, 1997). Durch die Ausübung bestimmter Unterrichtspraktiken können sich auch weitere Überzeugungen ändern, beispielsweise Überzeugungen zur Inklusion oder zur Klassenführung (Buehl & Beck, 2015, S. 69). Es gibt allerdings auch Studien, die darauf hindeuten, dass die Überzeugungen und das Handeln der Lehrkräfte nicht miteinander zusammenhängen (3) (Calderhead, 1996; siehe hierzu den folgenden Abschnitt 4.2.2). Zumeist handelt es sich um Studien zur Umsetzung von epistemologischen Überzeugungen im Unterricht (z.B. konstruktivistisch gegenüber transmissiv; Buehl & Beck, 2015, S. 70; siehe Kapitel 3.1). Auch bei Überzeugungen bezüglich pädagogischer Praktiken wie Gruppenarbeit konnten Abweichungen bezüglich des Handelns der Lehrkräfte festgestellt werden (Jorgensen, Grootenboer, Niesche & Lerman, 2010). Allerdings wurden in der Studie von Jorgensen et al. (2010) auch zwei Themen (Lernumgebung und intellektueller Anspruch) identifiziert, bei denen die Überzeugungen der Lehrkräfte mit ihren Handlungen kongruent waren. Schon an diesem Beispiel wird ersichtlich, dass eine „Entweder/oder“-Beziehung von Überzeugungen zu Handlungen der Komplexität der menschlichen Handlungen und des Unterrichtsgeschehens nicht gerecht werden kann und manche Überzeugungen nicht direkt mit dem Handeln zusammenhängen. Stattdessen ist vielmehr davon auszugehen, dass (4) diese Beziehung komplexer Natur ist und bei vielen Arten von Überzeugungen reziprok verläuft (Buehl & Beck, 2015; Richardson, 1996; Thompson, 1992). Da bei den Studi-

en, die eine Beziehung von Überzeugungen und Handlungen der Lehrkräfte belegen, nie eine vollständige Übereinstimmung von Überzeugungen und Handlungen vorlag und auch die Längsschnittstudien, die auf eine reziproke Verbindung zwischen Überzeugungen und Handlungen hindeuten (Buehl & Beck, 2015, S. 70), gezeigt haben, dass die Überzeugungen jeweils in unterschiedlichem Ausmaß in Handlungen übersetzt wurden, wird das Bild einer komplexen Beziehung gestärkt. Fives und Buehl (2012, S. 481) fassen es so zusammen: „It is not a matter of whether belief and practices are or are not congruent but rather the degree of congruence or incongruence between beliefs and practices.“

4.2.2 Mögliche Ursachen für die fehlende Übereinstimmung von Überzeugungen und unterrichtlichem Handeln

In dieser Arbeit wurden bislang Studien aufgeführt, die auf Zusammenhänge zwischen den Überzeugungen von Lehrkräften und ihrem Unterrichtshandeln schließen lassen (siehe Kapitel 3). Allerdings weisen Überblicksartikel zu Überzeugungen, welche deren Auswirkungen zusammenfassend betrachten, auch immer darauf hin, dass in manchen Studien keine solchen Zusammenhänge zwischen den Überzeugungen und dem Handeln von Lehrkräften im Unterricht nachgewiesen werden konnten (Buehl & Beck, 2015; Fives & Buehl, 2012; Pajares, 1992; Woolfolk Hoy et al., 2006).

Die Faktoren, die dazu führen, dass die Handlungen von Lehrkräften nicht immer mit ihren Überzeugungen kongruent sind, lassen sich in interne und externe Faktoren einteilen (Buehl & Beck, 2015). Die Eigenschaft von Überzeugungen, in Bündeln aufzutreten, kann auch dazu beitragen, mögliche Inkonsistenzen innerhalb von Überzeugungssystemen zu erklären (Furinghetti & Pehkonen, 2002, S. 45). Aufgrund dieser Struktur ist es möglich, dass innerhalb eines Überzeugungssystems widersprüchliche Überzeugungen gleichzeitig bestehen (siehe Abschnitt 2.4.1). Solche anderen Überzeugungen, die möglicherweise auch im Widerspruch zueinander stehen, zählen zu den internen Faktoren. Weitere interne Faktoren für die fehlende Übereinstimmung von Überzeugungen und Handlungen sind fehlendes Fach- oder pädagogisches Wissen sowie fehlende Erfahrung (Buehl & Beck, 2015). Erfahrenere Lehrkräfte weisen eher eine Übereinstimmung ihrer Überzeugungen mit ihrem Handeln auf (Basturkmen, 2012), während sich bei Lehrkräften in Übergangssituationen hierbei eher Abweichungen zeigen. Solche Übergangssituationen können das

Lehramtsstudium oder das Referendariat sein oder auch eine berufliche Weiterentwicklung, die sich noch nicht unmittelbar im Unterrichtsgeschehen bemerkbar macht, da es noch gilt, bestehende Routinen bei den bisherigen Unterrichtspraktiken zu verändern. Aus dieser Sicht kann die Nicht-Übereinstimmung von Überzeugungen und Handlungen von Lehrkräften ein natürlicher Bestandteil der beruflichen Weiterentwicklung sein. Ein Mangel an Selbstbewusstsein oder Selbstreflexion und die emotionale Beteiligung der Lehrkraft an der Handlungssituation stellen weitere interne Faktoren dar, die zu einer fehlenden Übereinstimmung von Überzeugungen und Handlungen beitragen können. Externe Faktoren, die mit einer Inkongruenz von Überzeugungen und Handlungen bei Lehrkräften zusammenhängen, sind der Klassenebene (z.B. Leistungsniveau, Einstellungen der Schülerinnen und Schüler oder Klassengröße), der Schulebene (Unterstützung durch Schulleitung, Kollegium oder Eltern, Ressourcen der Schule) oder höher angesiedelten Verwaltungsebenen (bildungspolitische Entscheidungen, Curriculum) zuzuordnen (Buehl & Beck, 2015).

Weitere Gründe für die Inkongruenz von Überzeugungen und Handlungen können (1) in der Art einer Überzeugung (Pajares, 1992), (2) ihrer Stellung innerhalb des Überzeugungssystems (Green, 1971; siehe Abschnitt 2.4.1) und (3) der Funktion der Überzeugung (Fives & Buehl, 2012; siehe Kapitel 4.1) liegen.

Um (1) die Art einer Überzeugung von der anderer zu unterscheiden, eignet sich die „Theories of Action“ von Argyris und Schön (1974) als Referenzrahmen. Danach entwickeln Menschen Modelle und dazugehörige Theorien über ihre Umwelt, um Komplexität zu reduzieren. Wenn man Personen aber nach diesen Theorien fragt, entsprechen die expliziten Äußerungen zu solchen Handlungsmaximen („espoused theories of action“) nicht unbedingt den impliziten Theorien („theories in use“), welche Ähnlichkeiten mit den subjektiven Theorien kürzerer Reichweite aufweisen (siehe Kapitel 3.2). Empirische Ergebnisse, die darauf hindeuten, dass sich die Überzeugungen von Personen nicht wie erwartet auf ihre Handlungen auswirken (Kane et al., 2002), könnten durch diese Abweichung von impliziten und expliziten Überzeugungen erklärt werden. Es gibt also einen Unterschied zwischen bewussten und handlungsleitenden Überzeugungen (Trautwein, 2013, S. 6). Eine weitere mögliche Ursache für fehlende Zusammenhänge der Überzeugungen mit dem Handeln von Lehrkräften könnte auch sein, dass die in einer Unterrichtssituation relevanten

Überzeugungen nicht vollständig erfasst werden können, wenn man einzelne Überzeugungen isoliert betrachtet und sie nicht mit anderen Überzeugungen verknüpfen kann, die für den Unterricht relevant sind. Seifried (2009, S. 119) weist darauf hin, dass dies besonders dann der Fall ist, wenn ausschließlich fachunabhängige Überzeugungen erfasst werden und die fachdidaktischen Aspekte von Überzeugungen nicht thematisiert werden. Als ein anderes Beispiel für eine mögliche Ursache fehlender Zusammenhänge zwischen Überzeugungen und Handlungen führen Fives et al. (2015, S. 261) an, dass bei einer Diskrepanz von konstruktivistischen Überzeugungen und dem Unterrichtshandeln von Lehrkräften der Grund häufig darin liegen kann, dass der Fokus zu sehr auf den Unterrichtsinhalt, also den konstruktivistischen Überzeugungen der Lehrenden, und nicht auf die konstruktivistischen Überzeugungen des Lernens, also die Lernenden gelegt wurde (auch Philipp, 2007).

(2) Zentrale Überzeugungen innerhalb eines Überzeugungssystems sollten stärker mit Handlungen verbunden sein als periphere Überzeugungen, da die zentralen stärker vertreten werden (siehe Abschnitt 2.4.1; Philipp, 2007). Phipps und Borg (2009) stellten in einer qualitativen Untersuchung fest, dass sich die zentralen Überzeugungen von Fremdsprachenlehrkräften über das Lernen von Schülerinnen und Schülern eher im Unterrichtshandeln widerspiegeln als Überzeugungen zum Erlernen von Fremdsprachen. Da sich Überzeugungen innerhalb eines Überzeugungssystems auch widersprechen können, hängt es vom Kontext bzw. der Situation ab, ob eine Überzeugung mit der Handlung der Lehrkraft kongruent ist. Situative Faktoren wie Unterrichtsstörungen können dazu führen, dass Lehrkräfte entgegen ihrer Überzeugung handeln, da Überzeugungen zwar handlungsleitend, aber nicht zwingend handlungswirksam sind (Seifried, 2009, S. 120).

(3) Überzeugungen können auch unterschiedliche Rollen bezüglich bestimmter Handlungen spielen, je nach der Funktion, welche die Überzeugung hat (Filter, Rahmen oder Lenkung; siehe Kapitel 4.1). Überzeugungen, die in der Zuordnung näher bei konkreten Handlungen liegen, lassen sich im Unterrichtshandeln von Lehrkräften leichter identifizieren (Buehl & Beck, 2015, S. 72). Daher können Überzeugungen, die eine bestimmte Aufgabe rahmen oder sofortiges Handeln lenken, leichter beobachtet werden als solche Überzeugungen, welche die Funktion des Filterns und Interpretierens von neuen Informationen übernehmen. Wenn eine Lehrkraft die Überzeugung vertritt, dass man schwache Schülerinnen und Schüler im Unterricht besonders fördern sollte, wird sie die

Planung ihres Unterrichts dementsprechend ausrichten (Rahmen). Bei der Umsetzung im Unterricht kann es jedoch trotzdem der Fall sein, dass keine entsprechenden Unterrichtshandlungen zum Einsatz kommen. Die Funktion der Lenkung wird hier also nicht von dieser Überzeugung ausgeübt. Handlungen, die zu ihren Überzeugungen inkonsistent sind, können bei Lehrkräften auch dann auftreten, wenn deren Situationsbeurteilung mit negativen Emotionen verbunden ist (Seifried, 2009, S. 119–120). Die emotionale Beteiligung einer Lehrkraft an der Handlungssituation kann in diesem Fall den Zusammenhang zwischen Überzeugungen und Unterrichtshandeln beeinflussen. So kann die Befürchtung einer Lehrkraft, ein Handlungsziel nicht zu erreichen (z.B. der Abschluss einer Unterrichtsstunde) dazu führen, dass sie entgegen ihren Überzeugungen handelt.

4.3 Handeln von Lehrkräften im Unterricht

Die Betrachtung der Überzeugungen Lehrender gewinnt an praktischer Relevanz, wenn man sie zu ihrem Handeln in Beziehung setzt. Lehrkräfte haben mit ihrer professionellen Kompetenz und ihrem unterrichtlichen Handeln einen erheblichen Einfluss auf die Lernergebnisse und die Lernentwicklung der Schülerinnen und Schüler (Lipowsky, 2006). In Kapitel 2.1 wurde dargelegt, dass Überzeugungen ein wichtiger Bestandteil der Kompetenz von Lehrkräften sind, in Kapitel 3 wurden empirisch relevante Bereiche der Überzeugungen von Lehrkräften dargelegt und in den Kapiteln 4.1 und 4.2 wurde deutlich, wie Überzeugungen sich auf das Handeln von Lehrkräften auswirken. Die Überzeugungen von Lehrkräften zeigen sich also in den didaktischen Entscheidungen darüber, welches Unterrichtshandeln sie ausüben wollen. Didaktische Entscheidungen sind der Kern der pädagogischen Professionalität (Scheunpflug, 2005, S. 60), daran erst wird die praktische Relevanz von Überzeugungen sichtbar. Dubberke et al. (2008, S. 194) bestätigen dies: „Theoretisch wird davon ausgegangen, dass die Überzeugungen der Lehrkräfte sich erst im Unterrichtsgeschehen – in der Gestaltung der Lernumgebungen – manifestieren müssen, um eine relevante Rolle für die Leistungsentwicklung zu spielen“ (Calderhead, 1996; Richardson, 1996; Thompson, 1992).

Es ist sinnvoll, unterschiedliche Überzeugungen und verschiedene Dimensionen des Unterrichtsgeschehens zu untersuchen, wenn man der Verbindung von Überzeugungen und Unterrichtshandeln bzw. dessen Vermittlung auf die Lernergebnisse von Schülerinnen und Schülern auf den Grund gehen will. Die Klas-

senführung und ein unterstützendes Unterrichtsklima bilden zusammen mit der kognitiven Aktivierung die drei Grunddimensionen der Unterrichtsqualität (Klieme, 2019). Ein unterstützendes Unterrichtsklima beinhaltet eine positive Beziehung zwischen den Lehrkräften und ihren Schülerinnen und Schülern. Außerdem werden dabei die individuellen Bedürfnisse und Interessen der Schülerinnen und Schüler berücksichtigt. Zur Klassenführung gehören klare Regeln, um eine aktive Beteiligung und die Aufmerksamkeit der Schülerinnen und Schüler sicherzustellen, sowie das Erkennen und die Prävention von Störungen. Im Folgenden werden ein unterstützendes Unterrichtsklima und die Klassenführung als zwei Bestandteile der Unterrichtsqualität näher beleuchtet, die im Rahmen dieser Arbeit als Vermittler zwischen den Überzeugungen der Lehrkräfte und den schulischen Ergebnissen der Schülerinnen und Schüler betrachtet werden.

4.3.1 Unterstützendes Unterrichtsklima

Zu einem unterstützenden Unterrichtsklima gehört das Eingehen auf die individuellen Bedürfnisse der Schülerinnen und Schüler. In der aktuellen Diskussion der Unterrichtsforschung wird die übergreifende Frage, was guter Unterricht ist, häufig zugespitzt auf die Wirksamkeit von „individueller Förderung“ (Hasselhorn, Decristan & Klieme, 2019; Klieme & Warwas, 2011) und ihre differentiellen Wirkungen. Angesichts einer heterogenen Schülerschaft ist die Frage, wie Unterricht an die vorgegebenen interindividuellen Unterschiede angepasst werden kann, wieder aktuell (Kunter & Ewald, 2016). In Anlehnung an die bestehende Forschungsliteratur wird in dieser Arbeit der Fokus auf „Adaptivität“ als Prozessmerkmal des Unterrichts gelegt (Helmke & Weinert, 1997, S. 137; siehe auch Dumont, 2019; Hardy, Decristan & Klieme, 2019).

Unter Adaptivität versteht man die Anpassung der Lernumgebung an die diagnostizierten Lernvoraussetzungen der Schülerinnen und Schüler (Corno & Snow, 1986). Corno und Snow (1986) unterscheiden zwei Arten des adaptiven Unterrichts: Als „Makro-Adaptivität“ bezeichnen sie es, wenn eine Bildungseinrichtung bzw. eine Lehrkraft aus einem Repertoire von Lehr-Lern-Arrangements die für die jeweilige Klasse angemessene Variante auswählt. Die Adaptionentscheidungen werden also vor Beginn einer Lehrereinheit getroffen. Von „Mikro-Adaptivität“ wird gesprochen, wenn die Abstimmung auf der Prozessebene im Verlauf der Lehrer-Schüler-Interaktion hergestellt wird, die Anpassung also anhand der Aufgaben und des Lernmaterials an die Lernvo-

raussetzungen der Schülerinnen und Schüler vorgenommen wird, beispielsweise durch individuell abgestimmte Rückmeldungen und Fragen. In der Dreiländer-Studie wurden verschiedene Facetten der Adaptivität erfasst, die aus heutiger Sicht als eine Kombination von Mikro- und Makro-Adaptivität verstanden werden können (siehe Kapitel 7.2.2).

Mehrere Studien zeigen, dass sich adaptiver Unterricht in der Grundschule positiv auf die Entwicklung des konzeptionellen Verstehens der Schülerinnen und Schüler auswirkt (Decristan et al., 2015; Möller, Hardy, Jön, Kleickmann & Blumberg, 2006; Murphy & Messer, 2010). Es konnten auch differentielle Effekte von adaptivem Unterricht nachgewiesen werden: So profitieren besonders sozial benachteiligte Schülerinnen und Schüler oder solche mit geringen sprachlichen Fähigkeiten von einer adaptiven Unterstützung durch die Lehrkräfte (Decristan et al., 2015; Hamre & Pianta, 2005). Darüber hinaus haben Studien gezeigt, dass Adaptivität im Unterricht sich nicht nur auf die kognitiven Ergebnisse der Schülerinnen und Schüler auswirkt, sondern auch auf metakognitive und affektive Merkmale (Beck et al., 2008; Corno & Snow, 1986). Für ausführliche empirische Belege der Wirkungen von Adaptivität im Unterricht bieten Hardy, Decristan und Klieme (2019) sowie Klieme und Warwas (2011) eine Übersicht, in der auch deutlich wird, dass es in der Forschungsliteratur mehr Belege zur Wirkung von adaptivem Unterricht auf kognitive Merkmale als auf affektive Merkmale wie Wohlfühlen im Unterricht gibt.

4.3.2 Klassenführung

Nach Helmke (2015, S. 173) wird eine effiziente Klassenführung übereinstimmend als ein „Schlüsselmerkmal der Unterrichtsqualität“ angesehen. Auch Lipowsky (2006, S. 56) unterstreicht, dass in theoretischen Modellen der Unterrichtsforschung eine effektive Klassenführung und die dadurch ermöglichte intensive Lernzeitnutzung das Ausmaß an Gelegenheiten zur inhaltlichen Auseinandersetzung mit dem Unterrichtsgegenstand bestimmen und somit basale Voraussetzungen für den Lernerfolg von Schülerinnen und Schülern sind. Die international vorherrschende Auffassung der Klassenführung umfasst präventive, proaktive und reaktive Elemente der Klassenführung (Helmke, 2015, S. 173). Die Reaktion auf Störungen im Unterricht ist also nur ein Teil der Klassenführung, die Etablierung von Regeln im Klassenraum und das Zeitmanagement sind weitere Bestandteile. Indem eine gelungene Klassenführung durch die Lehrkraft dafür sorgt, dass ein geordneter Rahmen für die Lehr- und

Lernaktivitäten vorhanden ist und damit verbunden den Schülerinnen und Schülern genügend aktive Lernzeit zur Verfügung steht, ist sie eine wichtige Voraussetzung für einen anspruchsvollen Unterricht. Die internationale Forschung beispielsweise von PISA 2003 oder DESI (2008) und Metaanalysen wie bei Hattie (2009) zeigen, dass die Klassenführung als Unterrichtsmerkmal stark mit dem Leistungsniveau von Schülerinnen und Schülern zusammenhängt (Helmke, 2015, S. 174; auch Dubberke et al., 2008, S. 195). Es gibt auch Befunde, wonach eine gute Klassenführung die Motivation der Lernenden fördern kann (Rakoczy, 2006). Da über die Klassenführung vor allem die aktive Lernzeit gesteigert werden kann, aber nicht direkt die Qualität des Lernens der Schülerinnen und Schüler, ist eine effiziente Klassenführung für ein erfolgreiches Lernen notwendig, aber nicht hinreichend (Helmke, 2015, S. 176–177). Eine effiziente Nutzung der Lernzeit signalisiert zudem die Wichtigkeit und Wertigkeit, welche die Lehrkraft dem Lernen zuschreibt (Helmke, 2007). Nach einem Ausschnitt des allgemeineren Angebots-Nutzungs-Modell der Unterrichtswirksamkeit von Helmke (2007, S. 45) wirkt sich neben dem Klassenkontext und der Unterrichtsqualität auch die Lehrerpersönlichkeit und damit auch deren Bestandteil der pädagogischen Orientierungen über die Unterrichtsqualität vermittelt auf die Klassenführung aus.

Um genügend aktive Lernzeit zur Verfügung zu stellen, müssen Unterrichtsstörungen von der Lehrkraft möglichst schnell erkannt und abgebaut werden. Eine geordnete Lernumgebung im Unterricht führt also zu möglichst langer aktiver Lernzeit in der Unterrichtsstunde (OECD, 2009, S. 107). Auch wenn eine maximale Nutzung der Lernzeit zum Erreichen gesetzter Leistungsziele beiträgt, können sich daraus auch Motivationsprobleme bei den Schülerinnen und Schülern ergeben (Helmke & Weinert, 1997, S. 135). Das Maximum an Lernzeitnutzung im Unterricht ist demzufolge kein optimales Ergebnis eines gelingenden Unterrichts, dennoch ist ein gewisser Grad an Disziplin ein Indikator für einen störungsarmen, gradlinigen und gut organisierten Unterrichtsablauf, also eine effiziente Klassenführung.

5 Fragestellungen und Hypothesen

In der vorliegenden Arbeit wird der Zusammenhang zwischen den allgemeinen pädagogischen Überzeugungen von Lehrkräften, ihrem pädagogischen Handeln und den schulischen Ergebnissen der Schülerinnen und Schüler in drei Forschungsfragen,²¹ jeweils für den Englisch- und den Mathematikunterricht, untersucht:

- 1) Gibt es einen Zusammenhang zwischen den Überzeugungen der Lehrkräfte und dem unterrichtlichen Handeln im Englisch- bzw. Mathematikunterricht?
- 2) Gibt es einen Zusammenhang zwischen den Überzeugungen der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler im Englisch- bzw. Mathematikunterricht, und wenn ja, wird dieser durch das Unterrichtshandeln in Form der (a) Adaptivität bzw. (b) der Disziplin vermittelt?
- 3) Gibt es einen Zusammenhang zwischen den Überzeugungen der Fachlehrkräfte und den Englisch- bzw. Mathematikleistungen von Schülerinnen und Schülern und wenn ja, wird dieser durch (a) die Adaptivität bzw. (b) die Disziplin des Englisch- bzw. Mathematikunterrichts vermittelt?

Die Überzeugungen von Lehrkräften wurden in der Drei-Länder-Studie in der Form von Einschätzungen der Lehrkräfte erfasst. Das Unterrichtshandeln von Lehrkräften ebenso wie das Wohlfühlen der Schülerinnen und Schüler wurde mittels Schülerbefragungen erfasst, die Fachleistungen über Leistungstests. Die folgenden Analysen werden fachbezogen für die Fächer Englisch und Mathematik durchgeführt (Seidel & Shavelson, 2007). Die Fragestellungen führen nach theoretischer Herleitung zu folgenden Hypothesen.

21 Die Ergebnisse zur ersten Forschungsfrage und zu der dazugehörigen Hypothese wurden für das Merkmal Adaptivität im Englischunterricht bereits von Dohrmann, Feldhoff, Steinhilber und Klieme (2019) veröffentlicht. In dieser Publikation werden zudem die Ergebnisse der Forschungsfragen 2a und 3a für den Englischunterricht dargestellt.

Zu den ersten beiden Fragestellungen lassen sich gerichtete Hypothesen spezifizieren. Die erste Fragestellung mündet in die Hypothesen eins bis drei, die vierte Hypothese bezieht sich auf die zweite Fragestellung.

- H1 Bei den Überzeugungen ist aus theoretischer Perspektive zu erwarten (vgl. Kapitel 4.1; Trautmann & Wischer, 2008), dass Lehrkräfte mit einer hohen Reformbereitschaft und einer positiven Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler aufmerksamer den unterschiedlichen Lernbedürfnissen insbesondere den der schwächeren Schülerinnen und Schüler gegenüberstehen und dementsprechend ihren Unterricht eher adaptiv gestalten, wohingegen Lehrkräfte, die der Überzeugung sind, dass Druck für Lernen notwendig ist, den unterschiedlichen Bedürfnissen der Schülerinnen und Schüler keine große Bedeutung beimessen und ihren Unterricht weniger adaptiv gestalten (vgl. Kapitel 4.1; Standop, 2017).
- H2 Lehrkräfte, die eine hohe Reformbereitschaft aufweisen, zeigen damit eine höhere Bereitschaft zur Anpassung an Veränderungen auf institutioneller Ebene (vgl. Kapitel 3.3) und sind offener für den Einsatz neuer Unterrichtsmethoden. Die Maximierung der zur Verfügung stehenden Lernzeit dagegen ist nicht ihre oberste Priorität. Daher messen sie der Disziplin und dem Ordnungssinn im Unterricht keine so große Bedeutung bei wie ihre Kolleginnen und Kollegen (vgl. Abschnitt 4.3.2), die weniger reformbereit sind und Wert darauf legen, an gewohnten Unterrichtsmethoden festzuhalten. Es wird also davon ausgegangen, dass Lehrkräfte mit einer hohen Reformbereitschaft und einer positiven Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler der Disziplin im Unterricht eine geringere Bedeutung beimessen. Lehrkräfte hingegen, die der Überzeugung sind, dass Druck für Lernprozesse notwendig ist, achten stärker auf Disziplin im Unterricht.

Neben der Adaptivität und der Disziplin im Unterricht werden weitere Merkmale eines unterstützenden Unterrichtsklimas sowie der Klassenführung im Unterricht herangezogen. Für deren Zusammenhänge mit den Überzeugungen der Lehrkräfte wird eine dritte Hypothese aufgestellt. Ihre theoretische Herleitung orientiert sich an den Hypothesen 1 (für Merkmale eines unterstützenden Unterrichtsklima) und 2 (für Merkmale der Klassenführung) und wird hier nicht im Detail ausgeführt, da im Zentrum der empirischen Analysen die Adaptivität als ein exemplarisches Unterrichtsmerkmal für ein unterstützendes

Unterrichtsklima und die Disziplin im Unterricht als ein Beispiel für die Klassenführung stehen.

- H3 Es wird davon ausgegangen, dass Lehrkräfte, die eine positive Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler aufweisen, die individuellen Stärken und Schwächen der Lernenden eher wahrnehmen (vgl. Kapitel 4.1) und jeden einzelnen Lernenden mit ihrem Unterricht erreichen möchten. Eine Lehrkraft hingegen, die von der Notwendigkeit von Druck im Unterricht überzeugt ist und Zwang, Kontrolle und Durchgreifen bei Schülerinnen und Schülern präferiert, würde die Lerngruppe eher als ein Ganzes wahrnehmen und sich an einer durchschnittlichen Schülerin bzw. an einem durchschnittlichen Schüler orientieren. Zudem kann angenommen werden, dass reformbereite Lehrkräfte einen schülerorientierten Unterricht bevorzugen und auf die individuellen Unterschiede der Schülerinnen und Schüler eingehen (Trautmann & Wischer, 2008). Daher wird jeweils bei den positiv formulierten Merkmalen aus dem Bereich des unterstützenden Unterrichtsklimas (Unterstützung H3a, Schülerorientiertheit H3c, Individualisierungsmaßnahmen H3d, Motivierungsbemühen H3e, Vertrauensgrad H3f) ein positiver Zusammenhang mit einer Reformbereitschaft und einer positiven Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler angenommen, bei der Strenge ein negativer (H3b). Bei der Überzeugung, dass bei Lernprozessen Druck notwendig ist, wird von einem negativen Zusammenhang mit dem jeweiligen Merkmal aus dem Bereich des unterstützenden Unterrichtsklimas ausgegangen (bzw. von einem positiven bei Strenge).

Die Klassenführung im Unterricht beinhaltet zwei Komponenten: erzieherische Verhaltenssteuerung sowie inhaltliche Strukturiertheit und Klarheit (Klieme, 2019). Die Unterrichtsmerkmale Strukturiertheit und Beschäftigungsradius beziehen sich auf eine inhaltliche Strukturiertheit des Unterrichts, während die Merkmale Disziplin und Zeitverschwendung eine erzieherische Verhaltenssteuerung, also eine soziale Strukturierung darstellen. Bei den hier untersuchten Merkmalen der inhaltlichen Strukturiertheit wird beispielsweise erfasst, ob die Lehrkraft beim Unterrichtsstoff Phasen der Wiederholung und Zusammenfassung vorsieht (Strukturiertheit) oder inwiefern Schülerinnen und Schüler, die sich nicht aktiv melden, in das Unterrichtsgeschehen miteinbezogen werden (Beschäftigungsradius). Diese Aspekte der Un-

terrichtsgestaltung sollten, im Gegensatz zur sozialen Strukturierung, unabhängig von den Überzeugungen der Lehrkräfte sein. Es wird also kein Zusammenhang von Strukturiertheit (H3g) und Beschäftigungsradius (H3h) mit den pädagogischen Überzeugungen der Lehrkräfte angenommen. Bei der Zeitverschwendung (H3i) wird jedoch angenommen, dass Lehrkräfte, die eine besonders hohe Reformbereitschaft und eine positive Einstellung gegenüber der Förderung zeigen, auch die Zeit im Unterricht weniger effektiv nutzen, als Lehrkräfte, die nicht so reformbereit und der Förderung gegenüber weniger positiv eingestellt sind (vgl. Abschnitt 4.3.2). Sie sind experimentierfreudiger hinsichtlich ihrer Methoden und greifen bei Unterrichtsstörungen nicht so hart durch wie diejenigen Lehrkräfte, die der Überzeugung sind, dass Druck beim Lernprozess notwendig ist.

Für alle elf Unterrichtsmerkmale sind die erwarteten Zusammenhänge mit den Überzeugungen der Lehrkräfte in Tabelle 5.1 aufgeführt. Die Merkmale Adaptivität und Disziplin zu den Hypothesen 1 und 2 sind in dieser Tabelle grau hinterlegt.

Tabelle 5.1: Erwartete Zusammenhänge zwischen den Überzeugungen von Lehrkräften und den Unterrichtsmerkmalen

	Hypo- these	Reformbe- reitschaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
unterstützendes Unterrichtsklima	H3a Unterstützung	+	–	+
	H3b Strenge	–	+	–
	H1 Adaptivität des Unterrichts	+	–	+
	H3c Positive Schülerorientiertheit/ Lehrer-Schüler-Beziehung	+	–	+
	H3d Individualisierungsmaßnahmen/ Förderungsorientiertheit	+	–	+
	H3e Motivierungsbemühen/Lob	+	–	+
	H3f Vertrauensgrad/ Schülerorientierung	+	–	+
Klassenführung	H3g Strukturiertheit des Unterrichts	o	o	o
	H3h Beschäftigungsradius	o	o	o
	H2 Disziplin und Ordnungssinn	–	+	–
	H3i Zeitverschwendung	+	–	+

Anmerkungen: +: positiver Zusammenhang erwartet, –: negativer Zusammenhang erwartet, o: kein Zusammenhang erwartet

H4.1 Es wird angenommen, dass eine stärkere Aufmerksamkeit für die und eine stärkere Ausrichtung an den unterschiedlichen (Lern-)Bedürfnissen der Schülerinnen und Schüler mit einem höheren Wohlfühlen im Unterricht einhergeht (vgl. Abschnitt 4.3.1). Schülerinnen und Schüler, die von einer Lehrkraft unterrichtet werden, die eine hohe Reformbereitschaft, eine negative Einstellung gegenüber der Notwendigkeit von Druck bei Lernprozessen oder eine positive Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler aufweist, fühlen sich im Unterricht wohler als ihre Mitschülerinnen und Mitschüler, die von einer Lehrkraft ohne diese Überzeugungsausprägungen unterrichtet werden.

- H4.2 Die Auswirkungen der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Unterricht (siehe H4.1) sind vermittelt über die Adaptivität des Unterrichts (siehe Abschnitt 4.3.1). Demnach müsste eine positive Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler sowie eine hohe Reformbereitschaft vermittelt über die Adaptivität in einem positiven und die positive Einstellung zur Notwendigkeit von Druck bei Lernprozessen in einem negativen Zusammenhang mit dem Wohlfühlen stehen.
- H4.3 Die Auswirkungen der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Unterricht (siehe H4.1) sind vermittelt über die Disziplin im Unterricht (siehe Abschnitt 4.3.2). Demnach müsste eine positive Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler sowie eine hohe Reformbereitschaft vermittelt über die Disziplin im Unterricht in einem positiven und die positive Einstellung gegenüber der Notwendigkeit von Druck bei Lernprozessen in einem negativen Zusammenhang mit dem Wohlfühlen stehen.²²

Wie bei der Darstellung des Forschungsstandes deutlich wurde (vgl. Abschnitt 4.3.1), stellt die Ausrichtung des Unterrichts an den Lernbedürfnissen der Schülerinnen und Schüler häufig eine Reaktion der Lehrkraft auf schwache Leistungen der Schülerinnen und Schüler dar, es handelt sich um eine remediale Beziehung. In querschnittlichen Untersuchungen wie der vorliegenden konterkariert dies die theoretisch erwartete positive Wirkung, so dass keine gerichtete Hypothese für die Fragestellung 3a (den Zusammenhang zwischen Leistung und Adaptivität) formuliert werden kann, und ebenso wenig für die allgemeinen pädagogischen Überzeugungen. Zur Disziplin im Unterricht wurde festgestellt, dass sie ein wichtiges Merkmal der Klassenführung ist und sich daher positiv auf die Leistungen in der Unterrichtsgruppe auswirken sollte (vgl. Abschnitt 4.3.2). Da vermittels der Klassenführung jedoch vor allem die aktive Lernzeit gesteigert werden kann, aber nicht direkt die Qualität des Lernens der Schülerinnen und Schüler, ist eine effiziente Klassenführung für ein erfolgreiches Lernen zwar notwendig, aber nicht hinreichend (Helmke, 2015, S. 176–177). Darüber hinaus könnte ein hohes Maß an Disziplin auch eine Reaktion auf Unterrichtsstörungen sein, beispielsweise dann, wenn der Unterricht sofort

22 Der Prozess der Vermittlung der Überzeugungen der Lehrkräfte auf die Ergebnisse der Schülerinnen und Schüler über das Unterrichtshandeln der Lehrkräfte ist Gegenstand der Kapitel 3 und 4 dieser Arbeit.

abgebrochen wird, sobald eine Unterrichtsstörung auftritt. Daher wird an dieser Stelle keine gerichtete Hypothese für die Fragestellung 3b, den Zusammenhang zwischen Leistung und Disziplin sowie die Vermittlung der pädagogischen Überzeugungen auf die Leistung mittels der Disziplin im Unterricht, formuliert.

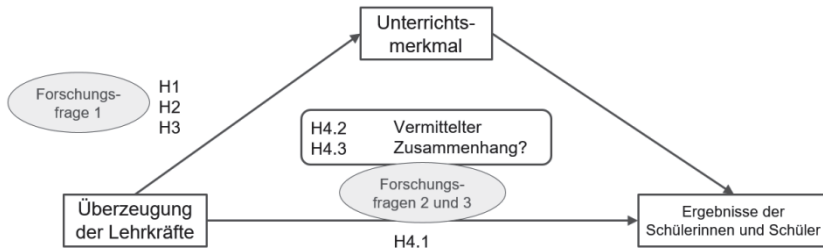


Abbildung 5.1: Schematische Darstellung der Forschungsfragen und Hypothesen

In Abbildung 5.1 sind alle Fragestellungen und Hypothesen der vorliegenden Arbeit noch einmal schematisch dargestellt. Die Pfeile stehen jeweils für einen Zusammenhang zwischen den Merkmalen. Da im Rahmen dieser Arbeit mit der Datengrundlage der Drei-Länder-Studie gearbeitet wird, bei der es sich um eine Querschnittsuntersuchung handelt, kann nicht von Wirkungen im strengen Sinn gesprochen werden, die hier untersucht werden. Daher wird hier bei den Hypothesen und den Ergebnissen von „Zusammenhängen“ gesprochen.

6 Die Datenbasis der Drei-Länder-Studie und ihre Aufbereitung

Die Analysen der vorliegenden Arbeit basieren auf Daten der Drei-Länder-Studie von Helmut Fend. Im Folgenden wird zunächst diese Datenbasis vorgestellt. Danach wird beschrieben, wie die Datenaufbereitung, die Neufassung der Fragebogenskalen und die Testskalierung durchgeführt wurden.

6.1 Die Datenbasis der Drei-Länder-Studie

6.1.1 Entstehungsgeschichte

Die Drei-Länder-Studie (Fend, 1982) wurde von Helmut Fend und seiner Projektgruppe „Wissenschaftliche Begleitung von Gesamtschulen“ an der Universität Konstanz im Auftrag der Kultusministerien der Bundesländer Nordrhein-Westfalen, Hessen und Niedersachsen durchgeführt. Die Aufgabe der Studie bestand darin, den Gesamtschulversuch in den drei Ländern wissenschaftlich zu begleiten. Der Gesamtschulversuch wurde im Jahr 1968 von der Bildungskommission des Deutschen Bildungsrates als „Experimentalprogramm“ zur Einführung von Gesamtschulen beschlossen (Deutscher Bildungsrat, 1969). Die wissenschaftliche Kontrolle des Gesamtschulversuchs war ein zentraler Bestandteil der Empfehlung der Bildungskommission. Erst nach einer wissenschaftlichen Begleitung des Gesamtschulversuchs sollte entschieden werden, ob Gesamtschulen flächendeckend eingeführt werden sollten (Fend, 1982). Deutlich wird dies beispielsweise daran, wie Fend (1979, S. V) den Auftrag, den seine Forschungsgruppe vom nordrhein-westfälischen Kultusministerium erhalten hatte, zusammenfasst: „Hilfestellungen für das Urteil über den Bewährungsgrad bzw. über die Korrekturbedürftigkeit eines Schulversuchs“. Der Schwerpunkt der Evaluation aus Sicht der Auftraggeber der Untersuchung bestand in einem Leistungsvergleich zwischen den neu eingerichteten Gesamtschulen und den Schulen des bisherigen Schulsystems („traditionelles System“ oder „herkömmliches Schulsystem“). Dazu wurde in der 6. und der 9. Jahrgangsstufe eine einmalige Querschnitterhebung durchgeführt. Die vorliegende Arbeit beschäftigt sich nur mit den Schülerinnen und Schülern der 9. Jahrgangsstufe. Der Zugang zu den Daten der vorliegenden Arbeit erfolgte im Projekt „Schule im Wandel“ des „DIPF | Leibniz-Institut für Bildungsforschung

und Bildungsinformation“. Dort wurde der Schwerpunkt aus Gründen der Vergleichbarkeit mit anderen Leistungsstudien auf die 9. Jahrgangsstufe festgelegt.²³

Die Drei-Länder-Studie wurde in den Jahren 1977 bis 1979 jeweils als Querschnittuntersuchung durchgeführt (Fend, 1982, S. 19–20). In den Bundesländern Nordrhein-Westfalen und Hessen wurden die Daten in den Jahren 1977 und 1978 erhoben, in Niedersachsen im Jahr 1979. Da bei den Erhebungen des Jahres 1977 der Schwerpunkt der Untersuchung allein auf den erzieherischen Wirkungen und der sozialen Umwelt sowie auf Chancengleichheit und der Offenheit der Bildungswege lag (Fend, 1979, S. VII), wurden keine Schulleistungstests eingesetzt. Diese kamen erst in den Untersuchungen der Jahre 1978 und 1979 zum Einsatz. Daher wird in der vorliegenden Arbeit auf die Daten der Drei-Länder-Studie aus den Jahren 1978 und 1979 zurückgegriffen. Auf diese bezieht sich auch die folgende Beschreibung der Stichprobenauswahl der Schulen.

6.1.2 Stichprobenauswahl der Schulen

Aufgrund der Zielsetzung eines Leistungsvergleichs zwischen Gesamtschulen und Schulen des traditionellen Schulsystems wurde bei der Zusammenstellung der Stichproben in den jeweiligen Bundesländern darauf geachtet, an einer ausreichenden Anzahl von Gesamtschulen zu erheben, um beispielsweise in Nordrhein-Westfalen rund 500 Schülerinnen und Schüler aus den Gesamtschulen zu erfassen (Lukesch, 1979a, S. 232). Für Niedersachsen war vom dortigen Kultusministerium festgelegt worden, dass sechs Gesamtschulen zu untersuchen seien (Fend, 1981, S. 58), also die Hälfte der damaligen Gesamtschulen in Niedersachsen. Da die Auswahl der Schulen in Nordrhein-Westfalen und Niedersachsen ähnlich ablief, wird die Stichprobenauswahl für diese beiden Bundesländer zunächst zusammen dargestellt, im darauf folgenden Abschnitt wird auf die spezielle Erhebungssituation in Hessen eingegangen.

23 Ein zusätzlicher Grund, den Fokus der vorliegenden Arbeit auf die 9. Jahrgangsstufe zu legen, besteht darin, dass in der Drei-Länder-Studie keine Daten der 6. Jahrgangsstufe von Schulen des „traditionellen Systems“ aus Niedersachsen vorliegen. Dort konnte in der 6. Jahrgangsstufe nur an Gesamtschulen erhoben werden, da zum Erhebungszeitpunkt bereits die Orientierungsstufe (zwei gemeinsame Schuljahre für alle Schülerinnen und Schüler der 5. und 6. Jahrgangsstufe) eingeführt worden war, welche von einem Großteil der Schülerinnen und Schüler besucht wurde. Aus diesem Grund wurden in den damaligen Berichten als Vergleichsgruppe für den Leistungsvergleich die Schulen des „traditionellen Systems“ aus Nordrhein-Westfalen angeführt (Fend, 1981, S. 56).

Wie aus Tabelle 6.1 zu ersehen ist, wurde die Untersuchung in Nordrhein-Westfalen an neun Gesamtschulen (Bergisch-Gladbach, Gelsenkirchen, Kierspe und Mülheim²⁴ sowie Dortmund, Fröndenberg, Köln-Rodenkirchen, Oberhausen und Wulfen) und in Niedersachsen an sechs Gesamtschulen (Aurich, Garbsen, Hannover-Linden, Hannover-Roderbruch, Hildesheim, Wolfsburg) durchgeführt. Die Auswahl dieser Schulen aus den insgesamt 29 Gesamtschulen in Nordrhein-Westfalen und den insgesamt 13 Gesamtschulen in Niedersachsen erfolgte nach mehreren Kriterien: Die Schulen mussten eine 9. Jahrgangsstufe aufgebaut haben, und die Auswahl der Schulen sollte unterschiedliche Standortbedingungen und unterschiedliche Modelle der Unterrichtsorganisation widerspiegeln. Für den Vergleich mit den Schulen des traditionellen Systems wurden in Nordrhein-Westfalen und in Niedersachsen je vier Regionen ausgewählt, in denen jeweils zwei Hauptschulen,²⁵ eine Realschule und ein Gymnasium ausgewählt wurden (in Nordrhein-Westfalen: Herne, Krefeld, Plettenburg, Siegburg; in Niedersachsen: Hannover, Göttingen, Salzgitter, Rotenburg (Wümme)). Lukesch (1979a, S. 233) begründet die Auswahl dieser Schulen am Beispiel von Nordrhein-Westfalen so:

„Bei der Auswahl der Schulen des traditionellen Systems wurde ebenfalls darauf geachtet, daß nach dem Regionalaspekt, der beruflichen Stellung, Erwerbstätigkeit und Religion dieselben Verhältnisse, wie sie für ganz Nordrhein-Westfalen gelten, gegeben waren. Entsprechend dem Schulartenanteil war angestrebt, etwa doppelt so viele Hauptschulen wie Realschulen und Gymnasien in die Untersuchung mit einzubeziehen.“

Fend (1982, S. 20) bezeichnet das Vorgehen der Erhebung von Vergleichsregionen – in seinen Worten „Gegenüberstellung des herkömmlichen Schulwesens zum Gesamtschulsystem“ – als „quasi-experimentelle Studie“ (Fend, 1982, S. 20). In Niedersachsen wurden diese Vergleichsregionen sogar jeweils einer Region zugeordnet, in denen die untersuchten Gesamtschulen lagen (Fend, 1981, S. 57):

„Anhand der statistischen Angaben (statistische Jahrbücher des Landes Niedersachsen) über die Bevölkerungsstruktur der Städte und Landkreise Niedersach-

24 Die ersten vier genannten Gesamtschulen sind Gesamtschulen der sogenannten Grundmodellphase, diese wurden aus Kontinuitätsgründen in die Erhebung einbezogen, da sie bereits in der früheren Konstanzer Sozialisationsstudie von 1977 untersucht worden waren (Lukesch, 1979a, S. 232). Einen Gesamtüberblick über die Studien in Nordrhein-Westfalen vermittelt Fend (1979).

25 In Krefeld wurde nur an einer Hauptschule erhoben, um die zahlenmäßig adäquate Repräsentation der Hauptschüler in der Stichprobe nicht zu überschreiten (Lukesch, 1979a, S. 233).

sens (insbesondere Anteile der Erwerbstätigen in den unterschiedlichen Bereichen) wurden Vergleichsregionen ausgewählt, die den Gesamtschulstandorten möglichst ähnlich sind.²⁶

Dieses besondere Verfahren bei der Auswahl der niedersächsischen Schulen des traditionellen Systems deckt sich größtenteils mit den vom Deutschen Bildungsrat (1969, S. 146) formulierten Vorschlägen zur Auswahl von Vergleichsstichproben zur wissenschaftlichen Kontrolle von Versuchen mit Gesamtschulen.

Tabelle 6.1: Anzahl der Schulen in der Drei-Länder-Studie nach Schulform und Bundesland

	Nordrhein-Westfalen	Hessen	Niedersachsen	Gesamt
Hauptschulen	7	3	8	18
Realschulen	4	–	4	8
Gesamtschulen	9	12	6	27
Gymnasien	4	–	4	8
Gesamt	24	15	22	61

In Hessen wurde die Erhebung an allen Gesamtschulen des Flächenversuchs Wetzlar durchgeführt, das heißt, alle 12 Gesamtschulen der hessischen Stichprobe lagen in einer Region. Ihre Untersuchung entspricht somit einer Vollerhebung der dortigen Gesamtschulen. Im Rahmen eines Schulversuchs wurden 1971 im Altkreis Wetzlar²⁷ alle Schulen von der Sekundarstufe I an in Gesamtschulen umgewandelt („Flächenversuch Wetzlar“).²⁸ In Hessen trat zudem der Sonderfall auf, dass fast alle Schulen des traditionellen Systems aus dem Dillkreis,²⁹ der als Vergleichsregion ausgewählt wurde,³⁰ die Teilnahme unmittel-

26 Dabei ergab sich folgende Zuordnung: drei Gesamtschulen aus dem Großraum Hannover (Garbsen, Hannover-Linden, Hannover-Roderbruch) – traditionelle Schulen aus Hannover, Gesamtschule Hildesheim – traditionelle Schulen aus Göttingen, Gesamtschule Wolfsburg – traditionelle Schulen aus Salzgitter, Gesamtschule Aurich – traditionelle Schulen aus Rotenburg (Wümme).

27 Als Altkreis Wetzlar wird der ehemalige Landkreis Wetzlar bezeichnet, der 1977 aufgelöst wurde und mittlerweile Teil des Lahn-Dill-Kreises ist.

28 Zum historischen Hintergrund des Schulversuchs im Altkreis Wetzlar siehe Mattes (2018).

29 Der Dillkreis wurde 1977 aufgelöst und ist heute Teil des Lahn-Dill-Kreises.

30 Haenisch (1979, S. 79) begründet die Auswahl des Dillkreises als Vergleichsregion für die Region Wetzlar so: „Bei der hessischen Leistungsstudie war geplant, die Gesamtschulen des Kreises Wetzlar mit einer repräsentativen Auswahl von Schularten des gegliederten Systems im Dillkreis zu vergleichen. Mit diesem Vorhaben sollte an die im Jahre 1977 durchgeführte Sozialisationsstudie angeknüpft werden, bei der ebenfalls diese beiden Vergleichskreise einbezogen waren. Die erneute Berücksichtigung der Dill-Kreis-Schulen re-

bar vor Untersuchungsbeginn verweigerten. Lediglich drei Hauptschulen (Beilstein, Breitscheid und Herborn) erfüllten ihre Zusage und nahmen teil (Lukesch, 1979b, S. 4). Dazu führt Haenisch (1979, S. 79) aus:

„Nachdem bereits alle Vorbereitungen abgeschlossen und die Untersuchungstermine mit den Schulen vereinbart waren, zogen kurz vor Untersuchungsbeginn die beteiligten Realschulen und Gymnasien des Dillkreises ihre Zusage zurück. Da außerdem auch noch eine Hauptschule kurzfristig zurücktrat, mußte letztlich darauf verzichtet werden, eine Kontrollgruppe hessischer TS-Schulen [Schulen des Traditionellen Systems, Anm. d. Verf.] für den Systemvergleich zu berücksichtigen.“³¹

Der Grund für die Teilnahmeverweigerung waren „Widerstände aus den Schulen und von Seiten der Elternvertreter [...], die sich einem Leistungsvergleich nicht stellen wollten“ (Lukesch, 1979b, S. 4). Diesem historischen Hintergrund ist es geschuldet, dass in der Gesamtstichprobe der Drei-Länder-Studie Gesamtschulen und Hauptschulen im Vergleich zu den anderen beiden Schularten besonders stark vertreten sind (siehe Tabelle 6.1).

6.1.3 Die Auswahl der Klassen in den Schulen der Stichprobe

An allen 61 Schulen wurde zumeist in zwei Klassen der 9. Jahrgangsstufe erhoben. Konkret wurden in den einzelnen Bundesländern folgende Klassenstichproben durchgeführt: In Nordrhein-Westfalen wurde an allen Schularten in zwei Klassen pro Schule erhoben (Lukesch, 1979a, S. 233). In Niedersachsen wurde an allen Schularten – außer an Hauptschulen – in zwei Klassen pro Schule erhoben³² (Fend, 1981, S. 54–58). In Hessen wurde an den Gesamtschulen in drei Klassen erhoben, bei den drei Hauptschulen wurde an zwei Hauptschulen in einer Klasse und an einer Hauptschule in zwei Klassen erhoben (Lukesch, 1979b, S. 6). Es wurden nur Stammklassen untersucht, das heißt, es

sultierte im Wesentlichen daraus, daß gerade von diesen die Durchführung einer Leistungsstudie gefordert wurde.“ Für die vorliegende Arbeit ist die Analogie zur Erhebung der Drei-Länder-Studie von 1977 nicht von Bedeutung.

31 Aus diesem Grund wurden in den damaligen Berichten als Vergleichsgruppe für den Leistungsvergleich der hessischen Gesamtschulen die Schulen des traditionellen Systems aus Nordrhein-Westfalen angeführt, an denen auch im Jahr 1978 erhoben wurde (Haenisch, 1979; Lukesch, 1979b).

32 In Niedersachsen wurde an einer Hauptschule pro Vergleichsregion in zwei Klassen, an der anderen Hauptschule der Vergleichsregion in einer Klasse erhoben, damit in jeder Vergleichsregion der traditionellen Schulen insgesamt drei Klassen befragt wurden. Somit entsprachen die Vergleichsregionen in ihren Schüleranteilen der Verteilung der Schülerinnen und Schüler des 9. Jahrgangs in ganz Niedersachsen (Fend, 1981, S. 57).

wurden an Schulen mit Leistungsdifferenzierung nicht die einzelnen Kurse befragt, sondern der Klassenverband, in dem der gemeinsame Unterricht in allen nicht leistungsdifferenzierten Fächern stattfand. Dies diente dazu, „die Probleme unterschiedlicher Kursysteme an Haupt-, Real- und Gesamtschulen zu umgehen“ (Lukesch, 1979a, S. 233). Die Auswahl der Klassen erfolgte per Zufallsauswahl aus allen Klassen der 9. Jahrgangsstufe der jeweiligen Schule. Bei der Ziehung der Klassen in den Gesamtschulen in Nordrhein-Westfalen wurde zusätzlich das Verhältnis von Hauptschulabschlussklassen insgesamt zu Hauptschulabschlussklassen mit Qualifikationsvermerk beachtet (Lukesch, 1979a, S. 233).

Insgesamt wurden an den Schulen der Drei-Länder-Studie 3493 Schülerinnen und Schüler der 9. Jahrgangsstufe und 1169 Lehrkräfte der 6. und 9. Jahrgangsstufe befragt. Eine ausführlichere Beschreibung der Stichprobe der Drei-Länder-Studie findet sich in Kapitel 7.1, dort werden die Angaben der Gesamtstichprobe der Drei-Länder-Studie zu der in der vorliegenden Arbeit verwendeten Teilstichprobe für die Sekundäranalysen in Beziehung gesetzt.

6.1.4 Die Erhebungsinstrumente

Anlass der Drei-Länder-Studie war zwar ein Schulsystemvergleich (Fend, 1982, S. 17), allerdings bezog das Untersuchungsinstrumentarium viel mehr als nur die Leistungsmerkmale der Schülerinnen und Schüler ein. Lukesch (1979a, S. 232) erklärt, warum eine Ausweitung der Untersuchung über einen reinen Leistungsvergleich zwischen Gesamtschulen und Schulen des traditionellen Systems hinaus sinnvoll war:

„Eine fundierte wissenschaftliche Untersuchung kann allerdings nicht bei einem deskriptiven Leistungsvergleich stehenbleiben, sondern es muß zusätzlich überlegt werden, wie eventuell aufweisbare Unterschiede auch erklärt werden können. Diese grundwissenschaftliche Fragestellung erfordert die Einbeziehung von möglichen Determinationsvariablen der Schulleistung, wie sie ausgehend von vorliegenden Modellen schulischen Lernens bzw. des Zusammenhangs zwischen Hintergrundfaktoren der Schülerpersönlichkeit und Leistungsergebnissen aufgewiesen werden können. Durch die Einbeziehung dieser Gegebenheiten lassen sich demnach zusätzliche Fragestellungen grundwissenschaftlichen Interesses formulieren, wodurch auch der wissenschaftliche Ertrag der Studie maximiert werden kann.“

Dies zeigt, dass es den Autorinnen und Autoren der Studie wichtig war, mit den gewonnenen Daten Fragen von grundlegendem erziehungswissenschaftlichem

Interesse beantworten zu können. Denn selbst wenn die Erhebungsinstrumente vor allem der Evaluierung von Gesamtschulen dienen sollten, betrafen sie neben den Leistungsergebnissen auch die Offenheit der Bildungswege, Fragen der Chancengleichheit sowie die soziale Umwelt und die erzieherischen Wirkungen von Schule (Fend, 1981, S. 5).

Fend (1982, S. 14) betont, dass man mit dem Konstanzer Forschungsprogramm zur Gesamtschule (das neben der Drei-Länder-Studie noch weitere Untersuchungen umfasste; siehe den Überblick bei Fend, 1982) über die Darstellung der Ergebnisse eines Schulversuchs³³ hinaus, Wesentliches „über Schule sagen kann“. Durch die Beschreibung von Indikatoren war es möglich, auch die „Verhältnisse im herkömmlichen Bildungswesen“ (Fend, 1982, S. 15) zu beleuchten. Dies geschieht laut Fend (1982, S. 16) mit Hilfe von „Indikatoren für die Bildungslaufbahn von Schülern, Indikatoren über Zustände der sozialen Gerechtigkeit oder Ungerechtigkeit, Indikatoren über die Effektivität des Bildungswesens sowie Indikatoren über die Humanität in den Institution, die unsere Kinder 10 bis 20 Jahre lang durchlaufen“. Die „vielleicht wichtigste“ mögliche Erkenntnis aus den Gesamtschuluntersuchungen wird laut Fend (1982, S. 16) dort angesprochen, wo es um „die Grundlagenfrage geht, welche Wirkungen das Bildungswesen auf die heranwachsende Generation hat und wie bedeutsam dabei einzelne Faktoren im Bildungssystem sind“. Zu diesen Faktoren zählt Fend (1982, S. 16) neben der Schulorganisation vor allem die einzelne Lehrkraft und das gesamte Kollegium einer Schule.

Als Erhebungsinstrumente wurden in der Drei-Länder-Studie neben kognitiven und affektiven Persönlichkeitsmerkmalen auch schülerbiografische Merkmale, familiäre Hintergrundvariablen, curriculare Merkmale sowie Merkmale der didaktischen Qualität des Unterrichts und des Schulklimas einbezogen, um die Leistungsunterschiede zwischen Schulsystemen, Schulen und einzelnen Schülerinnen und Schülern erklären zu können (Lukesch, 1979b, S. 2). Darüber hinaus sollte der Leistungsstand der verschiedenen Schularten in Bezug zu affektiven und schulklimatischen Merkmalen gesetzt werden, um eventuelle Unterschiede analysieren zu können.

33 Fend (1982, S. 14) stellt bereits kurze Zeit nach der Drei-Länder-Studie in seiner Bilanz des Gesamtschulversuchs fest: „Stünde das bildungspolitische Ziel der Beeinflussung des Entscheidungsprozesses über die Einführung der Gesamtschule als alleinige Ausrichtung im Vordergrund dieser Arbeit, dann könnte sie beim gegenwärtigen Stand der Gesamtschulentwicklung von vielen mit Recht als überholt betrachtet werden.“

Eine ausführliche Dokumentation aller Erhebungsinstrumente der Drei-Länder-Studie („Fend-Studie – Wissenschaftliche Begleitung von Gesamtschulen“) findet sich in der frei zugänglichen Datenbank des Forschungsdatenzentrums (FDZ) Bildung (2014b): <https://www.fdz-bildung.de/studiendetails.php?la=de&id=46>. Dort sind alle Skalen der Schülerinnen und Schüler der Erhebung 1978 einschließlich der gängigen Kennwerte aus der unveröffentlichten originalen Skalendokumentation³⁴ hinterlegt (Krüger-Haenisch, o. J.). Zusätzlich sind alle Einzelvariablen, die keiner Skala angehören, dokumentiert. Bei der Erhebung 1979 wurden die gleichen Instrumente wie 1978 eingesetzt, nur bei einzelnen Fragen aus den Bereichen soziobiografische Hintergrundvariablen und Schulbiographie gibt es geringe Unterschiede in der Formulierung oder der Reihenfolge. In der Datenbank des FDZ Bildung wurden die Instrumente den Kategorien „individueller und institutioneller Hintergrund“, „Schulische Wirkungen“, „Prozesse auf Schulebene“ und „Lehr-Lernarrangements“ zugeordnet. Für jede dieser Kategorien sind in Abbildung 6.1 die Oberbegriffe der Konstrukte so dargestellt, dass anhand der Größe und der Farbintensität der Darstellung ersichtlich ist, in welchem Umfang die jeweiligen Konstrukte durch Skalen abgedeckt sind. Es muss betont werden, dass die Sammlung in der Datenbank des FDZ Bildung die originalen Instrumente der Drei-Länder-Studie enthält, während für die vorliegende Arbeit und innerhalb des Projekts „Schule im Wandel“ die Instrumente neu gefasst und angepasst wurden (siehe Kapitel 6.3).

34 In der originalen Skalendokumentation sind, wie auch in der Datenbank des FDZ Bildung, die Skalen getrennt für die Schulen des traditionellen Systems und die Gesamtschulen aufgeführt. In allen Schularten wurden die gleichen Instrumente eingesetzt, allerdings gab es speziell für die Gesamtschulen spezifische Fragen zur Leistungsdifferenzierung. Abbildung 6.1 zeigt die Übersicht aus der Datenbank des FDZ Bildung für die Gesamtschulen.



Quelle: Forschungsdatenzentrum Bildung (2014a)

Abbildung 6.1: Erhebungsschwerpunkte der Instrumente für Schülerinnen und Schüler der Drei-Länder-Studie auf einen Blick

Für den Leistungsvergleich in der Drei-Länder-Studie sollte ein möglichst breitgefächertes Leistungsspektrum der Schülerinnen und Schüler erfasst werden. Deshalb wurden die Leistungen in den Fächern Deutsch, Mathematik, Englisch und Naturwissenschaften (Physik und Biologie) untersucht (Lukesch, 1979a, S. 235–237). Darüber hinaus wurden die kognitiven Fähigkeiten der Schülerinnen und Schüler mit dem CFT 2 (Cattell & Weiß, 1972) erfasst.³⁵ Die Befragung der Lehrkräfte umfasste ihre biografischen Merkmale und Einstellungen, die von ihnen benutzten methodisch-didaktischen Verfahren sowie Lerngelegenheitsbeurteilungen zu allen eingesetzten Fachleistungstests (Fend, 1981, S. 11). Alle Instrumente, die in der vorliegenden Arbeit zur Untersuchung der Schülerinnen und Schüler und der Lehrkräfte verwendet werden, werden im Kapitel 7.2 näher beschrieben. Eine Übersicht über alle Fragebogenskalen der Drei-Länder-Studie findet sich darüber hinaus in Kapitel 6.3. Für weiterführende Informationen zur Datenbasis und zu den Instrumenten wird auf das Skalenhandbuch aus dem Projekt „Schule im Wandel“ verwiesen (Schmid et al., 2019).

35 Die Abkürzung CFT steht für „Culture-Fair-Intelligence Test“ (Cattell & Weiß, 1972) und ist nicht mit der Abkürzung KFT für „Kognitiver Fähigkeitstest“ zu verwechseln.

6.1.5 Das Erhebungsdesign

Der Erhebungszeitraum erstreckte sich in Nordrhein-Westfalen vom 20. April bis zum 24. Mai 1978 (Lukesch, 1979a, S. 237) und in Hessen vom 5. bis zum 22. Juni 1978 (Lukesch, 1979b, S. 18). In Niedersachsen erfolgte die Erhebung zwischen dem 28. Mai und dem 4. Juli 1979 (Fend, 1981, S. 69). Die Erhebungen begannen demnach jeweils circa drei bis sieben Wochen nach den Osterferien und endeten etwa zwei bis sechs Wochen vor den Sommerferien der jeweiligen Bundesländer.

1. Woche		2. Woche	
1. Tag	2. Tag	3. Tag	4. Tag
Vorstellung (15 Min.)	Englisch (60 Min.)	Deutsch (45 Min.)	Intelligenztest (30 Min.)
Einweisung in die Fragebogenaus- füllung (30 Min.)	Pause (15 Min.)	Pause (15 Min.)	Pause (15 Min.)
Pause (15 Min.)	Fragebogen (30 Min.)	Fragebogen (45 Min.)	Fragebogen (30 Min.)
Mathematik (45 Min.)	Pause (15 Min.)	Pause (30 Min.)	Pause (15 Min.)
Pause (15 Min.)	Englisch (45 Min.)	Deutsch (30 Min.)	Naturwissen- schaft (60 Min.)
Mathematik (45 Min.)			

Quelle: Lukesch, 1979b, S. 21

Abbildung 6.2: Abfolge der Fragebögen und Tests in der Drei-Länder-Studie 1978/79

Wie aus Abbildung 6.2 ersichtlich wird, erstreckte sich die Erhebung an jeder Schule über vier Tage innerhalb von zwei Wochen (Fend, 1981, S. 70; Lukesch, 1979a, S. 251). Pro Tag waren für die Erhebung mindestens zweieinhalb Stunden eingeplant. Die reine Test- und Befragungszeit (ohne Pausen und Einweisung) betrug insgesamt sieben Stunden und 45 Minuten.

Die Befragung der Schülerinnen und Schüler war aufgeteilt auf sechs Fragebögen, die als „Schülerheft“ bezeichnet wurden und von Schülerheft 1 bis Schülerheft 6 durchnummeriert waren. Bei den Leistungstests gab es für die Fächer Englisch, Mathematik und Deutsch zwei Testhefte, in Naturwissenschaften und als Intelligenztest wurde ein Testheft eingesetzt. Die Testhefte wurden mit der

Jahrgangsstufe und einer laufenden Ziffer durchnummeriert, die Nummerierung entspricht jedoch nicht der Abfolge der Testung (siehe Abbildung 6.2): Deutsch 9/1 und 9/2, Mathematik 9/3 und 9/7, Englisch 9/4 und 9/5, Naturwissenschaften 9/6 sowie der kognitive Fähigkeitstest (ohne Nummer). Die Befragung der Lehrkräfte wurde auf zwei Fragebögen aufgeteilt, die analog zu den Schülerheften als „Lehrerheft 1“ und „Lehrerheft 2“ bezeichnet wurden.

Die die Erhebung durchführenden Personen bestanden pro Schule aus einer Teamleitung, die Mitglied im seinerzeitigen Forschungsprojekt war und drei (Nordrhein-Westfalen) bis fünf (Hessen) Hilfskräften (Lukesch, 1979a, S. 237, 1979b, S. 18). Auch in Niedersachsen wurde das Projektteam durch Studierende und Absolventen der Universität Konstanz unterstützt (Fend, 1981, S. 69). Darüber hinaus war in jeder Klasse eine Lehrkraft anwesend, die zu dieser Zeit in der jeweiligen Klasse Unterricht gehabt hätte (Lukesch, 1979b, S. 18) bzw. die Klasse kannte (Fend, 1981, S. 69).

6.2 Die Aufbereitung der Daten

Die Datensätze der Drei-Länder-Studie waren ursprünglich nicht in einer Datenstruktur verfügbar, die mit Hilfe eines aktuellen Statistikprogramms (beispielweise SPSS) hätte eingelesen werden können. Aufgrund der Übertragung vom originalen Datenformat aus den 1970er Jahren in ein Datenformat, das mit einem aktuellen Statistikprogramm lesbar ist, mussten die Datensätze umfangreich aufbereitet werden. Die Datensätze lagen jeweils getrennt für jedes Erhebungsjahr und für Lehrkräfte bzw. Schülerinnen und Schüler vor.

Vor allem die Aufbereitung der Daten der Schülerinnen und Schüler gestaltete sich sehr aufwändig, unter anderem aufgrund der Menge an Variablen. Die Aufbereitung der Schülerdatensätze betraf die Fragebogendaten sowie die Leistungsdaten in einer Größenordnung von circa 1360 Variablen (600 aus den Schülerheften und 760 Testitems). Die relevanten Variablen der Schülerhefte und die Testitems mussten aus über 1600 Variablen der beiden Rohdatensätze herausgefiltert³⁶ werden, die restlichen Daten bestanden aus Hilfsvariablen oder aggregierten Werten, die nicht übernommen, sondern auf der Grundlage der Originalitems im späteren Verlauf neu gebildet wurden. Bei den Daten der

36 Eine große Hilfe bei der Zuordnung der Variablen der Datensätze zu den Fragen der Fragebögen waren die Materialien und Unterlagen von Helmut Fend und Ulrich Steffens, die der Verfasserin freundlicherweise zur Verfügung gestellt wurden.

Lehrkräfte betraf die Aufbereitung knapp 940 Variablen (250 aus Lehrerheften und 690 aus der Beurteilung von Lerngelegenheiten), wobei diese aus knapp 1020 Variablen des Rohdatensatzes herausgefiltert wurden, da auch hier Hilfsvariablen oder bereits aggregierte Werte vorlagen, die im weiteren Verlauf auf der Grundlage der originalen Items neu gebildet wurden.

Zunächst erfolgte die Vergabe von eindeutigen Variablennamen, da die Variablen ursprünglich durchnummeriert waren und deshalb keinen Rückschluss mehr auf die eigentlichen Fragen bzw. Items der Leistungstests erlaubten; ein beispielhafter Auszug eines in SPSS transformierten Schülerdatensatzes im Rohformat findet sich in Abbildung 6.3. Da die Schülerdatensätze der Jahrgänge 1978 und 1979 nicht in der gleichen Reihenfolge sortiert waren und sich teilweise die Inhalte der Fragebögen dieser beiden Erhebungen voneinander unterschieden, mussten diese vor der Zusammenführung der Datensätze separat aufbereitet werden.

	Name	Typ	Spaltenformat	Dezimalstellen	Variablenlabel	Wertelabels	Fehlende Werte
498	v498	Numerisch	1	0	S6/78 PFE:	Keine	Keine
499	v499	Numerisch	1	0	S6/82 PFE:	Keine	Keine
500	v500	Numerisch	1	0	S6/85 PFE:	Keine	Keine
501	v501	Numerisch	1	0	S6/9 FLA:	Keine	Keine
502	v502	Numerisch	1	0	S6/20 FLA:	Keine	Keine
503	v503	Numerisch	1	0	S6/31 FLA:	Keine	Keine

Abbildung 6.3: Ausschnitt eines Rohdatensatzes der Drei-Länder-Studie in der SPSS-Variablenansicht

Im Folgenden wird beispielhaft für die beiden Schülerdatensätze das Vorgehen der Datenaufbereitung aufgeführt. Bei den Daten der Lehrkräfte gab es keine Unterschiede bei den Fragebögen zwischen den Jahrgängen 1978 und 1979 und keine Leistungstests. Das führte zu einer etwas anderen Reihenfolge der unten aufgeführten Arbeitsschritte, die hier nicht näher dargestellt wird.

Arbeitsschritte der Datenaufbereitung für Schülerinnen und Schüler:

- *Identifizierung von Dubletten und Löschen der doppelten Fälle:*
Die originalen Datensätze wiesen Dubletten auf, die für die weiteren Analysen entfernt werden mussten. Die Identifizierung von doppelten Fällen erfolgte in den Programmen R 3.4.2 (R Core Team, 2017) und SPSS 20.0. Eindeutige Dubletten wiesen bei allen biografischen Variablen die gleichen Werte auf wie ihr „Zwilling“. Es wurde immer ein

Fall beibehalten. Weitere Dubletten hatten lediglich die gleiche ID. Diese Fälle wurden manuell überprüft, zumeist enthielt der doppelte Fall keine weiteren Angaben neben der ID, war also ein „leerer“ Fall und konnte gelöscht werden. In einigen Fällen stellten sich die beiden Fälle mit gleicher ID als unterschiedliche Fälle heraus, da sich die übrigen Daten unterscheiden. Dort wurde für einen der beiden Fälle eine neue ID vergeben.

– *Abgleich der Inhalte der Messinstrumente mit der Datenstruktur:*

Für alle Daten aus den sechs Fragebögen (Schülerhefte 1 bis 6) und fünf Tests (Deutsch, Mathematik, Englisch, Naturwissenschaften, CFT) musste die Zuordnung zu den Inhalten hergestellt werden. Dazu war es notwendig, die originalen Fragebögen und Testhefte sowie den jeweiligen Datensatz sehr genau zu prüfen. Variablen im Datensatz ohne Bezug zu den originalen Instrumenten wurden gestrichen, der neue Datensatz enthielt nur Variablen, die eindeutig identifiziert werden konnten. Im Zuge der Kontextualisierung wurden Variablenamen vergeben (d.h. die Zuordnung zum Fragebogen erfolgte), das Variablenlabel beschrieben (d.h. der Fragentext wurde aus dem Fragebogen übernommen) und Wertelabels definiert (d.h. die Antwortmöglichkeiten wurden aus dem Fragebogen übernommen und den Ausprägungen zugeordnet).

– *Bereinigung der Datensätze:*

Bei den befragten Schülerinnen und Schulen aus Nordrhein-Westfalen und Hessen lagen beispielsweise Antworten vor, die aufgrund von Filterfragen nicht hätten gegeben werden dürfen. Für solche Antworten wurden fehlende Werte vergeben. Die Filterfragen waren für die Befragung der Schülerinnen und Schüler in Nordrhein-Westfalen und Hessen (nicht der in Niedersachsen und nicht für die Befragung der Lehrkräfte) bisweilen notwendig, da sie den gleichen Fragebogen vorgelegt bekamen, einige Fragen jedoch bundeslandspezifisch gestellt wurden. So hatten beispielsweise 70 Personen aus Hessen einen Fragenblock von sieben Fragen am Ende von Schülerheft 2 beantwortet (s225 bis s231), der nur für Ganztagschulen in Nordrhein-Westfalen gedacht war (was auf der Fragebogenseite auch deutlich zu lesen ist). Bei den neun Variablen bezüglich der Kurszuordnung von Schülerinnen und Schülern in leistungsdifferenziert unterrichteten Fächern in Schülerheft 1 (s137m bis s139d) ließen sich pro Variable höchstens sieben Personen keiner gültigen Antwortmöglichkeit zuordnen, da sie

eine Antwortmöglichkeit für das jeweilige anderen Bundesland gewählt hatten; Hessen und Nordrhein-Westfalen hatten unterschiedliche Kurssysteme, sie wurden aber jeweils in einer gemeinsamen Frage erfasst. Dass die Filterung bei den Antwortmöglichkeiten einer Frage fehleranfällig ist, wird auch bei der Frage nach der Schulmitteilung des zukünftigen Abschlusses deutlich (s154). Hier betrug die Fehlzuordnung aufgrund der Bundeslandzugehörigkeit jedoch ebenfalls nur Werte im einstelligen Bereich. Bei der Frage nach der Angst vor den schulischen Folgen eines Leistungsabfalls (s223 und s224) gab es getrennte Fragen für Schülerinnen und Schüler der Gesamtschule und der restlichen Schulformen, hier haben jeweils knapp 200 Personen die falsche Frage beantwortet. Darüber hinaus wurden Zahlenwerte, die außerhalb des Wertebereichs der Variablen lagen, als ungültiger Wert markiert. Dies trat lediglich bei zwei Variablen³⁷ auf und betraf jeweils nur circa 25 Personen. Insgesamt wurden alle Variablen aller Erhebungen daraufhin geprüft, ob ein hoher Zahlenwert immer einer hohen Ausprägung der Variable entsprach und gegebenenfalls rekodiert. Alle binären Variablen wurden auf solche Weise rekodiert, dass der Wert 1 „ja“ und der Wert 0 „nein“ abbildet.

– *Zusammenführen der Datensätze von 1978 und 1979:*

Um alle erhobenen Schülerinnen und Schüler der Drei-Länder-Studie in einer Datenbasis vorliegen zu haben, wurden die Datensätze von Nordrhein-Westfalen und Hessen 1978 und Niedersachsen 1979 zusammengeführt. Dafür mussten einige Variablennamen vereinheitlicht werden, da sie sich auf Variablennummern im Fragebogen bezogen und sich die Reihenfolge der Fragen in beiden Jahren zum Teil unterschied. Als Vorlage diente immer die Befragung von 1978. Darüber hinaus gab es spezifische Fragen, die nur in Niedersachsen bzw. nur in Nordrhein-Westfalen und Hessen gestellt wurden. Diese wurden mit den Anhängseln „ni“ bzw. „nh“ gekennzeichnet.

– *Neufassung der Fragebogenskalen:*

Zur Aufbereitung der Datensätze der Drei-Länder-Studie gehörte auch die Neufassung aller Fragebogenskalen. Hierzu wurden Faktorenanalysen durchgeführt, um die affektiven Schülerskalen, die Skalen zum elterlichen Verhalten und die Unterrichtsskalen zu erstellen und sie auf ihre Konsistenz hin zu prüfen (zur Ausführung siehe Kapitel 6.3).

37 s113_2 „Mit wem lebst Du zu Hause zusammen? Stiefvater“ und s301 „Ich habe zu Hause einen eigenen Schreibtisch.“

– *Testskalierung mit IRT:*

Die Skalierungen der Fachleistungstests und des CFT erfolgten auf der Grundlage der Item-Response-Theorie. Dafür waren in SPSS umfangreiche Kodierungsarbeiten notwendig. Anschließend wurden die berechneten Personenfähigkeiten aus R an den SPSS-Datensatz angehängt (Ausführung siehe Kapitel 6.4).

6.3 Die Neufassung der Fragebogenskalen

Da die Datenbasis der Drei-Länder-Studie im Rahmen dieser Arbeit komplett neu aufbereitet wurde, wurden auch die Fragebogenskalen der Schülerinnen und Schüler und der Lehrkräfte von der Projektgruppe „Schule im Wandel V“³⁸ neu erstellt. Diese Neubildung diente nicht einer kompletten Revision oder inhaltlichen Neuausrichtung der seinerzeit eingesetzten Skalen, sondern lediglich der Überprüfung, ob sich die Skalen, die 1978 und 1979 gebildet wurden, beim Einsatz heutiger statistischer Methoden ebenfalls ergeben und ob sie den heutigen Standards von Reliabilität, Dimensionalität, Konsistenz und Trennschärfe genügen (Rammstedt, 2010). Darüber hinaus war für die Dokumentation der neu aufbereiteten Datenbasis die Berechnung der statistischen Kennwerte aller Skalen und der dazugehörigen Items auf der Grundlage dieser Datensätze notwendig. Alle Skalen und die dazugehörigen Items der Schülerinnen und Schüler sowie der Lehrkräfte sind im Skalenhandbuch der aufbereiteten Datenbasis dokumentiert (Schmid et al., 2019). Im Folgenden wird das Vorgehen bei der Neufassung der Fragebogenskalen erklärt und sein Ergebnis in Übersichtstabellen vorgestellt.

Es wurden explorative Faktorenanalysen (Backhaus, Erichson, Plinke & Weber, 2010; Wolff & Bacher, 2010) durchgeführt, um festzustellen, auf wie vielen Dimensionen die Items einer ursprünglich als eindimensional behandelten Skala gegebenenfalls laden. Viele Skalen deckten, wie seinerzeit durchaus üblich, inhaltlich ein breites Spektrum ab. Daher war davon auszugehen, dass sich bei einigen Skalen mehrere Faktoren aus den Items ergeben würden (siehe beispielsweise Skala 1, Tabelle 6.3). Die explorativen Faktorenanalysen wurden in SPSS 20.0 durchgeführt. Die Extraktion der Faktoren erfolgte nach der Hauptkomponentenmethode mit orthogonaler Varimax-Rotation, wobei nach dem Kaiser-Kriterium alle Faktoren mit Eigenwerten größer 1 beibehalten

38 Das „V“ steht für „Vorstudie“.

wurden. Dort wo die Items einer Skala mehr als einen Faktor bildeten, wurde versucht, die Items der Skala so neu zusammenzustellen, dass sie statistisch auf einem gemeinsamen Faktor laden und zugleich inhaltlich sinnvoll interpretierbar sind. Darüber hinaus wurden die originalen Skalen dann verändert, wenn die Skalen bei den folgenden Reliabilitätsanalysen eine niedrige Konsistenz aufwiesen.³⁹ Auch in diesem Fall wurde das betreffende Item aus der Skala ausgeschlossen. In beiden Fällen wurde es als wichtig angesehen, die originalen Skalen möglichst wenig zu verändern, um nahe bei den Skalen der Dreiländer-Studie zu bleiben und mit den Veränderungen die Güte der Skala (gemessen durch Cronbachs Alpha) wenn möglich zu verbessern. Sowohl bei den Faktorenanalysen als auch bei den Reliabilitätsanalysen wurden fehlende Werte fallweise ausgeschlossen. Die Ausprägung der Skalen bei den einzelnen Personen wurde mit Hilfe von Summenscores, bereinigt um die Anzahl der Items, berechnet. Dadurch weist die gebildete Skala die gleichen Endpunkte des Antwortformats auf wie die Items (Mittelwertscores).⁴⁰ Es wurden nur für diejenigen Personen Werte bei einer Skala berechnet, die bei mindestens 75 Prozent der Items der Skala eine gültige Antwort gegeben hatten. Alle gebildeten Skalen, sowohl die ursprünglichen als auch die neu erstellten, wurden mit dem aktuellen Forschungsstand hinsichtlich der Messung dieser Konstrukte abgeglichen und von der Projektgruppe, wenn inhaltlich nötig, mit neuen Konstruktnamen versehen. Häufig wurden den ursprünglichen Namen der Konstrukte aus Gründen einer besseren Verständlichkeit lediglich Zusätze hinzugefügt.

Insgesamt konnte bei 21 von 52 Skalen eine Verbesserung der Skala hinsichtlich Eindimensionalität und Konsistenz dadurch erreicht werden, dass ein oder mehrere Items aus der Skala ausgeschlossen wurden. Bei vier Skalen wurde die ursprüngliche Skala in mehrere Unterasspekte aufgeteilt. In sechs Fällen wurden zwei Skalen, die inhaltlich ähnliche Bereiche abbildeten, zu einer neuen Skala zusammengefügt, welche nach dem Stand der heutigen Forschung ein gemeinsames Konstrukt abbildet. 14 Skalen wurden in ihrer ursprünglichen Form belassen. Nur bei sieben Skalen, die mehrfaktoriell waren oder eine niedrige

39 Dies lässt sich anhand des Gütekriteriums Cronbachs Alpha feststellen ($< .7$ zeigt eine zufriedenstellende, $< .8$ eine gute interne Konsistenz an; Rammstedt, 2010, S. 249) bzw. wenn die Trennschärfe eines oder mehrerer Items einer Skala besonders niedrig ist ($< .3$; Bortz & Döring, 2006, S. 220), das heißt wenn das Item mit der Skala keinen hohen Zusammenhang aufweist.

40 Wenn zum Beispiel die Items der Skala fünfstufig sind, ist der Wertebereich der Skala 1 bis 5.

Konsistenz aufwiesen, konnte keine Verbesserung erreicht werden. Bei den letzteren Skalen empfiehlt sich die Analyse von inhaltlich interessanten Einzelitems. Wie sich die Veränderungen an den Skalen auf die Datensätze der Schülerinnen und Schüler sowie der Lehrkräfte verteilen, zeigt Tabelle 6.2. Dort werden die Skalen der Schülerinnen und Schüler, die sich auf den Unterricht beziehen, analog zu den Darstellungen auf den folgenden Seiten separat ausgewiesen. Welche Items wie von den Veränderungen betroffen sind, lässt sich im Detail im Skalenhandbuch von Schmid et al. (2019) nachvollziehen.

Tabelle 6.2: Veränderungen an den Skalen der Drei-Länder-Studie in der Übersicht

Skalen der Schülerinnen und Schüler				
Vorgenommene Veränderung	affektive Skalen, elterliches Verhalten	Unterrichtsskalen	Skalen der Lehrkräfte	Anzahl
Ausschluss von Items	10*	7	4	21
Aufgeteilt in Subskalen	4	–	–	4
Neue Zusammenstellung	–	2	4	6
Keine Veränderung notwendig	5	5	4	14
Keine Verbesserung möglich	2	1	4	7
Gesamtanzahl	21	15	16	52
	36			

Anmerkung: *Die reduzierte Skala zur Lernmoral sks3r wird hier nicht mitgezählt, da die Lernmoral zusätzlich auf zwei Subskalen aufgeteilt wurde (siehe Tabelle 6.3).

Die Skalen der Schülerinnen und Schüler

Eine Übersicht über die affektiven Skalen der Schülerinnen und Schüler und der Skalen zum Verhalten der Eltern der Schülerinnen und Schüler findet sich in Tabelle 6.3. Eine Übersicht über die Unterrichtsskalen findet sich, jeweils für den Englisch- und den Mathematikunterricht, in Tabelle 6.4. In beiden Tabellen sind jeweils die originalen und die neuen Fassungen der Skalen abgetragen. Eine Legende und nähere Erläuterungen zu der Übersicht befinden sich im Anschluss an Tabelle 6.4.

Bei drei Skalen aus Tabelle 6.3 lag der Fall vor, dass bereits in den originalen Skalen Items ausgeschlossen worden waren, die Teil der Erhebung waren.

Diese Items konnten im jeweiligen Datensatz rekonstruiert werden. Nach der erneuten Überprüfung der Skalen stellte sich heraus, dass sich diese Items den reduzierten Skalen 6r*, 18.2* und 21.2* zuordnen ließen, obwohl sie in den originalen Skalen nicht verwendet wurden.

Für die drei Inhaltsbereiche, die jeweils durch mehrere Skalen abgedeckt werden, „Attribuierung schlechter Schulleistungen“ (Skalen 10 bis 11), „Elterliche Leistungs-/Förderungs-/Sanktionsorientierung“ (Skalen 13 bis 15) und „Elterliche Hausaufgabenunterstützung“ (Skalen 19 bis 21), wurde je eine Faktorenanalyse aller Items eines Inhaltsbereichs durchgeführt, wobei sich jeweils genau wieder diejenigen Faktoren ergaben, die in den einzelnen Skalen abgebildet sind.

Die Skalen der Lehrkräfte

Eine Übersicht über die Skalen der Lehrkräfte findet sich in Tabelle 6.5. In der Tabelle sind jeweils die originalen und die neuen Fassungen der Skalen abgetragen. Eine Legende und nähere Erläuterungen zu der Übersicht befinden sich wiederum im Anschluss an Tabelle 6.5.

Tabelle 6.3: Skalen der Drei-Länder-Studie für Schülerinnen und Schüler, originale und neue Fassungen – 1. Teil: affektive Skalen und elterliches Verhalten

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
1	sks1	Leistungsangst	21	0 = stimmt nicht 1 = stimmt	siehe Skala 1.1 und 1.2	.891	3	.386 - .620	
1.1	sks1_1	Prüfungsangst	7		„Wenn ich bei Klassenarbeiten merke, daß die Zeit knapp wird, verliere ich leicht die Nerven.“	.790	1	.423 - .601	
1.2	sks1_2	Leistungsangst (psychosomatisch)	8		„Wenn ich im Unterricht etwas vor der Klasse vortragen soll, habe ich Angst, mich zu verpressen oder zu stottern.“	.775	1	.417 - .545	
2	sks2	Selbstakzeptierung	9	0 = stimmt nicht 1 = stimmt	„Ich halte nicht viel von mir.“	.760	2	.371 - .494	Die Skala ist negativ gepolt.
2r	sks2r	Selbstakzeptierung/ Selbstwert – reduziert	5			.691	1	.411 - .474	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
3	sks3	Lernmoral	10		siehe Skala 3.1 und 3.2	.833	2	.418 - .613	
3r	sks3r	Lernmoral GFM – reduziert	9			.822	1	.407 - .620	Generalfaktormodell
3.1	sks3_1	Lernmoral bei Hausaufgaben	4	0 = stimmt nicht 1 = stimmt	„Ich mache zuerst die Hausaufgaben, um dann meine Freizeit mehr genießen zu können.“	.759	1	.467 - .612	Unterspekt Hausaufgaben
3.2	sks3_2	Lernmoral/-disziplin allgemein	6		„Wenn ich einmal in der Schule gefehlt habe, frage ich einen Schulkameraden, was wir in dieser Zeit gelernt haben und lerne es nach.“	.750	1	.400 - .548	Unterspekt allgemeine Lerndisziplin
4	sks4	Arbeitskapazität	10	0 = stimmt nicht 1 = stimmt	„Im Unterricht sind meine Gedanken oft ganz woanders.“	.793	2	.322 - .583	
4r	sks4r	Interferenzen bei der Selbstregulation	5			.756	1	.450 - .619	
5	sks5	Schulinvolvement - Einstellung zu Lehrern	6	0 = stimmt nicht 1 = stimmt	„Ich mag die meisten Lehrer bei uns recht gern.“	.691	1	.243 - .534	
5r	sks5r	Zufriedenheit mit Lehrern auf Schulebene (Lehrer-Schüler-Beziehung)	4			.717	1	.478 - .542	
6	sks6	Schulinvolvement - Wohlbefinden	8	0 = stimmt nicht 1 = stimmt	„Ich gehe ziemlich gerne in die Schule.“	.730	2	.311 - .496	
6r*	sks6r	Schulfreude/Well-Being	6			.747	1	.434 - .586	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
7	sk7	Anomia	8		„Ich kann tun, was ich will, ich bekomme immer schlechte Noten.“	.670	2	.292 - .481	
7r	sk7r	Anomia – reduziert	4	0 = stimmt nicht 1 = stimmt		.646	1	.371 - .500	
8	sk8	Begabungsselbstbild	9		„Häufig denke ich: ich bin nicht so klug wie die anderen.“	.711	2	.267 - .495	
8r	sk8r	Akademisches Selbstkonzept im sozialen Vergleich	5	0 = stimmt nicht 1 = stimmt		.683	1	.396 - .519	
9	sk9	Soziale Erwünschtheit	6	0 = stimmt nicht 1 = stimmt	„Ich sage immer die Wahrheit.“	.643	2	.278 - .543	keine Verbesserung durch Reduzierung möglich
10	sk10	Attribuierung schlechter Schulleistungen – Anstrengung	7	1 = Der Satz ist völlig falsch 2 = Der Satz ist eher falsch als richtig 3 = Ich bin unentschieden bzw. weiß es nicht 4 = Der Satz ist eher richtig als falsch 5 = Der Satz trifft genau zu	„Ich führe meine schlechteren Schulleistungen darauf zurück, daß ich mich nicht genügend anstrenge.“	.848	1	.462 - .671	
11	sk11	Attribuierung schlechter Schulleistungen – Begabung	5	1 = Der Satz ist völlig falsch 2 = Der Satz ist eher falsch als richtig 3 = Ich bin unentschieden bzw. weiß es nicht 4 = Der Satz ist eher richtig als falsch 5 = Der Satz trifft genau zu	„Ich führe meine schlechteren Schulleistungen darauf zurück, daß ich in diesen Fächern nicht so begabt bin.“	.787	1	.494 - .622	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
12	sks12	Attribuierung schlechter Schulleistungen – Umstände	5	1 = Der Satz ist völlig falsch 2 = Der Satz ist eher falsch als richtig 3 = Ich bin unentschieden bzw. weiß es nicht 4 = Der Satz ist eher richtig als falsch 5 = Der Satz trifft genau zu	„Ich führe meine schlechteren Schulleistungen darauf zurück, daß die Lehrer die Sache nicht gut erklären können.“	.718	1	.416 - .517	
13	sks13	Elterliche Leistungsorientierung	8	0 = stimmt nicht 1 = stimmt	„Meine Eltern meinen, daß ich besser sein soll als die anderen.“	.663	2	.265 - .425	keine Verbesserung durch Reduzierung mit Skalen 14 und 15 möglich
14	sks14	Elterliche Förderungsorientierung	5	0 = stimmt nicht 1 = stimmt	„Meine Eltern versuchen mich zu verstehen, wenn ich mal etwas falsch gemacht habe.“	.820	1	.567 - .669	
15	sks15	Elterliche Sanktionsorientierung	12	0 = stimmt nicht 1 = stimmt	„Meine Eltern schimpfen mit mir, wenn ich etwas angestellt habe.“	.797	2	.331 - .553	
15r	sks15r	Elterliche Sanktionsorientierung – reduziert	10			.798	1	.423 - .578	
16	sks16	Elterliche Zuwendungsintensität	10	0 = stimmt nicht 1 = stimmt	„Meine Eltern hören aufmerksam zu, wenn ich etwas erzähle.“	.788	2	.227 - .567	
16r	sks16r	Elterliche Zuwendungsintensität – reduziert	9			.799	1	.408 - .573	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
17	sks17	Elterliches Verhalten bei guten Noten	6	1 = nie 2 = manchmal 3 = häufig	Was machen Deine Eltern, wenn Du gute Noten nach Hause bringst? „Sie loben mich.“	.505	2	-.071 - .479	
	sks17r	Elterliches Verhalten bei guten Noten – reduziert	4			.773	1	.436 - .671	
18	sks18	Elterliches Verhalten bei schlechten Noten	9	1 = nie 2 = manchmal 3 = häufig	siehe Skala 18.1 und 18.2	.559	2	-.061 - .431	
	18.1	Elterliches Verhalten bei schlechten Noten – Sanktionierung	6		Was machen Deine Eltern, wenn Du schlechte Noten nach Hause bringst? „Sie machen mir Vorwürfe.“	.811	1	.440 - .635	
18.2*	sks18_2	Elterliches Verhalten bei schlechten Noten – positive Unterstützung	4		Was machen Deine Eltern, wenn Du schlechte Noten nach Hause bringst? „Sie versuchen, mir zu helfen.“	.751	1	.476 - .646	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
19	sks19	Elterliche Hausaufgabenunterstützung – Informiertheit	9	1 = Der Satz trifft überhaupt nicht zu 2 = Der Satz trifft eher nicht zu 3 = Der Satz trifft teilweise zu, also teils, teils 4 = Der Satz trifft in etwa zu 5 = Der Satz trifft genau zu	Wie gut wissen Deine Eltern über die Schule, über das, was Du in der Schule tust und über Deine Hausaufgaben Be-scheid?	.827	2	.404 - .683	
			6		„Sie interessieren sich für alles, was wir in der Schule tun.“	.833	1	.507 - .691	
20	sks20	Elterliche Hausaufgabenunterstützung - Konstruktive Hilfe	6	1 = Der Satz trifft überhaupt nicht zu 2 = Der Satz trifft eher nicht zu 3 = Der Satz trifft teilweise zu, also teils, teils 4 = Der Satz trifft in etwa zu 5 = Der Satz trifft genau zu	Wie gut wissen Deine Eltern über die Schule, über das, was Du in der Schule tust und über Deine Hausaufgaben Be-scheid? „Ich kann immer zu ihnen kommen um zu fragen, wie etwas geht.“	.824	1	.449 - .684	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
21	sk21	Elterliche Hausaufgabenunterstützung – Negativ sanktionierendes Antreiben	7		siehe Skala 21.1 und 21.2	.399	2	.029 - .249	
21.1	sk21_1	Elterliche Hausaufgabenunterstützung – Negativ sanktionierendes Antreiben – Kontrolle und Druck	4	1 = Der Satz trifft überhaupt nicht zu 2 = Der Satz trifft eher nicht zu 3 = Der Satz trifft teilweise zu, also teils, teils 4 = Der Satz trifft in etwa zu 5 = Der Satz trifft genau zu	Wie verhalten sich Deine Eltern bei Deinen Hausaufgaben? „Sie hacken dauernd auf mir herum, wenn ich die Hausaufgaben mache.“	.743	1	.526 - .594	wurde nur in Nordrhein-Westfalen und Hessen 1978 erhoben, nicht in Niedersachsen 1979
21.2*	sk21_2	Elterliche Hausaufgabenunterstützung – Negativ sanktionierendes Antreiben – Erwartung von Selbstständigkeit	4		Wie verhalten sich Deine Eltern bei Deinen Hausaufgaben? „Sie meinen, daß ich alt genug bin, die Hausaufgaben selbstständig zu erledigen.“	.707	1	.443 - .555	

Tabelle 6.4: Skalen der Drei-Länder-Studie für Schülerinnen und Schüler, originale und neue Fassungen – 2. Teil: Unterrichtsskalen

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Englischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
22	sks22	Unterstützung (Englisch- bzw. Mathematiklehrer)	15	1 = nie 2 = selten 3 = manchmal 4 = häufig 5 = sehr häufig	„Wenn ich bei einer Aufgabe nicht weiter weiß, hilft mir mein Lehrer.“	.933	1	.581 - .745	.931	1	.590 - .761	siehe Kapitel 7.2.2
23	sks23	Strenge (Englisch- bzw. Mathematiklehrer)	15	1 = nie 2 = selten 3 = manchmal 4 = häufig 5 = sehr häufig	„Wenn ich meinem Lehrer widerspreche, wird er ärgerlich.“	.934	2	.514 - .733	.931	2	.563 - .750	siehe Kapitel 7.2.2
						.931	1	.525 - .750	.930	1	.575 - .751	
23r	sks23r	(Englisch- bzw. Mathematiklehrer) – reduziert	14									
24	sks24	Voraussetzungsbezogene Anforderungen (Englisch- bzw. Mathematiklehrer)	6	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer gibt uns meist nur solche Aufgaben, die wir auch lösen können. ... gibt uns häufig solche Aufgaben, die wir nicht lösen können.“	.793	1	.458 - .599	.784	1	.421 - .605	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Englischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
25	sks25	Strukturiertheit der Präsentation (Englisch- bzw. Mathematiklehrer)	10	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer fasst häufig noch einmal den Stoff zusammen, damit wir ihn uns gut merken können. ... fasst nur sehr selten den Stoff zusammen.“	.876	1	.403 - .716	.855	1	.382 - .684	durch Reduzierung um ein Item wird die niedrigste Trennschärfe der Items für den Englisch- und den Mathematikunterricht leicht verbessert; siehe Kapitel 7.2.2
			9			.878	1	.479 - .713	.855	1	.384 - .691	
26	sks26	Beschäftigungsradius (Englisch- bzw. Mathematiklehrer)	5	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer ruft auch häufig Schüler auf, die sich nicht gemeldet haben ... ruft selten Schüler, die sich nicht gemeldet haben.“	.725	2	.426 - .579	.717	2	.412 - .578	zweiter Faktor entsteht durch umgepolte Items, wird einfaktoriell verwendet; siehe Kapitel 7.2.2

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Englischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
27	sks27	Gezielte Lernhilfen (Englisch- bzw. Mathematiklehrer)	4	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer sagt uns häufig, wie wir uns verbessern können ... sagt uns selten, wie wir uns verbessern können.“	.745	1	.495 - .560	.716	1	.456 - .555	
24_27	sks24_27	Adaptivität des Unterrichts (Englisch- bzw. Mathematiklehrer)	9	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	siehe Skala 24 und 27	.854	1	.465 - .668	.834	1	.427 - .626	Kombination aus Skala 24 „Voraussetzungsbezogene Anforderungen“ und 27 „Gezielte Lernhilfen“; siehe Kapitel 7.2.2

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Englischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
28	sks28	Positive Schülerorientiertheit/Lehrer-Schüler-Beziehung (Englisch- bzw. Mathe-matiklehrer) (original: Positive Schülerorientiertheit)	11	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer ist stets geduldig. ... ist immer ungeduldig.“	.912	1	.559 - .780	.910	2	.584 - .735	zweiter Faktor entsteht durch umgepolte Items; wird einfaktoriell verwendet; siehe Kapitel 7.2.2
29	sks29	Individualisierungsmaßnahmen/Förderungsorientiertheit (Englisch- bzw. Mathe-matiklehrer)	5	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer bemüht sich auch um die schwächeren Schüler. ... kümmert sich nur um die guten Schüler.“	.801	1	.340 - .699	.784	1	.359 - .684	Item mit niedriger Trennschärfe (<0.4) wurde herausgenommen; siehe Kapitel 7.2.2
			4			.834	1	.589 - .717	.809	1	.534 - .699	
29r	sks29r	Individualisierungsmaßnahmen/Förderungsorientiertheit – reduziert										

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Englischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
30	sks30	Monitoring (Englisch- bzw. Mathematiklehrer)	5	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer merkt sofort, wenn ein Schüler im Unterricht nicht aufpaßt. ... merkt selten, wenn ein Schüler im Unterricht nicht aufpaßt.“	.638	1	.339 - .445	.617	1	.308 - .471	keine Verbesserung durch Reduzierung möglich
31	sks31	Disziplin und Ordnungssinn (Englisch- bzw. Mathematiklehrer)	5	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Englisch-/Mathematiklehrer achtet immer genau auf Ruhe und Ordnung. ... läßt die Schüler machen, was sie wollen.“	.767	1	.412 - .618	.799	1	.480 - .635	siehe Kapitel 7.2.2

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Englischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
32	sks32	Motivierungs- bemühen (Englisch- bzw. Mathe- matiklehrer)	4	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere	„Unser Eng- lisch-/Mathe- matiklehrer sagt uns häufig, daß wir uns verbes- sert hätten. Ich ... geht nie darauf ein, wenn wir eine gute Lei- stung erbracht haben.“	.753	1	.412 - .643	.707	1	.327 - .595	Item mit niedriger Trennschärfe (<0.4) für den Mathematik- unterricht wurde her- ausgenom- men; analog auch für den Englischun- terricht; siehe Kapitel 7.2.2
			3	3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig		.763	1	.567 - .635	.741	1	.511 - .613	
33	sks33	Vertrauens- grad/Schüler- orientierung (Englisch- bzw. Mathe- matiklehrer) <i>(original: Vertrauens- grad)</i>	4	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Unser Eng- lisch-/Mathe- matiklehrer beachtet auch die Meinung der Schüler. Ich ... bestimmt immer allein, was zu machen ist.“	.742	1	.493 - .609	.728	1	.453 - .584	siehe Kapitel 7.2.2

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Engischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
34	sks34	Angebotsnutzung (Englisch- bzw. Mathematikunterricht)	7	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Ich selbst melde mich häufig im Englisch-/Mathematikunterricht. I ... melde mich selten im Englisch-/Mathematikunterricht.“	.787	1	.340 - .647	.757	1	.377 - .610	Item mit niedriger Trennschärfe (<0.4) wurde herausgenommen
	sks34r	Angebotsnutzung (Englisch- bzw. Mathematikunterricht) – reduziert	6			.789	1	.460 - .644	.747	1	.403 - .598	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Engischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
35	sk35	Wohlfühlen im Unterricht (Englisch bzw. Mathematik)	6	1 = Der untere Satz ist genau richtig 2 = Der untere Satz ist eher richtig als der obere 3 = Meine Meinung liegt zwischen beiden Sätzen 4 = Der obere Satz ist eher richtig als der untere 5 = Der obere Satz ist genau richtig	„Ich selbst fühle mich in Englisch/Mathematik meistens sehr wohl. ... fühle mich in Englisch/Mathematik sehr unwohl.“	.845	1	.495 - .763	.834	1	.428 - .758	zwei Items wurden aus inhaltlichen Gründen entfernt bei gleichbleibender Stabilität der Skala; siehe Kapitel 7.2.3
	sk35r	Wohlfühlen im Unterricht (Englisch bzw. Mathematik) – reduziert	4			.835	1	.576 - .755	.833	1	.575 - .782	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Itemkodierung	Beispielitem	Engischunterricht			Mathematikunterricht			Anmerkung
						Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	
36	sk36	Zeiter-schwendung/ Time on task (Englisch bzw. Mathematik) (original: Zeitverschwen-dung)	11	1 = Der Satz ist völlig falsch 2 = Der Satz ist eher falsch als richtig 3 = Ich bin unentschieden bzw. weiß es nicht 4 = Der Satz ist eher richtig als falsch 5 = Der Satz trifft genau zu	„Bei uns dauert es zu Beginn der Stunde lange, bis die Schüler ruhig werden und zu arbeiten beginnen.“	.889	2	.469 - .740	.878	2	.409 - .722	siehe Kapitel 7.2.2
36r	sk36r	Zeiter-schwendung/ Time on task (Englisch bzw. Mathematik) – reduziert	10			.887	1	.463 - .744	.871	1	.406 - .721	

Legende zu den Skalenübersichten in Tabelle 6.3 und Tabelle 6.4

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
X	sksX	Originale Skala in Fett (<i>optionaler kursiver Originalname</i>)	Anzahl	Itemkodierung mit Antwortmöglichkeiten	„Text eines Beispielitems“	[0;1]	1/2/3	.321 - .567	Erläuterungen; siehe Kap. Y
Xr	sksXr	Nicht fett: Name einer neu gebildeten Teilskala	Anzahl			[0;1]	1	.456 - .654	

Erläuterungen

- **Nr.:** Die Skalen sind in der Reihenfolge der unveröffentlichten Skalenhandbücher der Drei-Länder-Studie durchnummeriert (für Schülerinnen und Schüler: Krüger-Haenisch, o. J.). Die originale Skala ist immer fett gedruckt. Aufteilungen werden durch X.1, X.2 gekennzeichnet; Zusammenführungen durch einen Unterstrich, z.B. kombiniert X_Y die Skalen X und Y. Wenn die originale Skala um ein oder mehrere Items reduziert wurde, ist dies durch ein angehängtes „r“ gekennzeichnet (Xr). Ein * bedeutet, dass die Skala ein oder mehrere Items enthält, die nicht in der originalen Skala (des Skalenhandbuchs) enthalten waren, aber im Datensatz der Skala eindeutig zuzuordnen waren.
- **Name im Datensatz:** „sks“ bezeichnet Skalen der Schülerinnen und Schüler.
- **Name der Skala:** Die originale Skala aus dem Skalenhandbuch ist immer fett gedruckt. Falls der originalen Namen aus Gründen der Verständlichkeit geändert wurde, steht der originale Name kursiv in Klammern darunter.
- **Antwortmöglichkeiten:** Ein hoher Zahlenwert der Itemkodierung entspricht einer hohen Ausprägung auf der Skala.
- **Beispielitem:** Falls das Item umgepolt werden musste, damit ein hoher Zahlenwert bei der Itemkodierung einer hohen Ausprägung auf der Skala entspricht, ist dies durch den Nachtrag „(umgepolt)“ gekennzeichnet.
- **Cronbachs α :** Die interne Konsistenz sollte den Wert von .7 überschreiten. Ab einem Wert von .8 geht man von einer hohen internen Konsistenz aus (Rammstedt, 2010, S. 249).
- **Anzahl Faktoren:** Eindimensionalität ist anzustreben, das heißt die Anzahl „1“ ist hier optimal.
- **Itemtrennschärfe Min. - Max.:** Enthält den niedrigsten und den höchsten Wert der Korrelationen der Items der Skala mit allen anderen Items der Skala. Unter .4 kann man von einer niedrigen Trennschärfe sprechen, .3 sollte nicht unterschritten werden (Bortz & Döring, 2006, S. 220).
- **Anmerkung:** Nähere Erläuterungen zu Veränderungen an der Skala und Verweis auf ausführliche Beschreibung in folgenden Kapiteln.

Tabelle 6.5: Skalen der Drei-Länder-Studie für Lehrkräfte, originale und neue Fassungen

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
1	skl1	Berufsenge- men/Arbeits- zufriedenheit (original: <i>Berufs- engagement</i>)	5	1 = falsch 2 = eher falsch 3 = weder richtig noch falsch 4 = eher richtig 5 = richtig	„Ich gehe völlig in meiner Arbeit auf.“	.657	2	.351 - .510	keine Verbesserung durch Reduzierung möglich
2	skl2	Lehrer-Lehrer- Verhältnis in Bezug auf Ko- operation	6	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„In den letzten drei Monaten habe ich mehrfach mit Kollegen Unterrichtsentwürfe und Materialien ausgetauscht.“	.780	2	.167 - .684	
2r	skl2r	Lehrer-Lehrer- Verhältnis in Bezug auf Koope- ration – reduziert	5			.816	1	.523 - .676	
3	skl3	Leistungs- druck/hohe Leistungserwar- tung an die Schülerinnen und Schüler (original: <i>Leis- tungsdruck</i>)	4	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„An dieser Schule wird viel von den Schülern verlangt.“	.676	1	.411 - .559	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
4	skl4	Schülerzentriertheit/Partizipation und Bullying (original: Schülerzentriertheit)	5	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„In Lehrergesprächen werden häufig abfällige Bemerkungen über einzelne Schüler gemacht.“ (umgepolt)	.772	1	.511 - .587	
5	skl5	Förderung/Förderorientierung (original: Förderung)	3	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„Die meisten Lehrer bemühen sich, daß auch die schwächeren Schüler mitkommen.“	.696	1	.415 - .577	
6	skl6	Stellenwert von Disziplin in Schule/Kollegium (original: Disziplindruck)	4	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„Auf Fragen der Disziplin wird an dieser Schule sehr wenig Wert gelegt.“ (umgepolt)	.682	1	.412 - .533	
7	skl7	Konfliktlösung/Schulpartizipation (original: Konfliktlösung)	3	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„In den Konferenzen werden die Meinungen der Schülervertreter von den meisten Kollegen ernst genommen.“	.720	1	.504 - .600	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
4_7	skl4_7	Schülerzeugewandtheit (pädagogischer Ethos)	8	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	siehe Skala 4 und 7	.856	1	.520 - .666	Kombination aus Skala 4 „Schülerzeugewandtheit“ und 7 „Konfliktlösung“
8	skl8	Selbstverantwortung	2	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	„An unserer Schule geht häufig alles drunter und drüber.“	.762	1	.620	
6_8	skl6_8	Chaos und Verantwortungslosigkeit an der Schule	5	1 = völlig falsch 2 = eher falsch als richtig 3 = unentschieden, weiß nicht 4 = eher richtig als falsch 5 = stimmt genau	siehe Skala 6 (nicht umgepolt) und 8	.815	1	.483 - .723	Kombination aus Skala 6 „Stellenwert von Disziplin“ und 8 „Selbstverantwortung“
9	skl9	Reformbereitschaft	8	1 = Ich finde das sehr falsch 2 = Ich finde das falsch 3 = Ich finde das eher falsch als richtig 4 = Ich finde das eher richtig als falsch 5 = Ich finde das richtig 6 = Ich finde das sehr richtig	„Vor den Anforderungen der modernen Welt müssen die bislang gewohnten Schulformen und Unterrichtsmethoden einfach versagen.“	.807	2	.354 - .599	siehe Kapitel 7.2.1
9r	skl9r	Reformbereitschaft – reduziert	6			.807	1	.503 - .624	

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
10	sk10 10r	Notwendigkeit von Druck bei Lernprozessen (original: <i>Druckmotivierung</i>)	8	1 = Ich finde das sehr falsch 2 = Ich finde das falsch 3 = Ich finde das eher falsch als richtig 4 = Ich finde das eher richtig als falsch 5 = Ich finde das richtig 6 = Ich finde das sehr richtig	„Wenn man dem Lehrer die Zensur als Kontrollmittel nimmt, werden die meisten Schüler bald nichts mehr tun.“	.793	2	.302 - .602	siehe Kapitel 7.2.1
			6			.813	1	.519 - .609	
11	sk11 11r	Einstellung zur Förderung von schwachen Schülerinnen und Schülern (original: <i>Förderintention</i>)	8	1 = stimmt gar nicht 2 3 4 5 = stimmt völlig	„Die ständige Rücksichtnahme auf die schwachen Schüler beeinträchtigt erheblich meinen Unterricht.“ (umgepolt)	.662	2	.093 - .456	siehe Kapitel 7.2.1
			6			.691	1	.354 - .520	
12	sk12	Medienverwendung	18	1 = nie 2 = selten 3 = manchmal 4 = häufig 5 = sehr häufig	Wie häufig setzen Sie im Unterricht die folgenden Hilfsmittel ein? „vom Lehrer angefertigte Arbeitsblätter“	.813	5	-.128 - .602	ergibt keine Skala

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung
13	sk113	Fortbildung mittels Rezeption von Fachzeitschriften (original: Fortbildungssaktivität)	2	1 = nie 2 = selten 3 = gelegentlich 4 = regelmäßig	„Wie oft lesen Sie Fachzeitschriften, die in Zusammenhang mit Ihren Unterrichtsfächern stehen?“	.636	1	.468	
14	sk114	Fähigkeit der Englischlehrer in Bezug auf Sprachbeherrschung (original: Fähigkeit der Englischlehrer)	5	1 = weniger gut 2 3 4 5 = sehr gut	Geben Sie bitte an, wie Sie Ihre Fähigkeiten in der englischen Sprache einschätzen. „Ausprache“	.884	1	.670 - .777	
15	sk115	Disziplinierungsmaßnahmen	6	1 = nie 2 = selten 3 = gelegentlich 4 = häufig	Wie häufig wenden Sie die folgenden Maßnahmen im Unterricht in der hier untersuchten Klasse an? „Ermahnung“	.596	2	.282 - .417	keine Verbesserung durch Reduzierung möglich
16	sk116	Lobformen	6	1 = nie 2 = selten 3 = manchmal 4 = häufig 5 = sehr häufig	Wenn Sie mit einzelnen Schülern oder der ganzen Klasse zufrieden sind, welche Maßnahmen ergreifen Sie? „Verbales Lob eines einzelnen Schülers“	.516	2	.212 - .324	keine Verbesserung durch Reduzierung möglich

Legende zu der Skalenübersicht in Tabelle 6.5

Nr.	Name im Datensatz	Name der Skala	Anzahl Items	Antwortmöglichkeiten	Beispielitem	Cronbachs α	Anzahl Faktoren	Itemtrennschärfe Min. - Max.	Anmerkung		
X	skIX	Originale Skala in Fett <i>(optionaler kursiver Originalname)</i>	Anzahl	Itemkodierung mit Antwortmöglichkeiten	„Text eines Beispielitems“	[0;1]	1/2/3	.321 - .567	Erläuterungen; siehe Kap. Y		
Xr	skIXr	Nicht fett: Name einer neu gebildeten Teilskala	Anzahl			[0;1]	1	.456 - .654			
<u>Erläuterungen</u>											
-	Nr.: Die Skalen sind in der Reihenfolge der unveröffentlichten Skalenhandbücher der Drei-Länder-Studie durchnummeriert. Die originale Skala ist immer fett gedruckt. Aufteilungen werden durch X.1, X.2 gekennzeichnet; Zusammenführungen durch einen Unterstrich, z.B. kombiniert X_Y die Skalen X und Y. Wenn die originale Skala um ein oder mehrere Items reduziert wurde, ist das durch ein angehängtes „r“ gekennzeichnet (Xr). Ein * bedeutet, dass die Skala ein oder mehrere Items enthält, die nicht in der originalen Skala (des Skalenhandbuchs) enthalten waren, aber im Datensatz der Skala eindeutig zuzuordnen waren.										
-	Name im Datensatz: „skI“ bezeichnet Skalen der Lehrkräfte.										
-	Name der Skala: Die originale Skala aus dem Skalenhandbuch ist immer fett gedruckt. Falls der originalen Namen aus Gründen der Verständlichkeit geändert wurde, steht der originale Name kursiv in Klammern darunter.										
-	Antwortmöglichkeiten: Ein hoher Zahlenwert der Itemkodierung entspricht einer hohen Ausprägung auf der Skala.										
-	Beispielitem: Falls das Item umgepolt werden musste, damit ein hoher Zahlenwert bei der Itemkodierung einer hohen Ausprägung auf der Skala entspricht, ist dies durch den Nachtrag „(umgepolt)“ gekennzeichnet.										
-	Cronbachs α: Die interne Konsistenz sollte den Wert von .7 überschreiten. Ab einem Wert von .8 geht man von einer hohen internen Konsistenz aus (Rammstedt, 2010, S. 249).										
-	Anzahl Faktoren: Eindimensionalität ist anzustreben, das heißt die Anzahl „1“ ist hier optimal.										
-	Itemtrennschärfe Min. - Max.: Enthält den niedrigsten und den höchsten Wert der Korrelationen der Items der Skala mit allen anderen Items der Skala. Unter .4 kann man von einer niedrigen Trennschärfe sprechen, .3 sollte nicht unterschritten werden (Bortz & Döring, 2006, S. 220).										
-	Anmerkung: Nähere Erläuterungen zu Veränderungen an der Skala und Verweis auf ausführliche Beschreibung in folgenden Kapiteln.										

6.4 Testskalierung

Bei einer erneuten Erhebung mit den Instrumenten der Drei-Länder-Studie sollten auch die Leistungstests für Schülerinnen und Schüler eingesetzt werden, deshalb wurden diese im Rahmen dieser Arbeit erneut skaliert. Bei der Auswahl der geeigneten Testitems kamen inhaltliche, statistische und ökonomische Kriterien zum Tragen. Ziel der Neuskalierung der Leistungstest war es, mit einer kleineren Itembatterie einen ähnlich guten Test wie mit allen bisherigen Testitems zu konstruieren, damit bei einer möglichen Neuerhebung die vorliegenden Tests erneut zum Einsatz kommen könnten, aber weniger Testzeit benötigten. Dafür war es von großer Bedeutung, bei der Itemauswahl für die Neuzusammenstellung der Tests neben statistischen (siehe Kapitel 6.4.6) auch inhaltliche Kriterien einzubeziehen. Für die Fächer Englisch und Mathematik wurden daher fachdidaktische Experten und erfahrene Lehrpersonen der hessischen Lehrkräfteakademie zu Rate gezogen, um die curriculare Validität aus heutiger Sicht zu prüfen (siehe Kapitel 6.4.5).

6.4.1 IRT-Skalierung

Die Testskalierungen basieren auf der Item-Response-Theorie (IRT) (vgl. Aya-la, 2009; van der Linden & Hambleton, 1997), wobei das beobachtete Antwortverhalten auf verschiedene Testitems (manifeste Variablen) als abhängig von den nicht beobachteten dahinterliegenden Fähigkeiten (latente Variablen) angesehen wird (Moosbrugger, 2012, S. 228). In der IRT werden die Personenfähigkeiten und die Itemschwierigkeiten auf einer gemeinsamen Skala abgetragen. Die IRT ist besonders dann geeignet, wenn vor allem inhaltlich-theoretische Fragen der Konstruktvalidität entscheidend sind (Rauch & Hartig, 2012, S. 269) und wenn es um die Frage geht, ob und welche der Items einer Skala zur Messung einer Fähigkeit geeignet sind (Geiser & Eid, 2010, S. 312).

Es wurden eindimensionale Rasch-Modelle (auch: 1PL-Modell) geschätzt. Eine der Eigenschaften des Rasch-Modells ist es, dass man die Itemparameter schätzen kann, ohne die Personenparameter zu kennen und keine Verteilungsannahmen bezüglich der Personenfähigkeiten getroffen werden müssen (Rost, 1996, S. 133). Das Rasch-Modell geht davon aus, dass alle Items eines Tests dieselbe latente Dimension (hier: die Fähigkeit) mit unterschiedlichen Itemschwierigkeiten, jedoch identischer Trennschärfe erfassen (Geiser & Eid,

2010, S. 311). Moosbrugger (2012, S. 236) bezeichnet das Rasch-Modell auch als „das einfachste und zugleich vorteilhafteste Modell“ der IRT-Modelle.

Alle nicht bearbeiteten Aufgaben („non-response“) wurden bei der Skalierung als falsch gewertet, da davon ausgegangen wird, dass eine Person ein Item dann nicht beantwortet hatte, wenn es zu schwer war und die Person keine Antwort wusste (Rost, 2004, S. 325). Bei den nicht bearbeiteten Aufgaben ließ sich kein systematisches Muster erkennen, etwa dass die Items gegen Ende der Tests mehr fehlende Werte aufwiesen („not reached“). Die nicht bearbeiteten Aufgaben sind über den kompletten Test verteilt und die meisten Teilnehmenden haben die Tests vollständig ausgefüllt.

Für jede Person wurden Weighted Likelihood Estimates (WLEs) als Personenfähigkeitsparameter (Warm, 1989) berechnet. Ein Vorteil der WLEs ist, dass auch Personenparameter für Personen bestimmt werden können, die alle Items falsch oder alle Items richtig beantwortet haben (Moosbrugger, 2012, S. 244). Die IRT-Skalierungen erfolgten in R 3.4.2 (R Core Team, 2017) mit dem TAM-Paket (Robitzsch, Kiefer & Wu, 2017).

6.4.2 Beschreibung der Tests

Im Folgenden werden kurz die Inhaltsbereiche der Fachleistungstests und des kognitiven Fähigkeitstests dargestellt. Für weitere Informationen sei auf die angegebene Testliteratur und die Berichte der Drei-Länder-Studie, vor allem Fend (1982, S. 203) und Lukesch (1979a, 1979b) verwiesen; ausführliche Aufgabenbeschreibungen finden sich bei Fend (1981).

Mathematik: Der Mathematiktest war eine Eigenkonstruktion der Projektgruppe von Helmut Fend, da „kein handelsüblicher Test vor[lag], um genügend inhaltliche Bereiche mit variablen Schwierigkeiten zu erfassen“ (Fend, 1982, S. 203). Grundlage der Eigenkonstruktion waren die Mathematiktests der IEA-Studien (Husén, 1967), der MDA 10+ (Bartel, Hylla & Süllwold, 1965) und der Test für Zentrische Streckung, entwickelt von der Gruppe Leistungsmessung in Gesamtschulen (LiG, 1974, zitiert nach Fend, 1982, S. 203). Der Mathematiktest behandelt die Bereiche Algebra (vor allem Gleichungen und Ungleichungen), Geometrie, Zentrische Streckung, Prozentrechnung, Schlussrechnung (Dreisatz) und Dezimalbruchrechnung. Zu den meisten Items gibt es fünf, zu einigen sechs und zu einem Item gibt es acht Antwortmöglichkeiten.

Englisch: Der Englischtest bestand aus zwei Testverfahren: dem EnglischEinstufungstest EET9+, entwickelt von der Gruppe Leistungsmessung in Gesamtschulen (1974), und einem Teil des IEA-Tests, entwickelt von der Gruppe „International Evaluation of Achievement“. Der EET9+ soll möglichst umfassend die Englischkenntnisse von Schülerinnen und Schülern der 9. Jahrgangsstufe messen, wobei es für jedes Item vier Antwortmöglichkeiten gibt. Der Test umfasst die Bereiche Wortschatz (vor allem Strukturwörter) und Muster der mündlichen und schriftlichen Sprachausübung. Der IEA-Test prüft in Subtests (Vokabelkenntnisse, Aussprache, Grammatik, Wortbedeutung, Satzbedeutung, Leseverständnis) den Kenntnisstand von 14- bis 15-Jährigen, wobei jeweils drei, vier oder fünf Antwortmöglichkeiten gegeben sind.

Kognitive Fähigkeit: Die allgemeine kognitive Fähigkeit der Schülerinnen und Schüler wurde mittels des figuralen Grundintelligenztests CFT 2 (Cattell & Weiß, 1972) erfasst. Durch die sprachfreie Erfassung – die Problemstellungen sind zeichnerisch dargestellt – wird versucht, das allgemeine intellektuelle Problemlöseverhalten möglichst „milieuunabhängig“ (Weiß, 1972, S. 5) zu erfassen. Der erste Testteil des CFT 2 wurde als Speed-Test eingesetzt und enthält vier Subtests (Reihenfortsetzen, Klassifikationen, Matrizen und topologische Schlussfolgerungen), wobei jeweils fünf Antwortmöglichkeiten vorgegeben sind.

Deutsch: Für das Fach Deutsch kamen ein Test zum Leseverständnis und ein Test zum Rechtschreiben zum Einsatz. Es wurden der Test „Verständiges Lesen – VL 7-9“ (Anger, Bargmann & Voigt, 1971) und der Mannheimer Rechtschreibtest „MRT“ (Jäger & Jundt, 1973) verwendet. Der Leseverständnistest besteht aus sechs verschiedenen Lesetexten mit dazugehörigen Fragen, wobei jeweils vier Antwortmöglichkeiten vorgegeben sind. Der Rechtschreibtest enthält verschiedene Bereiche der Rechtschreibung, nämlich Dehnung, Kürzung, Konsonanten- und Vokalverwechslung, Groß- und Kleinschreibung und Fremdwörter. Auch hier musste die richtige von vier Antwortmöglichkeiten angekreuzt werden.

Naturwissenschaften: Der Test für die Bereiche Physik und Biologie wurde aus einer Auswahl des IEA-Tests zu Naturwissenschaften (für Physik: Comber & Keeves, 1973; für Biologie: Schultze, 1974) zusammengestellt. Für jedes Item stehen fünf Antwortmöglichkeiten zur Auswahl. Der Test für Naturwissenschaften wurde im Zuge dieser Arbeit nicht skaliert, da daran kein Forschungs-

interesse bestand. Darüber hinaus wurde in der seinerzeitigen Auswertung der Drei-Länder-Studie nur für den Teilbereich der Physik (22 Items) eine Skala gebildet, für den Bereich der Biologie (19 Items) hatte sich nach einer Itemanalyse keine Skala ergeben (Fend, 1981, S. 31; Lukesch, 1979a, S. 236).

6.4.3 Vorbereitungen für die Skalierung

Als Vorbereitung für die Skalierung mussten die Testitems so kodiert werden, dass eine binäre Kodierung mit „1“ für die richtige und „0“ für falsche Lösungen vorlag. Die korrekte Lösung der Testaufgaben konnte Lösungsschablonen entnommen werden, wenn bekannt war, aus welchem Test das Item entnommen war und ein Original-Testheft dieses Tests aus einer Bibliothek beschafft werden konnte. Aus den Beständen der Drei-Länder-Studie lagen keine Lösungsschablonen für die Tests vor. Für die Fächer Mathematik und Englisch konnte nicht für alle Aufgaben eine Lösungsschablone gefunden werden. Für das Fach Mathematik lagen nur für den Test MDA 10+ Lösungsschablonen vor. Daher konnten bei 19 Aufgaben von insgesamt 43 keine Lösungsschablonen zugeordnet werden. Bei diesen Aufgaben musste die richtige Lösung eigenständig berechnet werden. Für das Fach Englisch konnten nur bei einem Testheft (basierend auf dem IEA-Test) Lösungsschablonen zugeordnet werden. Beim anderen Testheft (EET 9+) musste die richtige Lösung eigenständig bestimmt werden. Diese wurde von einer Muttersprachlerin zusätzlich überprüft.

Damit die Schülerinnen und Schüler der Drei-Länder-Studie beim Bearbeiten der Testaufgaben nicht voneinander abschreiben konnten, wurden die Tests jeweils in Parallelformen durchgeführt, das heißt eine Hälfte der Schülerinnen und Schüler bearbeitete die sog. A-Form und eine Hälfte die sog. B-Form. Diese Parallelformen lagen entweder schon vom Verlag aus vor oder wurden durch Umstellen der Seiten und/oder Umstellen der Distraktoren erstellt. Zur weiteren Vorbereitung der Testskalierung mussten folglich die Testitems der A- und B-Formen zusammengeführt werden. Da jedes Item in einer A- und einer B-Form im Datensatz vorhanden war, mussten die zusammengehörigen Items zusammengespielt werden. Die A- und die B-Form unterschieden sich überwiegend durch eine andere Antwortreihenfolge, bei einem Englischtestheft lag der Unterschied jedoch in der Reihenfolge von Fragen *und* Antworten. Darüber hinaus besteht der Leseverständnistest im Fach Deutsch in den Formen A und B aus komplett unterschiedlichen Aufgaben und wurde daher nach Parallelformen getrennt skaliert. Da an einer erneuten Testskalierung im Bereich der Na-

turwissenschaften kein Forschungsinteresse bestand, wurde hier auf das Zusammenführen der Items aus den A- und B-Formen verzichtet.

In der Drei-Länder-Studie wurden sowohl Gesamtskalen- (in Mathematik und Deutsch – Lesen) als auch Subskalen-Auswertungen (Englisch IEA-Test und Deutsch – Rechtschreibung) vorgenommen (Fend, 1981, 123–125, 129), für die Reskalierung wurde in jedem Fach eine eindimensionale Gesamtskala angestrebt. In den Fächern Englisch und Mathematik war die Bildung einer Gesamtskala über beide Testhefte hinweg möglich, da in allen Testheften der gleiche Fragentyp (jeweils nur eine richtige Antwortmöglichkeit pro Item) zum Einsatz kam. Für die Fachleistungen in Deutsch ließ sich eine übergreifende Skalierung wegen der unterschiedlichen Parallelformen des Leseverständnistests nicht durchführen.

6.4.4 Curriculare Validität aus damaliger Sicht

In der Drei-Länder-Studie wurden die teilnehmenden Lehrkräfte nach den Lerngelegenheiten bezogen auf alle Inhalte der eingesetzten Fachleistungstests befragt. Leistungsunterschiede zwischen den Schülerinnen und Schülern an den Gesamtschulen und an den Schulen des traditionellen Schulsystems sollten daraufhin geprüft werden, ob sie auf unterschiedliche Lerngelegenheiten zurückgeführt werden könnten. Die curricularen Angebote der im Test erfassten Leistungsbereiche wurde über Angaben der jeweiligen Fachlehrer zum Zeitpunkt, zur Häufigkeit und zur Vollständigkeit der Durchnahme sowie zu der ihnen im Unterricht zugemessenen Bedeutung, zu ihrem Anteil an der Unterrichtszeit und zu ihrem Anteil am Lehrstoff im jeweiligen Schuljahr erfasst (Fend, 1981, S. 99; Klaghofer & Krüger-Haenisch, 1979, S. 448). Darüber hinaus wurden die Lehrkräfte gebeten, den Anteil der Items, welche die Schülerinnen und Schüler vermutlich im Durchschnitt richtig lösen würden, zu schätzen. Zusätzlich zu den Lerngelegenheiten wurde auch das Ausmaß der Unterrichtsstunden pro Fach und Woche über Angaben der Schülerinnen und Schüler erhoben (Lukesch, 1979a, S. 255–257, 1979b, S. 131–133).

Die Beurteilung der curricularen Validität erfolgte je nach Test, auf Item-, Subtest- oder Gesamttestebene. Die Auswertung der curricularen Validität der Erhebungen 1978 in Nordrhein-Westfalen und Hessen wird ausführlich bei Klaghofer und Krüger-Haenisch (1979) beschrieben, die für Niedersachsen

1979 bei Fend (1981, S. 94–132). Ein Vergleich der curricularen Validität zwischen den drei Bundesländern findet sich bei Fend (1982, S. 201–213).

Neben der curricularen Validität bezogen auf das implementierte Curriculum (der tatsächlich durchgenommene Stoff im Unterricht) wurde für Nordrhein-Westfalen und Hessen auch die curriculare Validität bezogen auf das intendierte Curriculum (Lehrplan) geprüft (Klaghofer & Krüger-Haenisch, 1979 und beispielhaft anhand von Bruchrechnen im Fach Mathematik und den beiden Englisch-Testheften: Fend, 1982, S. 205–211). In Nordrhein-Westfalen und Hessen kam darüber hinaus eine dritte Möglichkeit der Erfassung der curricularen Validität zum Einsatz: die Analyse von Klassenarbeiten in den Fächern Englisch und Mathematik in den untersuchten Unterrichtsgruppen (Klaghofer & Krüger-Haenisch, 1979, S. 392–447).

6.4.5 Curriculare Validität aus heutiger Sicht

Damit die Leistungstests bei einer möglichen Neuerhebung erneut zum Einsatz kommen konnten, wurden von der Projektgruppe externe Expertinnen und Experten der hessischen Lehrkräfteakademie gebeten,⁴¹ die curriculare Validität der Englisch- und Mathematiktests aus heutiger Sicht (bezogen auf das intendierte Curriculum) zu überprüfen. Dabei war von Interesse, welche der Testaufgaben den heutigen Unterrichtsaufgaben entsprechen bzw. welche Testaufgaben nicht mehr curricular valide oder nicht mehr zeitgemäß sind. Darüber hinaus sollte eingeschätzt werden, welche Testaufgaben von den Schülerinnen und Schülern heutzutage gut beantwortet werden könnten bzw. welche Testaufgaben heutzutage nicht mehr lösbar wären. Die Expertinnen und Experten gingen bei ihrer Beurteilung von ihrer eigenen Unterrichtsrealität und ihren Erfahrungen mit Abschlussprüfungen aus, wobei der Fokus auf den mittleren Bildungsgängen lag.

Das Ergebnis dieser Überprüfung ergab für den Englischtest, dass die acht Fragen des Aufgabentyps Phonem-Graphem-Zuordnung heute keine Entsprechung mehr im Unterricht finden, höchstens noch an Gymnasien und auch dort nur in Zusammenhang mit Hörverstehen. Sieben weitere Aufgaben sind aus anderen inhaltlichen Gründen nur für das Gymnasium geeignet. Bei 16 Fragen wurde festgestellt, dass die Lösung nicht eindeutig ist oder zu wenig Kontext

41 Ein großer Dank geht an Herrn Michael Tetzlaff von der Hessischen Lehrkräfteakademie, der diesen Kontakt vermittelt hat.

zum Lösen der Aufgabe vorgegeben wurde. Insgesamt erwiesen sich 78 der 115 Aufgaben als inhaltlich geeignet für eine erneute Testung an Schülerinnen und Schülern.

Beim Mathematiktest ergab die Überprüfung, dass der innermathematische Anteil und die (fach-)sprachlichen Anforderungen der Aufgaben sehr hoch sind. Darüber hinaus wird den Kompetenzen „Mathematisch Argumentieren“ und „Mathematisch Modellieren“ in dem Test eine deutlich geringere Bedeutung beigemessen als heute. Die Aufgaben zu den Inhaltsbereichen Primzahlen und Ungleichungen sind heute eher im Gymnasialbereich angesiedelt. Vier Aufgaben decken den Themenbereich der Zentrischen Streckung ab, dieses Thema wird heute an Hauptschulen gar nicht mehr und an Realschulen nur in geringem Rahmen behandelt. Insgesamt wurde bei 21 Aufgaben festgestellt, dass sie eher das heutige Schwierigkeitsniveau von Gymnasien abdecken. Darüber hinaus haben der Kontext oder die Begrifflichkeiten der Aufgaben teilweise wenig mit den Alltagserfahrungen der heutigen Lernenden zu tun (z.B. „Bedienungszuschlag“ statt Trinkgeld). Dennoch sind insgesamt 23 von 43 Aufgaben vom Inhalt her für eine erneute Testung an Schülerinnen und Schülern geeignet. Bei fünf dieser Aufgaben wurde angemerkt, dass sie ohne die Verwendung eines Taschenrechners, wie sie in der 9. Klasse Jahrgangsstufe zumindest an Realschulen und Gymnasien heute üblich sei, schwierig zu lösen wären.

6.4.6 Die Auswahl der Testitems

Wie bereits zu Beginn von Kapitel 6.4 erwähnt, war das Ziel der Testskalierung die Neuzusammenstellung der Tests mit einer kleineren Itembatterie. Nachdem im vorherigen Abschnitt (6.4.5) die Auswahl der Items nach inhaltlichen Kriterien für die Fächer Englisch und Mathematik im Mittelpunkt stand, wird im Folgenden beschrieben, nach welchen statistischen Kriterien die Items für alle Fächer ausgewählt wurden. Mit Hilfe der Itemfits sollten diejenigen Items identifiziert werden, die keine gute Übereinstimmung mit dem Rasch-Modell aufweisen (Wu & Adams, 2007, S. 85). Bei den Fächern Englisch und Mathematik folgte die Auswahl der Items primär der fachdidaktischen Expertise. Im nächsten Schritt kamen dann die im Folgenden beschriebenen statistischen Kriterien zum Tragen.

Bei der IRT-Skalierung (siehe Abschnitt 6.4.1) wurden lokale Modellgeltungskontrollen durchgeführt, indem die Passung („fit“) der einzelnen Items zum Rasch-Modell anhand der Infits untersucht wurde (Ames & Penfield, 2015). Der Infit (auch „weighted fit mean-square“) ist ein residuenbasierter deskriptiver Kennwert bezogen auf die Passung eines Items zu den anderen Testitems, der bei Werten über 1 (hier: > 1.2) eine zu geringe Itemdiskrimination („Underfit“) und bei Werten unter 1 (hier: < 0.8) eine zu starke Itemdiskrimination („Overfit“) anzeigt⁴² (Linacre, 2002; Wu & Adams, 2007). Die Fit-Statistik testet, ob die Steigung der beobachteten itemcharakteristischen Funktion (ICC) mit der theoretischen ICC des Rasch-Modells übereinstimmt (Wu & Adams, 2007, S. 78). Tabelle 6.6 zeigt die Anzahl der Items für jeden Test, die außerhalb der hier verwendeten Schwellenwerte des Infits liegen. Dort wird deutlich, dass die Anzahl der Items mit zu geringer oder zu starker Itemdiskrimination durch die Neuskalierung (Zeile „Auswahl“) für die Leistungstests erheblich reduziert werden konnte.

Auch die Anzahl der Items mit einer Trennschärfe nach der klassischen Testtheorie von $< .3$ (Bortz & Döring, 2006, S. 220) konnte durch die Neuskalierung der Leistungstests reduziert werden (siehe Tabelle 6.6). Anhand der Spannweite der Lösungswahrscheinlichkeiten aller Items eines Tests wird deutlich, dass auch nach der Neuskalierung noch Items in den Tests enthalten sind, welche die Randbereiche der Itemschwierigkeiten (sehr leichte und sehr schwere Items) abdecken. Zur Reliabilitätsbestimmung des Modells wurden die WLE- und EAP-Reliabilität verwendet (Adams, 2005). Als ausreichende Reliabilität gilt ein Wert von $> .7$, der in allen Tests außer in Mathematik durchgehend erreicht werden konnte (siehe Rost, 2004, S. 382 zur Ähnlichkeit der Interpretation von EAP-Reliabilität und Cronbachs Alpha). Die Test-Versionen für das Fach Mathematik weisen eine zufriedenstellende interne Konsistenz auf (Cronbachs Alpha $> .7$; siehe Rammstedt, 2010, S. 249) und alle weiteren Tests eine hohe interne Konsistenz (Cronbachs Alpha $> .8$; siehe Werte in Tabelle 6.6).

42 Die Wahl der Schwellenwerte bei Itemfit-Statistiken ist rein konventionell: „Many textbooks or other resources make recommendations on the range of acceptable mean-square values (...) for residual based fit statistics. There are probably no right or wrong answers“ (Wu & Adams, 2007, S. 85). In großen Schulleistungsstudien haben sich die Schwellenwerte 0.8 und 1.2 durchgesetzt (siehe Wright & Linacre, 1994; Wu & Adams, 2007), die auch hier verwendet wurden.

Tabelle 6.6: Kennwerte der verschiedenen Versionen aller Tests der Drei-Länder-Studie aus der IRT-Skalierung

Fach	N	Test-Version	Item Anzahl	WLE-Reliabilität	EAP-Reliabilität	Cronbachs α	Spannweite Infit	Infit < 0.8 > 1.2	Spannweite TS	Anzahl Items: TS < .3	Spannweite LSW
Mathematik	3382	Original	43	.723	.692	.77	0.898 - 1.101	–	.030 - .493	21	.062 - .791
		Reduziert	19	.719	.701	.77	0.910 - 1.129	–	.242 - .557	1	.166 - .791
		Auswahl	24	.703	.684	.75	0.908 - 1.150	–	.176 - .531	5	.117 - .791
Englisch	3313	Original	115	.959	.872	.97	0.750 - 1.702	10	-.174 - .675	8	.182 - .896
		Auswahl	34	.860	.791	.88	0.795 - 1.478	3	-.112 - .637	3	.182 - .896
CFT	3150	Original	46	.847	.734	.85	0.796 - 1.145	1	.047 - .594	8	.071 - .901
Lesen Form A	1710	Original	48	.915	.924	.93	0.833 - 1.144	–	.378 - .583	–	.377 - .788
		Auswahl	42	.906	.918	.93	0.884 - 1.125	–	.406 - .595	–	.377 - .786
Lesen Form B	1648	Original	47	.901	.910	.92	0.862 - 1.147	–	.308 - .578	–	.303 - .773
		Auswahl	43	.895	.905	.91	0.881 - 1.137	–	.368 - .581	–	.366 - .747
Rechtschreibung	3225	Original	117	.946	.838	.95	0.823 - 1.342	2	.036 - .563	22	.137 - .908
		Reduziert	115	.946	.838	.95	0.823 - 1.209	2	.075 - .564	21	.137 - .908
		Auswahl	88	.934	.829	.94	0.873 - 1.167	–	.265 - .547	5	.260 - .908

Anmerkungen: TS = Trennschärfe; LSW = Lösungswahrscheinlichkeit

Kognitiver Fähigkeitstest

Der kognitive Fähigkeitstest (CFT) wurde zwar IRT-skaliert, allerdings wurden hier keine Items ausgeschlossen, da dies ein feststehender Test zur Messung der kognitiven Fähigkeit ist, während die Fachleistungstests häufig eine Zusammenstellung aus mehreren Quellen darstellen (siehe Abschnitt 6.4.2). Der CFT-Test weist eine hohe Reliabilität mit sehr guten Infit-Werten (zwischen 0.80 und 1.15) auf.

Mathematiktest

Der Mathematiktest liegt in mehreren Versionen vor: eine Skalierung basiert auf allen 43 Items, die im Test erhoben wurden, eine weitere Skalierung basiert auf den 19 Items, die bei den damaligen Berechnungen verwendet wurden, da „nach der testtheoretischen Analyse [...] nur 19 Aufgaben für die endgültige Berechnung beibehalten werden“ (Lukesch, 1979b, S. 17) konnten. Für die Auswahl von Items für eine erneute Skalierung waren allerdings alle 43 Items die Ausgangsbasis, da hierfür nicht nur statistische, sondern vor allem inhaltliche Fragen ausschlaggebend waren. Von den 43 ursprünglichen Items wurden 19 Items ausgeschlossen, da sie den inhaltlichen und teststatistischen Kriterien nicht genügten. Die verbliebenen 24 Items weisen eine akzeptable Reliabilität mit sehr guten Infit-Werten (zwischen 0.91 und 1.15) auf.

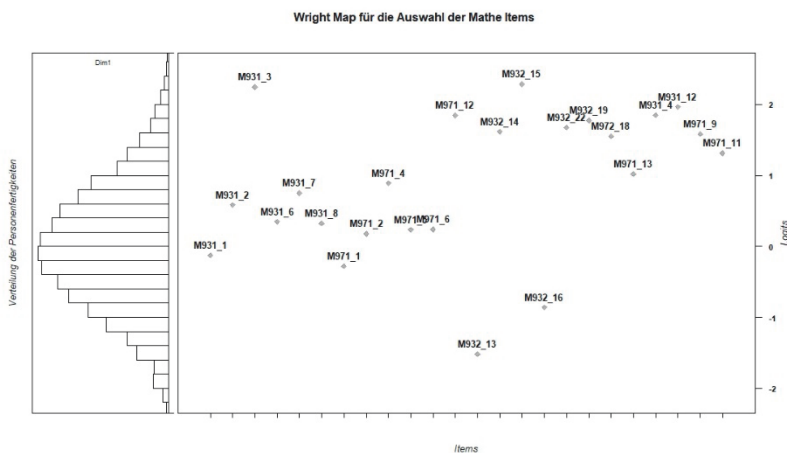


Abbildung 6.4: Wright Map für die 24 ausgewählten Items in Mathematik

Eine Übersicht über das Resultat der IRT-skalierten Testdaten der 24 ausgewählten Items für das Fach Mathematik findet sich in Abbildung 6.4. Mit Hilfe einer „Wright Map“ werden dort gleichzeitig die Item- und die Personenparameter dargestellt, also die Schwierigkeit des Items und die Fertigkeit der Personen. Viele der ausgewählten Items liegen eher in der oberen Hälfte des Bereichs der Itemschwierigkeiten. Dies deckt sich mit der Feststellung zur curricularen Validität aus heutiger Sicht, dass viele Items heute eher das Schwierigkeitsniveau von Gymnasien abdecken (siehe Abschnitt 6.4.5). Unter den 24 ausgewählten Items befinden sich acht, bei denen die inhaltliche Expertise eher

von einem gymnasialen Schwierigkeitsniveau ausgeht. Drei weitere Items der Auswahl sind ohne Taschenrechner nicht leicht zu lösen (siehe Abschnitt 6.4.5).

Englishtest

Da der ursprüngliche Englishtest mit 115 Items sehr umfangreich war, wurden dort besonders viele Items ausgeschlossen, um zu einer Auswahl für die erneute Skalierung zu gelangen und eine kürzere Erhebungszeit zu erreichen. Die Kriterien der Itemselektion waren: ein positiver Bescheid der fachdidaktischen Expertise, keine negative oder eine sehr niedrige Trennschärfe und eine Lösungswahrscheinlichkeit von über 50 Prozent. Der gesamte Test sollte zwei Texte zum Leseverstehen mit allen dazugehörigen Items enthalten, ansonsten sollte das Fragenformat der Items nur Lückentext bzw. C-Test umfassen. Es wurde ein Englishtest mit 34 Items zusammengestellt, der weiterhin eine hohe Reliabilität und gute Infit-Werte (zwischen 0.80 und 1.48) aufweist. Ein Item weist eine negative Trennschärfe auf, wurde aber aus inhaltlichen Gründen und auf Rat der fachdidaktischen Expertinnen und Experten im Test belassen. Es handelt sich hier um ein außergewöhnlich schwierig zu lösendes Item mit einer Lösungswahrscheinlichkeit von 20 Prozent. Abbildung 6.5 zeigt beispielhaft für das Fach Englisch die itemcharakteristischen Funktionen zweier Items und die empirischen Antwortkurven, die für alle Items aller Tests erstellt wurden.

Das erste Item E943_22 ist ein Grammatik-Item aus dem IEA-Test, das nicht in die Auswahl aufgenommen wurde. Die empirische Antwortkurve (in Schwarz mit Anzeige der Spannweiten) weist nicht die erwartete Form (in Blau) auf, was auch am Infit-Wert von 1.702 und einer negativen Trennschärfe sichtbar wird. Die Lösungswahrscheinlichkeit bei diesem Item lag bei 42,1 Prozent und es wurde von den Expertinnen und Experten als teilweise passend bewertet. Das zweite Item E95_20 hingegen wurde in die Auswahl aufgenommen, es handelt sich hierbei um ein C-Test Item aus dem Test EET 9+. Die empirische Antwortkurve weist die erwartete Form auf. Das Item hat einen guten Infit-Wert von 0.945 und eine gute Trennschärfe (.513). Von den Expertinnen und Experten wurde es als passend bewertet und hat eine Lösungswahrscheinlichkeit von 59,5 Prozent.

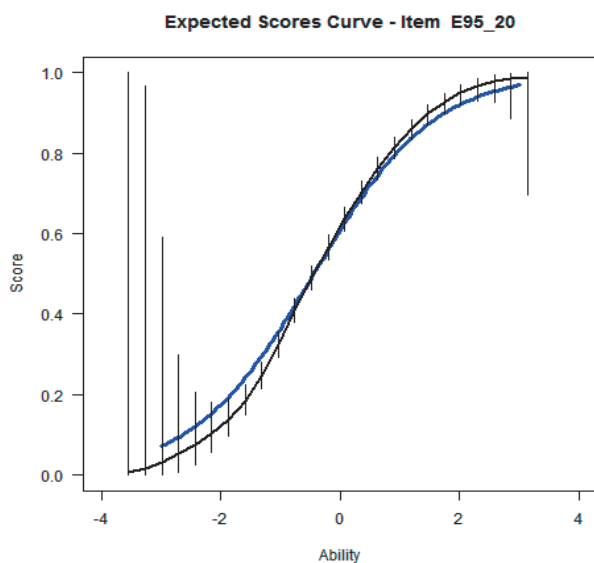
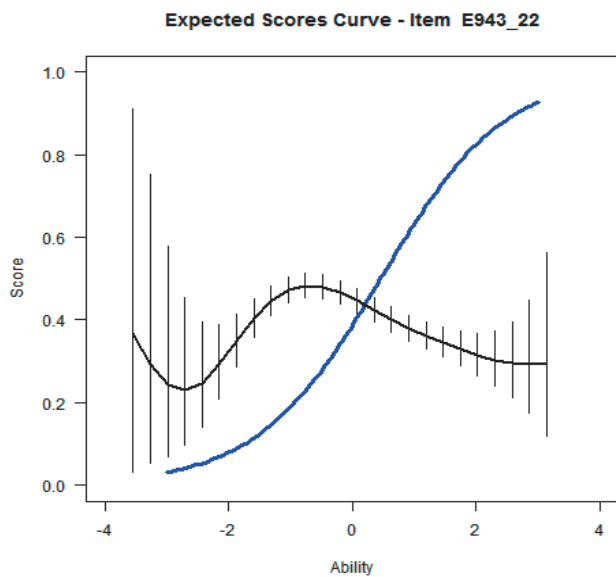


Abbildung 6.5: Itemcharakteristische Funktionen und empirische Antwortkurven für zwei Items des Englischtests

Eine Übersicht über das Resultat der IRT-skalierten Testdaten der ausgewählten Items für das Fach Englisch findet sich in Abbildung 6.6. Wie schon für das Fach Mathematik erfolgt mit einer „Wright Map“ die gleichzeitige Darstellung der Item- und Personenparameter. Die ausgewählten Items verteilen sich über den kompletten Bereich der Itemschwierigkeiten. In der Auswahl befinden sich drei Items, die heute eher für das Schwierigkeitsniveau von Gymnasien geeignet sind (siehe Abschnitt 6.4.5).

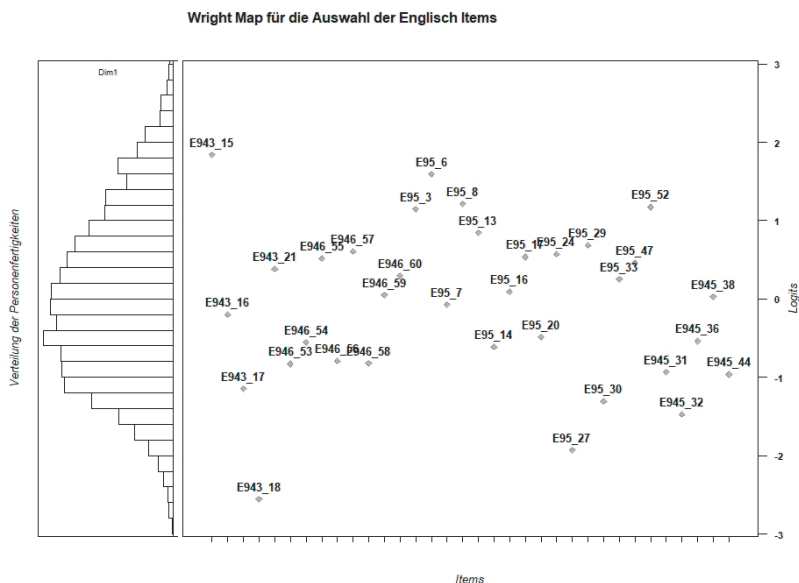


Abbildung 6.6: Wright Map für die 34 ausgewählten Items in Englisch

Deutschtests

Die Leistungstests für das Fach Deutsch werden in der vorliegenden Arbeit nicht in die weiteren Analysen einbezogen, da für dieses Fach keine Angaben der Schülerinnen und Schüler über ihren Unterricht vorliegen. Aus diesem Grund wurden auch keine Einschätzungen hinsichtlich der curricularen Validität aus heutiger Sicht eingeholt. Der Vollständigkeit halber wird an dieser Stelle dennoch kurz auf die Skalierung eingegangen. Die unterschiedlichen Parallelformen des Leseverständnistests wurden getrennt voneinander skaliert. Die interne Konsistenz der beiden Testformen ist auch bei einer Auswahl von Items

sehr hoch und die Infit-Werte sind sehr gut (zwischen 0.88 und 1.13 bzw. 0.88 und 1.14).

Beim Rechtschreibe-Test im Fach Deutsch wurden zwei Items aus der ursprünglichen Skalierung ausgeschlossen, da sie im jeweiligen Datensatz (eventuell aufgrund einer fehlerhaften Eingabemaske) in je vier Dummy-Variablen aufgeteilt waren. Nach Rekonstruktion dieser beiden Items und der Überführung der Dummys in je ein Testitem konnte der Test, wie er ursprünglich gedacht war, mit allen 117 Items skaliert werden. Dennoch wurde zum Vergleich auch die reduzierte Version mit 115 Items erstellt. Von den 117 ursprünglichen Items wurden 29 aufgrund teststatistischer Kriterien ausgeschlossen. Die verbliebenen 88 Items weisen eine sehr hohe Reliabilität mit sehr guten Infit-Werten (zwischen 0.87 und 1.17) auf.

7 Daten und Methoden der Sekundäranalysen

Im Folgenden soll zunächst beschrieben werden, auf welcher Stichprobe die vorliegenden Sekundäranalysen beruhen. Anschließend werden die verwendeten Instrumente und die Auswertungsmethoden vorgestellt.

7.1 Die Stichprobe

Die folgenden Analysen umfassen alle Schulen der Drei-Länder-Studie 1978/79, mit denen eine Mehrebenenstruktur der Schülerinnen und Schülern sowie ihrer Lehrkräfte abbildbar ist. Bei einer Gesamtschule in Nordrhein-Westfalen waren keine Angaben von Lehrkräften vorhanden, es lagen nur Angaben von den Schülerinnen und Schülern vor. Diese Schule wurde aus den Analysen ausgeschlossen. Bei einer anderen Hauptschule desselben Bundeslandes waren zwar Angaben sowohl von Lehrkräften als auch von Schülerinnen und Schülern vorhanden, allerdings fehlten bei letzteren die Angaben zur Personen-ID ihrer jeweiligen Fachlehrkraft. Da bei dieser Schule keine Zuordnung der Schülerinnen und Schülern zu ihrer jeweiligen Fachlehrkraft möglich war, wurde auch diese Schule bei den weiteren Analysen nicht berücksichtigt. Somit konnten von 61 Schulen 59 in die Analysen einbezogen werden. Die Schulen verteilen sich auf 17 Hauptschulen, 8 Realschulen, 26 Gesamtschulen und 8 Gymnasien (siehe Tabelle 7.1).

Tabelle 7.1: Anzahl der Schulen in der Sekundäranalyse nach Schulform und Bundesland

	Nordrhein- Westfalen	Hessen	Niedersachsen	Gesamt
Hauptschulen	6	3	8	17
Realschulen	4	–	4	8
Gesamtschulen	8	12	6	26
Gymnasien	4	–	4	8
Gesamt	22	15	22	59

Anmerkung: Änderungen im Vergleich zur Ausgangsstichprobe der Drei-Länder-Studie in fett

Bei allen Untersuchungsbereichen und auch bei den Leistungstests aller untersuchten Fächer (u.a. Englisch und Mathematik) sowie dem kognitiven Fähigkeitstest (CFT) waren alle Schülerinnen und Schüler der Stichprobe einbezogen, nur die Skalen zum Verhalten der Lehrkräfte im Unterricht und die dazu-

gehörigen fächerbezogenen Variablen der Schülerinnen und Schülern wurden jeweils immer nur von der Hälfte der Schülerinnen und Schüler bearbeitet: die eine Hälfte der Schülerinnen und Schüler beantwortete die Items zum Englischunterricht, die andere Hälfte beantwortete die textlich identischen Items mit Bezug zum Mathematikunterricht.⁴³

Die Anzahl der Schülerinnen und Schüler, die für Analysen des Unterrichts unter Einbezug der Daten der Lehrkräfte verwendet werden kann, beträgt 1043 für Englisch und 1188 für Mathematik. Es gibt mehrere Gründe, warum insgesamt 1262 Schülerinnen und Schüler bei den hier vorliegenden Analysen keine Berücksichtigung finden: 1. Bei bestimmten Schülerinnen und Schülern konnte nicht zugeordnet werden, ob sie die Fragen zum Unterricht für Englisch oder für Mathematik beantwortet hatten (weil in der entsprechenden Zuordnungsvariable keine Angabe vorlag, N=206). 2. Es lag keine Angabe für die zugeordnete Personen-ID ihrer Fachlehrkraft für Englisch (N=158) bzw. Mathematik vor (N=163).⁴⁴ 3. Die zugeordnete Lehrkraft hatte an der Befragung nicht teilgenommen⁴⁵ (dies betrifft im Englischunterricht 381 Schülerinnen und Schüler, deren 65 Lehrkräfte nicht teilgenommen hatten, im Mathematikunterricht 342 Schülerinnen und Schüler bzw. 56 Lehrkräfte).⁴⁶

Wie der Tabelle 7.2 zu entnehmen ist, verteilten sich die Schülerinnen und Schüler, die für den Englisch- und den Mathematikunterricht untersucht werden können, unterschiedlich auf die Schulformen. Bei den Daten für den Englischunterricht besuchte je ein Fünftel der Schülerinnen und Schüler das Gymnasium und die Hauptschule, für den Mathematikunterricht liegen mehr Daten von Lernenden an Hauptschulen vor als von denen an Gymnasien.

43 In den Berichten der Drei-Länder-Studie finden sich keine genauen Informationen darüber, wie die Aufteilung der Schülerinnen und Schüler auf die Schülerhefte mit Items ihren Englisch- bzw. Mathematikunterricht betreffend erfolgte.

44 Darunter befindet sich die gesamte erhobene Schülerschaft einer Hauptschule in Nordrhein-Westfalen, für die keine Angabe über die Zuordnung ihrer Fachlehrkräfte vorliegt, siehe oben.

45 Darunter befindet sich die komplette erhobene Schülerschaft einer Gesamtschule in Nordrhein-Westfalen, bei der grundsätzlich keine Angaben von Lehrkräften vorliegen, siehe S. 121.

46 Zwölf weitere Schülerinnen und Schüler (Englischunterricht: 4, Mathematikunterricht: 8) mussten aus den Analysen herausgenommen werden, da bei deren zugeordneter Fachlehrkraft die Schul-ID nicht mit der Schul-ID der Schülerin bzw. des Schülers übereinstimmt.

Tabelle 7.2:

Anzahl und Anteile der Schülerinnen und Schüler nach Schulform

	Englischunterricht		Mathematikunterricht		Gesamtstichprobe Drei-Länder-Studie	
	N	Prozent	N	Prozent	N	Prozent
Haupt- schulen	201	19,3	300	25,3	821	23,5
Realschulen	156	15,0	167	14,1	454	13,0
Gesamt- schulen	470	45,1	540	45,5	1749	50,1
Gymnasien	216	20,7	181	15,2	469	13,4
Gesamt	1043	100	1188	100	3493	100

Von den Schülerinnen und Schülern der für die hier vorliegende Sekundäranalyse verwendeten Stichprobe für den Englischunterricht waren 47 Prozent weiblich und 45 Prozent männlich. Bei sieben Prozent der Schülerinnen und Schüler sowohl für den Englisch- als auch für den Mathematikunterricht fehlt diese Angabe. In der Stichprobe des Mathematikunterrichts waren 46 Prozent weiblich und 47 Prozent männlich (siehe Anhang, Tabelle A.1). 87 bzw. 89 Prozent der Schülerinnen und Schüler des Englisch- bzw. Mathematikunterrichts sprachen zu Hause überwiegend Deutsch, und nur drei Prozent sprachen zu Hause nicht Deutsch (siehe Anhang, Tabelle A.2). Bei dieser Frage waren die fehlenden Angaben ähnlich hoch wie bei der Geschlechtervariablen. Bei den Angaben zum Sozialstatus der Familie ordneten sich etwa zwei Drittel der Schülerinnen und Schüler in beiden Fächergruppen den mittleren drei Kategorien (von sieben) zu, etwa 11 Prozent machten keine Angabe (siehe Anhang, Tabelle A.3). Nach Kleining und Moore (1968) bilden die drei mittleren Kategorien die obere Unterschicht bzw. die untere und die obere Mittelschicht ab.

Die Datenbasis für die Analysen zum Englischunterricht bilden die 1043 Schülerinnen und Schüler und die Englischlehrkraft ihrer jeweiligen Unterrichtsgruppe. Die Unterrichtsgruppen sind nicht identisch mit den Klassen, da die Untersuchung auch Schulen enthält, an denen leistungsdifferenziert unterrichtet wurde. Infolgedessen umfasst diese Stichprobe 147 Lehrkräfte, woraus sich genauso viele Unterrichtsgruppen (damit ist je nach Differenzierung der Kurs oder die Klasse gemeint) ergeben. Die Datenbasis für die Analysen zum Mathematikunterricht bilden die 1188 Schülerinnen und Schüler und die 164 Mathematiklehrkräfte ihrer Unterrichtsgruppen.

Die für die folgenden Analysen verwendete Stichprobe von 147 Englisch- und 164 Mathematiklehrkräften setzt sich wie folgt zusammen: Knapp über die Hälfte der untersuchten Englisch- wie Mathematiklehrkräfte unterrichteten an Gesamtschulen, etwas über ein Viertel an Hauptschulen. Die restlichen Lehrkräfte verteilen sich beinah hälftig auf Realschulen und Gymnasien, wobei die Mathematiklehrkräfte an Gymnasien die kleinste Gruppe stellen (mit acht Prozent, siehe Tabelle 7.3). In Tabelle 7.3 ist zudem die Verteilung der Gesamtstichprobe der Lehrkräfte in der Drei-Länder-Studie nach Schulformen abgetragen. Eine Beschreibung der Gesamtstichprobe der Lehrkräfte findet sich im Skalenhandbuch der Drei-Länder-Studie von Schmid et al. (2019).

Tabelle 7.3: Anzahl und Anteile der Lehrkräfte nach Schulform

	Englischlehrkräfte		Mathematiklehrkräfte		Gesamtstichprobe Drei-Länder-Studie	
	N	Prozent	N	Prozent	N	Prozent
Haupt- schulen	37	25,2	47	28,7	251	21,5
Realschulen	17	11,6	21	12,8	114	9,8
Gesamtschulen	78	53,1	83	50,6	719	61,5
Gymnasien	15	10,2	13	7,9	85	7,3
Gesamt	147	100	164	100	1169	100

Von den Englischlehrkräften der verwendeten Stichprobe waren 52 Prozent weiblich und 41 Prozent männlich. Sieben Prozent der befragten Englischlehrkräfte hatten die entsprechende Frage nicht beantwortet. Bei den Mathematiklehrkräften waren 23 Prozent weiblich, 58 Prozent männlich und 19 Prozent haben die entsprechende Frage nicht beantwortet (siehe Anhang, Tabelle A.4). Der Großteil der untersuchten Lehrkräfte war zum Zeitpunkt der Befragung (1978/79) zwischen 26 und 35 Jahren alt: 63 Prozent der Englischlehrkräfte und 47 Prozent der Mathematiklehrkräfte. Der Anteil der Personen, die zu dieser Frage keine Angaben gemacht hatten, ist dabei ähnlich hoch wie beim Geschlecht: fünf Prozent bei den Englischlehrkräften und 14 Prozent bei den Mathematiklehrkräften (siehe Anhang, Tabelle A.5). Dass es sich bei den untersuchten Lehrkräften eher um junge Personen handelte, wird auch daran ersichtlich, dass bei den Englischlehrkräften 37 Prozent erst seit maximal fünf Jahren als Lehrkraft arbeiteten (fünf Prozent keine Angabe) und bei den Mathematik-

lehrkräften 31 Prozent (neun Prozent keine Angabe; siehe Anhang, Tabelle A.6). Genau so lange, also maximal fünf Jahre, unterrichtete auch etwa die Hälfte der Englisch- bzw. Mathematiklehrkräfte schon an der damaligen Schule (fünf bzw. neun Prozent keine Angabe; siehe Anhang, Tabelle A.7).

7.2 Die Instrumente

7.2.1 Überzeugungen der Lehrkräfte

Um die Überzeugungen der Lehrkräfte zu erfassen, wurden die Lehrkräfte gebeten, zu einer Reihe von Äußerungen über Schule und Erziehung Stellung zu nehmen. Die deskriptiven Statistiken der drei hier behandelten Überzeugungen der Lehrkräfte finden sich in Tabelle 7.7 am Ende dieses Kapitels.

Reformbereitschaft

Die Skala „Reformbereitschaft“ stammt ursprünglich aus dem Konstanzer Fragebogen für Lehrereinstellungen (KLE),⁴⁷ der von Ende der 1960er bis Anfang der 1970er Jahre entstanden ist (Koch, 1972). In der Drei-Länder-Studie wurde eine Kurzform der Skala verwendet, die Cloetta und Hohner (1976) entwickelt hatten. Im KLE ist diese Skala umgekehrt gepolt („negative Reformbereitschaft“)⁴⁸ und erfasst die „negative“, das heißt konservative und bremsende Einstellung zur Reform von Institutionen allgemein und des Schulsystems insbesondere, und zwar in organisatorischer und inhaltlicher Hinsicht, gegenüber einer reform- und innovationsfreundlichen Haltung“ (Cloetta & Hohner, 1976, S. 3). Ausführlicher beschreibt Koch (1972, S. 67) die Extrempositionen der Skala:

„Lehrer unterscheiden sich hinsichtlich ihrer Vorstellungen über die Reform- und Veränderungsbedürftigkeit von gesellschaftlichen Institutionen allgemein und der Schule im besonderen. Die einen sind der Meinung, daß man bestehende und bewährte Einrichtungen und Verhaltensformen nicht aufs Spiel setzen, sondern pflegen und erhalten sollte und daß ein Fortschritt allein durch allmähliche und sorgsam bedachte Schritte erzielt werden könne. Die anderen betonen, daß Überkommenes ständiger und radikaler Kritik unterworfen werden müsse und

47 Der „Konstanzer Fragebogen für Lehrereinstellungen (KLE)“ wurde später auch unter dem Namen „Konstanzer Fragebogen für Schul- und Erziehungseinstellungen (KSE)“ veröffentlicht (Koch, 1972, S. 55).

48 Bei Koch (1972, S. 67) findet auch die Bezeichnung „Negative Reformeinstellung vs. Veränderungsbereitschaft“ Verwendung.

daß man neue Ideen auch dann ausprobieren sollte, wenn die Konsequenzen nicht von vorneherein überschaubar sind.“

Bei Fend (1984) wird die Reformbereitschaft als eine erzieherische Grundhaltung betrachtet, die in engem Zusammenhang mit Konservatismus steht. Da ein wichtiger Aspekt des Konservatismus die „Orientierung am Gegebenen und die Verehrung des Traditionalen und historisch Gewordenen“ sei, müsste sich das auch auf die Reformskepsis bezüglich schulischer Veränderungen ausweiten (Fend, 1984, S. 124).

Tabelle 7.4: Itemformulierungen und Itemtrennschärfen der Skala „Reformbereitschaft“

Itemformulierung	Trennschärfe für Englisch- lehrkräfte	Trennschärfe für Mathematik- lehrkräfte
„Institutionen, wie z.B. das Schulsystem, müssen ständig radikal in Frage gestellt werden.“	.486	.617
„Wer ständig auf Neuerungen aus ist, erkennt die wahren Aufgaben der Schule.“ (Umgepolt)	.469	.565
„Bei allen Bemühungen um Schulreform wird heutzutage oft vergessen, daß es hier Dinge gibt, die bewahrt werden müssen.“ (Umgepolt)	.610	.669
„Vor den Anforderungen der modernen Welt müssen die bislang gewohnten Schulformen und Unterrichtsmethoden einfach versagen.“	.444	.588
„Schulmethoden und Schuleinrichtungen, die sich einmal als brauchbar erwiesen haben, sollte man nicht einfach aufgeben.“ (Umgepolt)	.652	.622
„Für die Schule ist es immer noch besser, nicht ganz vollkommene Einrichtungen beizubehalten, als sich den Ungewißheiten nicht erprobter Neuerungen auszusetzen.“ (Umgepolt)	.612	.688

Die Formulierungen der sechs Items der Skala „Reformbereitschaft“ sowie die dazugehörigen Trennschärfen für die Englisch- und Mathematiklehrkräfte sind in Tabelle 7.4 aufgeführt. Es stehen sechs Antwortmöglichkeiten zur Verfügung, die drei Stufen der Ablehnung und drei Stufen der Zustimmung enthalten: (1) „Ich finde das sehr falsch“, (2) „Ich finde das falsch“, (3) „Ich finde das eher falsch als richtig“, (4) „Ich finde das eher richtig als falsch“, (5) „Ich finde das richtig“ und (6) „Ich finde das sehr richtig“. Diese Skala weist mit einem

Cronbachs α von .79 für die Englischlehrkräfte (N = 148)⁴⁹ und einem Cronbachs α von .84 für die Mathematiklehrkräfte (N = 156) für beide Fachgruppen eine hohe interne Konsistenz auf.

Notwendigkeit von Druck bei Lernprozessen

Die Skala der „Notwendigkeit von Druck bei Lernprozessen“⁵⁰ stammt, ebenso wie die Skala „Reformbereitschaft“, aus dem Konstanzer Fragebogen für Lehrereinstellungen (KLE; Koch, 1972). Cloetta und Hohner (1976) haben auch hierzu eine Kurzform der Skala entwickelt, welche die „positive Bewertung der Wirksamkeit und Berechtigung von Druck, Kontrolle und Disziplin im Schulunterricht gegenüber der Ablehnung dieser Methoden“ erfasst (Cloetta & Hohner, 1976, S. 3). Ausführlicher beschreibt wiederum Koch (1972, S. 66–67) die Extrempositionen der Skala:

„Lehrer unterscheiden sich in ihren Vorstellungen darüber, wie schulische Leistungen am besten zu erreichen sind. Manche versuchen, die Schüler durch Kontrolle, Strafen und eine unnachgiebige Haltung zum Erbringen von Leistungen und zu einem akzeptablen Verhalten zu bewegen. Andere sind der Meinung, daß auf die Dauer Leistungen nur dort erbracht werden, wo eine innere Bereitschaft dazu besteht, die am besten durch das Wecken sachbezogener Interessen, Ansporn und eine an die Bedürfnisse der Schüler anknüpfende Unterrichtsgestaltung gefördert wird. – Die Kontroverse bezieht sich also auf die optimale Motivierung von Schülern.“

Fend (1984) betrachtet dieses Konstrukt, ähnlich wie das der Reformbereitschaft, als eine weitere erzieherische Grundhaltung, die in engem Zusammenhang mit Konservatismus stehe. Dort wird diese Skala auch mit „Skrupel im Umgang mit Macht“ und „Einstellung zu Strafe und zu der Notwendigkeit von physischem Zwang“ assoziiert (Fend, 1984, S. 124).

49 Die Angaben der Fallzahlen für Cronbachs α beziehen sich auf die Gesamtstichprobe der Lehrkräfte in der Drei-Länder-Studie. Dabei werden diejenigen Lehrkräfte berücksichtigt, die bei allen Items der Skala einen gültigen Wert aufweisen.

50 Der originale Name der Skala mit acht Items (siehe Tabelle 6.5: skl10) lautet in der Drei-Länder-Studie und bei Cloetta und Hohner (1976) „Druckmotivierung“, bei Fend (1984, S. 168) sowie bei Koch (1972, S. 66) wird die Skala als „Druck-Zug“ bezeichnet bzw. als „Druckorientierung“ (Fend, 1984, S. 124).

Tabelle 7.5: Itemformulierungen und Itemtrennschärfen der Skala „Notwendigkeit von Druck bei Lernprozessen“

Itemformulierung	Trennschärfe für Englischlehrkräfte	Trennschärfe für Mathematiklehrkräfte
„Wenn man dem Lehrer die Zensur als Kontrollmittel nimmt, werden die meisten Schüler bald nichts mehr tun.“	.590	.541
„Ganz ohne Zwang wird auch der beste Lehrer nichts erreichen.“	.736	.573
„Bei den meisten Kindern muß man von Zeit zu Zeit einmal hart durchgreifen.“	.594	.631
„Ohne strikte Disziplin ist eine erfolgreiche Arbeit in den meisten Klassen nicht möglich.“	.522	.510
„Schüler finden die Ordnung ihres Arbeitens und Zusammenlebens bei rechter Anleitung selbst und benötigen daher keine äußerliche Disziplin.“ (Umgepolt)	.502	.484
„Man dürfte wohl nicht weit kommen, wenn man die Arbeit in der Schule allein auf den individuellen Interessen der Schüler aufbauen wollte.“	.535	.464

Die Formulierungen der sechs Items der Skala „Notwendigkeit von Druck bei Lernprozessen“ sowie die dazugehörigen Trennschärfen für die Englisch- und Mathematiklehrkräfte sind in Tabelle 7.5 aufgeführt. Es stehen sechs Antwortmöglichkeiten zur Verfügung, die drei Stufen der Ablehnung und drei Stufen der Zustimmung enthalten: (1) „Ich finde das sehr falsch“, (2) „Ich finde das falsch“, (3) „Ich finde das eher falsch als richtig“, (4) „Ich finde das eher richtig als falsch“, (5) „Ich finde das richtig“ und (6) „Ich finde das sehr richtig“. Diese Skala weist mit einem Cronbachs α von .82 für die Englischlehrkräfte ($N = 147$) und einem Cronbachs α von .78 für die Mathematiklehrkräfte ($N = 158$) für beide Fachgruppen eine hohe interne Konsistenz auf.

Einstellung zur Förderung schwacher Schülerinnen und Schüler

Die Skala „Einstellung zur Förderung schwacher Schülerinnen und Schüler“⁵¹ wurde von Fend (1984, S. 134) verwendet, um zu prüfen, ob „sich konservative Lehrer vornehmlich um die starken Schüler kümmern, kritisch-liberale eher um

51 Der originale Name der Skala mit acht Items (siehe Tabelle 6.5: sk111) lautet „Förderintention“, bei Fend (1984, S. 135) wird eine leichte Abwandlung der hier verwendeten Skala mit sechs Items als „Eliteorientierung“ bezeichnet bzw. als „Förderungsorientierung vs. Elite“ (Fend, 1984, S. 168).

die schwachen“. Da in den Publikationen zur Drei-Länder-Studie keine Quelle dieses Konstrukts vermerkt ist, besteht die Möglichkeit, dass es sich bei dieser Skala um eine Eigenentwicklung der damaligen Projektgruppe handelt.

Tabelle 7.6: Itemformulierungen und Itemtrennschärfen der Skala „Einstellung zur Förderung von schwachen Schülerinnen und Schülern“

Itemformulierung (alle umgepolt)	Trennschärfe für Englischlehrkräfte	Trennschärfe für Mathematiklehrkräfte
„Es ist wichtiger, den vorgeschriebenen Lehrstoff durchzunehmen, als zu warten, bis jeder Schüler alles verstanden hat.“	.404	.312
„Wenn Schüler nicht eine bestimmte Intelligenz haben, nützen auch die besten Methoden nichts.“	.566	.497
„Eigentliche Befriedigung kann ein Lehrer nur empfinden, wenn er auch gute Schüler mit hervorragenden Leistungen unterrichten kann.“	.384	.392
„Die ständige Rücksichtnahme auf die schwachen Schüler beeinträchtigt erheblich meinen Unterricht.“	.473	.423
„Schwache Schüler sollten aus normalen Klassen ausgegliedert und getrennt unterrichtet werden.“	.391	.417
„Leistungsstarke Schüler lernen zu wenig, wenn sich der Lehrer zu häufig mit den schwachen Schülern beschäftigt.“	.620	.490

Die Skala „Einstellung zur Förderung von schwachen Schülerinnen und Schülern“ war ursprünglich so gepolt, dass hohe Werte eine negative Einstellung zur Förderung von Schülerinnen und Schülern repräsentieren und niedrige Werte eine positive Einstellung. Aus heutiger Sicht scheint es angemessener, die Skala so zu interpretieren, dass eine positive Einstellung gemessen wird, daher wurden alle sechs Items der Skala umgepolt. Die Formulierungen der Items dieser Skala sowie die dazugehörigen Trennschärfen für die Englisch- und Mathematiklehrkräfte sind in Tabelle 7.6 aufgeführt. Es stehen sechs Antwortmöglichkeiten zur Verfügung mit den Extrempolen: (1) „stimmt gar nicht“ und (5) „stimmt völlig“. Die Antwortmöglichkeiten (2), (3) und (4), die zwischen den Polen der Ablehnung und Zustimmung liegen, sind im Fragebogen nicht durch eine Beschreibung ergänzt. Die Skala weist mit einem Cronbachs α von .74 für Englischlehrkräfte ($N = 158$) und einem Cronbachs α von .69 für die

Mathematiklehrkräfte (N = 165) für beide Fachgruppen eine zufriedenstellende interne Konsistenz auf.

In Tabelle 7.7 finden sich die deskriptiven Statistiken der Überzeugungen der Lehrkräfte. Für die Englisch- und Mathematiklehrkräfte sind jeweils ihre Anzahl sowie der Mittelwert, die Standardabweichung und die Spannweite der jeweiligen Skala angegeben.

Tabelle 7.7: Deskriptive Statistiken der Überzeugungen von Lehrkräften in der Sekundäranalyse

	Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung schwacher Schülerinnen und Schüler	
	Englisch-lehrkräfte	Mathematik-lehrkräfte	Englisch-lehrkräfte	Mathematik-lehrkräfte	Englisch-lehrkräfte	Mathematik-lehrkräfte
MW	3,28	3,02	3,80	3,80	3,41	3,53
SD	0,864	0,852	0,970	0,871	0,733	0,648
Min.	1	1	1	1,5	1	1,67
Max.	6	5,5	5,67	5,67	5	5
N	126	135	124	137	130	144

Anmerkung: MW = Mittelwert, SD = Standardabweichung, Min. = Minimum, Max. = Maximum

7.2.2 Unterrichtsmerkmale

Bei den Analysen zum Zusammenhang der Überzeugungen der Lehrkräfte mit ihrem Unterrichtshandeln kommen verschiedene Aspekte von Unterrichtsmerkmalen zum Einsatz. Dabei werden Konstrukte betrachtet, die ein unterstützendes Unterrichtsklima abbilden, sowie solche Konstrukte, welche die Klassenführung im Unterricht abbilden. Die einzelnen Skalen mit je einem Beispielitem sind in Tabelle 7.8: Verwendete Unterrichtsmerkmale der Sekundäranalyse aufgeführt. Alle Items der Skalen finden sich im Skalenhandbuch, das vom Projekt „Schule im Wandel“ veröffentlicht wurde (Schmid et al., 2019). Die Basis aller Unterrichtsskalen in der Drei-Länder-Studie waren Eigenentwicklungen der damaligen Projektgruppe, die eigens für die Vergleichsstudie 1978/79 angefertigt worden waren. Lediglich zwei Skalen, „Unterstützung“ und „Strenge“, sind von Heinrich (1974) übernommen worden, auch um die neu entwickelten Skalen zu validieren (siehe Krüger-Haenisch, o. J., S. III–IV sowie Haenisch, 1979, S. 43).

Tabelle 7.8: Verwendete Unterrichtsmerkmale der Sekundäranalyse

Unterrichtsdimension	Name der Skala	Beispielitem	Nr.
unterstützendes Unterrichtsklima	Unterstützung	„Wenn ich bei einer Aufgabe nicht weiter weiß, hilft mir mein Lehrer.“	sks22
	Strenge	„Wenn ich meinem Lehrer widerspreche, wird er ärgerlich.“	sks23r
	Adaptivität des Unterrichts	siehe Tabelle 7.9	sks24_27
	Positive Schülerorientiertheit/Lehrer-Schüler-Beziehung	„Unser Englisch-/Mathematiklehrer ist stets geduldig. ... ist immer ungeduldig.“	sks28
	Individualisierungsmaßnahmen/Förderungsorientiertheit	„Unser Englisch-/Mathematiklehrer bemüht sich auch um die schwächeren Schüler. ... kümmert sich nur um die guten Schüler.“	sks29r
	Motivierungsbemühen/Lob	„Unser Englisch-/Mathematiklehrer sagt uns häufig, daß wir uns verbessert hätten. ... geht nie darauf ein, wenn wir eine gute Leistung erbracht haben.“	sks32r
	Vertrauensgrad/Schülerorientierung	„Unser Englisch-/Mathematiklehrer beachtet auch die Meinung der Schüler. ... bestimmt immer allein, was zu machen ist.“	sks33
Klassenführung	Strukturiertheit des Unterrichts	„Unser Englisch-/Mathematiklehrer fasst häufig noch einmal den Stoff zusammen, damit wir ihn uns gut merken können. ... fasst nur sehr selten den Stoff zusammen.“	sks25r
	Beschäftigungsradius	„Unser Englisch-/Mathematiklehrer ruft auch häufig Schüler auf, die sich nicht gemeldet haben ... ruft selten Schüler, die sich nicht gemeldet haben.“	sks26
	Disziplin und Ordnungssinn	siehe Tabelle 7.10	sks31
	Zeitverschwendung/Time on task	„Bei uns dauert es zu Beginn der Stunde lange, bis die Schüler ruhig werden und zu arbeiten beginnen.“	sks36r

Anmerkung: Weitere Informationen und statistische Kennzahlen zu den aufgeführten Skalen können mit Hilfe der Nummer der Skala der Tabelle 6.4 entnommen werden.

Da der Skala „Adaptivität des Unterrichts“ als ein Aspekt eines unterstützenden Unterrichtsklimas und der Skala „Disziplin und Ordnungssinn“ als ein Aspekt

der Klassenführung bei den Analysen in dieser Arbeit besondere Bedeutung zukommt, werden diese beiden im Folgenden näher beschrieben.

Adaptivität

Die Wahrnehmung der Adaptivität des Englisch- bzw. Mathematikunterrichts durch die Schülerinnen und Schüler wurde über Aussagen zu ihrem Fachunterricht erfasst.

Tabelle 7.9: Itemformulierungen und Itemtrennschärfen der Skala „Adaptivität des Englisch- bzw. Mathematikunterrichts“

Itemformulierung: „Unser Englischlehrer“ bzw. „Unser Mathematiklehrer“	Trennschärfe für Englisch- unterricht	Trennschärfe für Mathematik- unterricht
„... sagt uns häufig, wie wir uns verbessern können ... sagt uns selten, wie wir uns verbessern können.“	.543	.514
„... sagt uns häufig, wie man leichter lernt ... sagt uns nie, wie man leichter lernt.“	.580	.552
„... erklärt es uns zusätzlich noch einmal an der Tafel, wenn wir etwas nicht verstanden haben. ... geht meist mit dem Stoff weiter, auch wenn wir ihn noch nicht ganz verstanden haben.“	.668	.569
„... sagt uns häufig, warum wir bei einer Prüfungsar- beit einige Aufgaben falsch beantwortet haben. ... sagt uns selten, warum wir bei einer Prüfungsarbeit eine Aufgabe falsch beantwortet haben.“	.521	.530
„... benotet uns genau richtig. ... benotet uns zu schlecht.“	.563	.585
„... weiß genau, was jeder von uns leisten kann. ... schätzt viele Schüler in ihrer Leistung falsch ein.“	.601	.626
„... merkt sofort, wenn ein Schüler im Unterricht nicht mitkommt. ... merkt selten, wenn ein Schüler im Unterricht Schwierigkeiten hat.“	.639	.596
„... weiß sofort genau, bei welchen Aufgaben wir Schwierigkeiten haben. ... merkt erst später, bei welchen Aufgaben wir Schwierigkeiten haben.“	.605	.616
„... gibt uns meist nur solche Aufgaben, die wir auch lösen können. ... gibt uns häufig solche Aufgaben, die wir nicht lösen können.“	.465	.427

Die für die vorliegende Sekundäranalyse neu gebildete Skala enthält sowohl Items zur individualisierten Rückmeldung an die Schülerinnen und Schüler (Mikro-Adaptivität) als auch Items zur Anpassung der Aufgaben bzw. des Stoffes an die Voraussetzungen der gesamten Lerngruppe (Makro-Adaptivität; siehe Abschnitt 4.3.1). Die neun Items der Skala bestehen aus gegensätzlichen Aussagen, die zustimmend bzw. ablehnend auf fünf Stufen von „Der obere Satz ist genau richtig“ bis „Der untere Satz ist genau richtig“ beantwortet werden konnten (siehe Tabelle 7.9).

55.	merkt sofort, wenn ein Schüler im Unterricht nicht mit-	2 ▶
		kommt	1 ▶
			0 ▶
			1 ▶
	merkt selten, wenn ein Schüler im Unterricht Schwierig-	2 ▶
		keiten hat	

Abbildung 7.1: Beispielitem für die Skala „Adaptivität“ aus dem Fragebogen „Schülerheft 4“ der Drei-Länder-Studie

Im folgenden findest Du immer zwei gegensätzliche Aussagen über Deinen Englischlehrer. Du sollst Dir jedesmal überlegen, welcher der beiden Sätze Deiner Meinung nach eher richtig ist und dann das Kästchen anstreichen, das Deiner Meinung am ehesten entspricht.

Die Zeichen bedeuten:

Der obere Satz ist genau richtig	→ → →	2
Der obere Satz ist eher richtig als der untere	→ →	1
Meine Meinung liegt zwischen beiden Sätzen	→ →	0
Der untere Satz ist eher richtig als der obere	→ → →	1
Der untere Satz ist genau richtig	→ → →	2

Abbildung 7.2: Instruktion und Erläuterung des Antwortformats aus dem Fragebogen „Schülerheft 4“ der Drei-Länder-Studie für das Fach Englisch

Aus technischen Gründen wurden die beiden gegensätzlichen Aussagen jeweils untereinander im Fragebogen platziert (siehe Abbildung 7.1 und Abbildung 7.2). Die fünf Antwortkategorien lauten: (1) „Der untere Satz ist genau richtig“, (2) „Der untere Satz ist eher richtig als der obere“, (3) „Meine Meinung liegt zwischen beiden Sätzen“, (4) „Der obere Satz ist eher richtig als der untere“ und (5) „Der obere Satz ist genau richtig“.

Die Skala mit neun Items weist für den Englischunterricht ($\alpha = .85$, $N = 1256$) und für den Mathematikunterricht ($\alpha = .83$, $N = 1361$) eine hohe interne Konsistenz auf. Ein Item, das zum Beispiel den Aspekt der Mikro-Adaptivität beleuchtet, lautet: „Unser Englisch-/Mathematiklehrer merkt sofort, wenn ein Schüler im Unterricht nicht mitkommt“ vs. „... merkt selten, wenn ein Schüler im Unterricht Schwierigkeiten hat“. Ein Beispielitem für den Aspekt der Makro-Adaptivität lautet: „Unser Englisch-/Mathematiklehrer gibt uns meist nur solche Aufgaben, die wir auch lösen können“ vs. „... gibt uns häufig solche Aufgaben, die wir nicht lösen können“.

Disziplin und Ordnungssinn

Die Wahrnehmung der Disziplin im Englisch- bzw. Mathematikunterricht durch die Schülerinnen und Schüler wurde über Aussagen zu ihrem Fachunterricht erfasst. Die originale Skala aus der Drei-Länder-Studie erfasst neben der Disziplin im Unterricht auch den Ordnungssinn der Englisch- bzw. Mathematiklehrkräfte aus Sicht der Schülerinnen und Schüler.

Die Formulierungen der fünf Items der Skala „Disziplin und Ordnungssinn“ sowie die dazugehörigen Trennschärfen für den Englisch- und Mathematikunterricht sind in Tabelle 7.10 aufgeführt. Wie bei der Skala „Adaptivität“ stehen folgende fünf Antwortkategorien zur Verfügung: (1) „Der untere Satz ist genau richtig“, (2) „Der untere Satz ist eher richtig als der obere“, (3) „Meine Meinung liegt zwischen beiden Sätzen“, (4) „Der obere Satz ist eher richtig als der untere“ und (5) „Der obere Satz ist genau richtig“. Die Skala weist für den Englischunterricht ($\alpha = .77$, $N = 1381$) und für den Mathematikunterricht ($\alpha = .80$, $N = 1528$) eine hohe interne Konsistenz auf.

Tabelle 7.10: Itemformulierungen und Itemtrennschärfen der Skala „Disziplin im Englisch- bzw. Mathematikunterricht“

Itemformulierung:	Trennschärfe für Englisch- unterricht	Trennschärfe für Mathematik- unterricht
„... spricht sofort einen Schüler an, wenn dieser ihn beim Unterricht stört. ... kümmert sich nicht darum, wenn ihn ein Schüler beim Unterricht zu stören versucht.“	.525	.594
„... bricht sofort den Unterricht ab, wenn ein Schüler Unsinn macht. ... läßt sich nicht stören, wenn ein Schüler Unsinn macht.“	.412	.480
„... achtet immer darauf, dass wir uns gegenseitig nicht unterhalten. ... kümmert sich nicht darum, wenn wir uns gegenseitig unterhalten.“	.583	.598
„... achtet sehr darauf, daß wir immer aufpassen. ... achtet kaum darauf, ob wir aufpassen.“	.562	.611
„... achtet immer genau auf Ruhe und Ordnung. ... läßt die Schüler machen, was sie wollen.“	.618	.635

7.2.3 Schulische Ergebnisse der Schülerinnen und Schüler

Die abhängigen Variablen der Analysen der zweiten Fragestellung sind das Wohlfühlen der Schülerinnen und Schüler im Englisch- bzw. Mathematikunterricht und ihre Englisch- bzw. Mathematikleistung.

Wohlfühlen im Englisch- bzw. Mathematikunterricht

Das Wohlfühlen der Schülerinnen und Schüler wurde über die Zustimmung bzw. Ablehnung gegensätzlicher Aussagen zu ihrem Fachunterricht erfragt. Die Skala stellt eine Neuentwicklung der Arbeitsgruppe um Helmut Fend aus dem Jahr 1978 dar (Krüger-Haenisch, o. J.). Die Formulierungen der vier Items der Skala „Wohlfühlen im Englisch- bzw. Mathematikunterricht“ sowie die dazugehörigen Trennschärfen für den Englisch- und Mathematikunterricht sind in Tabelle 7.11 aufgeführt. Wie bei der Skala „Adaptivität“ stehen folgende fünf Antwortkategorien zur Verfügung: (1) „Der untere Satz ist genau richtig“, (2) „Der untere Satz ist eher richtig als der obere“, (3) „Meine Meinung liegt zwischen beiden Sätzen“, (4) „Der obere Satz ist eher richtig als der untere“ und

(5) „Der obere Satz ist genau richtig“. Die Skala weist für den Englischunterricht ($\alpha = .84$, $N = 1364$) und für den Mathematikunterricht ($\alpha = .83$, $N = 1407$) eine hohe interne Konsistenz auf.

Tabelle 7.11: Itemformulierungen und Itemtrennschärfen der Skala „Wohlfühlen im Englisch- bzw. Mathematikunterricht“

Itemformulierung: „Ich selbst“	Trennschärfe für Englisch- unterricht	Trennschärfe für Mathematik- unterricht
„... fühle mich in Englisch/Mathematik meistens sehr wohl. ... fühle mich in Englisch/Mathematik sehr unwohl.“	.755	.782
„... mag den Englisch-/Mathematiklehrer sehr gerne. ... kann den Englisch-/Mathematiklehrer nicht leiden.“	.676	.627
„... finde den Englisch-/Mathematikunterricht zur Zeit interessant. ... was wir zur Zeit im Englisch-/Mathematikunterricht machen, interessiert mich nicht.“	.662	.674
„... habe sehr viel vom Englisch-/Mathematikunterricht, weil ich gut mitkomme. ... habe nicht viel vom Englisch-/Mathematikunterricht, weil er für mich zu schwierig ist.“	.576	.575

Anmerkung: Weitere Informationen zu dieser Skala können unter der Nummer sks35r der Tabelle 6.4 entnommen werden.

Englischleistungen

Für die Erfassung der Englischleistungen der Schülerinnen und Schüler wurde der auf 34 Items reduzierte Test aus der Drei-Länder-Studie eingesetzt. Die hier dargestellte Sekundäranalyse verwendet eine IRT-Skalierung auf der Basis des eindimensionalen Rasch-Modells und berechnet WLE-Schätzer als Personenfähigkeitsparameter (Warm, 1989). Die WLE-Reliabilität des Tests beträgt .86 (EAP-Reliabilität .79). Eine ausführliche Beschreibung des Vorgehens bei der Testskalierung sowie der Inhaltsbereiche des Tests befindet sich in Kapitel 6.4.

Mathematikleistungen

Für die Erfassung der Mathematikleistungen der Schülerinnen und Schüler wurde der auf 24 Items reduzierte Test aus der Drei-Länder-Studie eingesetzt. Die IRT-Skalierung und Berechnung der Personenfähigkeitsparameter für die hier dargestellte Sekundäranalyse erfolgten analog zum Fach Englisch. Die WLE-Reliabilität des Tests beträgt .70 (EAP-Reliabilität .68). Eine ausführliche

Beschreibung des Vorgehens bei der Testskalierung sowie der Inhaltsbereiche des Tests befindet sich in Kapitel 6.4.

7.2.4 Soziobiografische Hintergrundmerkmale und Komposition der Unterrichtsgruppe

Als Kontrollvariablen bei den folgenden Sekundäranalysen werden vier soziobiografische Merkmale der Schülerinnen und Schüler, die Komposition der Unterrichtsgruppe hinsichtlich dieser Merkmale und die Schulart als Indikator für das institutionelle Lernmilieu berücksichtigt.

Geschlecht: Bei der Kodierung des Geschlechts wurde den Mädchen der Wert 1, den Jungen der Wert 0 zugewiesen.

Zu Hause gesprochene Sprache: Auf der Basis von Angaben der Schülerinnen und Schüler wurde eine Aufteilung in zwei Gruppen vorgenommen. Unterschieden wurde zwischen Schülerinnen und Schülern, bei denen (1) die zu Hause gesprochene Sprache überwiegend Deutsch ist, und solchen, bei denen (0) die zu Hause gesprochene Sprache überwiegend nicht Deutsch ist.

Sozialstatus: Der Sozialstatus der Familie der Schülerinnen und Schüler wurde über den Beruf des Vaters (bzw. des Erziehungsberechtigten) mit Hilfe der „Sozialen Selbsteinstufung“ nach Kleining und Moore (1968) erfasst. Die Schülerinnen und Schüler erhielten eine Liste mit verschiedenen Berufen, die in sieben Kategorien (nummeriert von A bis G) eingeteilt waren, und wurden dann aufgefordert, diejenige Kategorie auszuwählen, zu der der Beruf ihres Vaters am ehesten passte. Somit weist die Variable ein siebenstufiges Antwortformat auf.

Allgemeine kognitive Fähigkeit: Die allgemeine kognitive Fähigkeit der Schülerinnen und Schüler wurde in der Drei-Länder-Studie mit den 46 Items des figuralen Grundintelligenztests CFT 2 (Cattell & Weiß, 1972) erfasst. Die hier dargestellte Sekundäranalyse verwendet eine IRT-Skalierung auf der Basis des eindimensionalen Rasch-Modells und berechnet WLE-Schätzer als Personenfähigkeitsparameter (Warm, 1989). Die WLE-Reliabilität des Tests beträgt .85 (EAP-Reliabilität .73). Eine ausführliche Beschreibung des Vorgehens bei der Testskalierung sowie der Inhaltsbereiche des Tests findet sich in Kapitel 6.4.

Komposition: Die sprachliche, soziale und kognitive Zusammensetzung sowie der Mädchenanteil wurden auf der Ebene der Unterrichtsgruppe in Form des Mittelwerts aggregiert, um die Komposition der Unterrichtsgruppe kontrollieren zu können.

Schulart: Bei den Schularten wurden Hauptschulen, Realschulen, Gesamtschulen und Gymnasien in Form von Dummy-Variablen unterschieden. Die Realschulen bilden die Referenzkategorie. Bei den Sekundäranalysen der Dreiländer-Studie ist es besonders wichtig, die Schulart zu kontrollieren, da diese Studie aufgrund ihres Erhebungsziels keine ausgeglichene Verteilung der Schularten aufweist (siehe Abschnitt 6.1.2).

7.3 Auswertung

Bei den nachfolgenden Analysen zur ersten Fragestellung (siehe Kapitel 5) wird die Wahrnehmung verschiedener Unterrichtsmerkmale aus den Bereichen des unterstützenden Unterrichtsklimas und der Klassenführung durch die Schülerinnen und Schüler als abhängige Variable untersucht. Prädiktoren zur Erklärung der Unterschiede bei der Unterrichtswahrnehmung durch die Schülerinnen und Schüler sind die Überzeugungen der Lehrkräfte zu Reformbereitschaft, zur Notwendigkeit von Druck bei Lernprozessen und zur Förderung schwacher Schülerinnen und Schüler.

In den Analysen der zweiten und der dritten Fragestellung werden das Wohlfühlen im Unterricht und die Leistungen der Schülerinnen und Schüler der neunten Jahrgangsstufe als Zielkriterien schulischen Lernens untersucht. Prädiktoren zur Erklärung der Unterschiede bei affektiven und kognitiven Ergebnissen stellen auch hier die Überzeugungen der Lehrkräfte zu Reformbereitschaft, der Notwendigkeit von Druck bei Lernprozessen und der Förderung schwacher Schülerinnen und Schüler dar. Darüber hinaus wird untersucht, ob der Zusammenhang zwischen den Überzeugungen von Lehrkräften und den schulischen Ergebnissen der Schülerinnen und Schüler durch das Unterrichtshandeln vermittelt wird. Dieser vermittelte Zusammenhang wird exemplarisch für die Adaptivität des Unterrichts und die Disziplin im Unterricht betrachtet.

Um diese Fragestellungen zu beantworten, werden Daten benötigt, die sowohl Angaben über die Schülerinnen und Schüler als auch Angaben über ihre Lehrkräfte enthalten. Diese Daten weisen demnach eine hierarchische Struktur auf,

da mehrere Schülerinnen und Schüler jeweils von einer Lehrkraft in Englisch bzw. Mathematik unterrichtet wurden. In dieser Struktur ist davon auszugehen, dass sich die Schülerinnen und Schüler einer Unterrichtsgruppe systematisch ähnlicher sind als die verschiedener Unterrichtsgruppen (Hochweber & Hartig, 2012, S. 3). Darüber hinaus lauten die zweite und die dritte Fragestellung, ob die Variation in den schulischen Ergebnissen von Schülerinnen und Schülern durch die Überzeugungen ihrer Lehrkraft erklärt werden kann und welche Rolle der Unterricht dabei einnimmt – die erklärende Variable liegt hier auf einer höheren Ebene der Aggregation (Lehrkraft, Unterricht) als die zu erklärenden Ergebnisse der Schülerinnen und Schüler.

Die Daten wurden daher mit Hilfe von Mehrebenenanalysen (Hox, 2010; Snijders & Bosker, 1999) auf den zwei hierarchisch geschachtelten Ebenen Schülerin bzw. Schüler (Level 1) und Unterrichtsgruppe (Level 2) mit der Software Mplus 7.11 (Muthén & Muthén, 1998-2013) analysiert. Die Mplus-Analysen wurden teilweise automatisiert durchgeführt, dies erfolgte mit dem R-Paket MplusAutomation (R-Paket Version 0.7, Hallquist & Wiley, 2018). Für jedes Modell werden vollstandardisierte Parameterschätzwerte angegeben. Für diese Standardisierung werden sowohl die Varianzen der abhängigen als auch der unabhängigen Variablen verwendet (Muthén & Muthén, 2015, S. 737).

Die Unterrichtsmerkmale Adaptivität des Englisch- bzw. Mathematikunterrichts bzw. Disziplin im Englisch- bzw. Mathematikunterricht wurden in den Mehrebenenmodellen der zweiten und dritten Fragestellung sowohl als individuelle Wahrnehmung des Unterrichts durch die Schülerinnen und Schüler (Ebene 1) als auch aggregiert in Form des Mittelwerts auf der Ebene der Unterrichtsgruppe als Unterrichtsmerkmal verwendet. Gruehn (2000) hat gezeigt, dass Schülerangaben valide Informationen über das Unterrichtsgeschehen liefern und somit als Unterrichtsmerkmal verwendet werden können. Das auf die Unterrichtsgruppe aggregierte Urteil der Schülerinnen und Schüler über den Unterricht wird im Sinne einer geteilten Wahrnehmung interpretiert, welches am ehesten die tatsächlichen Verhältnisse reflektiert (Hochweber & Hartig, 2012, S. 5).

Bei allen Analysen werden auf individueller Ebene soziobiografische Hintergrundmerkmale (Geschlecht, zu Hause gesprochene Sprache und Sozialstatus) und kognitive Grundfähigkeiten der Schülerinnen und Schüler kontrolliert. Auf

der Ebene der Unterrichtsgruppe werden die Komposition dieser Merkmale sowie die Schulart kontrolliert.

Die Datenbasis für die Analysen zum Englischunterricht bilden 1043 Schülerinnen und Schüler sowie deren Englischlehrkraft der jeweiligen Unterrichtsgruppe (siehe Kapitel 7.1). Die Unterrichtsgruppen sind nicht identisch mit den Klassen, da die Untersuchung auch Schulen enthält, an denen leistungsdifferenziert unterrichtet wurde. Infolge dessen enthält diese Stichprobe 147 Lehrkräfte, woraus sich ebenso viele Unterrichtsgruppen (damit ist je nach Differenzierung der Kurs oder die Klasse gemeint) ergeben. Ein durchschnittliches Cluster auf dieser Ebene enthält 7,1 befragte Schülerinnen und Schüler pro Englisch-Unterrichtsgruppe. Die Datenbasis für die Analysen zum Mathematikunterricht bilden 1188 Schülerinnen und Schüler und 164 Mathematiklehrkräfte (siehe Kapitel 7.1). Daraus ergibt sich eine durchschnittliche Clustergröße von 7,2 befragten Schülerinnen und Schülern pro Mathematik-Unterrichtsgruppe. Die Anzahl der Unterrichtsgruppen als Cluster ist in den hier durchgeführten Analysen mit 147 bzw. 164 sehr hoch (Richtwert: mindestens 40 bis 50, Lüdtke, Robitzsch, Trautwein & Kunter, 2009; mindestens 50, Maas & Hox, 2005), die Anzahl der Schülerinnen und Schüler in jeder Unterrichtsgruppe ist aufgrund der Aufteilung der Klassen in zwei Fächergruppen eher klein. Nach Lüdtke et al. (2009) sollte die Clustergröße bei 15 bis 20, nach Snijders und Bosker (1999, S. 152) bei 5 bis 20 Schülerinnen und Schülern liegen. Allerdings konnten Maas und Hox (2005) in ihrer Simulationsstudie zeigen, dass eher eine geringe Anzahl an Clustern zu Verzerrungen bei der Schätzung der Effekte von Level-2-Variablen führt als eine geringe Clustergröße. Sie stellen auch bei einer Clustergröße von 5 keine Verzerrungen der Parameterschätzer und nur sehr geringe Verzerrungen der Standardfehler von Level-2-Variablen fest. Auch Hox (2010, S. 234) betont die Wichtigkeit einer ausreichenden Clusterzahl gegenüber der Clustergröße: „For accuracy and high power a large number of groups appears more important than a large number of individuals per group.“ Diesen Vorgaben entsprechen die hier vorliegenden Daten.

7.3.1 Mehrebenen-Modellierung

Wie schon zu Beginn dieses Kapitels 7.3 beschrieben, werden in dieser Arbeit Mehrebenenanalysen eingesetzt. Im Folgenden werden kurz die Vorteile bzw. die Notwendigkeit dieser Analysemethode dargestellt. Bei der Auswahl einer geeigneten Methode für die geplanten Analysen musste beachtet werden, dass

die Untersuchung eine gleichzeitige Betrachtung von Lehrkräften sowie Schülerinnen und Schülern beinhaltet. Hier liegt eine hierarchische Mehrebenenstruktur der Daten vor: Schülerinnen und Schüler sind genestet innerhalb von Unterrichtsgruppen. Innerhalb jeder Unterrichtsgruppe ist anzunehmen, dass die Beobachtungen der Schülerinnen und Schüler nicht unabhängig voneinander sind (Gruppenhomogenität); damit wird eine wichtige Annahme der konventionellen Regressionsanalyse verletzt (Bickel, 2007). Nach Opdenakker und van Damme (2000) führt das Ignorieren einer solchen Mehrebenenstruktur bei Analysen dazu, dass die Varianz auf der untersten Ebene (hier: Schülerinnen und Schüler) überschätzt wird. Darüber hinaus weisen dann auch die Koeffizienten und deren Standardfehler Verzerrungen auf. Martínez (2012) macht darüber hinaus darauf aufmerksam, dass insbesondere die Klassenebene (hier: Ebene der Unterrichtsgruppe) in Analysen nicht unberücksichtigt bleiben darf, da sonst nicht nur der Einfluss der Schule an sich unterschätzt wird, sondern auch ein verzerrtes Bild bezüglich der Mechanismen, durch die das schulische Umfeld die Leistungen der Schülerinnen und Schüler beeinflusst, entsteht. In der Sozialforschung ist es wichtig, kontextabhängige Faktoren bei den Analysen zu berücksichtigen: Die Werte der abhängigen Variablen können sowohl auf individueller Ebene als auch auf einer Kontext- oder Gruppenebene von unabhängigen Variablen beeinflusst werden. Die Möglichkeit, individuelle und kontextuelle Effekte bei der gleichen Analyse zu berücksichtigen, ist ein wesentlicher Grund dafür, dass die Mehrebenen-Modellierung in der Untersuchung der Ergebnisse von Schülerinnen und Schülern große Bedeutung erlangt hat (Bickel, 2007, S. 3). Um die Mehrebenenstruktur in Mplus zu modellieren, wurden die Regressionsbeziehungen auf jeder Ebene des Modells spezifiziert („type = twolevel“; siehe Geiser, 2010, S. 199–233 für Mehrebenen-Regressionsmodelle und MacKinnon, 2008, S. 259–266 für Mehrebenen-Mediationsmodelle).

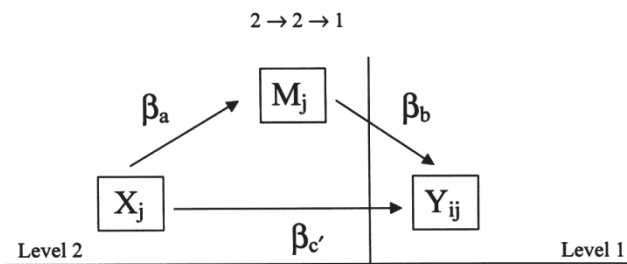
Mehrebenen-Regressionsanalysen

Mithilfe von Mehrebenen-Regressionsanalysen (Langer, 2010) wurden Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und den abhängigen Variablen der verschiedenen Unterrichtsmerkmale aus den Bereichen unterstützendes Unterrichtsklima und Klassenführung untersucht (siehe Fragestellung 1, Kapitel 5). Auch die zweite und dritte Fragestellung zu den Zusammenhängen von Überzeugungen, Unterrichtsmerkmalen und Ergebnissen der Schülerinnen und Schüler wurde zuerst mithilfe von Mehrebenen-Regressionsanalysen untersucht. Bei einer Mehrebenen-Regressionsanalyse ist die abhän-

gige Variable (wie die Englischleistungen von Schülerinnen und Schülern) immer auf der Ebene 1 angesiedelt, die Prädiktoren können sich auf allen Ebenen befinden. Dabei wird die Varianz der abhängigen Variable auf die beiden hierarchischen Ebenen, hier Schülerin bzw. Schüler und Unterrichtsgruppe, zerlegt und es werden regressionsanalytische Prozeduren auf die hierarchischen Datenstrukturen angewendet (Hochweber & Hartig, 2012). Detaillierte statistische Ausführungen zur Mehrebenen-Regressionsanalyse finden sich bei Hox (2010) sowie bei Heck und Thomas (2015).

Mehrebenen-Mediationsanalysen

Um Mediatorbeziehungen der zweiten und dritten Fragestellung (siehe Kapitel 5) zu prüfen, ist eine Mehrebenen-Mediationsanalyse die geeignete Methode. Die Mehrebenen-Mediationsanalyse ermöglicht die Untersuchung von Beziehungen zwischen und innerhalb von Analyseebenen, beispielsweise den Mediationseffekt auf Gruppenebene (MacKinnon, 2008). Darüber hinaus löst sie die durch den Verstoß gegen die Unabhängigkeitsannahme verursachten statistischen Probleme, die die Gruppenhomogenität in hierarchischen Daten hervorruft (siehe oben). Die Standardfehlerschätzungen für das Mehrebenen-Mediationsmodell sind genauer als die für ein Modell mit nur einer Ebene, bei dem allein das Individuum die Analyseeinheit ist, da sie die Abhängigkeit der innerhalb von Gruppen gemessenen Personen berücksichtigen (MacKinnon, 2008, S. 243; auch Krull & MacKinnon, 2001). Für genauere statistische Ausführungen zu Mediationsanalysen wird an dieser Stelle auf MacKinnon (2008) verwiesen.



Quelle: Krull & MacKinnon, 2001, S. 254

Abbildung 7.3: Mediationsmodell

Es gibt verschiedene Varianten der Mehrebenen-Mediation: Die Mediation kann auf der unteren oder auf einer der oberen Ebenen stattfinden. Bei den hier vorliegenden Analysen findet die Mediation durch die Unterrichtsvariable Adaptivität oder Disziplin (M_j) auf der Ebene der Unterrichtsebene (Level 2) statt (siehe Abbildung 7.3). Da die unabhängige Variable (X_j), die jeweilige Überzeugung der Lehrkräfte, auch auf Level 2 liegt und die abhängige Variable (Y_{ij}), das jeweilige Ergebnis der Schülerinnen und Schüler, auf Level 1, handelt es sich um eine sogenannte 2-2-1-Mediation (MacKinnon, 2008, S. 272).

7.3.2 Alphafehler-Adjustierung

Bei den hier durchgeführten Analysen zur Beantwortung der ersten Fragestellung wurden für den Englisch- und den Mathematikunterricht insgesamt jeweils 33 Signifikanztests durchgeführt, da für alle elf Unterrichtsmerkmale in der Drei-Länder-Studie untersucht wird, ob eine der drei betrachteten Überzeugungen von Lehrkräften mit Unterrichtsmerkmalen zusammenhängt. Diese Menge an Signifikanztests kann zu einer Kumulierung des Alphafehlers führen, daher ist eine Adjustierung des α -Niveaus und somit eine Korrektur der p-Werte für die Ergebnisse der ersten Fragestellung sinnvoll. Hierbei wurde das Verfahren der Kontrolle der „False Discovery Rate“ (Falscherkennungsrate) von Benjamini und Hochberg (1995) eingesetzt. Zur Berechnung der adjustierten p-Werte wurde Hemmerich (2019) herangezogen.

7.3.3 Zentrierung

Bei Mehrebenenanalysen spielt die Frage der Zentrierung – vor allem auf L1 – eine wichtige Rolle, um Metriken von Konstrukten einen interpretierbaren Nullpunkt zuzuweisen (Enders & Tofighi, 2007). Bei den folgenden Analysen der zweiten und dritten Fragestellung wurde die individuelle Wahrnehmung der Adaptivität bzw. der Disziplin im Fachunterricht auf L1 am Gruppenmittelwert („group-mean“) zentriert. Das auf L2 aggregierte Unterrichtsmerkmal Adaptivität bzw. Disziplin ist von zentralem Interesse und bildet nun im Modell den unbereinigten „Between-group“-Regressionskoeffizienten ab (Lüdtke et al., 2009), das heißt, der Koeffizient ist nicht um Unterschiede bei der individuellen Wahrnehmung der Adaptivität bzw. der Disziplin des Fachunterrichts bereinigt. Dieses Vorgehen ist angebracht, da angenommen wird, dass alle individuellen Wahrnehmungen das gleiche Konstrukt auf der Ebene des Unterrichts messen und dass das individuelle Rating beispielsweise der Adaptivität einer Schülerin

bzw. eines Schülers in ihrer/seiner Unterrichtsgruppe von der tatsächlichen Adaptivität in dieser Unterrichtsgruppe abhängt. Darüber hinaus können Unterschiede zwischen den Unterrichtsgruppen bei den individuellen Wahrnehmungen der Lernumgebung bereits als ein Effekt des Gruppenkonstrukts „Adaptivität des Unterrichts“ bzw. „Disziplin im Unterricht“ angesehen werden, der durch die Auskunft von Schülerinnen und Schüler über ihre Lernumgebung gemessen wird.

Alle anderen Variablen auf der Ebene der Schülerinnen und Schüler (L1) wurden an ihrem Gesamtmittelwert („grand-mean“) zentriert; auch die Dummy-Variablen für das Geschlecht und die zu Hause gesprochene Sprache (Enders & Tofighi, 2007). Dieses Vorgehen ist angebracht, da das primäre Interesse auf den Effekten der Variablen der Unterrichtsgruppe bzw. der Lehrkraft liegt und die Variablen auf der individuellen Ebene als Kontrollvariablen dienen (Enders & Tofighi, 2007; Lüdtke et al., 2009).

Alle Variablen auf der Ebene der Unterrichtsgruppe (L2) wurden an ihrem Gesamtmittelwert zentriert, darunter auch die Dummy-Variablen für die Schulart (Enders & Tofighi, 2007). Dies betrifft die Überzeugungen der Lehrkräfte, die Kompositionsmerkmale der Unterrichtsgruppe und das aggregierte Unterrichtsmerkmal Adaptivität des Englisch- bzw. Mathematikunterrichts oder Disziplin im Englisch- bzw. Mathematikunterricht.

7.3.4 Umgang mit fehlenden Werten

Zum Umgang mit den fehlenden Werten aller Variablen der folgenden Mehrebenenanalysen wurde ein modellbasiertes Verfahren gewählt. Der Modellschätzung liegt ein ML-Verfahren („Maximum Likelihood“) zugrunde, wobei die Parameter und ihre Standardfehler auf der Grundlage der beobachteten Daten geschätzt werden (FIML: Full Information Maximum Likelihood, siehe Lüdtke, Robitzsch, Trautwein & Köller, 2007; ausführlicher bei Enders, 2010). Die FIML-Methode enthält in der Likelihood-Funktion sowohl die festen Regressionsparameter als auch die Varianzkomponenten der Fehlerterme (Langer, 2010, S. 750).

An dieser Methode kann kritisiert werden, dass sie nur auf den beobachteten Daten basiert, das heißt der Mechanismus, aufgrund dessen die Daten fehlen, wird ignoriert (Schafer & Graham, 2002, S. 162). Darüber hinaus unterstellen

Likelihood-Verfahren ein „missing at random“ (MAR⁵²). Diese Annahme ist eher unrealistisch, allerdings sehen Schafer und Graham (2002, S. 164) die so verursachte Abweichung als nicht groß genug an, um die Ergebnisse einer Analyse, die auf MAR beruht, tatsächlich zu entkräften. Deshalb empfehlen sie Likelihood-Verfahren für den Umgang mit fehlenden Werten. Auch Lüdtke und Kollegen (2007, S. 112) resümieren, „dass bei Verletzung der Annahmen MAR und multivariate Normalverteilung verzerrte Schätzungen resultieren können, letztendlich aber auch in dieser Datensituation die FIML-Methode erheblich bessere Ergebnisse zeigt als klassische Verfahren zur Behandlung fehlender Werte.“

Bei den Variablen, die bei den folgenden Sekundäranalysen verwendet wurden, lagen die Anteile der fehlenden Werte zwischen 1,9 und 11,1 Prozent auf der Ebene der Schülerinnen und Schüler. Tabelle 7.12 gibt die fehlenden Datenraten für jede der Analysevariablen wieder. Da die Schulform Bestandteil der Personen-ID der Schülerinnen und Schüler ist, gab es für diese Variable keine fehlenden Werte.

52 „Missing at random“ (MAR) bedeutet, dass „nach Kontrolle von zusätzlichen beobachteten Variablen [...] das Auftreten von fehlenden Werten weder von der Ausprägung der Variablen selbst noch der Ausprägung anderer Variablen des Datensatzes ab[hängt]“ (Lüdtke, Robitzsch, Trautwein & Köller, 2007, S. 104).

Tabelle 7.12: Anteil fehlender Angaben bei Variablen der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie

		Englischunterricht		Mathematikunterricht	
		N	Prozent	N	Prozent
Unterrichtsmerkmale	Unterstützung	62	5,9	132	11,1
	Strenge	36	3,5	89	7,5
	Adaptivität	57	5,5	32	2,7
	Schülerorientiertheit	60	5,8	33	2,8
	Individualisierungsmaßnahmen	66	6,3	40	3,4
	Motivierungsbemühen	94	9,0	76	6,4
	Vertrauensgrad	61	5,8	26	2,2
	Strukturiertheit	57	5,5	35	2,9
	Beschäftigungsradius	66	6,3	48	4,0
	Disziplin	63	6,0	30	2,5
	Zeitverschwendung	39	3,7	104	8,8
Kontrollvariablen	Geschlecht	76	7,3	87	7,3
	Sprache: Deutsch	95	9,1	88	7,4
	Sozialstatus	110	10,5	126	10,6
	CFT	88	8,4	107	9,0
Schulische Ergebnisse	Wohlfühlen im Unterricht	56	5,4	99	8,3
	Leistungstest	32	3,1	23	1,9
Gesamt		1043	100	1188	100

Von den 147 Englisch- und den 164 Mathematiklehrkräften fehlen bei ihren Überzeugungen jeweils zwischen 12 und 18 Prozent der Angaben (siehe Tabelle 7.13). Bei den Kompositionsvariablen der Unterrichtsgruppe (sprachliche, soziale und kognitive Zusammensetzung sowie Mädchenanteil) gibt es für den Englischunterricht ein bis zwei Unterrichtsgruppen, für die diese Angabe nicht berechnet werden konnte (Mathematikunterricht: ein fehlender Wert bei sozialer, zwei bei kognitiver Zusammensetzung; in der Tabelle nicht ausgewiesen). In diesen Fällen liegt die Clustergröße der Unterrichtsgruppe bei fünf oder weniger Schülerinnen und Schülern; wenn bei ihnen allen keine Angabe zu der betroffenen Variable vorliegt, weist auch die aggregierte Variable für die Unterrichtsgruppe einen fehlenden Wert auf (zur Clustergröße siehe Seite 138). Dies betrifft auch die Aggregation der Unterrichtsmerkmale: Bei zehn der ins-

gesamt elf Unterrichtsmerkmale weisen ein bis zwei Unterrichtsgruppen für Englisch keinen Wert auf (Mathematikunterricht: fünf Unterrichtsmerkmale).

Tabelle 7.13: Anteil fehlender Angaben bei Überzeugungen der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie

	Englischlehrkräfte		Mathematiklehrkräfte	
	N	%	N	%
Reformbereitschaft	21	14,3	29	17,7
Notwendigkeit von Druck bei Lernprozessen	23	15,6	27	16,5
Förderung von schwachen Schülerinnen und Schülern	17	11,6	20	12,2
Gesamt	147	100	164	100

8 Empirische Analysen

Im empirischen Teil der vorliegenden Arbeit werden die Forschungsfragen, die im ersten Teil dieser Arbeit hergeleitet und in Kapitel 5 aufgestellt wurden, beantwortet. Zunächst wird in Abschnitt 8.1 die Datengrundlage für die vorliegenden Analysen beschrieben, welche die deskriptive Verteilungen und Korrelationen der wichtigsten Analysevariablen enthält. In Abschnitt 8.2.1 wird die erste Forschungsfrage untersucht, die nach dem Zusammenhang zwischen den Überzeugungen von Lehrkräften und dem Unterrichtshandeln fragt. Danach werden Zusammenhänge zwischen den Überzeugungen von Lehrkräften und den schulischen Ergebnissen behandelt (Abschnitt 8.2.2). Diese beziehen sich für das Wohlfühlen der Schülerinnen und Schüler auf die zweite und für die Fachleistungen auf die dritte Forschungsfrage. In den dortigen Mehrebenen-Regressionen erhält man auch erste Hinweise auf eine mögliche Vermittlung des Einflusses von Überzeugungen über den Unterricht auf die Ergebnisse der Schülerinnen und Schüler. Vollständig beantworten lässt sich die Frage der Vermittlung von Überzeugungen über den Unterricht erst im Abschnitt 8.2.3. Dies wird dort mittels Mediationsmodellen überprüft.

8.1 Beschreibung der Datengrundlage

Bei der Darstellung der deskriptiven Ergebnisse werden die Verteilungen der für diese Arbeit untersuchten Variablen sowie deren Korrelationen mit den Überzeugungen der Lehrkräfte beschrieben.

8.1.1 Deskriptive Verteilung der Überzeugungen, der Unterrichtsmerkmale und der schulischen Ergebnisse

In Tabelle 8.1 sind die Mittelwerte und Standardabweichungen sowie die Fallzahlen und Spannweiten der untersuchten Variablen für die Überzeugungen der Lehrkräfte, die Unterrichtsmerkmale und die schulischen Ergebnisse der Schülerinnen und Schüler im Englischunterricht abgebildet. Darüber hinaus findet sich dort der ICC-Wert (Messung der Intraklassenkorrelation) für die abhängigen Variablen. Tabelle 8.2 enthält die gleichen Angaben für den Mathematikunterricht. Die Mittelwerte der Überzeugungen der Lehrkräfte liegen für beide Fächer im mittleren Ausprägungsbereich. In beiden Fächern ist der Mittelwert

der Skala „positive Einstellung zur Förderung von schwachen Schülerinnen und Schülern“ höher ausgeprägt als der von „Reformbereitschaft“, wobei bei der Skala der Einstellung zur Förderung der Maximalwert der Ausprägungen zudem niedriger ist als bei der Reformbereitschaft. Da die Skala „Einstellung zur Notwendigkeit von Druck bei Lernprozessen“ negativ gepolt ist und ihr Mittelwert bei 3,8 liegt, hat sie einen ähnlichen Ausprägungsbereich wie die positiv geprägte Reformbereitschaft. Bei den Unterrichtsmerkmalen ist in beiden Fächern die Skala „Strenge“ am niedrigsten und „Individualisierungsmaßnahmen“ am höchsten ausgeprägt.

Tabelle 8.1: Deskriptive Verteilung der Überzeugungen von Lehrkräften, der Unterrichtsmerkmale und der schulischen Ergebnisse von Schülerinnen und Schülern im Englischunterricht

		N	Min.	Max.	MW	SD	ICC	
Überzeugungen von Lehrkräften ⁵³	Reformbereitschaft	882	1	6	3.20	0.922	–	
	Notwendigkeit von Druck	869	1	5.67	3.82	1.021	–	
	Einstellung zu Förderung	900	1	5	3.41	0.675	–	
Unterrichtsmerkmale	unterstützendes Unterrichtsklima	Unterstützung	981	1	5	2.76	0.864	.265
		Strenge	1007	1	5	2.01	0.818	.237
		Adaptivität des Unterrichts	986	1	5	3.50	0.819	.293
		Positive Schülerorientiertheit/Lehrer-Schüler-Beziehung	983	1	5	3.38	0.913	.416
		Individualisierungsmaßnahmen/Förderungsorientiertheit	977	1	5	3.76	0.960	.305
		Motivierungsbemühen/Lob	949	1	5	3.25	1.016	.285
		Vertrauensgrad/Schülerorientierung	982	1	5	3.57	0.974	.342
		Klassenführung	Strukturiertheit des Unterrichts	986	1	5	3.58	0.899
	Beschäftigungsradius		977	1	5	3.73	0.830	.182
	Disziplin und Ordnungssinn		980	1	5	3.68	0.777	.205
	Zeitverschwendung		1004	1	5	3.13	0.945	.293
	Schulische Ergebnisse	Wohlfühlen im Englischunterricht	987	1	5	3.24	1.074	.195
Englischleistungen		1011	–3.264	4.110	.209	1.182	.593	

Anmerkungen: N_{gesamt} = 1043; MW: Mittelwert; SD: Standardabweichung

Die Intraklassenkorrelationen der Unterrichtsmerkmale zeigen, dass eine Mehrebenenanalyse bei der Untersuchung dieser Merkmale angebracht ist. Am höchsten ist die Intraklassenkorrelation bei der positiven Schülerorientiertheit, bei der über 40 Prozent der Varianz durch die Ebene der Unterrichtsgruppe

53 Die Mittelwerte und Standardabweichungen der Überzeugungen der Lehrkräfte können sich in Tabelle 8.1 und Tabelle 8.2 leicht von den Werten in Tabelle 7.7 unterscheiden, da die Berechnungsgrundlage hier die Datensätze der Schülerinnen und Schüler mit den angespielten Daten der Lehrkräfte sind.

erklärt werden können. Dieses Unterrichtsmerkmal bezieht sich auf die allgemeine Wahrnehmung der Lehrkraft durch die Schülerinnen und Schüler der Unterrichtsgruppe (Beispielitem: „Unser Lehrer ist beliebt bei den Schülern.“ vs. „...ist unbeliebt bei den Schülern.“). Es zeigen sich keine Unterschiede zwischen den beiden Fächern. Für das Fach Englisch zeigt sich beim Unterrichtsmerkmal „Strenge“ eine höhere Intraklassenkorrelation mit 32 Prozent Varianz auf der Ebene der Unterrichtsgruppe, im Vergleich zu 24 Prozent im Mathematikunterricht. Bei der Strukturiertheit des Unterrichts ist die Intraklassenkorrelation in Mathematik deutlich höher als in Englisch (35 zu 27 Prozent). Die niedrigsten Werte beim ICC mit .18 bzw. .19 weist die Variable Beschäftigungsradius auf. Dies zeigt, dass eher wenig Varianz der Angaben der Schülerinnen und Schüler zu Aussagen wie „Der Lehrer ruft auch häufig Schüler auf, die sich nicht gemeldet haben“ durch die Ebene der Unterrichtsgruppe erklärt werden kann. Die individuelle Wahrnehmung der einzelnen Schülerin oder des einzelnen Schülers erklärt hier den größten Teil der Varianz des Beschäftigungsradius.

Bei den Ergebnissen der Schülerinnen und Schüler liegt die Varianzerklärung auf der Unterrichtsebene für das Wohlfühlen im Unterricht bei 20 Prozent für Englisch und etwas darüber für Mathematik. Bei den Fachleistungen liegt dieser Wert mit 59 Prozent für Englisch und 45 Prozent für Mathematik deutlich höher. Für eine detailliertere Aufschlüsselung der Varianzkomponenten in den Daten der Drei-Länder-Studie sei an dieser Stelle auf die Arbeit von Wurster und Feldhoff (2019) verwiesen. Diese gliedern die Varianzanteile auf bis zu vier Ebenen (Schulform, Schule, Unterrichtsgruppe, Schülerinnen und Schüler) auf, wodurch sich die Varianzerklärung auf der Unterrichtsebene etwas reduziert, aber immer noch bedeutsam ist.

Tabelle 8.2: Deskriptive Verteilung der Überzeugungen von Lehrkräften, der Unterrichtsmerkmale und der schulischen Ergebnisse von Schülerinnen und Schülern im Mathematikunterricht

		N	Min.	Max.	MW	SD	ICC	
Überzeugungen von Lehrkräften	Reformbereitschaft	969	1	5.50	3.01	0.882	–	
	Notwendigkeit von Druck	973	1.50	5.67	3.79	0.894	–	
	Einstellung zu Förderung	1061	1.67	5	3.51	0.610	–	
Unterrichtsmerkmale	unterstützendes Unterrichtsklima	Unterstützung	1056	1	5	2.81	0.853	.326
		Strenge	1099	1	5	1.98	0.791	.324
		Adaptivität des Unterrichts	1156	1	5	3.56	0.786	.242
		Positive Schülerorientiertheit/Lehrer-Schüler-Beziehung	1155	1	5	3.42	0.866	.432
		Individualisierungsmaßnahmen/Förderungsorientiertheit	1148	1	5	3.82	0.906	.259
		Motivierungsbemühen/Lob	1112	1	5	3.17	0.953	.239
		Vertrauensgrad/Schülerorientierung	1162	1	5	3.55	0.954	.332
	Klassenführung	Strukturiertheit des Unterrichts	1153	1	5	3.66	0.820	.267
		Beschäftigungsradius	1140	1	5	3.75	0.789	.187
		Disziplin und Ordnungssinn	1158	1	5	3.66	0.751	.227
		Zeitverschwendung	1084	1	5	2.80	0.860	.294
Schulische Ergebnisse	Wohlfühlen im Mathematikunterricht	1089	1	5	3.49	1.006	.235	
	Mathematikleistungen	1165	–3.063	3.440	.037	0.958	.451	

Anmerkungen: N_{gesamt} = 1188; MW: Mittelwert; SD: Standardabweichung

8.1.2 Korrelationen der drei Überzeugungen der Lehrkräfte

Nach der Betrachtung der deskriptiven Verteilungen aller interessierenden Variablen werden nun Korrelationen wiedergegeben, zuerst zwischen den drei Überzeugungen der Lehrkräfte im Englisch- und Mathematikunterricht (siehe Tabelle 8.3). Im Englischunterricht hängt die Reformbereitschaft der Lehrkräfte stark negativ mit der Einstellung zur Notwendigkeit von Druck bei Lernprozes-

sen zusammen (Pearsons $r = -.629, p < .000$). Positiv hingegen hängt die Reformbereitschaft mit der Einstellung zur Förderung schwacher Schülerinnen und Schüler (Pearsons $r = .504, p < .000$) zusammen. Die Einstellung zur Förderung hängt wiederum negativ mit der Einstellung zur Notwendigkeit von Druck zusammen (Pearsons $r = -.555, p < .000$).

Auch im Mathematikunterricht hängen die Reformbereitschaft und die Einstellung zur Notwendigkeit von Druck am stärksten zusammen ($r = -.651, p < .000$), die anderen beiden Zusammenhänge liegen auf einem etwas niedrigeren Niveau als im Englischunterricht: Reformbereitschaft und Einstellung zur Förderung: $r = .462, p < .000$; Notwendigkeit von Druck und Einstellung zur Förderung: $r = -.467, p < .000$. Die Richtung der Zusammenhänge weist überall das erwartete Vorzeichen auf. Die drei Merkmale für die Überzeugungen für Lehrkräfte korrelieren sehr hoch. Deswegen werden bei den folgenden Analysen jeweils getrennte Modelle für Reformbereitschaft, Notwendigkeit von Druck bei Lernprozessen und der Einstellung zur Förderung schwacher Schülerinnen und Schüler berechnet, um Verzerrungen aufgrund von Multikollinearität zu vermeiden.

Tabelle 8.3: Korrelationen der Überzeugungen von Lehrkräften

		Englischunterricht		Mathematikunterricht	
		Notwendigkeit von Druck	Einstellung zu Förderung	Notwendigkeit von Druck	Einstellung zu Förderung
Reformbereitschaft	r	-.629	.504	-.651	.462
	p	.000	.000	.000	.000
	N	861	827	957	939
Notwendigkeit von Druck	r	/	-.555	/	-.467
	p	/	.000	/	.000
	N	/	814	/	937

8.1.3 Korrelationen der Überzeugungen mit Unterrichtsmerkmalen und schulischen Ergebnissen

Bevor im nächsten Kapitel mehrebenenanalytisch und unter der Berücksichtigung von Kovariaten der Zusammenhang zwischen den Überzeugungen der Lehrkräfte und den Unterrichtsmerkmalen untersucht wird, werden im Folgenden die Korrelationen dieser Merkmale berichtet. In Tabelle 8.4 sind die Korrelationen von Überzeugungen und Unterrichtsmerkmalen im Englischunterricht,

in Tabelle 8.5 die für den Mathematikunterricht aufgeführt. Dort sind jeweils die Unterrichtsmerkmale Adaptivität und Disziplin grau hinterlegt. Diese beiden Unterrichtsmerkmale werden zur Beantwortung der zweiten und dritten Fragestellung dieser Arbeit herangezogen, um die Vermittlung der Zusammenhänge der Überzeugungen der Lehrkräfte mit den schulischen Ergebnissen der Schülerinnen und Schüler zu erklären.

Reformbereitschaft

Die Reformbereitschaft der Lehrkräfte im Englischunterricht hängt sowohl mit den von den Schülerinnen und Schülern wahrgenommenen Unterrichtsmerkmalen als auch mit ihren schulischen Ergebnissen zusammen. Bei den Unterrichtsmerkmalen Strenge und Disziplin zeigt sich – wie erwartet – ein negativer Zusammenhang, bei den weiteren Unterrichtsmerkmalen ein positiver. Lediglich bei den Unterrichtsmerkmalen Beschäftigungsradius und Zeitverschwendung treten keine Zusammenhänge auf. Mit dem Wohlfühlen der Schülerinnen und Schüler zeigt sich ein positiver, mit den Englischleistungen ein negativer Zusammenhang (siehe Tabelle 8.4). Im Mathematikunterricht zeigen sich deutlich weniger Zusammenhänge. So hängt die Reformbereitschaft der Lehrkräfte nur mit dem Wohlfühlen der Schülerinnen und Schüler, nicht jedoch mit deren Mathematikleistungen zusammen. Bei den Unterrichtsmerkmalen im Fach Mathematik zeigen sich vor allem bei denen aus dem Bereich des unterstützenden Unterrichtsklimas Zusammenhänge mit der Reformbereitschaft der Lehrkräfte: ein positiver Zusammenhang mit Unterstützung, Schülerorientiertheit, Individualisierungsmaßnahmen und Vertrauensgrad sowie ein negativer Zusammenhang mit Strenge (siehe Tabelle 8.5). Zudem tritt ein negativer Zusammenhang mit der Disziplin im Unterricht auf.

Tabelle 8.4: Korrelationen der Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen und schulischen Ergebnissen der Schülerinnen und Schüler im Englischunterricht

		Reformbereit- schaft		Notwendigkeit von Druck		Einstellung zu Förderung	
		r	p	r	p	r	p
unterstützendes Unterrichtsklima	Unterstützung	.245	.000	-.213	.000	.189	.000
	Strenge	-.106	.002	.170	.000	-.136	.000
	Adaptivität des Unterrichts	.121	.000	-.130	.000	.130	.000
	Positive Schülerorientiertheit/Lehrer-Schüler-Beziehung	.199	.000	-.243	.000	.150	.000
	Individualisierungsmaßnahmen/Förderungsorientiertheit	.105	.003	-.128	.000	.121	.000
	Motivierungsbemühen/Lob	.194	.000	-.245	.000	.160	.000
	Vertrauensgrad/Schülerorientierung	.199	.000	-.215	.000	.170	.000
Klassenführung	Strukturiertheit des Unterrichts	.116	.001	-.107	.002	.127	.000
	Beschäftigungsradius	.042	.231	-.040	.251	.092	.007
	Disziplin und Ordnungssinn	-.156	.000	.108	.002	.022	.515
	Zeitverschwendung	.067	.050	-.040	.248	.019	.567
	Wohlfühlen im Englischunterricht	.132	.000	-.116	.001	.176	.000
Englischleistungen		-.196	.000	-.056	.106	-.045	.188

Anmerkungen: signifikante Korrelationskoeffizienten r: fett ($p < .05$); $N_{\text{Min.}} = 786$, $N_{\text{Max.}} = 872$

Notwendigkeit von Druck bei Lernprozessen

Im Englischunterricht zeigen sich für die Einstellung der Lehrkräfte zur Notwendigkeit von Druck bei Lernprozessen beinahe die gleichen Zusammenhänge wie für die Reformbereitschaft, jedoch mit umgedrehten Vorzeichen. Allerdings findet sich kein Zusammenhang mit den Englischleistungen (siehe Tabelle 8.4). Im Mathematikunterricht findet sich, wie bei der Reformbereitschaft, nur ein Zusammenhang mit dem Wohlfühlen der Schülerinnen und Schüler – diesmal aber ein negativer –, nicht jedoch mit den Mathematikleistungen. Viele der Unterrichtsmerkmale des Mathematikunterrichts weisen einen Zusammenhang mit der Notwendigkeit von Druck bei Lernprozessen auf: einen negativen Zusammenhang bei Unterstützung, Schülerorientiertheit, Individualisierung, Motivierungsbemühen, Vertrauensgrad und Strukturiertheit sowie einen positiven Zusammenhang bei Strenge und Disziplin (siehe Tabelle 8.5).

Einstellung zur Förderung von schwachen Schülerinnen und Schülern

Die Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler hängt im Englischunterricht mit fast allen Unterrichtsmerkmalen – bis auf Disziplin und Zeitverschwendung – zusammen. Alle Zusammenhänge sind positiv, nur bei der Strenge gibt es einen negativen Zusammenhang. Wie bei der Reformbereitschaft zeigt sich für die Einstellung zur Förderung ein positiver Zusammenhang mit dem Wohlfühlen der Schülerinnen und Schüler im Englischunterricht, allerdings gibt es – wie schon bei der Notwendigkeit von Druck – keinen Zusammenhang mit den Englischleistungen (siehe Tabelle 8.4). Im Mathematikunterricht gibt es wiederum keinen Zusammenhang der Einstellung zur Förderung schwacher Schülerinnen und Schüler mit dem Wohlfühlen, jedoch einen negativen Zusammenhang mit den Mathematikleistungen. Bei den Unterrichtsmerkmalen im Mathematikunterricht weisen nur wenige einen positiven Zusammenhang mit der Einstellung zur Förderung auf: Vertrauensgrad, Strukturiertheit und Zeitverschwendung (siehe Tabelle 8.5).

Tabelle 8.5: Korrelationen der Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen und den schulischen Ergebnissen der Schülerinnen und Schüler im Mathematikunterricht

		Reformbereit- schaft		Notwendigkeit von Druck		Einstellung zu Förderung	
		r	p	r	p	r	p
unterstützendes Unterrichtsklima	Unterstützung	.105	.002	-.130	.000	.054	.095
	Strenge	-.074	.026	.215	.000	-.021	.505
	Adaptivität des Unterrichts	.019	.554	-.062	.056	.044	.162
	Positive Schülerorientier- theit/Lehrer-Schüler-Beziehung	.083	.011	-.173	.000	.033	.286
	Individualisierungsmaßnah- men/Förderungsorientiertheit	.069	.035	-.133	.000	.028	.364
	Motivierungsbemühen/Lob	.055	.099	-.070	.034	.015	.641
	Vertrauensgrad/Schüler- orientierung	.077	.017	-.144	.000	.082	.008
Klassenführung	Strukturiertheit des Unterrichts	.060	.064	-.099	.002	.064	.040
	Beschäftigungsradius	.004	.910	-.015	.639	-.002	.954
	Disziplin und Ordnungssinn	-.083	.010	.163	.000	-.005	.869
	Zeitverschwendung	.042	.206	-.047	.163	.078	.015
Wohlfühlen im Mathematik- unterricht		.124	.000	-.134	.000	.053	.100
Mathematikleistungen		-.035	.282	.019	.569	-.188	.000

Anmerkungen: signifikante Korrelationskoeffizienten r: fett ($p < .05$); $N_{\text{Min.}} = 865$, $N_{\text{Max.}} = 1039$

Zusammenfassend lässt sich sagen, dass es auf korrelativer Ebene bedeutsame Zusammenhänge zwischen den Überzeugungen der Lehrkräfte sowohl mit den Unterrichtsmerkmalen als auch mit Ergebnissen der Schülerinnen und Schüler gibt. Auch zeigen sich hier schon Unterschiede zwischen dem Englisch- und dem Mathematikunterricht. Diese Ergebnisse sind deskriptiver Natur und erlauben noch keine Beantwortung der Forschungsfragen. Fragen nach Zusammenhängen auf der Grundlage bivariater Korrelationen reichen dafür nicht aus. Erst in Mehrebenen-Analysen und durch Hinzunahme von Kontrollvariablen, die den Kontext berücksichtigen, lassen sich die Zusammenhänge genauer bestimmen (siehe Abschnitt 7.3.1). Die Ergebnisse der Mehrebenen-Analysen werden im nächsten Abschnitt dargestellt.

8.2 Ergebnisse zur Beantwortung der Forschungsfragen

In Abschnitt 8.2.1 wird die erste Forschungsfrage untersucht, die nach dem Zusammenhang zwischen den Überzeugungen der Lehrkräfte und ihrem von den Schülerinnen und Schülern wahrgenommenen Unterrichtshandeln fragt. Danach werden Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und den schulischen Ergebnissen behandelt (8.2.2). Das Wohlfühlen der Schülerinnen und Schüler bezieht sich auf die zweite und die Fachleistungen auf die dritte Forschungsfrage. In den Mehrebenen-Regressionsanalysen gibt es erste Hinweise auf eine mögliche Vermittlung des Einflusses der Überzeugungen über den Unterricht auf die Ergebnisse der Schülerinnen und Schüler. Die Frage der Vermittlung der Überzeugungen über den Unterricht auf das Wohlfühlen sowie die Fachleistungen der Schülerinnen und Schüler im Mathematik- und Englischunterricht, die auch Teil der zweiten und dritten Forschungsfrage ist, wird im Abschnitt 8.2.3 mittels Mediationsmodellen überprüft.

8.2.1 Zusammenhänge zwischen Überzeugungen und Unterricht

Zunächst werden die Ergebnisse zur ersten Forschungsfrage nach dem Zusammenhang zwischen den Überzeugungen der Lehrkräfte und ihrem unterrichtlichen Handeln im Englisch- bzw. Mathematikunterricht betrachtet. Aus Kollinearitätsgründen (siehe Korrelationen der Überzeugungen in Abschnitt 8.1.2) werden die drei Überzeugungen separat betrachtet. Zu den Zusammenhängen zwischen den Überzeugungen der Lehrkräfte und den Unterrichtsmerkmalen wurden insgesamt 66 Mehrebenenmodelle berechnet (drei Überzeugungen \times elf Unterrichtsmerkmale \times zwei Fächer). Im Folgenden werden exemplarisch die Ergebnisse für die Unterrichtsmerkmale Adaptivität (Abschnitt 8.2.1.1) und Disziplin (Abschnitt 8.2.1.2) ausführlich dargestellt, da diese im Verlauf der weiteren Analysen auch hinsichtlich ihrer Vermittlungsfunktion zwischen den Überzeugungen der Lehrkräfte und den Ergebnissen der Schülerinnen und Schüler betrachtet werden. Die Adaptivität des Unterrichts ist ein bedeutsames Merkmal eines unterstützenden Unterrichtsklimas (siehe Abschnitt 4.3.1), und die Disziplin im Unterricht ist ein wichtiges Merkmal der Klassenführung (siehe Abschnitt 4.3.2). Die Ergebnisse der übrigen 60 Modelle zur ersten Fragestellung sind tabellarisch mit Angaben zu dem zentralen Parameter, dem β -Koeffizient für die Überzeugung der Lehrkraft, in Abschnitt 8.2.1.3 dargestellt. Dort werden ergänzend auch die Ergebnisse zur Adaptivität und zur Disziplin zusammengefasst.

8.2.1.1 Zusammenhänge zwischen Überzeugungen und adaptivem Unterricht

Im folgenden Abschnitt wird der Zusammenhang zwischen den Überzeugungen von Lehrkräften und adaptivem Unterricht aus Sicht der Schülerinnen und Schüler untersucht. Englisch- und Mathematikunterricht werden getrennt betrachtet. Es wird erwartet, dass Lehrkräfte mit einer hohen Reformbereitschaft und einer positiven Einstellung zur Förderung schwacher Schülerinnen und Schüler ihren Unterricht adaptiver gestalten, wohingegen Lehrkräfte, die der Überzeugung sind, dass Druck für Lernen notwendig ist, ihren Unterricht weniger adaptiv gestalten (siehe Kapitel 5, Hypothese 1).

Englischunterricht

Es zeigen sich im Englischunterricht signifikante Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und der von den Schülerinnen und Schülern wahrgenommenen Adaptivität des Unterrichts (vgl. Tabelle 8.6). Je höher die Bereitschaft einer Lehrkraft zu Reformen ist (Modell 1) und je positiver ihre Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler (Modell 3), desto adaptiver gestaltet sie ihren Unterricht. Je stärker die Befürwortung von Druck bei Lernprozessen ausgeprägt ist (Modell 2), desto weniger adaptiv gestaltet eine Lehrkraft den Unterricht. Bei den Kontrollvariablen zeigt sich für die Gesamtschule nur dann ein signifikanter Effekt, wenn man den Zusammenhang von Reformbereitschaft sowie Notwendigkeit von Druck und Adaptivität betrachtet; hier ist adaptiver Unterricht in Gesamtschulen stärker ausgeprägt als in Realschulen. Mädchen nehmen den Englischunterricht bei den vorliegenden Analysen grundsätzlich adaptiver wahr als Jungen. In den Modellen 1 bis 3 liegt der Anteil der erklärten Varianz der „Adaptivität des Englischunterrichts“ auf der Ebene der Unterrichtsgruppe bei 13 (Modell 1) bis 18 Prozent (Modelle 2 und 3).

Tabelle 8.6: Mehrebenenanalysen zur Adaptivität des Englischunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte

	Modell 1			Modell 2			Modell 3		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in									
Mädchen	.114	(.043)	.008	.114	(.043)	.008	.113	(.043)	.009
Sprache Deutsch	.051	(.070)	.469	.051	(.070)	.464	.050	(.070)	.472
Sozialstatus	-.025	(.037)	.502	-.025	(.037)	.503	-.025	(.037)	.504
CFT	-.033	(.038)	.387	-.032	(.038)	.404	-.034	(.038)	.374
R ²		.017			.017			.017	
Ebene: Unterrichtsgruppe									
Reformbereitschaft	.178	(.090)	.049	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	-.275	(.085)	.001	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	.278	(.104)	.007
Komposition: Mädchen	-.004	(.095)	.967	.005	(.096)	.961	-.021	(.094)	.823
Sprachkomposition Deutsch	.096	(.114)	.399	.091	(.114)	.426	.071	(.114)	.532
Soziale Komposition	-.089	(.098)	.363	-.089	(.094)	.346	-.110	(.091)	.224
CFT-Komposition	.029	(.108)	.792	-.009	(.113)	.938	.037	(.101)	.713
Hauptschulen ¹	.037	(.143)	.794	.092	(.145)	.523	.019	(.146)	.898
Gesamtschulen ¹	.307	(.146)	.035	.355	(.146)	.015	.292	(.156)	.060
Gymnasien ¹	.005	(.112)	.963	.007	(.107)	.950	-.016	(.113)	.885
R ²		.130			.184			.179	

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1043; ¹Realschule ist Referenzkategorie

Mathematikunterricht

Bei Betrachtung der Korrelationen der Überzeugungen der Lehrkräfte mit der von den Schülerinnen und Schülern wahrgenommenen Adaptivität des Mathematikunterrichts zeigten sich bereits in Abschnitt 8.1.3 (siehe dort Tabelle 8.5) keine signifikanten Zusammenhänge. Dieses Ergebnis bleibt auch bei Berücksichtigung der Mehrebenenstruktur und unter Einbezug sämtlicher Kontrollva-

riablen bestehen (vgl. Tabelle 8.7). Auch zeigen die Kontrollvariablen keine signifikanten Effekte. Jedoch weist die kognitive Komposition der Unterrichtsgruppe in allen drei Modellen einen positiven Zusammenhang mit der Adaptivität des Mathematikunterrichts auf. Der Anteil der aufgeklärten Varianz auf der Ebene der Unterrichtsgruppe liegt hier zwischen 10 und 14 Prozent.

Tabelle 8.7: Mehrebenenanalysen zur Adaptivität des Mathematikunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte

	Modell 1			Modell 2			Modell 3		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in									
Mädchen	.023	(.042)	.586	.023	(.042)	.588	.023	(.042)	.588
Sprache Deutsch	.032	(.042)	.437	.032	(.042)	.440	.032	(.042)	.442
Sozialstatus	-.025	(.036)	.485	-.025	(.036)	.484	-.025	(.036)	.488
CFT	-.053	(.040)	.186	-.053	(.040)	.187	-.052	(.040)	.195
R ²		.005			.005			.005	
Ebene: Unterrichtsgruppe									
Reformbereitschaft	.018	(.114)	.874	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	-.150	(.126)	.234	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	.163	(.095)	.084
Komposition: Mädchen	-.139	(.121)	.249	-.166	(.121)	.170	-.138	(.118)	.244
Sprachkomposition Deutsch	-.092	(.078)	.240	-.101	(.074)	.173	-.114	(.076)	.131
Soziale Komposition	-.012	(.128)	.924	-.028	(.129)	.830	-.005	(.125)	.965
CFT-Komposition	.187	(.094)	.046	.193	(.092)	.037	.218	(.093)	.019
Hauptschulen ¹	-.017	(.164)	.916	-.035	(.168)	.834	-.041	(.167)	.807
Gesamtschulen ¹	.182	(.161)	.258	.140	(.173)	.419	.145	(.166)	.381
Gymnasien ¹	-.008	(.126)	.947	-.028	(.134)	.835	-.029	(.128)	.822
R ²		.102			.128			.140	

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1188; ¹Realschule ist Referenzkategorie

8.2.1.2 Zusammenhänge zwischen Überzeugungen und Disziplin im Unterricht

In folgendem Kapitel wird die Frage des Zusammenhangs zwischen den Überzeugungen von Lehrkräften und der Disziplin im Unterricht aus der Sicht der Schülerinnen und Schüler untersucht. Zunächst wird der Englisch-, dann der Mathematikunterricht betrachtet. Dabei wird erwartet, dass Lehrkräfte mit einer hohen Reformbereitschaft und einer positiven Einstellung zur Förderung schwacher Schülerinnen und Schüler der Disziplin im Unterricht eher eine niedrige Bedeutung beimessen. Unterrichtsgruppen hingegen, deren Lehrkräfte der Überzeugung sind, dass Druck für Lernprozesse notwendig ist, weisen eine höhere Disziplin im Unterricht auf (siehe Kapitel 5, Hypothese 2).

Englischunterricht

Im Englischunterricht besteht ausschließlich für die Reformbereitschaft von Lehrkräften ein signifikanter Zusammenhang mit der von den Schülerinnen und Schülern wahrgenommenen Disziplin im Unterricht (vgl. Tabelle 8.8). Je höher die Bereitschaft einer Lehrkraft zu Reformen ist (Modell 1), desto weniger diszipliniert wird der Unterricht wahrgenommen. Bei den Kontrollvariablen zeigt sich für die Hauptschule ein signifikanter Effekt, wenn man den Zusammenhang von Reformbereitschaft und Disziplin im Unterricht betrachtet; die Disziplin im Unterricht ist an Hauptschulen stärker ausgeprägt als an Realschulen. Bei Modell 1 liegt der Anteil der erklärten Varianz der „Disziplin im Englischunterricht“ auf der Ebene der Unterrichtsgruppe bei 19 Prozent.

Tabelle 8.8: Mehrebenenanalysen zur Disziplin im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte

	Modell 1			Modell 2			Modell 3		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in									
Mädchen	.019	(.040)	.641	.019	(.040)	.633	.019	(.040)	.626
Sprache Deutsch	.075	(.053)	.160	.075	(.053)	.161	.075	(.053)	.159
Sozialstatus	-.022	(.037)	.555	-.022	(.037)	.552	-.022	(.037)	.549
CFT	-.039	(.041)	.348	-.039	(.041)	.353	-.037	(.041)	.366
R ²		.008			.008			.008	
Ebene: Unterrichtsgruppe									
Reformbereitschaft	-.235	(.107)	.028	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	.144	(.128)	.259	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	.069	(.128)	.590
Komposition: Mädchen	.073	(.133)	.584	.083	(.131)	.525	.099	(.130)	.447
Sprachkomposition Deutsch	.156	(.206)	.447	.165	(.206)	.424	.160	(.208)	.441
Soziale Komposition	.034	(.135)	.801	.035	(.136)	.795	.028	(.136)	.837
CFT-Komposition	.229	(.129)	.077	.233	(.132)	.077	.204	(.125)	.103
Hauptschulen ¹	.308	(.140)	.028	.271	(.146)	.063	.291	(.156)	.062
Gesamtschulen ¹	.210	(.136)	.124	.165	(.141)	.244	.175	(.152)	.250
Gymnasien ¹	.054	(.148)	.717	.074	(.149)	.621	.090	(.153)	.557
R ²		.186			.145			.121	

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1043; ¹Realschule ist Referenzkategorie

Mathematikunterricht

Im Unterschied zum Englischunterricht liegt im Mathematikunterricht kein signifikanter Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und der Disziplin im Unterricht vor. Allerdings besteht hier ein signifikanter Zusammenhang zwischen der Einstellung der Lehrkraft zur Notwendigkeit von Druck im Unterricht und der von den Schülerinnen und Schülern wahrgenom-

menen Disziplin im Unterricht (vgl. Tabelle 8.9). Je höher die Einstellung einer Lehrkraft zur Notwendigkeit von Druck ausgeprägt ist (Modell 2), desto disziplinierter ist der Unterricht gestaltet. Bei Modell 2 liegt der Anteil der erklärten Varianz der „Disziplin im Englischunterricht“ auf der Ebene der Unterrichtsgruppe bei 15 Prozent. Bei den Kontrollvariablen zeigen sich in den vorliegenden Modellen keine signifikanten Effekte, wenn man den Zusammenhang der Überzeugungen mit der Disziplin im Unterricht betrachtet. Um die Vergleichbarkeit der Ergebnisse über alle Kombinationen von Fächern, Überzeugungen und Unterrichtsmerkmalen beizubehalten, wurden trotzdem in alle Modelle jeweils die gleichen Kontrollvariablen aufgenommen, auch wenn sie keine Effekte zeigten.

Tabelle 8.9: Mehrebenenanalysen zur Disziplin im Mathematikunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte

	Modell 1			Modell 2			Modell 3		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in									
Mädchen	.053	(.036)	.144	.053	(.036)	.141	.053	(.036)	.144
Sprache Deutsch	-.005	(.032)	.868	-.005	(.032)	.872	-.006	(.032)	.865
Sozialstatus	.000	(.038)	.991	-.001	(.038)	.982	.000	(.038)	.993
CFT	.010	(.042)	.820	.009	(.043)	.830	.010	(.043)	.811
R ²		.003			.003			.003	
Ebene: Unterrichtsgruppe									
Reform-bereitschaft	-.152	(.096)	.113	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	.305	(.111)	.006	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	.012	(.097)	.897
Komposition: Mädchen	-.112	(.118)	.343	-.066	(.120)	.582	-.123	(.121)	.310
Sprachkomposition Deutsch	-.010	(.076)	.891	.004	(.071)	.956	-.019	(.083)	.822
Soziale Komposition	-.216	(.126)	.086	-.180	(.124)	.147	-.206	(.128)	.106
CFT-Komposition	-.062	(.110)	.571	-.068	(.107)	.527	-.054	(.118)	.650
Hauptschulen ¹	-.105	(.168)	.531	-.060	(.166)	.718	-.106	(.170)	.533
Gesamtschulen ¹	-.109	(.165)	.507	-.061	(.167)	.714	-.162	(.161)	.313
Gymnasien ¹	.035	(.132)	.793	.064	(.135)	.633	.015	(.135)	.909
R ²		.099			.145			.079	

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1188; ¹Realschule ist Referenzkategorie

8.2.1.3 Zusammenhänge zwischen Überzeugungen und allen Unterrichtsmerkmalen

Die erste Fragestellung dieser Arbeit untersucht den Zusammenhang zwischen drei Überzeugungen von Lehrkräften und Unterrichtsmerkmalen aus der Sicht von Schülerinnen und Schülern. In den vorangehenden Abschnitten 8.2.1.1 und

8.2.1.2 wurde dieser Zusammenhang für die Unterrichtsmerkmale Adaptivität und Disziplin ausführlich betrachtet. Die entsprechenden Analysen wurden auch für die anderen neun Unterrichtsmerkmale der Drei-Länder-Studie durchgeführt (siehe Kapitel 5, Hypothesen 3a bis 3i), der Übersichtlichkeit halber werden die Ergebnisse dieser Analysen aber im vorliegenden Kapitel komprimiert dargestellt. In den Überblickstabellen der folgenden Abschnitte werden für den Englisch- und Mathematikunterricht der β -Koeffizient und der p-Wert der jeweiligen Überzeugung abgetragen. Diese Werte stammen aus den Mehrebenenanalysen für jedes Unterrichtsmerkmal, die mit allen Kontrollvariablen analog zu den bisherigen Analysen durchgeführt wurden. Die kompletten Modelle dieser Analysen mit allen Koeffizienten sind hier nicht abgebildet, können aber auf Anfrage zur Verfügung gestellt werden.

Englischunterricht

Die Ergebnisse der Analysen der Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und der Unterrichtsmerkmale im Englischunterricht finden sich in Tabelle 8.10. In der mittleren Spalte sind die Ergebnisse zur Reformbereitschaft (entspricht Modell 1 der ausführlichen Darstellungen zur Adaptivität des Unterrichts und zur Disziplin in Tabelle 8.6 bis Tabelle 8.9), in der nächsten Spalte sind die Ergebnisse zur Notwendigkeit von Druck bei Lernprozessen aus Modell 2 und in der letzten Spalte sind die Ergebnisse zur Förderung schwacher Schülerinnen und Schüler aus Modell 3.

Tabelle 8.10: Übersicht über die Ergebnisse aus Mehrebenenanalysen zu Unterrichtsmerkmalen des Englischunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte

<i>Unabhängige Variable:</i> Überzeugung der Lehrkraft		Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung	
<i>Abhängige Variable:</i> Unterrichtsmerkmal		β	p	β	p	β	p
unterstützendes Unterrichtsklima	Unterstützung	.413	.000	-.384	.000	.293	.005
	Strenge	-.193	.035	.266	.008	-.282	.007
	Adaptivität des Unterrichts ⁵⁴	.178	.049	-.275	.001	.278	.007
	Positive Schülerorientiertheit/ Lehrer-Schüler-Beziehung	.252	.003	-.302	.001	.290	.002
	Individualisierungsmaßnahmen/ Förderungsorientiertheit	.290	.056	-.238	.009	.280	.006
	Motivierungsbemühen/Lob	.301	.003	-.425	.000	.311	.002
Klassenführung	Vertrauensgrad/ Schülerorientierung	.260	.001	-.350	.000	.335	.000
	Strukturiertheit des Unterrichts	.129	.148	-.174	.056	.256	.009
	Beschäftigungsradius	.064	.528	-.149	.132	.196	.097
	Disziplin und Ordnungssinn ⁵⁵	-.235	.028	-.149	.259	.069	.590
	Zeitverschwendung	.164	.109	-.116	.256	.047	.661

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; unter Kontrolle von Geschlecht, zu Hause gesprochener Sprache, Sozialstatus und allgemeiner kognitiver Fähigkeit der Schülerinnen und Schüler sowie der Komposition dieser Merkmale in der Unterrichtsgruppe und der Schultart; N = 1043

Die Reformbereitschaft der Lehrkräfte weist im Englischunterricht einen positiven Zusammenhang mit Unterstützung, Adaptivität, Schülerorientiertheit, Motivierungsbemühen und Vertrauensgrad auf. Ein negativer Zusammenhang liegt für Strenge und Disziplin vor. Die gleichen Zusammenhänge finden sich auch bei einer positiven Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler: Hier bestehen zusätzlich noch positive Zusammenhänge mit Individualisierungsmaßnahmen und Strukturiertheit. Lediglich der

54 Die ausführlichen Modelle zur Adaptivität des Englischunterrichts befinden sich in Tabelle 8.6.

55 Die ausführlichen Modelle zur Disziplin im Englischunterricht befinden sich in Tabelle 8.8.

Zusammenhang mit der Disziplin lässt sich hier nicht nachweisen. Bei der Notwendigkeit von Druck bei Lernprozessen ist das Bild der Effekte – wie erwartet – größtenteils gespiegelt: Hier liegen negative Zusammenhänge zu Unterstützung, Adaptivität, Schülerorientiertheit, Individualisierungsmaßnahmen, Motivierungsbemühen und Vertrauensgrad vor. Bei der Strenge im Unterricht gibt es einen positiven Zusammenhang. Für das Unterrichtsmerkmal Beschäftigungsradius findet sich wie erwartet keinerlei Zusammenhang mit den Überzeugungen der Lehrkräfte. Insgesamt zeigen alle gefundenen Effekte in die erwartete Richtung. Es wird deutlich, dass sich die Überzeugungen der Englischlehrkräfte in für die Schülerinnen und Schüler wahrnehmbaren Handlungen niederschlagen, vor allem im Bereich des unterstützenden Unterrichtsklimas, als der erlebten konstruktiven Unterstützung. Bei den Unterrichtsmerkmalen, die dem Bereich der Klassenführung zuzuordnen sind, gibt es nur vereinzelte Zusammenhänge mit den Überzeugungen der Lehrkräfte.

Die Mehrzahl der dargestellten Zusammenhänge bleibt auch nach einer Korrektur des p -Wertes nach Benjamini und Hochberg (1995) signifikant. Diese Korrektur der p -Werte ist aufgrund der Menge an Signifikanztests sinnvoll, da somit einer Kumulierung des Alphafehlers vorgebeugt werden kann (siehe Kapitel 7.3.2). Die Zusammenhänge zwischen Reformbereitschaft und Strenge bzw. zwischen Reformbereitschaft und Adaptivität weisen nach einer Benjamini-Hochberg-Korrektur einen p -Wert von $p = .055$ bzw. $p = .074$ auf und sind somit nicht mehr auf dem Fünf-Prozent-Niveau, sondern nur noch auf dem Zehn-Prozent-Niveau signifikant.

Bezüglich der Kontrollvariablen wird hier nur von Effekten berichtet, die sich in den Modellen bei allen drei Überzeugungen als relevant für Unterrichtsmerkmale im Englischunterricht herausgestellt haben: In den Analysen (hier nicht detailliert dargestellt) wird deutlich, dass Mädchen den Englischunterricht nicht nur adaptiver wahrnehmen als Jungen (siehe Tabelle 8.6), sondern auch bei den Unterrichtsmerkmalen Individualisierungsmaßnahmen, Vertrauensgrad, Strukturiertheit und Beschäftigungsradius höhere Werte angaben. Kinder mit einem höheren Sozialstatus nehmen hingegen einen niedrigeren Beschäftigungsradius wahr. Schülerinnen und Schüler an Gesamtschulen wiederum berichten von einem höheren Vertrauensgrad und Beschäftigungsradius als die Referenzgruppe an Realschulen.

Mathematikunterricht

In der Tabelle 8.11 sind die Ergebnisse der Mehrebenenanalysen zu Zusammenhängen von Überzeugungen und Unterrichtsmerkmalen im Mathematikunterricht dargestellt. Im Vergleich zum Englischunterricht sind hier weniger Koeffizienten signifikant, und bezüglich der Reformbereitschaft der Lehrkräfte lässt sich kein einziger Effekt auf Unterrichtsmerkmale nachweisen. Die Einstellung von Lehrkräften, dass Druck ein besonders notwendiger Bestandteil von Lernprozessen sei, hängt negativ mit der Unterstützung im Unterricht und positiv mit Strenge und Disziplin zusammen. Eine positive Einstellung zur Förderung schwacher Schülerinnen und Schüler hängt signifikant positiv mit der Schülerorientierung zusammen. Auch hier zeigen die aufgetretenen Effekte in die erwartete Richtung. Dennoch muss festgestellt werden, dass die hier untersuchten Überzeugungen von Lehrkräften im Mathematikunterricht eine geringere Rolle bei den Unterrichtsmerkmalen aus der Sicht der Schülerinnen und Schüler spielen als im Englischunterricht.

Tabelle 8.11: Übersicht über die Ergebnisse aus Mehrebenenanalysen zu Unterrichtsmerkmalen des Mathematikunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte

<i>Unabhängige Variable:</i> Überzeugung der Lehrkraft		Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung	
<i>Abhängige Variable:</i> Unterrichtsmerkmal		β	p	β	p	β	p
unterstützendes Unterrichtsklima	Unterstützung	.169	.087	-.256	.006	.130	.155
	Strenge	.017	.837	.225	.024	-.059	.427
	Adaptivität des Unterrichts ⁵⁶	.018	.874	-.150	.234	.163	.084
	Positive Schülerorientiertheit/ Lehrer-Schüler-Beziehung	.080	.366	-.194	.060	.113	.180
	Individualisierungsmaßnahmen/ Förderungsorientiertheit	.056	.591	-.209	.078	.084	.373
	Motivierungsbemühen/Lob	.102	.351	-.181	.116	.090	.413
	Vertrauensgrad/ Schülerorientierung	.079	.401	-.196	.073	.186	.027
Klassenführung	Strukturiertheit des Unterrichts	.060	.609	-.169	.185	.189	.057
	Beschäftigungsradius	-.052	.634	-.020	.855	.072	.471
	Disziplin und Ordnungssinn ⁵⁷	-.152	.113	.305	.006	.012	.897
	Zeitverschwendung	.135	.165	-.168	.111	.094	.321

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; unter Kontrolle von Geschlecht, zu Hause gesprochener Sprache, Sozialstatus und allgemeiner kognitiver Fähigkeit der Schülerinnen und Schüler sowie der Komposition dieser Merkmale in der Unterrichtsgruppe und der Schultart; N = 1188

Nach einer Benjamini-Hochberg-Korrektur der p-Werte (siehe Kapitel 7.3.2) ist im Mathematikunterricht kein Zusammenhang mehr zwischen den Überzeugungen der Lehrkräfte und Unterrichtsmerkmalen auf dem Fünf-Prozent-Niveau signifikant. Der p-Wert des Koeffizienten der Notwendigkeit von Druck bei Lernprozessen liegt nach einer solchen Korrektur für die Unterstützung bei $p = .099$, für die Strenge bei $p = .223$ und für die Disziplin bei

56 Die ausführlichen Modelle zur Adaptivität des Mathematikunterrichts befinden sich in Tabelle 8.7.

57 Die ausführlichen Modelle zur Disziplin im Mathematikunterricht befinden sich in Tabelle 8.9.

$p = .099$. Der p -Wert für den Zusammenhang der Einstellung zur Förderung schwacher Schülerinnen und Schüler mit „Vertrauensgrad/Schülerorientierung“ liegt dann bei $p = .223$.

Bei den Kontrollvariablen (hier nicht detailliert dargestellt) ist auffällig, dass in fast allen Modellen zu den Unterrichtsmerkmalen des Mathematikunterrichts die kognitive Komposition der Unterrichtsgruppe einen signifikanten Effekt aufweist. Lediglich mit der Strukturiertheit und der Disziplin (siehe Tabelle 8.9) lässt sich kein Zusammenhang feststellen. Die Merkmale aus dem Bereich des unterstützenden Unterrichtsklimas⁵⁸ sind höher ausgeprägt, wenn die mittlere kognitive Fähigkeit der Unterrichtsgruppe hoch ist (außer das Merkmal der Strenge, hier verläuft der Effekt gegenläufig). Auch die Ausprägung des Beschäftigungsradius ist in Unterrichtsgruppen mit einer höheren kognitiven Fähigkeit höher, die Zeitverschwendung hingegen geringer. Mädchen nehmen im Mathematikunterricht weniger Strenge und Zeitverschwendung wahr. Schülerinnen und Schüler, die zu Hause Deutsch sprechen, empfinden im Mathematikunterricht einen höheren Vertrauensgrad. Betrachtet man allerdings den Kompositionseffekt der zu Hause Deutsch sprechenden Kinder, dann weisen deren Unterrichtsgruppen einen niedrigeren Wert beim Beschäftigungsradius auf. Ebenso wie für den Englischunterricht wurde hier nur von Effekten der Kontrollvariablen berichtet, die sich in den Modellen bei allen drei Überzeugungen als relevant für Unterrichtsmerkmale im Mathematikunterricht herausgestellt haben.

8.2.2 Zusammenhänge zwischen Überzeugungen und schulischen Ergebnissen

Nach der Untersuchung von Zusammenhängen zwischen den Überzeugungen von Lehrkräften und ihrem Unterrichtshandeln aus der Sicht der Schülerinnen und Schüler (siehe Abschnitt 8.2.1) ist die zentrale Frage dieser Arbeit, ob und wie die Überzeugungen der Lehrkräfte mit kognitiven und nichtkognitiven Ergebnissen der Schülerinnen und Schüler zusammenhängen. Diese Frage wird im folgenden Abschnitt mit Hilfe von Mehrebenen-Regressionsanalysen (siehe Abschnitt 7.3.1) untersucht. Als abhängige Variablen werden das Wohlfühlen für die zweite Forschungsfrage und die Fachleistungen der Schülerinnen und Schüler für die dritte Forschungsfrage getrennt für den Englisch- und Mathe-

58 Unterstützung, Adaptivität, Schülerorientiertheit, Individualisierungsmaßnahmen, Motivierungsbemühen, Vertrauensgrad

matikunterricht betrachtet. Als vermittelndes Unterrichtsmerkmal in den Regressionsanalysen wird zuerst die Adaptivität des Unterrichts als bedeutsames Unterrichtsmerkmal aus dem Bereich des unterstützenden Unterrichtsklimas in den Blick genommen. Anschließend werden Regressionsanalysen für die Vermittlung durch die Disziplin im Unterricht als wichtiges Merkmal von Klassenführung kurz dargestellt.

8.2.2.1 Schulische Ergebnisse in Abhängigkeit von Überzeugungen und Adaptivität

Zur Beantwortung der zweiten und dritten Forschungsfrage wird zunächst der Zusammenhang zwischen den Überzeugungen der Lehrkräfte und den schulischen Ergebnissen der Schülerinnen und Schüler im Englischunterricht betrachtet (Modelle a: Tabelle 8.12 und Tabelle 8.13). Sofern ein solcher Zusammenhang gegeben ist, wird die Unterrichtsvariable Adaptivität dem Modell hinzugefügt, um festzustellen, ob der Zusammenhang über dieses Unterrichtsmerkmal vermittelt wird (Modelle b: Tabelle 8.12 und Tabelle 8.13). Anschließend wird das Gleiche für den Mathematikunterricht durchgeführt (Tabelle 8.14 und Tabelle 8.15).

*Englischunterricht*⁵⁹

Bei Betrachtung des Wohlfühlens der Schülerinnen und Schüler im Englischunterricht zeigt sich ein positiver Zusammenhang mit der Reformbereitschaft (Modell 1a, Tabelle 8.12), der bei Kontrolle der Adaptivität verschwindet (Modell 1b). Ein analoges Muster, nur mit umgekehrtem Vorzeichen, findet sich für die Einstellung zur Notwendigkeit von Druck bei Lernprozessen (Modelle 2a, 2b). Der positive Zusammenhang zwischen der Förderorientierung der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler bleibt auch bei Kontrolle der Adaptivität bestehen (Modelle 3a, 3b). Die Zusammenhänge der Überzeugungen von Lehrkräften mit dem Wohlfühlen der Schülerinnen und Schüler im Unterricht fallen demnach wie erwartet aus (siehe Kapitel 5, Hypothese 4.1).

Auf der Ebene der Unterrichtsgruppe lassen sich durch die Überzeugungen der Lehrkraft unter Kontrolle von Kompositionseffekten und Schulart 20 bis 21

59 Die Ergebnisse der schulischen Ergebnisse in Abhängigkeit von Überzeugungen und Adaptivität im Englischunterricht wurden bereits von Dohrmann et al. (2019) veröffentlicht.

Prozent (Modelle 1a, 2a) bzw. bei der Einstellung zur Förderung schwacher Schülerinnen und Schülern 28 Prozent (Modell 3a) der Unterschiede des Wohlfühlens im Englischunterricht erklären. Unter Hinzunahme der Adaptivität des Unterrichts in den Modellen b steigt der Anteil auf 73 bis 78 Prozent.

Die Varianzaufklärung und die Effektstärken zeigen, dass die Adaptivität des Englischunterrichts eine hohe Erklärungskraft für das Wohlfühlen der Schülerinnen und Schüler besitzt (Modelle b: $\beta = .83, p < .001$). Die Modelle a zeigen, dass auch die Überzeugungen der Lehrkräfte – ohne Betrachtung von Adaptivität – einen deutlichen Effekt auf das Wohlfühlen der Schülerinnen und Schüler aufweisen ($\beta = .27, p = .020$; $\beta = -.25, p = .006$; $\beta = .36, p < .001$).

Tabelle 8.12: Mehrebenenanalysen zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte und der Adaptivität des Englischunterrichts

	Modell 1a			Modell 1b			Modell 2a			Modell 2b			Modell 3a			Modell 3b		
	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in																		
Adaptivität	–	–	–	.429	(.040)	.000	–	–	–	.429	(.040)	.000	–	–	–	.428	(.040)	.000
Mädchen	.119	(.047)	.010	.075	(.039)	.057	.119	(.047)	.010	.075	(.039)	.058	.119	(.046)	.010	.075	(.039)	.057
Sprache Deutsch	–.061	(.049)	.214	–.080	(.042)	.055	–.061	(.049)	.218	–.080	(.042)	.055	–.062	(.049)	.210	–.081	(.042)	.052
Sozialstatus	.003	(.034)	.919	.014	(.029)	.627	.004	(.034)	.908	.014	(.029)	.619	.003	(.034)	.919	.014	(.029)	.620
CFT	.027	(.037)	.463	.039	(.033)	.242	.028	(.037)	.448	.039	(.033)	.243	.027	(.037)	.463	.038	(.033)	.246
R ²	.019			.197			.019			.198			.019			.197		
Ebene:																		
Unterrichtsgruppe																		
Reformbereitschaft	.226	(.097)	.020	.111	(.081)	.170	–	–	–	–	–	–	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	–	–	–	–.245	(.088)	.006	–.071	(.084)	.399	–	–	–	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Adaptivität	–	–	–	.833	(.065)	.000	–	–	–	.831	(.066)	.000	–	–	–	.827	(.067)	.000
Komposition: Mädchen	.063	(.105)	.549	.033	(.085)	.700	.065	(.104)	.530	.028	(.084)	.739	.039	(.105)	.712	.023	(.086)	.790
Sprachkomposition D	.106	(.119)	.373	.034	(.082)	.679	.096	(.119)	.420	.030	(.082)	.716	.075	(.120)	.532	.019	(.086)	.822
Soziale Komposition	–.104	(.134)	.436	–.031	(.110)	.781	–.106	(.130)	.418	–.032	(.109)	.772	–.127	(.121)	.293	–.050	(.106)	.636
CFT-Komposition	–.044	(.124)	.719	–.067	(.088)	.451	–.072	(.122)	.555	–.070	(.088)	.421	–.037	(.120)	.758	–.066	(.089)	.459
Hauptschulen ¹	–.134	(.154)	.385	–.171	(.111)	.123	–.078	(.153)	.611	–.153	(.111)	.168	–.151	(.158)	.337	–.186	(.112)	.097
Gesamtschulen ¹	.247	(.146)	.091	.007	(.109)	.949	.299	(.146)	.041	.029	(.113)	.798	.238	(.159)	.133	.010	(.110)	.929
Gymnasien ¹	.015	(.119)	.903	–.002	(.104)	.988	.005	(.119)	.965	–.010	(.103)	.919	–.010	(.119)	.931	–.009	(.100)	.932
R ²	.197			.744			.213			.733			.276			.775		

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1043; ¹Realschule ist Referenzkategorie

Die Analysen der Englischleistungen der Schülerinnen und Schüler zeigen deutliche Unterschiede im Vergleich zum Wohlfühlen. Modell 1a (vgl. Tabelle 8.13) dokumentiert einen negativen Zusammenhang der Reformbereitschaft von Lehrkräften mit den Englischleistungen von Schülerinnen und Schülern, der auch unter Einbezug des Unterrichtsmerkmals Adaptivität bestehen bleibt (Modell 1b). Angesichts der uneindeutigen Erwartungen zur dritten Forschungsfrage (siehe Seite 63) ist bemerkenswert, dass die Adaptivität bei Kontrolle der Reformbereitschaft einen signifikant positiven Effekt hat (Modell 1b), während sie in einem Modell ohne Einbezug von Überzeugungen (kann auf Anfrage zur Verfügung gestellt werden) weder positiv noch negativ mit den Englischleistungen zusammenhängt. Die Überzeugungen der Lehrkräfte „Einstellung zur Notwendigkeit von Druck bei Lernprozessen“ und „positive Einstellung zur Förderung schwacher Schülerinnen und Schüler“ zeigen keine Zusammenhänge mit den Englischleistungen (Modelle 2a, 3a); dementsprechend wurde auf Modellvarianten unter Einbeziehung der Adaptivität hier verzichtet.⁶⁰

Auf der Ebene der Unterrichtsgruppe werden bei alleiniger Betrachtung der Überzeugungen der Lehrkräfte 55 bis 57 Prozent der Varianz erklärt, bei zusätzlicher Berücksichtigung der Adaptivität in Modell 1b steigt der Anteil der erklärten Varianz lediglich um drei Prozentpunkte.

Die Schulart sowie die kognitive und soziale Komposition der Unterrichtsgruppe sind ebenso wie die individuelle kognitive Fähigkeit und das Geschlecht im Zusammenhang mit den Englischleistungen erwartungsgemäß von Bedeutung. Ein Zusammenhang mit dem Wohlfühlen zeigt sich nur beim Geschlecht; das positivere Erleben der Mädchen ist interessanterweise – wie der Vergleich der Modelle a und b in Tabelle 8.12 zeigt – vermittelt über eine stärkere Wahrnehmung von Adaptivität (vgl. Tabelle 8.6).

60 Dieses Vorgehen erfolgt analog zu Dohrmann et al. (2019).

Tabelle 8.13: Mehrebenenanalysen zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von den Überzeugungen der Lehrkräfte und der Adaptivität des Englischunterrichts

	Modell 1a			Modell 1b			Modell 2a			Modell 3a		
	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in												
Adaptivität	–	–	–	.085	(.036)	.019	–	–	–	–	–	–
Mädchen	.213	(.040)	.000	.204	(.040)	.000	.213	(.040)	.000	.213	(.040)	.000
Sprache Deutsch	–.001	(.041)	.982	–.003	(.040)	.936	–.001	(.041)	.979	–.001	(.041)	.986
Sozialstatus	.045	(.039)	.247	.048	(.039)	.221	.045	(.039)	.251	.045	(.039)	.250
CFT	.174	(.036)	.000	.177	(.035)	.000	.174	(.036)	.000	.175	(.036)	.000
R ²	.078			.082			.078			.078		
Ebene: Unterrichtsgruppe												
Reformbereitschaft	–.152	(.064)	.017	–.178	(.063)	.005	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	–	–	–	.091	(.074)	.220	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	–	–	–	–.076	(.079)	.336
Adaptivität	–	–	–	.132	(.062)	.034	–	–	–	–	–	–
Komposition: Mädchen	.124	(.084)	.142	.120	(.084)	.153	.125	(.084)	.136	.135	(.082)	.098
Sprachkomposition D	–.086	(.112)	.440	–.099	(.110)	.369	–.081	(.109)	.457	–.076	(.109)	.483
Soziale Komposition	.181	(.085)	.033	.186	(.084)	.026	.179	(.085)	.035	.190	(.086)	.027
CFT-Komposition	.420	(.078)	.000	.415	(.077)	.000	.416	(.080)	.000	.408	(.080)	.000
Hauptschulen ¹	–.568	(.098)	.000	–.565	(.099)	.000	–.583	(.094)	.000	–.564	(.096)	.000
Gesamtschulen ¹	–.404	(.099)	.000	–.432	(.096)	.000	–.426	(.093)	.000	–.414	(.093)	.000
Gymnasien ¹	.155	(.075)	.038	.153	(.073)	.036	.165	(.070)	.018	.171	(.070)	.015
R ²	.561			.589			.566			.552		

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); vx-standardisierte Koeffizienten; SE: Standardfehler, N = 1043; ¹Realschule ist Referenzkategorie

Mathematikunterricht

Bei der Betrachtung des Wohlfühlens der Schülerinnen und Schüler im Mathematikunterricht zeigt sich ein positiver Zusammenhang mit der Reformbereitschaft (Modell 1a, Tabelle 8.14), der bei Kontrolle der Adaptivität bestehen bleibt (Modell 1b). Ein analoges Muster, nur mit umgekehrtem Vorzeichen, findet sich für die Einstellung zur Notwendigkeit von Druck bei Lernprozessen (Modell 2a, 2b). Für die Überzeugung der Lehrkräfte „Positive Einstellung zur Förderung schwacher Schülerinnen und Schüler“ zeigt sich kein Zusammenhang mit dem Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht (Modell 3a); dementsprechend wurde auf das Modell unter Einbeziehung der Adaptivität hier verzichtet.

Auf der Ebene der Unterrichtsgruppe lassen sich durch die Überzeugungen der Lehrkraft unter Kontrolle von Kompositionseffekten und Schulart 19 bis 23 Prozent der Unterschiede des Wohlfühlens im Mathematikunterricht erklären. Unter Hinzunahme der Adaptivität des Unterrichts in den Modellen 2a und 2b steigt der Anteil auf 70 bis 72 Prozent. Die Höhe der Varianzaufklärung liegt hierbei in einer etwa gleichen Größenordnung wie bei den Analysen für den Englischunterricht.

Wie im Englischunterricht besitzt die Adaptivität auch im Mathematikunterricht eine hohe Erklärungskraft für das Wohlfühlen der Schülerinnen und Schüler. Dies wird durch die Varianzaufklärung und die Effektstärken ersichtlich (Modelle b: $\beta = .81$, $p < .001$; $\beta = .80$, $p < .001$). Die Modelle 1a und 2a zeigen, dass auch die Überzeugungen der Lehrkräfte (ohne Betrachtung der Adaptivität) einen deutlichen Effekt auf das Wohlfühlen der Schülerinnen und Schüler aufweisen ($\beta = .21$, $p = .043$; $\beta = -.27$, $p = .007$). Bei der Einzelbetrachtung der Merkmale Adaptivität, Reformbereitschaft und Einstellung der Lehrkräfte zur Notwendigkeit von Druck bei Lernprozessen finden sich Gemeinsamkeiten zwischen dem Englisch- und dem Mathematikunterricht. Bezüglich der Vermittlungsprozesse der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Unterricht unterscheiden sich die Fächer jedoch.

Tabelle 8.14: Mehrebenenanalysen zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte und der Adaptivität des Mathematikunterrichts

	Modell 1a			Modell 1b			Modell 2a			Modell 2b			Modell 3a		
	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in															
Adaptivität	–	–	–	.407	(.040)	.000	–	–	–	.408	(.040)	.000	–	–	–
Mädchen	–.065	(.033)	.052	–.083	(.030)	.006	–.065	(.033)	.051	–.083	(.030)	.006	–.066	(.033)	.049
Sprache Deutsch	–.023	(.037)	.529	–.035	(.033)	.290	–.024	(.037)	.523	–.035	(.033)	.285	–.024	(.037)	.516
Sozialstatus	.014	(.041)	.735	.021	(.033)	.524	.014	(.041)	.724	.022	(.033)	.508	.015	(.041)	.718
CFT	.107	(.040)	.007	.126	(.033)	.000	.107	(.040)	.007	.126	(.033)	.000	.108	(.039)	.006
R ²	.016			.190			.017			.191			.017		
Ebene:															
Unterrichtsgruppe															
Reformbereitschaft	.214	(.106)	.043	.183	(.070)	.009	–	–	–	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	–	–	–	–.268	(.100)	.007	–.152	(.069)	.028	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	–	–	–	–	–	–	.204	(.105)	.052
Adaptivität	–	–	–	.813	(.047)	.000	–	–	–	.803	(.049)	.000	–	–	–
Komposition: Mädchen	–.152	(.123)	.215	–.039	(.082)	.636	–.181	(.118)	.125	–.054	(.083)	.512	–.134	(.120)	.262
Sprachkomposition D	–.030	(.085)	.727	–.002	(.052)	.975	–.037	(.080)	.640	–.003	(.050)	.955	–.049	(.085)	.567
Soziale Komposition	–.010	(.131)	.938	–.033	(.086)	.704	–.046	(.131)	.724	–.057	(.089)	.523	–.013	(.126)	.918
CFT-Komposition	.237	(.110)	.031	.101	(.067)	.132	.236	(.108)	.029	.103	(.069)	.133	.267	(.104)	.011
Hauptschulen ¹	–.045	(.175)	.797	–.027	(.118)	.816	–.080	(.175)	.647	–.046	(.121)	.701	–.071	(.172)	.677
Gesamtschulen ¹	.224	(.172)	.192	.090	(.117)	.442	.205	(.172)	.234	.108	(.117)	.358	.240	(.164)	.144
Gymnasien ¹	.017	(.129)	.898	.039	(.089)	.660	.000	(.129)	.998	.038	(.090)	.673	.013	(.124)	.918
R ²	.190			.719			.233			.704			.216		

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx -standardisierte Koeffizienten; SE: Standardfehler; N = 1188; ¹Realschule ist Referenzkategorie

Für alle drei hier untersuchten Überzeugungen der Lehrkräfte zeigen sich keine Zusammenhänge mit den Mathematikleistungen (vgl. Tabelle 8.15); dementsprechend wurde auf die Erstellung von Modellvarianten unter Einbeziehung der Adaptivität hier verzichtet.

Wie schon für die Leistungen im Fach Englisch spielen die Schulart und die kognitive Zusammensetzung der Unterrichtsgruppe sowie das Geschlecht auch im Zusammenhang mit den Mathematikleistungen erwartungsgemäß eine starke Rolle. Allerdings gibt es im Gegensatz zu Englisch bei den Schularten keine signifikanten Leistungsunterschiede im Fach Mathematik zwischen den Gymnasien und den Realschulen. Auch weist, anders als im Englischunterricht, die soziale Komposition der Unterrichtsgruppe keinen signifikanten Effekt auf die Mathematikleistungen auf. Die Kontrollvariablen auf der Ebene der Unterrichtsgruppe zeigen insgesamt einen hohen Zusammenhang mit den Leistungen im Fach Mathematik, sodass sie in jedem der hier betrachteten Modelle circa 50 Prozent der Varianz erklären können.

Beim Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht gibt es auch einen positiven Zusammenhang mit der kognitiven Zusammensetzung der Unterrichtsgruppe, der über die Adaptivität vermittelt ist, wie ein Vergleich der Modelle 1a bzw. 2a und 1b bzw. 2b in Tabelle 8.14 zeigt. Bei alleiniger Betrachtung der Adaptivität als abhängiger Variable tritt kein Zusammenhang mit der kognitiven Zusammensetzung der Unterrichtsgruppe auf (vgl. Tabelle 8.6). Außerdem gibt es in einigen Modellen Hinweise darauf, dass Mädchen sich im Mathematikunterricht weniger wohl fühlen als Jungen (vgl. Modelle 1b, 2b, 3a in Tabelle 8.14). Dies lässt sich hier nicht – wie im Fach Englisch – über ihre Wahrnehmung des adaptiven Unterrichts erklären. Anzunehmen ist zudem, dass die kognitive Fähigkeit auf individueller Ebene, also der einzelnen Schülerin oder des einzelnen Schülers einen Zusammenhang mit ihrem oder seinem Wohlfühlen im Mathematikunterricht aufweist.

Tabelle 8.15: Mehrebenenanalysen zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von den Überzeugungen der Lehrkräfte

	Modell 1a			Modell 2a			Modell 3a		
	β	(SE)	p	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in									
Mädchen	-.087	(.038)	.022	-.087	(.038)	.022	-.087	(.038)	.022
Sprache Deutsch	.014	(.029)	.639	.014	(.029)	.637	.014	(.029)	.632
Sozialstatus	.018	(.028)	.525	.018	(.028)	.522	.017	(.028)	.536
CFT	.256	(.034)	.000	.256	(.034)	.000	.256	(.034)	.000
R ²		.074			.074			.074	
Ebene: Unterrichtsgruppe									
Reform-bereitschaft	-.101	(.087)	.247	–	–	–	–	–	–
Notwendigkeit von Druck	–	–	–	.080	(.091)	.380	–	–	–
Einstellung zu Förderung	–	–	–	–	–	–	-.149	(.077)	.053
Komposition: Mädchen	-.062	(.091)	.498	-.054	(.094)	.567	-.070	(.094)	.456
Sprach-komposition D	.047	(.057)	.411	.049	(.058)	.400	.065	(.060)	.275
Soziale Komposition	.048	(.094)	.608	.063	(.096)	.515	.049	(.095)	.608
CFT-Komposition	.322	(.076)	.000	.323	(.077)	.000	.300	(.080)	.000
Hauptschulen ¹	-.683	(.082)	.000	-.677	(.085)	.000	-.672	(.086)	.000
Gesamtschulen ¹	-.209	(.092)	.023	-.219	(.092)	.017	-.209	(.096)	.029
Gymnasien ¹	.114	(.074)	.124	.115	(.075)	.125	.127	(.076)	.096
R ²		.505			.493			.502	

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1188; ¹Realschule ist Referenzkategorie

8.2.2.2 Schulische Ergebnisse in Abhängigkeit von Überzeugungen und Disziplin

Für das Merkmal Disziplin im Unterricht wurde zur Beantwortung der zweiten und dritten Forschungsfrage ebenso vorgegangen wie beim Unterrichtsmerkmal Adaptivität. Zunächst wurde wie im vorigen Abschnitt 8.2.1.1 der Zusammen-

hang zwischen den Überzeugungen der Lehrkräfte und den schulischen Ergebnissen der Schülerinnen und Schüler betrachtet (Modelle a, Tabelle 8.12 bis Tabelle 8.15). Sofern ein solcher Zusammenhang gegeben war, wurde die Unterrichtsvariable Adaptivität dem Modell hinzugefügt, um festzustellen, ob der Zusammenhang über dieses Unterrichtsmerkmal vermittelt wird (Modelle b, Tabelle 8.12 bis Tabelle 8.15). Dann wurde untersucht, ob diese Vermittlung durch das Merkmal Disziplin im Unterricht als Indikator für die Klassenführung erfolgt. Dies betrifft die folgenden Zusammenhänge: alle drei Überzeugungen und Wohlfühlen im Englischunterricht, Reformbereitschaft und Englischleistungen, Reformbereitschaft und Einstellung zur Notwendigkeit von Druck und Wohlfühlen im Mathematikunterricht. Bei den Leistungen im Mathematikunterricht konnten keine Zusammenhänge mit den Überzeugungen der Lehrkräfte festgestellt werden.

Da das Vorgehen in diesem Abschnitt dem im vorherigen entspricht, werden die Ergebnisse hier der Übersichtlichkeit halber in Textform berichtet. Die Tabellen zu sämtliche Analysen dieses Kapitels sind hier nicht abgebildet, können aber auf Anfrage zur Verfügung gestellt werden.

Englischunterricht

Bei der Betrachtung des Wohlfühlens der Schülerinnen und Schüler im Englischunterricht konnte durch Hinzunahme des Unterrichtsmerkmals Disziplin der Zusammenhang mit den Überzeugungen der Lehrkräfte nicht erklärt werden. Es gab keine Veränderungen bei den Koeffizienten der Modelle 1a bis 3a aus Tabelle 8.12, wenn die Variable Disziplin ins Modell aufgenommen wurde. Auch der bereits festgestellte negative Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und den Englischleistungen der Schülerinnen und Schüler (siehe Tabelle 8.13, Modell 1a) lässt sich durch das Unterrichtsmerkmal Disziplin nicht erklären. Zwar ist der Koeffizient von Reformbereitschaft unter Einbezug von Disziplin nicht mehr auf dem Fünf-Prozent-Niveau signifikant, aber auch bei der Variable Disziplin ist kein signifikanter Effekt festzustellen. Alle anderen Koeffizienten weisen keine nennenswerten Veränderungen auf.

Mathematikunterricht

Das Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht weist, analog zum Englischunterricht, keinen Zusammenhang mit der Disziplin im Unterricht auf. Durch das Unterrichtsmerkmal Disziplin lässt sich somit der

bereits festgestellte Zusammenhang (siehe Tabelle 8.14, Modelle 1a und 2a) zwischen der Reformbereitschaft der Lehrkräfte und ihrer Einstellung zur Notwendigkeit von Druck mit dem Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht nicht erklären. Bei den Kontrollvariablen zeigen sich nach der Hinzunahme des Unterrichtsmerkmals Disziplin in die Modelle keine nennenswerten Veränderungen. Da im vorigen Kapitel bereits festgestellt wurde, dass im Fach Mathematik keine Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und den Fachleistungen der Schülerinnen und Schüler auftreten (vgl. Tabelle 8.15), wird auch hier keine Modellvariante unter Einbeziehung des Unterrichtsmerkmals Disziplin dargestellt.

8.2.3 Direkte und indirekte Zusammenhänge zwischen Überzeugungen von Lehrkräften und schulischen Ergebnissen von Schülerinnen und Schülern

Im Folgenden werden die zweite und dritte Forschungsfrage nach dem Zusammenhang von Überzeugungen und Ergebnissen von Schülerinnen und Schülern durch Mehrebenen-Mediationsanalysen beantwortet. Diese beiden Fragen wurden bereits im vorangegangenen Abschnitt 8.2.2 untersucht, dort allerdings mittels Mehrebenen-Regressionsanalysen. Die Mehrebenen-Mediationsanalyse ermöglicht darüber hinaus die Untersuchung von Beziehungen zwischen und innerhalb von Analyseebenen, beispielsweise den Mediationseffekt auf Gruppenebene (MacKinnon, 2008). Ein Vorteil gegenüber der Mehrebenen-Regression ist die Berechnung des indirekten Effekts für die Mediatorvariable. So lässt sich bestimmen, ob eine statistische Mediation, also die Vermittlung eines Effekts, vorliegt (siehe Abschnitt 7.3.1). Dies betrifft vor allem die Hypothesen 4.2 zur Vermittlung der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler durch die Adaptivität des Unterrichts sowie die Hypothese 4.3 zu dieser Vermittlung durch die Disziplin im Unterricht (siehe Kapitel 5). Zur Fragestellung 3, die Fachleistungen der Schülerinnen und Schüler betreffend, wurden keine gerichteten Hypothesen formuliert (siehe S. 63).

8.2.3.1 Das Unterrichtsmerkmal Adaptivität als Mediator

Englischunterricht

Die Mediationsanalyse der Zusammenhänge zwischen der Überzeugung Reformbereitschaft, der Adaptivität des Unterrichts und dem Wohlfühlen der Schülerinnen und Schüler im Englischunterricht wird im Folgenden exemplarisch detailliert dargestellt (siehe Abbildung 8.1, Tabelle 8.16, Tabelle 8.17). Bei den Analysen der weiteren Überzeugungen, Unterrichts- und Zielvariablen war das Vorgehen analog, es wird an dieser Stelle aber jeweils nur über das Ergebnis der Mediationsanalyse, also über die indirekten, direkten und totalen Effekte, berichtet. Alle ausführlichen Ergebnisse sind hier nicht abgebildet, können aber auf Anfrage zur Verfügung gestellt werden.

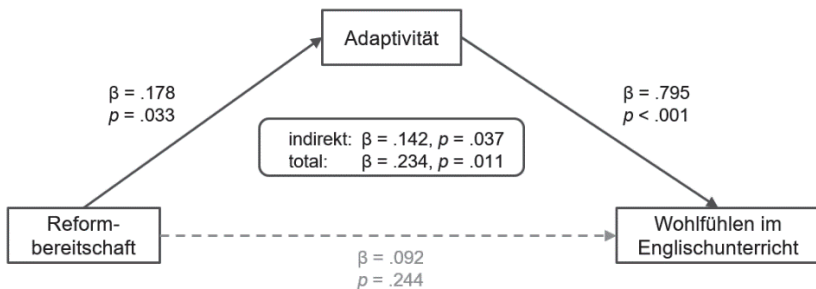


Abbildung 8.1: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Adaptivität als Mediator

Im Mediationsmodell wird gleichzeitig untersucht, ob es einen Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und der Adaptivität des Unterrichts sowie dem Wohlfühlen der Schülerinnen und Schüler und der Adaptivität und dem Wohlfühlen gibt (siehe Abbildung 8.1). Darüber hinaus wird mit Hilfe des indirekten Effekts betrachtet, ob der Zusammenhang der Reformbereitschaft der Lehrkräfte mit dem Wohlfühlen der Schülerinnen und Schüler durch die Adaptivität des Unterrichts vermittelt wird. Im folgenden Abschnitt werden die Pfade der Abbildung 8.1 mit Hilfe des ausführlichen Modells aus Tabelle 8.16 ausführlich erläutert. Alle weiteren Modelle der Abschnitte 8.2.3.1 und 8.2.3.2 werden komprimiert wiedergegeben.

Wenn die Adaptivität des Unterrichts in Mediationsanalysen als Mediator betrachtet wird, fungiert diese Variable gleichzeitig als eine abhängige Variable

(im Verhältnis zu Reformbereitschaft) und als unabhängige Variable (im Verhältnis zum Wohlfühlen der Schülerinnen und Schüler im Unterricht). Die Koeffizienten aller einbezogenen Variablen dieser beiden Pfade „Reformbereitschaft – Adaptivität“ sowie „Adaptivität – Wohlfühlen“ sind in Tabelle 8.16 abgetragen. Hier zeigt sich (in der fünften Spalte), was bereits in Abschnitt 8.2.1.1 festgestellt wurde: Es gibt einen signifikant positiven Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und der Adaptivität des Englischunterrichts ($\beta = .18, p = .033$; siehe auch Abbildung 8.1). Auch der positive Zusammenhang zwischen der Adaptivität des Englischunterrichts und dem Wohlfühlen der Schülerinnen und Schüler (siehe Abschnitt 8.2.2.1) findet sich hier in der zweiten Spalte wieder ($\beta = .80, p < .001$; siehe auch Abbildung 8.1).

Tabelle 8.16: Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte mit Adaptivität als Mediator

	AV: Wohlfühlen			AV: Adaptivität		
	β	(SE)	p	β	(SE)	p
Ebene: Schüler/in						
Adaptivität	.429	(.040)	.000			
Mädchen	.075	(.039)	.057			
Sprache Deutsch	-.080	(.042)	.054			
Sozialstatus	.014	(.029)	.624			
CFT	.039	(.033)	.243			
R ²		.197				
Ebene: Unterrichtsgruppe						
Reformbereitschaft	.092	(.079)	.244	.178	(.084)	.033
Adaptivität	.795	(.074)	.000	–	–	–
Komposition: Mädchen	.029	(.082)	.721	.106	(.070)	.126
Sprachkomposition D	.034	(.079)	.667	.112	(.068)	.101
Soziale Komposition	-.029	(.106)	.782	-.099	(.078)	.205
CFT-Komposition	-.063	(.086)	.461	.080	(.094)	.398
Hauptschulen ¹	-.166	(.107)	.120	-.049	(.123)	.691
Gesamtschulen ¹	.008	(.106)	.941	.210	(.119)	.076
Gymnasien ¹	-.005	(.100)	.964	-.034	(.097)	.730
R ²		.759			.137	

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1043; ¹Realschule ist Referenzkategorie

Will man einen Mediationseffekt feststellen, dann ist die Betrachtung des indirekten und des direkten Effekts in Tabelle 8.17 (auch dargestellt in der Abbildung 8.1) zentraler als die einzelnen Pfade. Da sich der hier untersuchte Media-

tionseffekt der Adaptivität auf der Ebene der Unterrichtsgruppe befindet, gibt es keine Effekte auf der Ebene der einzelnen Schülerin bzw. des einzelnen Schülers.

Tabelle 8.17: Totaler, indirekter und direkter Effekt der Adaptivität beim Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte

Reformbereitschaft → Wohlfühlen			
	β	(SE)	p
Ebene: Schüler/in			
Totaler Effekt	.000	(.000)	1.000
Indirekter Effekt	.000	(.000)	1.000
Ebene: Unterrichtsgruppe			
Totaler Effekt	.234	(.092)	.011
Indirekter Effekt	.142	(.068)	.037
Direkter Effekt	.092	(.079)	.244

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; SE: Standardfehler; N = 1043

In diesem Fall liegt ein Mediationseffekt vor, das heißt, der Effekt der Reformbereitschaft der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Englischunterricht wird durch die Variable Adaptivität im Unterricht vermittelt. Da hier ein signifikanter indirekter Effekt vorliegt, der direkte Effekt dagegen nicht signifikant ist, kann von einem totalen Mediatoreffekt⁶¹ gesprochen werden (Urban & Mayerl, 2018, S. 335). Dies bedeutet, dass der Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler innerhalb des Modells vollständig über die Mediatorvariable Adaptivität vermittelt wird.

Ein totaler Mediatoreffekt für die Adaptivität des Unterrichts liegt auch für den Zusammenhang zwischen der Einstellung der Lehrkräfte zur Notwendigkeit von Druck und dem Wohlfühlen der Schülerinnen und Schüler vor (siehe Abbildung 8.2). Dieser Effekt ist negativ, da die Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit der Adaptivität des Unterrichts negativ zusammenhängt. Die Schülerinnen und Schüler fühlen sich in einem weniger adaptiven Unterricht weniger wohl.

61 Ein totaler Mediatoreffekt ist nicht zu verwechseln mit dem totalen Effekt, der lediglich der Summe des indirekten und direkten Effekts entspricht (vgl. Tabelle 8.17).

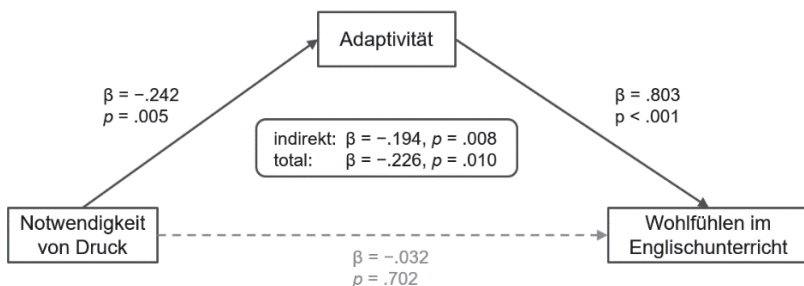


Abbildung 8.2: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Adaptivität als Mediator

Für die Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler sehen die Ergebnisse ähnlich aus wie für die Reformbereitschaft: Auch hier liegt ein positiver Mediationseffekt für die Adaptivität des Unterrichts auf das Wohlfühlen der Schülerinnen und Schüler vor (siehe Abbildung 8.3). Allerdings liegt hier kein totaler, sondern ein partieller Mediatoreffekt vor, da es einen signifikanten direkten Effekt der Einstellung zur Förderung schwacher Schülerinnen und Schüler auf das Wohlfühlen der Schülerinnen und Schüler im Englischunterricht gibt. Dies bestätigt die Ergebnisse aus Abschnitt 8.2.2.1, wo in Modell 3b der Tabelle 8.12 der Effekt der Einstellung zur Förderung nicht komplett durch die Variable Adaptivität erklärt werden konnte.

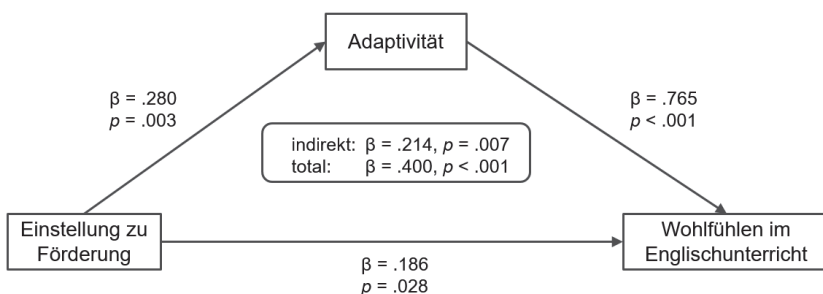


Abbildung 8.3: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator

Für eine kurze Übersicht sind in Tabelle 8.18 alle indirekten, direkten und totalen Effekte aller Mediationsanalysen zu den Überzeugungen der Lehrkräfte im Englischunterricht abgetragen. Dies beinhaltet die Ergebnisse der Abbildung

8.1 bis Abbildung 8.3 sowie der exemplarischen Tabelle 8.17. Dort sind zum Vergleich auch die Ergebnisse der Mediationsanalysen mit den Englischleistungen der Schülerinnen und Schüler als abhängiger Variable enthalten, die in der folgenden Abbildung 8.4, Abbildung 8.5 und Abbildung 8.6 dargestellt werden.

Tabelle 8.18: Übersicht über die Ergebnisse der Mediationsanalysen zu Adaptivität im Englischunterricht

Unabhängige Variable: Überzeugung	Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung	
	β	p	β	p	β	p
Abhängige Variable: Schulische Ergebnisse						
Wohlfühlen						
Totaler Effekt	.234	.011	-.226	.010	.400	.000
Indirekter Effekt	.142	.037	-.194	.008	.214	.007
Direkter Effekt	.092	.244	-.032	.702	.186	.028
Englischleistungen						
Totaler Effekt	-.169	.007	.112	.117	-.090	.249
Indirekter Effekt	.032	.071	-.038	.059	.040	.083
Direkter Effekt	-.201	.001	.151	.037	-.130	.119

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; unter Kontrolle von Geschlecht, zu Hause gesprochener Sprache, Sozialstatus und allgemeiner kognitiver Fähigkeit der Schülerinnen und Schüler sowie der Komposition dieser Merkmale in der Unterrichtsgruppe und der Schultart; N = 1043

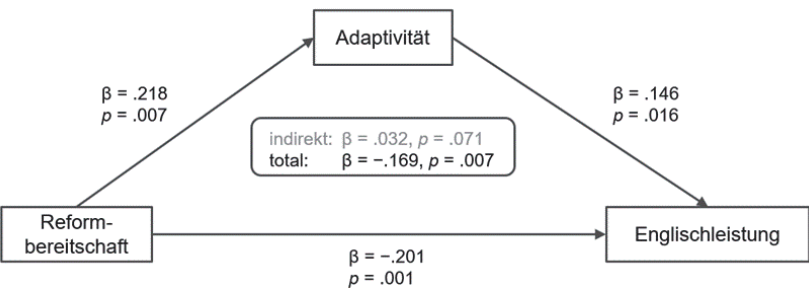


Abbildung 8.4: Grafisches Mediationsmodell zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Adaptivität als Mediator

Bei den Englischleistungen als abhängiger Variable gibt es keine indirekten Effekte für die Adaptivität als Vermittler der Überzeugungen der Lehrkräfte. Es bestehen allerdings direkte Effekte: ein negativer für die Reformbereitschaft (siehe Abbildung 8.4) und ein positiver für die Notwendigkeit von Druck bei Lernprozessen (siehe Abbildung 8.5). Das Muster für Reformbereitschaft entspricht den Regressionsanalysen aus Abschnitt 8.2.2.1. Am Beispiel der Notwendigkeit von Druck bei Lernprozessen kann man auch sehen, dass ein nicht-signifikanter totaler Effekt, wie in Tabelle 8.18 oder wie in Modell 2a in den Analysen in Tabelle 8.13, nicht automatisch bedeutet, dass kein Zusammenhang zwischen dieser Überzeugung und der abhängigen Variable Englischleistungen bestünde: Es gibt zwar keinen indirekten Effekt und somit auch keinen Mediationseffekt für die Adaptivität, aber einen signifikant positiven direkten Effekt.

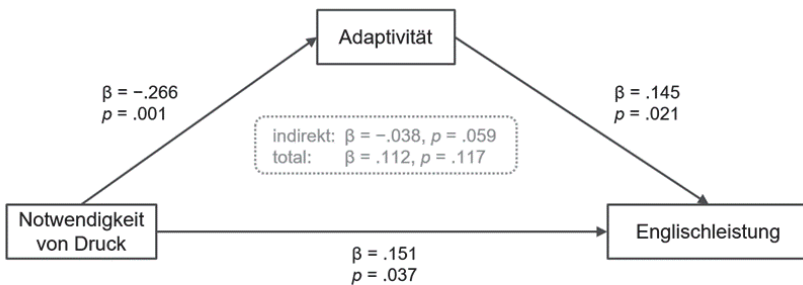


Abbildung 8.5: Grafisches Mediationsmodell zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Adaptivität als Mediator

Das in Abbildung 8.6 dargestellte Modell für die Struktur des Zusammenhangs der Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler und deren Englischleistungen bestätigt die Ergebnisse aus Modell 3a in Tabelle 8.13, dass kein Zusammenhang zwischen der Einstellung zur Förderung und den Englischleistungen vorliegt. Darüber hinaus liegt weder ein indirekter noch ein direkter Effekt für den Zusammenhang dieser beiden Variablen vor.

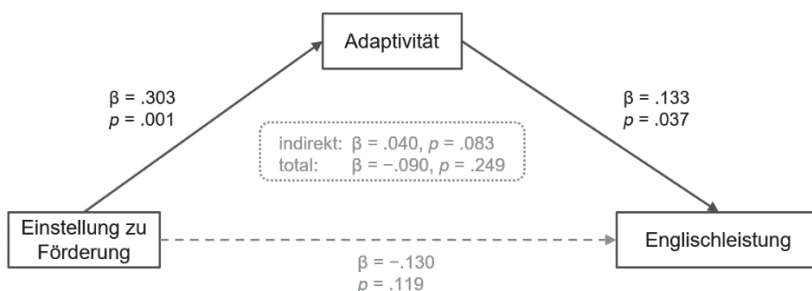


Abbildung 8.6: Grafisches Mediationsmodell zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator

Mathematikunterricht

Im Mathematikunterricht sind, im Gegensatz zum Englischunterricht, keine Mediationseffekte für die Adaptivität des Unterrichts zu verzeichnen (siehe Tabelle 8.19). Vereinzelt treten direkte Effekte von Überzeugungen auf schulische Ergebnisse auf: ein signifikant positiver für die Reformbereitschaft und das Wohlfühlen (siehe Abbildung 8.7) und ein signifikant negativer für die Notwendigkeit von Druck bei Lernprozessen und das Wohlfühlen (siehe Abbildung 8.8). Auch hier bestätigen sich die Ergebnisse aus Abschnitt 8.2.2.1.

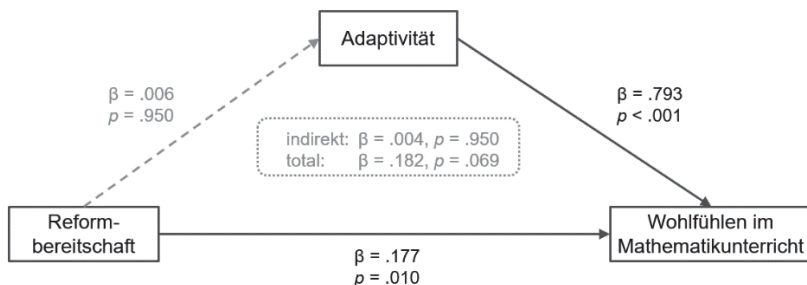


Abbildung 8.7: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Adaptivität als Mediator

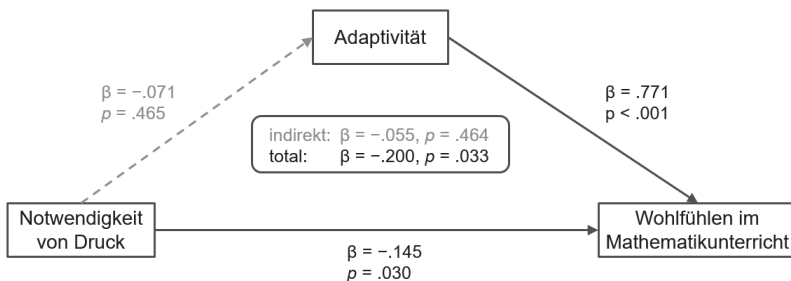


Abbildung 8.8: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Adaptivität als Mediator

Für den Zusammenhang der Einstellung der Lehrkräfte mit der Förderung schwacher Schülerinnen und Schüler liegen zwar weder indirekte noch direkte Effekte auf das Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht vor (siehe Abbildung 8.9), der totale Effekt lässt sich aber dadurch erklären, dass es sowohl von der Förderung zur Adaptivität als auch von der Adaptivität zum Wohlfühlen einen direkten Effekt gibt. Dieser Zusammenhang zwischen der Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler und der Adaptivität des Mathematikunterrichts ist erst bei der Betrachtung im Mediationsmodell aufgetreten, im Regressionsmodell aus Abschnitt 8.2.2.1 (siehe Tabelle 8.7, Modell 3) ist der Zusammenhang nicht signifikant.

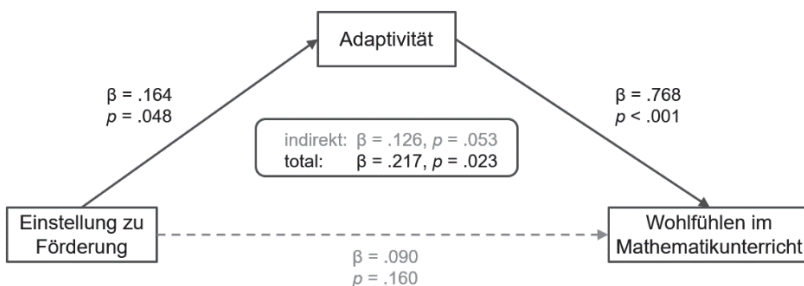


Abbildung 8.9: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator

In Tabelle 8.19 sind alle Ergebnisse der Mediationen zur Adaptivität im Mathematikunterricht zusammengefasst aufgeführt.

Tabelle 8.19: Übersicht über die Ergebnisse der Mediationsanalysen zur Adaptivität im Mathematikunterricht

<i>Unabhängige Variable:</i> Überzeugung	Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung	
	β	p	β	p	β	p
<i>Abhängige Variable:</i> Schulische Ergebnisse						
Wohlfühlen						
Totaler Effekt	.182	.069	-.200	.033	.217	.023
Indirekter Effekt	.004	.950	-.055	.464	.126	.053
Direkter Effekt	.177	.010	-.145	.030	.090	.160
Mathematikleistungen						
Totaler Effekt	-.201	.241	.084	.350	-.141	.072
Indirekter Effekt	.001	.895	-.008	.524	.019	.312
Direkter Effekt	-.103	.234	.091	.291	-.160	.033

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; unter Kontrolle von Geschlecht, zu Hause gesprochener Sprache, Sozialstatus und allgemeiner kognitiver Fähigkeit der Schülerinnen und Schüler sowie der Komposition dieser Merkmale in der Unterrichtsgruppe und der Schultart; N = 1188

Bei den Mathematikleistungen zeigen sich wie in den Analysen in Abschnitt 8.2.2.1 – bei denen kein Zusammenhang zwischen den Überzeugungen und den Mathematikleistungen (siehe Tabelle 8.15) festgestellt werden konnte – erwartungsgemäß keine indirekten Effekte der Adaptivität. Dennoch findet sich für die Einstellung von Lehrkräften zur Förderung schwacher Schülerinnen und Schüler ein direkter negativer Effekt auf die Mathematikleistungen (siehe Abbildung 8.10).⁶²

62 Da bei den Mediationsmodellen von der Reformbereitschaft der Lehrkräfte und ihrer Einstellung zur Notwendigkeit von Druck bei Lernprozessen auf die Mathematikleistungen der Schülerinnen und Schüler weder signifikante indirekte noch direkte und auch keine signifikanten totalen Effekte aufgetreten sind (siehe Tabelle 8.19) und die anderen beiden abzubildenden Pfade in den Mediationsmodellen ebenso keine signifikanten Zusammenhänge aufzeigen, wird an dieser Stelle auf die Abbildung der Mediationsmodelle verzichtet. Die ausführlichen Ergebnisse der Analysen können auf Anfrage zur Verfügung gestellt werden.

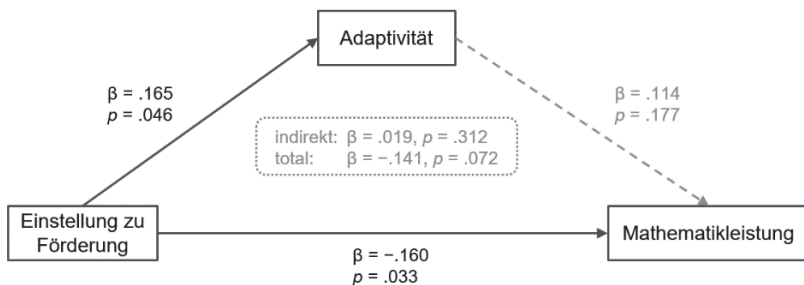


Abbildung 8.10: Grafisches Mediationsmodell zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator

8.2.3.2 Das Unterrichtsmerkmal Disziplin als Mediator

Englischunterricht

Durch die Mediationsanalysen mit Disziplin als Mediator konnte bestätigt werden, was sich bereits in Abschnitt 8.2.2.2 andeutete: Disziplin im Unterricht fungiert nicht als Mediator beim Zusammenhang der Überzeugungen der Lehrkräfte mit dem Wohlfühlen der Schülerinnen und Schüler im Englischunterricht (siehe Tabelle 8.20). Die direkten Effekte der Überzeugungen auf das Wohlfühlen sind uns bereits aus Abschnitt 8.2.2 bekannt und können in den folgenden drei Abbildungen nochmal nachvollzogen werden. Für die Englischleistungen sind weder signifikante indirekte noch direkte Effekte festzustellen. Daher wird auf die Abbildungen dieser Mediationsmodelle verzichtet.⁶³

63 Darüber hinaus sind in den Mediationsmodellen zur Englischleistung keine weiteren signifikanten totalen Effekte aufgetreten (siehe Tabelle 8.20) und die anderen beiden abzubildenden Pfade der Mediationsmodelle zeigen auch keine signifikanten Zusammenhänge auf. Die ausführlichen Ergebnisse der Analysen können auf Anfrage zur Verfügung gestellt werden.

Tabelle 8.20: Übersicht über die Ergebnisse der Mediationsanalysen zu Disziplin im Englischunterricht

Unabhängige Variable: Überzeugung	Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung	
	β	p	β	p	β	p
Abhängige Variable: Schulische Ergebnisse						
Wohlfühlen						
Totaler Effekt	.234	.015	-.248	.005	.358	.000
Indirekter Effekt	-.021	.304	.012	.453	.010	.472
Direkter Effekt	.255	.008	-.261	.005	.348	.000
Englischleistungen						
Totaler Effekt	-.145	.027	.088	.237	-.063	.424
Indirekter Effekt	-.013	.375	.010	.468	.015	.345
Direkter Effekt	-.132	.065	.078	.307	-.078	.322

Anmerkungen: signifikante Regressionskoeffizienten β: fett ($p < .05$); yx-standardisierte Koeffizienten; unter Kontrolle von Geschlecht, zu Hause gesprochener Sprache, Sozialstatus und allgemeiner kognitiver Fähigkeit der Schülerinnen und Schüler sowie der Komposition dieser Merkmale in der Unterrichtsgruppe und der Schulart; N = 1043

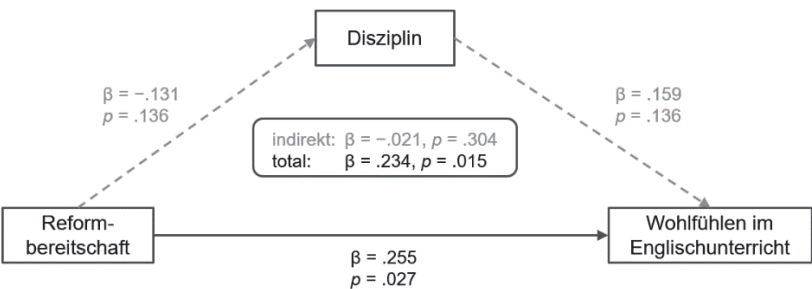


Abbildung 8.11: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Disziplin als Mediator

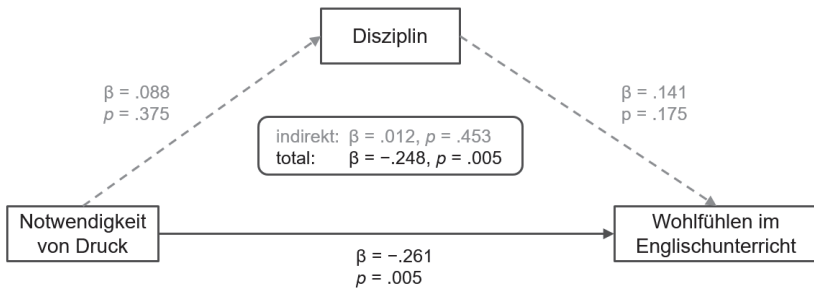


Abbildung 8.12: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Disziplin als Mediator

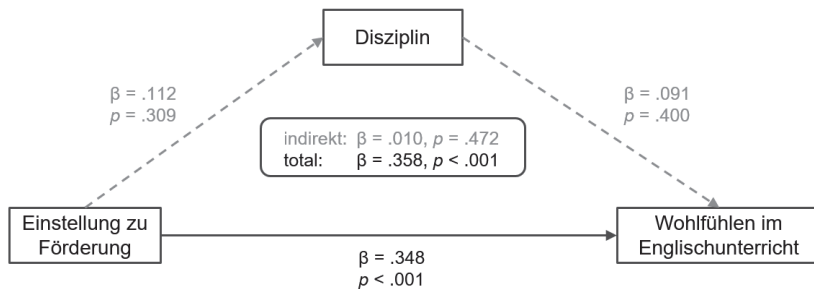


Abbildung 8.13: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Disziplin als Mediator

Mathematikunterricht

Auch bei der Betrachtung der Disziplin im Mathematikunterricht treten, wie im Englischunterricht, keine Mediationseffekte auf (siehe Tabelle 8.21). Es gibt aber direkte Effekte der Reformbereitschaft und der Einstellung zur Notwendigkeit von Druck bei Lernprozessen auf das Wohlfühlen der Schülerinnen und Schüler (siehe Abbildung 8.14 und Abbildung 8.15). Diese sind uns bereits aus den bisherigen Analysen in Abschnitt 8.2.2 und auch aus den Mediationsanalysen des vorigen Abschnitts (siehe Tabelle 8.19) bekannt.

Tabelle 8.21: Übersicht über die Ergebnisse der Mediationsanalysen zur Disziplin im Mathematikunterricht

Unabhängige Variable: Überzeugung	Reformbereitschaft		Notwendigkeit von Druck bei Lernprozessen		Einstellung zur Förderung	
	β	p	β	p	β	p
<i>Abhängige Variable:</i> Schulische Ergebnisse						
Wohlfühlen						
Totaler Effekt	.205	.054	-.262	.009	.201	.054
Indirekter Effekt	-.018	.413	.036	.240	.001	.919
Direkter Effekt	.223	.033	-.297	.003	.201	.054
Mathematikleistungen						
Totaler Effekt	-.094	.284	.071	.436	-.148	.054
Indirekter Effekt	.003	.775	-.006	.729	.000	.985
Direkter Effekt	-.097	.271	.077	.411	-.148	.054

Anmerkungen: signifikante Regressionskoeffizienten β : fett ($p < .05$); yx-standardisierte Koeffizienten; unter Kontrolle von Geschlecht, zu Hause gesprochener Sprache, Sozialstatus und allgemeiner kognitiver Fähigkeit der Schülerinnen und Schüler sowie der Komposition dieser Merkmale in der Unterrichtsgruppe und der Schulart; N = 1188

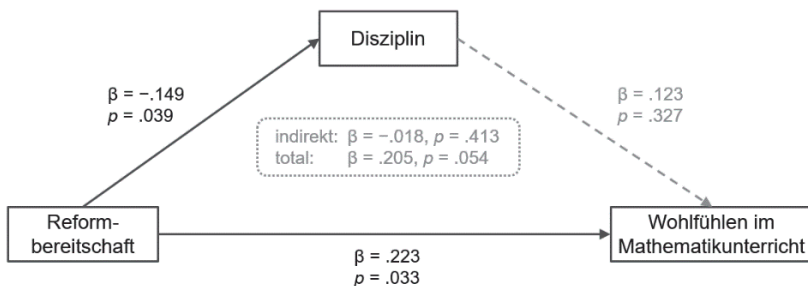


Abbildung 8.14: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Disziplin als Mediator

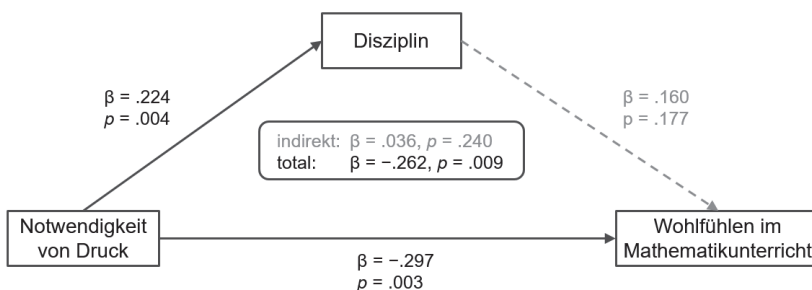


Abbildung 8.15: Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Disziplin als Mediator

Auch für die Mathematikleistungen als abhängiger Variable gibt es, wie beim Wohlfühlen im Mathematikunterricht, keine Mediationseffekte für die Disziplin im Unterricht (siehe Tabelle 8.21). Die folgenden beiden Abbildungen dienen lediglich der Verdeutlichung des Zusammenhangs von Reformbereitschaft (Abbildung 8.16) und der Einstellung zur Notwendigkeit von Druck (Abbildung 8.17) mit der Disziplin im Unterricht in den Mediationsmodellen.⁶⁴ Der negative Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und der Disziplin im Mathematikunterricht tritt hier zum ersten Mal auf, in den Regressionsanalysen aus Abschnitt 8.2.1.2 (Tabelle 8.9, Modell 1) konnte dieser Zusammenhang nicht nachgewiesen werden. Der positive Zusammenhang der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit der Disziplin im Mathematikunterricht tritt hingegen bereits in Abschnitt 8.2.1.2 (Tabelle 8.9, Modell 2) auf.

64 Auf die Abbildung des Mediationsmodells zur Vermittlung der Einstellung zur Förderung auf die Mathematikleistungen über die Disziplin im Unterricht wird an dieser Stelle verzichtet, da dort weder ein signifikanter indirekter oder direkter noch ein signifikanter totaler Effekt aufgetreten ist (siehe Tabelle 8.21) und auch die anderen beiden abzubildenden Pfade des Mediationsmodells keine signifikanten Zusammenhänge aufzeigen. Die ausführlichen Ergebnisse der Analyse können auf Anfrage zur Verfügung gestellt werden.

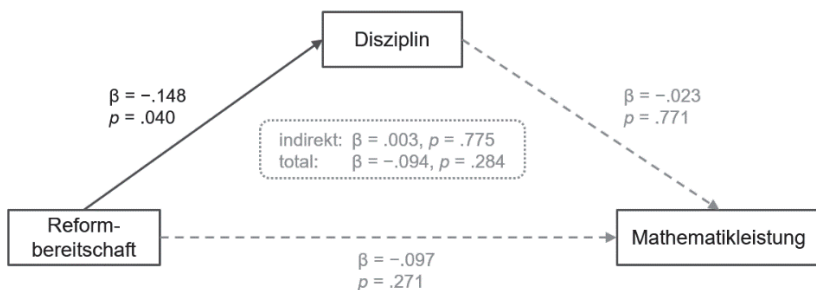


Abbildung 8.16: Grafisches Mediationsmodell zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Disziplin als Mediator

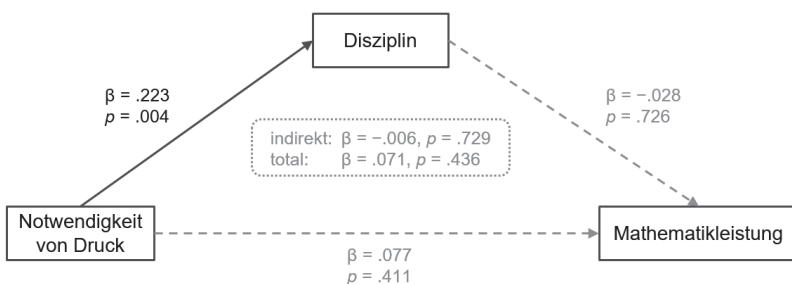


Abbildung 8.17: Grafisches Mediationsmodell zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Disziplin als Mediator

9 Zusammenfassung und Diskussion der Ergebnisse

Ausgangspunkt dieser Arbeit war es, zu klären, ob und wie allgemeine pädagogische Überzeugungen von Lehrkräften mit den schulischen Ergebnissen von Schülerinnen und Schülern zusammenhängen und über welche Unterrichtsmerkmale ein solcher Zusammenhang vermittelt sein könnte. Für den Englischunterricht lässt sich diese Frage eindeutig beantworten: Allgemeine pädagogische Überzeugungen hängen mit Unterrichtsmerkmalen aus dem Bereich des unterstützenden Unterrichtsklimas zusammen. Exemplarisch für die Adaptivität des Unterrichts kann darüber hinaus eine Vermittlung der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Englischunterricht festgestellt werden.

In Kapitel 9.1 werden die Ergebnisse, die im vorigen Kapitel 8.2 ausführlich dargestellt wurden, vor dem Hintergrund der in Kapitel 5 aufgestellten Fragestellungen und Hypothesen betrachtet. Nach der Diskussion der Befunde in Kapitel 9.2 wird in Kapitel 9.3 das Fazit der Arbeit festgehalten.

9.1 Beantwortung der Fragestellungen und Prüfung der Hypothesen

Die in Kapitel 5 formulierten Forschungsfragen und aufgestellten Hypothesen, die in Kapitel 8.2 ausführlich bearbeitet wurden, werden im Folgenden beantwortet und geprüft.

Fragestellung 1

Die erste Fragestellung bezieht sich auf den Zusammenhang zwischen den Überzeugungen der Lehrkräfte und ihrem unterrichtlichen Handeln im Englisch- bzw. Mathematikunterricht. Diese Zusammenhänge unterscheiden sich stark zwischen den beiden Unterrichtsfächern. Im Mathematikunterricht zeigen sich nur sehr wenige Zusammenhänge zwischen den einzelnen Überzeugungen und den untersuchten Unterrichtsmerkmalen (siehe Tabelle 9.2). Im Englischunterricht zeigen sich bei einem Großteil der Unterrichtsmerkmale Zusammenhänge mit den Überzeugungen (siehe Tabelle 9.1). Im Folgenden wird dargelegt, was daraus für die drei ersten Hypothesen geschlossen werden kann.

Tabelle 9.1: Zusammenhang von Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen im Englischunterricht: Ergebnisse der Mehrebenen-Regressionsanalysen im Überblick

		Reformbereit- schaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
unterstützendes Unterrichtsklima	Unterstützung	+	–	+
	Strenge	–	+	–
	Adaptivität des Unterrichts	+	–	+
	Positive Schülerorientiertheit/ Lehrer-Schüler-Beziehung	+	–	+
	Individualisierungsmaßnahmen/ Förderungsorientiertheit	☒	–	+
	Motivierungsbemühen/Lob	+	–	+
	Vertrauensgrad/ Schülerorientierung	+	–	+
Klassenführung	Strukturiertheit des Unterrichts	☒	☒	+
	Beschäftigungsradius	☒	☒	☒
	Disziplin und Ordnungssinn	–	☒	☒
	Zeitverschwendung	☒	☒	☒

Anmerkungen: +: signifikanter positiver Zusammenhang, –: signifikant negativer Zusammenhang, ☒: kein signifikanter Zusammenhang ($p < .05$)

Gemäß der ersten Hypothese wurde erwartet, dass Lehrkräfte mit einer hohen Reformbereitschaft und einer positiven Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler aufmerksamer gegenüber den unterschiedlichen Lernbedürfnissen insbesondere der schwächeren Schülerinnen und Schüler sind und ihren Unterricht eher adaptiv gestalten, während Lehrkräfte, die der Überzeugung sind, dass Druck für Lernen notwendig ist, den unterschiedlichen Bedürfnissen der Schülerinnen und Schüler keine große Bedeutung beimessen und ihren Unterricht weniger adaptiv gestalten. Die erste Hypothese kann für den Englischunterricht bestätigt werden (siehe Tabelle 9.1), für den Mathematikunterricht dagegen nicht (siehe Tabelle 9.2), da hier keine Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und der Adaptivität ihres Unterrichts auftreten.

Tabelle 9.2: Zusammenhang von Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen im Mathematikunterricht: Ergebnisse der Mehrebenen-Regressionsanalysen im Überblick

		Reform- bereitschaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
unterstützendes Unterrichtsklima	Unterstützung	☒	–	☒
	Strenge	☒	+	☒
	Adaptivität des Unterrichts	☒	☒	☒
	Positive Schülerorientiertheit/ Lehrer-Schüler-Beziehung	☒	☒	☒
	Individualisierungsmaßnahmen/ Förderungsorientiertheit	☒	☒	☒
	Motivierungsbemühen/Lob	☒	☒	☒
	Vertrauensgrad/ Schülerorientierung	☒	☒	+
Klassenführung	Strukturiertheit des Unterrichts	☒	☒	☒
	Beschäftigungsradius	☒	☒	☒
	Disziplin und Ordnungssinn	☒	+	☒
	Zeitverschwendung	☒	☒	☒

Anmerkungen: +: signifikanter positiver Zusammenhang, –: signifikant negativer Zusammenhang, ☒: kein signifikanter Zusammenhang ($p < .05$)

Die zweite Hypothese bezog sich auf die Disziplin im Unterricht. Es wurde erwartet, dass Lehrkräfte mit einer hohen Reformbereitschaft und einer positiven Einstellung zur Förderung schwacher Schülerinnen und Schüler der Disziplin im Unterricht eher eine niedrige Bedeutung beimessen, diejenigen Lehrkräfte dagegen, die der Überzeugung sind, dass Druck für Lernprozesse notwendig ist, stärker auf Disziplin im Unterricht achten. Die zweite Hypothese kann für die Reformbereitschaft der Englischlehrkräfte bestätigt werden (siehe Tabelle 9.1). Bei den Überzeugungen der Englischlehrkräfte zu Druck und Förderung zeigen sich keine Zusammenhänge mit der Disziplin im Unterricht, sodass hier die zweite Hypothese zurückzuweisen ist. Im Mathematikunterricht weisen Lehrkräfte, die eher die Einstellung vertreten, dass Druck für Lernprozesse notwendig ist, höhere Werte bei der Disziplin im Unterricht auf, somit ist für diese Überzeugung die zweite Hypothese bestätigt (siehe Tabelle 9.2). Zwischen der Reformbereitschaft sowie der Einstellung zur Förderung schwacher

Schülerinnen und Schüler zeigen sich keine Zusammenhänge mit der Disziplin im Mathematikunterricht, sodass die zweite Hypothese für diese beiden Überzeugungen widerlegt ist.

Die dritte Hypothese betrifft die Zusammenhänge der Überzeugungen und der weiteren Unterrichtsmerkmale aus den Bereichen des unterstützenden Unterrichtsklimas und der Klassenführung, die noch nicht in den beiden ersten Hypothesen berücksichtigt wurden. Zuerst wird die dritte Hypothese für folgende Merkmale aus dem Bereich des unterstützenden Unterrichtsklimas betrachtet: Unterstützung, Strenge, positive Schülerorientiertheit, Individualisierungsmaßnahmen, Motivierungsbemühen und Vertrauensgrad. Bei den jeweils positiv formulierten Merkmalen (Unterstützung H3a, Schülerorientiertheit H3c, Individualisierungsmaßnahmen H3d, Motivierungsbemühen H3e, Vertrauensgrad H3f) wurde ein positiver Zusammenhang mit der Reformbereitschaft und der Einstellung zur Förderung schwacher Schülerinnen und Schüler angenommen, bei der Strenge entsprechend der Polung ein negativer (H3b). Bei der Überzeugung, dass Druck bei Lernprozessen nötig ist, wurde von einem negativen Zusammenhang mit diesen Merkmalen ausgegangen (bzw. positiv bei Strenge). Im Englischunterricht (siehe Tabelle 9.1) gibt es keinen Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und den Individualisierungsmaßnahmen, daher wird für diesen Fall die Hypothese 3d widerlegt. Alle anderen Zusammenhänge für die Unterrichtsmerkmale aus dem Bereich des unterstützenden Unterrichtsklimas mit den Überzeugungen von Lehrkräften weisen die erwartete Richtung auf und bestätigen die Hypothesen 3a bis 3f.

Im Mathematikunterricht (siehe Tabelle 9.2) liegt ein entgegengesetztes Zusammenhangsmuster vor: Von den erwarteten Zusammenhängen der Unterrichtsmerkmale aus dem Bereich des unterstützenden Unterrichtsklimas mit den Überzeugungen der Lehrkräfte werden die Hypothesen 3a bis 3f bis auf drei Fälle widerlegt. Lediglich die Einstellung der Lehrkräfte zur Notwendigkeit von Druck bei Lernprozessen weist einen Zusammenhang in der erwarteten Richtung zur Unterstützung (H3a) und zur Strenge auf (H3b). Zudem gibt es einen Zusammenhang zwischen der Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler und dem Vertrauensgrad (H3f). Nur für diese Zusammenhänge einzelner Überzeugungen und Unterrichtsmerkmale können die Hypothesen 3a, 3b und 3f bestätigt werden.

Die dritte Hypothese bezieht sich auch auf folgende Unterrichtsmerkmale aus dem Bereich der Klassenführung: Strukturiertheit, Beschäftigungsradius und Zeitverschwendung. Bei den Unterrichtsmerkmalen Strukturiertheit (H3g) und Beschäftigungsradius (H3h) wurde davon ausgegangen, dass kein Zusammenhang dieser Merkmale der inhaltlichen Unterrichtsstrukturiertheit mit den pädagogischen Überzeugungen der Lehrkräfte besteht. Bei der Zeitverschwendung (H3i) wurde angenommen, dass Lehrkräfte, die eine besonders hohe Reformbereitschaft und eine positive Einstellung zur Förderung zeigen, auch mehr Zeit im Unterricht verschwenden als solche Lehrkräfte, die nicht so reformbereit und einer Förderung gegenüber weniger positiv eingestellt sind. Sie greifen bei Unterrichtsstörungen nicht immer so hart durch wie diejenigen Lehrkräfte, die der Überzeugung sind, dass Druck beim Lernprozess notwendig ist. Betrachtet man nun die Unterrichtsmerkmale der dritten Hypothese aus dem Bereich der Klassenführung, so zeigen die Ergebnisse für den Englisch- und Mathematikunterricht überwiegend Gemeinsamkeiten: Für beide Fächer werden die Hypothesen 3g und 3h größtenteils bestätigt, da kein Zusammenhang der Überzeugungen der Lehrkräfte mit der Strukturiertheit und dem Beschäftigungsradius auftritt. Lediglich für die Einstellung der Englischlehrkräfte zur Förderung schwacher Schülerinnen und Schüler kann ein Zusammenhang mit der Strukturiertheit des Englischunterrichts festgestellt werden, was in diesem Fall die Hypothese 3g widerlegt. Auch die angenommenen Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und der Zeitverschwendung in ihrem Unterricht treten für beide Unterrichtsfächer nicht auf, somit kann die Hypothese 3i nicht bestätigt werden.

Zusammenfassend kann festgestellt werden, dass sich die übergreifende Annahme eines Zusammenhangs der Überzeugungen von Lehrkräften mit den Unterrichtsmerkmalen für ein unterstützendes Unterrichtsklima im Englischunterricht bestätigen lässt. Für den Mathematikunterricht und für die Unterrichtsmerkmale aus dem Bereich der Klassenführung, welche die soziale Strukturierung abbilden, wird diese Annahme größtenteils widerlegt.

Fragestellungen 2 und 3

Die zweite Fragestellung bezieht sich auf den Zusammenhang zwischen den Überzeugungen der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler im Englisch- bzw. Mathematikunterricht und der Vermittlung der Überzeugungen durch das Unterrichtshandeln der Lehrkräfte in Form von Adaptivität oder Disziplin. Die dritte Fragestellung bezieht sich auf den Zu-

sammenhang zwischen den Überzeugungen von Lehrkräften und den Englisch- bzw. Mathematikleistungen von Schülerinnen und Schülern und der Vermittlung der Überzeugungen durch Adaptivität oder Disziplin. Diese beiden Fragestellungen werden in der folgenden Zusammenfassung der Ergebnisse hinsichtlich der aufgestellten Hypothesen gemeinsam betrachtet.

Zunächst kann festgehalten werden, dass es im Englischunterricht einen Zusammenhang zwischen den Überzeugungen der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler in der jeweils erwarteten Richtung gibt: Schülerinnen und Schüler, die von einer Lehrkraft unterrichtet werden, welche eine hohe Reformbereitschaft, eine negative Einstellung zur Notwendigkeit von Druck bei Lernprozessen oder eine positive Einstellung zur Förderung schwacher Schülerinnen und Schüler aufweist, fühlen sich im Unterricht wohler als ihre Mitschülerinnen und Mitschüler. Dies bestätigt die Hypothese 4.1. Auch im Mathematikunterricht zeigen sich die erwarteten Zusammenhänge, allerdings findet sich bei der Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler kein signifikanter Zusammenhang mit dem Wohlfühlen der Schülerinnen und Schüler. Bei zwei von drei Überzeugungen lässt sich die Hypothese 4.1 also auch im Mathematikunterricht bestätigen.

Für den Zusammenhang der Überzeugungen der Lehrkräfte mit den Leistungen der Schülerinnen und Schüler wurden keine gerichteten Hypothesen formuliert. Es gibt einen negativen Zusammenhang zwischen der Reformbereitschaft der Lehrkräfte und den Englischleistungen der Schülerinnen und Schüler. Für die beiden anderen Überzeugungen im Englischunterricht und alle drei Überzeugungen von Mathematiklehrkräften treten keine Zusammenhänge mit den Leistungen ihrer Schülerinnen und Schüler auf.

Die Hypothesen 4.2 und 4.3 behandeln den Vermittlungsprozess der pädagogischen Überzeugungen von Lehrkräften über die Adaptivität oder die Disziplin des Unterrichts zum Wohlfühlen der Schülerinnen und Schüler. Für die Vermittlung mit den Lernergebnissen der Schülerinnen und Schüler wurden keine Hypothesen aufgestellt.

Bei der Hypothese 4.2 wurde erwartet, dass die Auswirkungen der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Unterricht über die Adaptivität des Unterrichts vermittelt sind. Demnach müsste eine positive Einstellung gegenüber der Förderung schwacher Schülerinnen

und Schüler sowie eine hohe Reformbereitschaft vermittelt über die Adaptivität in einem positiven und die Einstellung zur Notwendigkeit von Druck bei Lernprozessen in einem negativen Zusammenhang mit dem Wohlfühlen als nicht-kognitivem Ergebnis der Schülerinnen und Schüler stehen. Diese Hypothese kann für den Englischunterricht bestätigt werden (siehe indirekte Effekte in Tabelle 9.3), für den Mathematikunterricht aber nicht (siehe Tabelle 9.4). Auch für die Englischleistungen und die Mathematikleistungen der Schülerinnen und Schüler wurde keine Vermittlung der pädagogischen Überzeugungen der Lehrkräfte über die Adaptivität ihres Unterrichts festgestellt (siehe Tabelle 9.3 und Tabelle 9.4).

Tabelle 9.3: Adaptivität im Englischunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick

	Reformbereitschaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
Wohlfühlen			
Indirekter Effekt	+	–	+
Direkter Effekt	☒	☒	+
Englischleistungen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	–	+	☒

Anmerkungen: +: signifikanter positiver Effekt, –: signifikant negativer Effekt, ☒: kein signifikanter Effekt ($p < .05$)

Die untersuchten Mediationsmodelle zum Vermittlungseffekt der Adaptivität des Englisch- und Mathematikunterrichts liefern über die Beantwortung der Forschungsfragen hinaus noch weitere interessante Erkenntnisse: So weisen die verbliebenen direkten Effekte darauf hin, dass Beziehungen zwischen den pädagogischen Überzeugungen der Lehrkräfte und den Ergebnissen der Schülerinnen und Schüler bestehen, die sich nicht durch die Adaptivität des Unterrichts erklären lassen. Dies betrifft im Englischunterricht den negativen Zusammenhang zwischen der Reformbereitschaft und den Englischleistungen (wie er in den Regressionsanalysen in Tabelle 8.13 bereits festgestellt wurde) und den positiven Zusammenhang zwischen der Einstellung zur Notwendigkeit von Druck bei Lernprozessen und den Englischleistungen (siehe Tabelle 9.3). Dieser positive Zusammenhang zwischen der Notwendigkeit von Druck und

den Englischleistungen tritt in den Mediationsmodellen zum ersten Mal auf. Darüber hinaus bleibt bei der Betrachtung des Wohlfühlens der Schülerinnen und Schüler im Englischunterricht ein direkter Zusammenhang mit der Einstellung der Lehrkräfte gegenüber der Förderung schwacher Schülerinnen und Schüler bestehen, der nicht vollständig durch die Mediation über Adaptivität erklärt werden kann.

Tabelle 9.4: Adaptivität im Mathematikunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick

	Reformbereitschaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
Wohlfühlen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	+	–	☒
Mathematikleistungen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	☒	☒	–

Anmerkungen: +: signifikanter positiver Effekt, –: signifikant negativer Effekt, ☒: kein signifikanter Effekt ($p < .05$)

Auch im Mathematikunterricht treten in den Mediationsmodellen zur Untersuchung der Adaptivität des Unterrichts als Vermittlerin der pädagogischen Überzeugungen der Lehrkräfte auf die Ergebnisse der Schülerinnen und Schüler direkte Effekte auf, auch wenn keine Mediationen festgestellt werden konnten (siehe Tabelle 9.4). Beim Wohlfühlen der Schülerinnen und Schüler treten genau die direkten Effekte auf, die in Hypothese 4.1 angenommen und bis auf die Einstellung zur Förderung schwacher Schülerinnen und Schüler auch bestätigt werden konnten. Bei den Mathematikleistungen und der Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler besteht ein direkter negativer Zusammenhang, der in den Regressionsanalysen (in Tabelle 8.15) nicht signifikant war.

Bei Hypothese 4.3 wurde erwartet, dass die Auswirkungen der Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler im Unterricht über die Disziplin im Unterricht vermittelt sind. Demnach müsste eine positive Einstellung gegenüber der Förderung schwacher Schülerinnen und Schüler

sowie eine hohe Reformbereitschaft vermittelt über die Disziplin im Unterricht in einem positiven und die Einstellung zur Notwendigkeit von Druck bei Lernprozessen in einem negativen Zusammenhang mit dem Wohlfühlen stehen. Diese Hypothese kann weder für den Englischunterricht noch für den Mathematikunterricht bestätigt werden (siehe indirekte Effekte in Tabelle 9.5 und Tabelle 9.6). Ebenso wurde für die Englischleistungen und die Mathematikleistungen der Schülerinnen und Schüler keine Vermittlung der pädagogischen Überzeugungen der Lehrkräfte über die Disziplin in ihrem Unterricht festgestellt.

Tabelle 9.5: Disziplin im Englischunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick

	Reformbereitschaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
Wohlfühlen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	+	-	+
Engischleistungen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	☒	☒	☒

Anmerkungen: +: signifikanter positiver Effekt, -: signifikant negativer Effekt, ☒: kein signifikanter Effekt ($p < .05$)

Die direkten Effekte, die in Tabelle 9.5 und Tabelle 9.6 dargestellt sind, sind aus den vorangegangenen Analysen bekannt. Hierbei handelt es sich um die Zusammenhänge zwischen den pädagogischen Überzeugungen der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler im Englisch- bzw. teilweise im Mathematikunterricht, die gemäß Hypothese 4.1 bestätigt wurden und mit Hilfe der Mediationsmodelle erklärt werden sollten.

Tabelle 9.6: Disziplin im Mathematikunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick

	Reformbereitschaft	Notwendigkeit von Druck bei Lernprozessen	Einstellung zur Förderung
Wohlfühlen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	+	–	☒
Mathematikleistungen			
Indirekter Effekt	☒	☒	☒
Direkter Effekt	☒	☒	☒

Anmerkungen: +: signifikanter positiver Effekt, –: signifikant negativer Effekt, ☒: kein signifikanter Effekt ($p < .05$)

9.2 Diskussion der Befunde

Die drei in der vorliegenden Arbeit untersuchten pädagogischen Überzeugungen von Lehrkräften – Einstellung zur Förderung schwacher Schülerinnen und Schüler, Einstellung zur Notwendigkeit von Druck bei Lernprozessen und Reformbereitschaft – werden in aktuellen Studien selten untersucht (siehe Kapitel 3.3). Diese Arbeit leistet einen Beitrag dazu, die allgemeinen pädagogischen Überzeugungen von Lehrkräften in der Schul- und Unterrichtsforschung wieder stärker in den Blick zu nehmen.

Bei allen hier ermittelten Zusammenhängen zwischen den Überzeugungen von Lehrkräften und ihrem Unterrichtshandeln sowie den Ergebnissen ihrer Schülerinnen und Schüler vermittelt über das Unterrichtshandeln ist festzuhalten, dass sich die Befunde für den Englisch- und den Mathematikunterricht oft voneinander unterscheiden. Generell finden sich die erwarteten Zusammenhänge mit den pädagogischen Überzeugungen der Lehrkräfte häufiger im Englischunterricht und nur in Ausnahmefällen im Mathematikunterricht. Dies ist nicht darauf zurückzuführen, dass sich die Überzeugungen der Fachlehrkräfte stark voneinander unterscheiden: Betrachtet man die Mittelwerte der jeweiligen Überzeugungsskalen fallen keine gravierenden Unterschiede zwischen den Lehrkräften der Fächer Englisch und Mathematik auf (siehe Tabelle 7.7 in Abschnitt 7.2.1). Auch das ist ein Befund dieser Arbeit: Obwohl sich die Lehrkräfte der Fächer Englisch und Mathematik im Mittel nicht in ihren pädagogischen Überzeugun-

gen unterscheiden, unterscheiden sie sich stark darin, wie sich diese Überzeugungen in ihrem Unterricht bemerkbar machen. Dass es für Merkmale des Mathematikunterrichts weniger Zusammenhänge mit den untersuchten Überzeugungen der Lehrkräfte gibt, könnte in der Struktur des Unterrichtsfachs liegen: Der Mathematikunterricht läuft offenbar standardisierter und strukturierter ab als der Englischunterricht. Aufgrund des Curriculums scheint es stärkere inhaltliche Vorgaben für Aufgaben zu geben als beim Erlernen einer Fremdsprache, wo Aufgaben zum Sprachgebrauch und sprachliche Interaktionen der Schülerinnen und Schüler stärker variieren können. Im Mathematikunterricht scheint es weniger Gelegenheiten zu geben, dass allgemeine pädagogische Überzeugungen als Filter oder Rahmen (siehe Kapitel 4.1) wirken und das Unterrichtshandeln der Lehrkräfte bestimmen.

Hervorzuheben ist außerdem, dass nur wenige Zusammenhänge zwischen den pädagogischen Überzeugungen der Lehrkräfte und Unterrichtsmerkmalen aus dem Bereich der Klassenführung aufgetreten sind. Bei den Merkmalen der inhaltlichen Strukturiertheit des Unterrichts (Strukturiertheit des Unterrichts und Beschäftigungsradius) wurde dies auch erwartet. Allerdings war davon auszugehen, dass es einen Zusammenhang zwischen pädagogischen Überzeugungen, speziell der Einstellung zur Notwendigkeit von Druck bei Lernprozessen, und Merkmalen der erzieherischen Verhaltenssteuerung, wie Disziplin und Zeitverschwendung, geben könnte. Der dort aufgefundene Zusammenhang zwischen der Einstellung der Mathematiklehrkräfte zur Notwendigkeit von Druck bei Lernprozessen und der Disziplin im Mathematikunterricht hat sich nach der Adjustierung der Alphafehler als nicht signifikant herausgestellt. Bestätigt hat sich allerdings der Zusammenhang für die Reformbereitschaft der Englischlehrkräfte mit der Disziplin in ihrem Unterricht. Dies stützt das angeführte Argument, dass im Englischunterricht allgemeine pädagogische Überzeugungen eher als Filter oder Rahmen zum Tragen kommen als im Mathematikunterricht: Reformorientierte Englischlehrkräfte scheinen in ihrem Unterricht weniger auf Disziplin und Ordnungssinn zu achten. Insgesamt weisen Englischlehrkräfte, die sehr reformbereit sind, eine positive Einstellung zur Förderung schwacher Schülerinnen und Schüler haben und nicht der Meinung sind, dass Druck bei Lernprozessen unabdingbar ist, ein stärker unterstützendes Unterrichtsklima auf als ihre gegensätzlich eingestellten Fachkolleginnen und Fachkollegen.

Dass im Mathematikunterricht und bei Merkmalen der Klassenführung in beiden Fächern weniger Zusammenhänge zwischen den allgemeinen pädagogischen Überzeugungen und dem Unterrichtshandeln aufzufinden sind, könnte auch an der Funktion der allgemeinen pädagogischen Überzeugungen liegen. In Kapitel 4.1 wurde festgestellt, dass die hier untersuchten Überzeugungen in Bezug auf das Unterrichtshandeln von Lehrkräften am ehesten als Filter oder als Rahmen dienen können. Allerdings wurde bei den möglichen Ursachen für eine fehlende Übereinstimmung von Überzeugungen und unterrichtlichem Handeln (siehe Abschnitt 4.2.2) auch herausgearbeitet, dass Überzeugungen, die direkt das Handeln lenken, grundsätzlich leichter in Handlungen zu beobachten sind als solche Überzeugungen, welche die Funktion des Filterns und Interpretierens von neuen Informationen übernehmen. Überzeugungen, die in der Zuordnung ihrer Funktion näher bei konkreten Handlungen liegen als allgemeine pädagogische Überzeugungen mit ihrer Funktion des Filters oder Rahmens, lassen sich im Unterrichtshandeln von Lehrkräften leichter identifizieren (Buehl & Beck, 2015, S. 72).

Dass bezüglich der Ergebnisse der Schülerinnen und Schüler weniger Zusammenhänge der pädagogischen Überzeugungen mit den Fachleistungen und mehr Zusammenhänge mit dem Wohlfühlen der Schülerinnen und Schüler im Unterricht aufzufinden sind, überrascht nicht. Lipowsky (2006, S. 49) merkt dazu an, dass man grundsätzlich berücksichtigen muss, dass der Wirkungspfad zwischen handlungsfernen Komponenten der Expertise von Lehrkräften (wie Überzeugungen) und den Leistungen von Schülerinnen und Schülern „durch eine Reihe anderer Variablen beeinflusst wird, sodass mit schwachen Zusammenhängen zwischen diesen beiden Polen zu rechnen ist.“ Gerade im Hinblick auf die Fachleistungen kommt wohl ein Mechanismus zum Tragen, der in Kapitel 4.2 thematisiert wurde: Es kann sein, dass die in einer Unterrichtssituation relevanten Überzeugungen nicht vollständig erfasst werden können, wenn man einzelne Überzeugungen isoliert betrachtet und sie nicht mit den anderen Überzeugungen verknüpfen kann, die für den Unterricht relevant sind. Dies kann nach Seifried (2009, S. 119) besonders dann der Fall sein, wenn – wie in dieser Arbeit – mit den drei allgemeinen pädagogischen Überzeugungen ausschließlich fachunabhängige Überzeugungen erfasst werden, die fachdidaktischen Aspekte der Überzeugungen dagegen nicht.

Überzeugungen wirken auch über weitere, in den Analysen nicht berücksichtigte Unterrichtsprozesse auf die Lernergebnisse von Schülerinnen und Schüler

ein, worauf die Ergebnisse dieser Untersuchung bereits hingedeutet haben. Die Auswahl der zu untersuchenden Mediatoren erfolgte in dieser Arbeit theorie- und nicht datengetrieben. So konnte der Vermittlungseffekt von allgemeinen pädagogischen Überzeugungen der Lehrkräfte auf das Wohlfühlen der Schülerinnen und Schüler für die Adaptivität im Englischunterricht nachgewiesen werden, für die Disziplin im Englischunterricht jedoch nicht. Für zukünftige Analysen könnte deshalb die Betrachtung weiterer Merkmale eines unterstützenden Unterrichtsklimas interessant sein, um zu sehen, ob auch hier, wie bei der Adaptivität, Mediationseffekte vorliegen. Weiterhin müsste untersucht werden, ob es nicht doch auch im Mathematikunterricht Unterrichtsmerkmale gibt, welche die wenigen, aber dennoch vorhandenen direkten Effekte zwischen den Überzeugungen und den Ergebnissen der Schülerinnen und Schüler im Mathematikunterricht erklären können. Denkbar wären hier Merkmale aus dem Bereich der kognitiven Aktivierung, die in dieser Untersuchung nicht berücksichtigt werden konnten, da in der Drei-Länder-Studie dieser Bereich der Unterrichtsqualität nicht erfasst wurde.

Die vorliegenden Analysen unterliegen einigen Einschränkungen. Aufgrund des Querschnittsdesigns der Drei-Länder-Studie lässt sich die Wirkrichtung der Zusammenhänge in den Mediationsmodellen nicht bestimmen, da sich mit Querschnittsdaten keine empirischen Tests von Kausalität durchführen lassen und damit auch keine „harten“ Tests von Mediatoreffekten möglich sind (Urban & Mayerl, 2018, S. 340). Also kann weder das Verhältnis remedialer zu förderlichen Effekten beim adaptiven Unterrichtshandeln noch die Interdependenz zwischen der Reformbereitschaft der Lehrkräfte und dem Leistungsniveau ihrer Schülerschaft abschließend geklärt werden. Ob Lehrkräfte, die eine hohe Ausprägung bei der Einstellung zur Förderung schwacher Schülerinnen und Schüler aufweisen, aufgrund ihrer Einstellung ihren Unterricht adaptiver gestalten oder ob sie damit auf schwache Leistungen der Schülerinnen und Schüler reagieren (oder beides), lässt sich in einer querschnittlichen Betrachtung der Zusammenhänge und der Mediation zwischen Überzeugung, Adaptivität und Ergebnissen der Schülerinnen und Schüler nicht ermitteln. Der festgestellte negative Zusammenhang zwischen der Reformbereitschaft der Englischlehrkräfte und den Englischleistungen ihrer Schülerinnen und Schüler könnte daher rühren, dass Lehrkräfte in Unterrichtsgruppen mit einem niedrigen Leistungsniveau eher dazu neigen, die gewohnten Unterrichtsmethoden zu hinterfragen, und offen für Innovationen sind. Lehrkräfte in Unterrichtsgruppen mit einem hohen Leistungsniveau sehen dagegen weniger eine Notwendigkeit zu Verän-

derungen. Um gerichtete Wirkungen der Überzeugungen der Lehrkräfte, vermittelt über ihr Handeln, auf die Lernergebnisse der Schülerinnen und Schüler zu untersuchen, sind Längsschnittanalysen notwendig.

Die Zusammensetzung der Stichprobe ist nicht repräsentativ. Sie stellt kein Abbild der gesamten Schülerschaft der Jahre 1978 und 1979 dar, sie wurde vielmehr zusammengestellt mit dem Fokus auf dem Vergleich von Schülerinnen und Schülern der 9. Jahrgangsstufe dreier Bundesländer in der damaligen Zusammensetzung von Schularten mit einer besonderen Betonung der Untersuchung an Gesamtschulen (siehe Kapitel 6.1). Allerdings ist besonders die Stichprobenzusammensetzung des Bundeslands Niedersachsen in der Drei-Länder-Studie als innovativ zu bewerten, da mit der Berücksichtigung strukturell ähnlicher Vergleichsregionen der wichtige Aspekt des regionalen Kontexts bei der Ziehung der Schulen berücksichtigt wurde. Zudem liegt für die Region Wetzlar in Hessen eine Vollerhebung aller Schulen des damaligen Flächenversuchs vor, eine Datenlage, die heute nahezu unvorstellbar ist, da sich die Frage der Repräsentativität – zumindest für diese Region – überhaupt gar nicht erst stellt. Auch wenn die gesamte Stichprobe der Drei-Länder-Studie nicht den Kriterien einer repräsentativen Stichprobenauswahl entspricht, liegt in den Daten ein großes Potential hinsichtlich der allgemeinen Untersuchung pädagogischer und schulischer Zusammenhänge, gerade für einen historischen Vergleich im Sinne einer Analyse des schulischen Wandels oder auch der Kontinuität von Schule und Unterricht.

9.3 Fazit

Zunächst werden noch einmal die wichtigsten Ergebnisse der vorliegenden Arbeit übersichtlich zusammengefasst:

Ziel dieser Arbeit war es, die Bedeutung der allgemeinen pädagogischen Überzeugungen von Lehrkräften für ihr Handeln, vor allem in Form der Adaptivität und der Disziplin des Unterrichts, sowie für die Lernergebnisse der Schülerinnen und Schüler (Wohlfühlen und Fachleistungen) für den Englisch- und Mathematikunterricht zu untersuchen.

Der Zusammenhang zwischen der *Adaptivität* des Englischunterrichts und den Überzeugungen der Lehrkräfte konnte empirisch bestätigt werden. Hohe Ausprägungen bei der Reformbereitschaft von Englischlehrkräften und eine posi-

ve Einstellung zur Förderung schwacher Schülerinnen und Schüler gehen mit einem adaptiveren Unterricht einher. Lehrkräfte, die die Notwendigkeit von Druck beim Lernen sehen, gestalten ihren Englischunterricht weniger adaptiv. Für den Mathematikunterricht konnte diese Annahme nicht bestätigt werden, hier zeigen sich keine Zusammenhänge zwischen den Überzeugungen der Lehrkräfte und der Adaptivität ihres Unterrichts. Die Annahme eines Zusammenhangs zwischen den pädagogischen Überzeugungen und der *Disziplin* im Unterricht konnte nur teilweise empirisch abgesichert werden. Er hat sich nur für die Reformbereitschaft von Englischlehrkräften und für die Einstellung zur Notwendigkeit von Druck bei Mathematiklehrkräften gezeigt.

Insgesamt wird die Frage des Zusammenhangs der hier untersuchten Überzeugungen und den Unterrichtsmerkmalen dahingehend bilanziert, dass dieser vornehmlich für Merkmale eines unterstützenden Unterrichtsklimas im Englischunterricht empirisch abgesichert werden kann. Nur hier zeigen sich durchgehend Zusammenhänge zwischen den pädagogischen Überzeugungen der Lehrkräfte und den jeweiligen Unterrichtsmerkmalen. Für Unterrichtsmerkmale der Klassenführung und für den Mathematikunterricht liegen solche Zusammenhänge nicht systematisch vor.

Die Annahme eines Zusammenhangs der pädagogischen Überzeugungen der Lehrkräfte mit dem Wohlfühlen der Schülerinnen und Schüler im Unterricht wurde weitgehend empirisch gestützt. Schülerinnen und Schüler, die von Lehrkräften unterrichtet wurden, die eine hohe Reformbereitschaft im Englisch- bzw. im Mathematikunterricht oder eine positive Einstellung zur Förderung schwacher Schülerinnen und Schüler im Englischunterricht aufwiesen, fühlten sich im Unterricht wohler als ihre Mitschülerinnen und Mitschüler. Schülerinnen und Schüler, deren Lehrkraft Druck beim Lernen als notwendig erachtete, fühlten sich sowohl im Englisch- als auch im Mathematikunterricht weniger wohl.

Bei der Reformbereitschaft und der Notwendigkeit von Druck in Lernprozessen ist weiterhin davon auszugehen, dass dieser Zusammenhang zumindest im Englischunterricht durch *Adaptivität* vermittelt ist. Der Zusammenhang zwischen einer Förderorientierung der Lehrkräfte und dem Wohlfühlen der Schülerinnen und Schüler bleibt hingegen auch bei einer teilweisen Vermittlung über die Adaptivität des Englischunterrichts bestehen (partieller Mediatoreffekt, siehe Abbildung 8.3 in Abschnitt 8.2.3). Es ist anzunehmen, dass die Vermitt-

lungsprozesse hier nicht nur über die Adaptivität, sondern auch über Unterrichtsmerkmale ablaufen, die im Rahmen der vorliegenden Analysen nicht betrachtet wurden. Die Annahme zur Vermittlung der pädagogischen Überzeugungen von Lehrkräften über die *Disziplin* im Unterricht auf das Wohlfühlen von Schülerinnen und Schülern lässt sich empirisch nicht belegen. Im Mathematikunterricht konnten keine Vermittlungsprozesse pädagogischer Überzeugungen auf das Wohlfühlen von Schülerinnen und Schülern nachgewiesen werden.

Hinsichtlich der Englisch- oder Mathematikleistungen wurden keine gerichteten Annahmen über die Wirkung der Überzeugungen formuliert. Tatsächlich fand sich für zwei der drei untersuchten Überzeugungen von Lehrkräften, nämlich die Einstellung zur Förderung schwacher Schülerinnen und Schüler und die Einstellung zur Notwendigkeit von Druck bei Lernprozessen, kein statistisch signifikanter Zusammenhang mit den Englischleistungen der Schülerinnen und Schüler (siehe Tabelle 8.13). Je höher aber die Reformbereitschaft einer Lehrkraft ist, desto niedriger ist das Leistungsniveau ihrer Unterrichtsgruppe – und zwar unabhängig davon, ob zugleich die wahrgenommene Adaptivität als erklärende Variable berücksichtigt wird oder nicht. Möglicherweise liegt hier eine umgekehrte Kausalität zugrunde, die mit dem Unterrichtshandeln gar nichts zu tun hat: Je schwächer die Lernergebnisse sind, desto mehr könnten Lehrkräfte gewillt sein, externe Veränderungen zu akzeptieren. Dieser Zusammenhang beruht übrigens nicht auf Einflüssen der Schulart. Zum einen wurde in den Modellen die Schulart kontrolliert, zum anderen zeigt sich dasselbe Muster von Zusammenhängen mit den Englischleistungen, wenn man ausschließlich die Gesamtschulen betrachtet⁶⁵ (bei den übrigen Schularten wäre dafür die Stichprobe zu klein). Bemerkenswert ist außerdem der Befund, dass ein adaptives Unterrichtshandeln nicht ausschließlich als remediale Reaktion auf unzureichende Leistungen interpretiert werden kann. Bei gleichen Ausgangsbedingungen (einschließlich Schulart) der Lernenden und den gleichen Einstellungen der Lehrkraft zu Reformen im Schulwesen besteht sogar ein positiver Zusammenhang zwischen der Adaptivität und den Englischleistungen. Ohne Kontrolle anderer Faktoren, so zeigen ergänzende Analysen (siehe ebenfalls Fußnote 65), besteht kein Zusammenhang.

65 Diese Analysen sind in der vorliegenden Arbeit nicht abgebildet und werden von der Autorin auf Anfrage zur Verfügung gestellt.

Auch bei den Mathematikleistungen fanden sich für die untersuchten Überzeugungen der Lehrkräfte keine statistisch signifikanten Zusammenhänge (siehe Tabelle 8.15). Allerdings deutet sich bei der Betrachtung der Einstellung der Lehrkräfte zur Förderung schwacher Schülerinnen und Schüler an, dass ein negativer Zusammenhang mit den Mathematikleistungen besteht (direkter Effekt in Tabelle 9.4). Je positiver also die Einstellung einer Lehrkraft zur Förderung schwacher Schülerinnen und Schüler ist, desto niedriger ist das Leistungsniveau ihrer Unterrichtsgruppe. Auch hier könnte eine umgekehrte Kausalität vorliegen, da Lehrkräfte in Unterrichtsgruppen, die eine niedrigere Mathematikleistung aufweisen, eher zu der Einstellung tendieren, dass schwächere Schülerinnen und Schüler gefördert werden müssen.

Zusammengefasst zeigen die Ergebnisse der Sekundäranalyse der Drei-Länder-Studie, dass im Englischunterricht die pädagogischen Überzeugungen der Lehrkräfte Ende der 1970er Jahre – vermittelt über das Unterrichtshandeln in Form von Adaptivität – mit den Lernergebnissen der Schülerinnen und Schüler vor allem im affektiven Bereich zusammenhingen. Trotz der bereits genannten Einschränkungen (siehe Kapitel 9.2) lässt sich weiterhin feststellen, dass allgemeine pädagogische Überzeugungen von Lehrkräften mit einem unterstützenden Unterrichtsklima im Englischunterricht zusammenhingen. Zudem waren diese pädagogischen Überzeugungen für das Wohlfühlen der Schülerinnen und Schüler und teilweise auch für deren Schulleistungen im Englisch- und im Mathematikunterricht bedeutsam. Diese Erkenntnis ist auch jenseits des historischen Kontextes von grundlagentheoretischer Bedeutung. Mit Bezug auf das Konzept der „Grammar of Schooling“ (Tyack & Tobin, 1994), also einer relativen Beständigkeit und Beharrungskraft der Schule und vor allem des in ihr stattfindenden Unterrichts, ist zu vermuten, dass derartige Zusammenhänge im Bereich der allgemeinen pädagogischen Überzeugungen von Lehrkräften auch heute noch bestehen.

Für die pädagogische Praxis bedeutet dies, dass Lehrkräfte sich in den verschiedenen Phasen ihrer Ausbildung damit beschäftigen sollten, wie ihre eigenen Überzeugungen mit ihrem Unterricht zusammenhängen und so Auswirkungen vor allem auf die affektiven Lernergebnisse ihrer Schülerinnen und Schüler haben. Eine solche Reflexion der eigenen Überzeugungen ist ein wichtiger Bestandteil der Professionalisierung von Lehrkräften.

Literatur

- Abelson, R. P. (1979). Differences between Belief and Knowledge Systems. *Cognitive Science* (3), 355–366. https://doi.org/10.1207/s15516709cog0304_4
- Adams, R. J. (2005). Reliability as a Measurement Design Effect. *Studies in Educational Evaluation*, 31 (2), 162–172. <https://doi.org/10.1016/j.stueduc.2005.05.008>
- American Educational Research Association. (2005). Auf die Lehrer kommt es an. *Pädagogik*, 57 (7-8), 70–74.
- Ames, A. J. & Penfield, R. D. (2015). An NCME Instructional Module on Item-Fit Statistics for Item Response Theory Models. *Educational Measurement: Issues and Practice*, 34 (2), 1–10. <https://doi.org/10.1111/emip.12067>
- Anger, H., Bargmann, R. & Voigt, M. (1971). *Verständiges Lesen VL 7-9. Schulleistungs- und Begabungstest für 7. bis 9. Klassen* (Deutsche Schultests). Weinheim: Beltz.
- Argyris, C. & Schön, D. A. (1974). *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.
- Ayala, R. J. de. (2009). *The Theory and Practice of Item Response Theory*. New York: Guilford Press.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2010). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (Springer-Lehrbuch, 13. Aufl.). Berlin: Springer. <https://doi.org/10.1007/978-3-662-46076-4>
- Bandura, A. (1997). *Self-efficacy. The Exercise of Control*. New York: Freeman.
- Bartel, H., Hylla, E. & Süllwold, F. (1965). *Mathematische Denkaufgaben MDA 10+* (Deutsche Schultests). Weinheim: Beltz.
- Basturkmen, H. (2012). Review of Research into the Correspondence between Language Teachers' Stated Beliefs and Practices. *System*, 40 (2), 282–295. <https://doi.org/10.1016/j.system.2012.05.001>
- Bauer, K.-O., Kopka, A. & Brindt, S. (1996). *Pädagogische Professionalität und Lehrerarbeit. Eine qualitativ empirische Studie über professionelles Handeln und Bewusstsein* (Eine Veröffentlichung des Instituts für Schulentwicklungsforschung der Universität Dortmund). Weinheim: Juventa.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), 469–520. <https://doi.org/10.1007/s11618-006-0165-2>
- Beck, E., Baer, M., Guldimann, T., Bischoff, S., Brühwiler, C., Müller, P. et al. (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 63). Münster: Waxmann.
- Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical So-*

- ciety. *Series B (Methodological)*, 57 (1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bickel, R. (2007). *Multilevel Analysis for Applied Research. It's Just Regression!* (Methodology in the social sciences). New York: Guilford Press.
- Biedermann, H., Brühwiler, C. & Krattenmacher, S. (2012). Lernangebote in der Lehrerausbildung und Überzeugungen zum Lehren und Lernen. Beziehungsanalysen bei angehenden Lehrpersonen. *Zeitschrift für Pädagogik*, 58 (4), 460–475.
- Bishop, A. J., Seah, W. T. & Chin, C. (2003). Values in Mathematics Teaching—The Hidden Persuaders? In A. J. Bishop, M. A. K. Clements, C. Keitel-Kreidt, J. Kilpatrick & F. K.-S. Leung (Hrsg.), *Second International Handbook of Mathematics Education* (Springer International Handbooks of Education, Bd. 10, S. 717–765). Dordrecht: Springer. <https://doi.org/10.1007/978-3-540-33306-7>
- Blömeke, S. (2010). Zur Bedeutung der Wertebasierung von Lehrkräften für deren Professionalität aus Sicht der Lehrberufsforschung. *Zeitschrift für Religionspädagogik*, 9 (1), 210–220.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung*. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2010). *TEDS-M 2008 – Professionelle Kompetenz und Lernangelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich*. Münster: Waxmann.
- Blömeke, S., Müller, C., Felbrich, A. & Kaiser, G. (2008). Epistemologische Überzeugungen zur Mathematik. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung* (S. 219–246). Münster: Waxmann.
- Blossing, U. & Ekholm, M. (2005). *School Reforms and Local Response in the Long Run. A Twenty Year Longitudinal Study of 35 Swedish "grund" Schools*. Stockholm: Second OECD-conference on Evidence Based Policy Research.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation. Für Human- und Sozialwissenschaftler* (4., überarb. Aufl.). Heidelberg: Springer. <https://doi.org/10.1007/978-3-540-33306-7>
- Böse, S., Neumann, M., Becker, M., Maaz, K. & Baumert, J. (2013). Beurteilung der Berliner Schulstrukturreform durch Schulleiterinnen und Schulleiter, Lehrkräfte und Eltern. In K. Maaz, J. Baumert, M. Neumann, M. Becker & H. Dumont (Hrsg.), *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* (S. 209–261). Münster: Waxmann.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyk-

- lopädie der Psychologie, Pädagogische Psychologie, Bd. 3, S. 177–212). Göttingen: Hogrefe.
- Buehl, M. M. & Beck, J. S. (2015). The Relationship between Teachers' Beliefs and Teachers' Practices. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 66–84). New York: Routledge.
- Calderhead, J. (1996). Teachers: Beliefs and Knowledge. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of Educational Psychology* (S. 709–726). New York: Macmillan.
- Campbell, E. (2003). *Ethical Teacher* (Professional learning). Maidenhead: Open University Press.
- Cattell, R. B. & Weiß, R. H. (1972). *Grundintelligenztest CFT 2 Skala 2*. Braunschweig: Westermann.
- Clarkson, P. C. & Bishop, A. J. (2000). Values and Mathematics Education. In A. Ahmed, J. M. Kraemer & H. Williams (Hrsg.), *Cultural Diversity in Mathematics (Education)*. CIEAEM 51 (S. 239–243). Chichester: Horwood.
- Cloetta, B. & Hohner, H.-U. (1976). *Die Kurzfassung des Konstanzer Fragebogens für Schul- und Erziehungseinstellung (KSE-KF). Erprobung bei Lehrern an konventionellen Schulen und Gesamtschulen* (Arbeitsbericht Nr. 24). Universität Konstanz, Zentrum I Bildungsforschung, Sonderforschungsbereich 23.
- Comber, L. C. & Keeves, J. P. (1973). *Science Education in Nineteen Countries. An Empirical Study* (International Studies in Evaluation, I). New York: Wiley.
- Corno, L. & Snow, R. E. (1986). Adapting Teaching to Individual Differences Among Learners. In M. C. Wittrock (Hrsg.), *Handbook of Research on Teaching* (S. 605–629). Macmillan.
- Cross Francis, D., Rapacki, L. & Eker, A. (2015). The Individual, the Context, and Practice. A Review of the Research on Teachers' Beliefs Related to Mathematics. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 336–352). New York: Routledge.
- Davis, H. A. (2006). Exploring the Contexts of Relationship Quality between Middle School Students and Teachers. *The Elementary School Journal*, 106 (3), 193–223. <https://doi.org/10.1086/501483>
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 39 (2), 223–238.
- Decristan, J., Hondrich, A. L., Büttner, G., Hertel, S., Klieme, E., Kunter, M. et al. (2015). Impact of Additional Guidance in Science Education on Primary Students' Conceptual Understanding. *The Journal of Educational Research*, 108 (5), 358–370. <https://doi.org/10.1080/00220671.2014.899957>
- DESI-Konsortium (Hrsg.). (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz.

- Deutscher Bildungsrat. (1969). *Empfehlungen der Bildungskommission. Einrichtung von Schulversuchen mit Gesamtschulen*. Verabschiedet auf der 19. Sitzung der Bildungskommission am 30./31. Januar 1969. Bad Godesberg.
- Dohrmann, J., Feldhoff, T., Steinert, B. & Klieme, E. (2019). Überzeugungen von Lehrkräften, Adaptivität des Unterrichts und Lernergebnisse im Fach Englisch. *Zeitschrift für Pädagogik*, 65 (1), 56–72.
- 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. <https://doi.org/10.1024/1010-0652.22.34.193>
- Dumont, H. (2019). Neuer Schlauch für alten Wein? Eine konzeptuelle Betrachtung von individueller Förderung im Unterricht. *Zeitschrift für Erziehungswissenschaft*, 22 (2), 249–277. <https://doi.org/10.1007/s11618-018-0840-0>
- Enders, C. K. (2010). *Applied Missing Data Analysis* (Methodology in the social sciences). New York: Guilford Press.
- Enders, C. K. & Tofighi, D. (2007). Centering Predictor Variables in Cross-sectional Multilevel Models: A New Look at an Old Issue. *Psychological Methods*, 12 (2), 121–138. <https://doi.org/10.1037/1082-989X.12.2.121>
- Ertmer, P. A., Ottenbreit-Leftwich, A. T. & Tondeur, J. (2015). Teachers' Beliefs and Uses of Technology to Support 21st-century Teaching and Learning. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 403–418). New York: Routledge.
- Felbrich, A., Schmotz, C. & Kaiser, G. (2010). Überzeugungen angehender Primarstufenlehrkräfte im internationalen Vergleich. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Primarlehrkräfte im internationalen Vergleich* (S. 297–325). Münster: Waxmann.
- Feldhoff, T. (2011). *Schule organisieren. Der Beitrag von Steuergruppen und Organisationalem Lernen zur Schulentwicklung* (Educational Governance, Bd. 15). Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-93384-9>
- Fend, H. (1979). Gesamtüberblick zu den Konstanzer Vergleichsuntersuchungen. In H. Haenisch, H. Lukesch, R. Klaghofer & E.-M. Krüger-Haenisch (Hrsg.), *Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik* (Schule und Weiterbildung, Bd. 8, S. V–XVII). Paderborn: Schöningh.
- Fend, H. (1981). *Leistungsvergleich zwischen Gesamtschulen und Schulen des traditionellen Schulsystems. Abschlußbericht des Einzelprojekts: Leistungsvergleich zwischen Gesamtschulen und Schulen des traditionellen Schulsystems im Rahmen der*

- „Wissenschaftlichen Bestandsaufnahme der Arbeit der niedersächsischen Gesamtschulen“. Eine Information des Niedersächsischen Kultusministers. Konstanz.
- Fend, H. (1982). *Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs*. Weinheim: Beltz.
- Fend, H. (1984). *Die Pädagogik des Neokonservatismus* (Suhrkamp-Taschenbuch Wissenschaft, Bd. 475). Frankfurt am Main: Suhrkamp.
- Fiebert, M. & Solzbacher, C. (2014). „Bescheidenheit und Festigkeit des Charakters...“ Das Konstrukt Lehrerhaltung aus historisch-systematischer Perspektive. In C. Schwer & C. Solzbacher (Hrsg.), *Professionelle pädagogische Haltung. Historische, theoretische und empirische Zugänge zu einem viel strapazierten Begriff* (S. 17–47). Bad Heilbrunn: Klinkhardt.
- Fischer, N. & Ehmke, T. (2019). Empirische Erfassung eines „messy constructs“. Überzeugungen angehender Lehrkräfte zu sprachlich-kultureller Heterogenität in Schule und Unterricht. *Zeitschrift für Erziehungswissenschaft*, 22 (2), 411–433. <https://doi.org/10.1007/s11618-018-0859-2>
- Fishbein, M. & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
- Fives, H. & Buehl, M. M. (2012). Spring Cleaning for the “Messy” Construct of Teachers’ Beliefs: What Are They? Which Have Been Examined? What Can They Tell Us? In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer & M. Zeidner (Hrsg.), *APA Educational Psychology Handbook, Vol 2: Individual Differences and Cultural and Contextual Factors* (S. 471–499). Washington: American Psychological Association. <https://doi.org/10.1037/13274-019>
- Fives, H. & Gill, M. G. (Hrsg.). (2015). *International Handbook of Research on Teachers’ Beliefs* (Educational Psychology Handbook). New York: Routledge. <https://doi.org/10.4324/9780203108437>
- Fives, H., Lacatena, N. & Gerard, L. (2015). Teachers’ Beliefs about Teaching (and Learning). In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers’ Beliefs* (Educational Psychology Handbook, S. 249–265). New York: Routledge. <https://doi.org/10.4324/9780203108437>
- Forschungsdatenzentrum Bildung (Hrsg.). (2014a). *Fragebogenerhebung (Skalenkollektion): Vergleichsuntersuchung zwischen Gesamtschulen und Schulen des traditionellen Schulsystems - GS: 9. Klassen/Schuljahr 1978 (Fend-Studie)*. <https://www.fdz-bildung.de/erhebung.php?id=67> [22.04.2021].
- Forschungsdatenzentrum Bildung (Hrsg.). (2014b). *Fend-Studie – Wissenschaftliche Begleitung von Gesamtschulen*. <https://www.fdz-bildung.de/studiendetails.php?la=de&id=46> [22.04.2021].
- Frenzel, A. C., Götz, T. & Pekrun, R. (2008). Ursachen und Wirkungen von Lehreremotionen: Ein Modell zur reziproken Beeinflussung von Lehrkräften und Klassenmerkmalen. In M. Gläser-Zikuda & J. Seifried (Hrsg.), *Lehrerexpertise. Analyse und Bedeutung unterrichtlichen Handelns* (S. 187–209). Münster: Waxmann.

- Furinghetti, F. & Pehkonen, E. (2002). Rethinking Characterizations of Beliefs. In G. C. Leder, E. Pehkonen & G. Törner (Hrsg.), *Beliefs: A Hidden Variable in Mathematics Education?* (Mathematics Education Library, Bd. 31, S. 39–57). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47958-3_3
- Gay, G. (2015). Teachers' Beliefs about Cultural Diversity. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 436–452). New York: Routledge.
- Geiser, C. (2010). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften / GWV Fachverlage GmbH Wiesbaden. <https://doi.org/10.1007/978-3-531-92042-9>
- Geiser, C. & Eid, M. (2010). Item-Response-Theorie. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 311–332). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92038-2_14
- Gill, M. G. & Hardin, C. (2015). A “Hot” Mess: Unpacking the Relation between Teachers' Beliefs and Emotions. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 230–245). New York: Routledge.
- Goldin, G., Rösken, B. & Törner, G. (2009). Beliefs – No Longer a Hidden Variable in Mathematical Teaching and Learning Processes. In J. Maaß & W. Schlöglmann (Hrsg.), *Beliefs and Attitudes in Mathematics Education. New Research Results* (S. 1–18). Rotterdam: Sense Publishers. https://doi.org/10.1163/9789087907235_002
- Green, T. F. (1971). *The Activities of Teaching*. New York: McGraw-Hill.
- Groebe, N., Wahl, D., Schlee, J. & Scheele, B. (1988). *Das Forschungsprogramm Subjektive Theorien. Eine Einführung in die Psychologie des reflexiven Subjekts*. Tübingen: Francke.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 12). Münster: Waxmann.
- Haenisch, H. (1979). *Schulleistungsvergleiche zwischen Gesamtschulen in Hessen und Schulen des gegliederten Schulsystems am Ende des 6. Schuljahres* (Arbeitsbericht Nr. 12). Konstanz: Universität Konstanz, Zentrum I Bildungsforschung.
- Hall, G. E. & Hord, S. M. (1987). *Change in Schools. Facilitating the Process*. Albany: State University of New York Press.
- Hallquist, M. N. & Wiley, J. F. (2018). MplusAutomation. An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling*, 25 (4), 621–638. <https://doi.org/10.1080/10705511.2017.1402334>
- 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. <https://doi.org/10.1111/j.1467-8624.2005.00889.x>

- Handal, B. (2003). Teachers' Mathematical Beliefs: A Review. *The Mathematics Educator*, 13 (2), 47–57.
- Harder, P. (2014). *Werthaltungen und Ethos von Lehrern. Empirische Studie zu Annahmen über den guten Lehrer* (Schriften aus der Fakultät Sozial- und Wirtschaftswissenschaften der Otto-Friedrich-Universität Bamberg, Bd. 12). Bamberg: University of Bamberg Press.
- Hardy, I., Decristan, J. & Klieme, E. (2019). Adaptive Teaching in Research on Learning and Instruction. *Journal of Educational Research Online*, 11 (2), 169–191.
- Hartinger, A., Kleickmann, T. & Hawelka, B. (2006). Der Einfluss von Lehrervorstellungen zum Lernen und Lehren auf die Gestaltung des Unterrichts und auf motivationale Schülervariablen. *Zeitschrift für Erziehungswissenschaft*, 9 (1), 110–126. <https://doi.org/10.1007/s11618-006-0008-1>
- Hasselhorn, M., Decristan, J. & Klieme, E. (2019). Individuelle Förderung. In O. Köller, M. Hasselhorn, F. W. Hesse, K. Maaz, J. Schrader, H. Solga et al. (Hrsg.), *Das Bildungswesen in Deutschland. Bestand und Potenziale* (S. 375-402). Bad Heilbrunn: Klinkhardt.
- Hattie, J. (Oktober 2003). *Teachers Make a Difference: What Is the Research Evidence?* Paper presented at the ACER Research Conference “Building Teacher Quality: What Does the Research Tell Us?”. Melbourne: Australian Council for Educational Research. http://research.acer.edu.au/research_conference_2003/4/ [22.04.2021].
- Hattie, J. (2009). *Visible Learning. A Synthesis of over 800 Meta-Analyses Relating to Achievement*. London: Routledge. <https://doi.org/10.4324/9780203887332>
- Hausen, J., Schmid, C., Wurster, S., Dohrmann, J., Feldhoff, T., Steinert, B. et al. (2019). *Skalenhandbuch Schule im Wandel (SchiWa) – Vorstudie. Dokumentation der Erhebungsinstrumente für Schulleitungen, Schülerinnen und Schüler sowie für Lehrkräfte*. <http://nbn-resolving.de/urn:nbn:de:0111-pedocs-174367>
- Heck, R. H. & Thomas, S. L. (2015). *An Introduction to Multilevel Modeling Techniques. MLM and SEM Approaches Using Mplus* (Quantitative Methodology Series, 3. Aufl.). New York: Routledge. <https://doi.org/10.4324/9781315746494>
- Heinrich, H. C. (1974). Skalen zur Erfassung von Formen der Bekräftigung in der Erziehung durch Lehrer. *Zeitschrift für experimentelle und angewandte Psychologie*, 21 (4), 530–545.
- Helmke, A. (2007). Aktive Lernzeit optimieren: Was wissen wir über effiziente Klassenführung? *Pädagogik*, 59 (5), 44–48.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (6. Auflage). Seelze-Velber: Klett/Kallmeyer.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (Enzyklopädie der Psychologie, Pädagogische Psychologie, Bd. 3, S. 71–176). Göttingen: Hogrefe.
- Hemmerich, W. (2019). *Rechner zur Adjustierung des α -Niveaus: StatistikGuru*. <https://statistikguru.de/rechner/adjustierung-des-alphaniveaus.html> [22.04.2021].

- Hochweber, J. & Hartig, J. (2012). Mehrebenenanalyse. *Enzyklopädie Erziehungswissenschaft Online*, 1–33.
- Holtappels, H. G. (1997). *Grundschule bis mittags. Innovationsstudie über Zeitgestaltung und Lernkultur*. Weinheim: Juventa.
- Hox, J. J. (2010). *Multilevel Analysis. Techniques and Applications*. New York: Routledge. <https://doi.org/10.4324/9780203852279>
- Husén, T. (Hrsg.). (1967). *International Study of Achievement in Mathematics. A Comparison of Twelve Countries* (Volume I & II). Stockholm: Almqvist & Wiksell.
- Jäger, R. & Jundt, E. (1973). *Mannheimer Rechtschreib-Test (M-R-T)*. Göttingen: Hogrefe, Westermann.
- Jorgensen, R., Grootenboer, P., Niesche, R. & Lerman, S. (2010). Challenges for Teacher Education: The Mismatch between Beliefs and Practice in Remote Indigenous Contexts. *Asia-Pacific Journal of Teacher Education*, 38 (2), 161–175. <https://doi.org/10.1080/13598661003677580>
- Kagan, D. M. (1990). Ways of Evaluating Teacher Cognition: Inferences Concerning the Goldilocks Principle. *Review of Educational Research*, 60 (3), 419–469. <https://doi.org/10.3102/00346543060003419>
- Kagan, D. M. (1992). Implications of Research on Teacher Beliefs. *Educational Psychologist*, 27 (1), 65–90. https://doi.org/10.1207/s15326985ep2701_6
- Kane, R., Sandretto, S. & Heath, C. (2002). Telling Half the Story: A Critical Review of Research on the Teaching Beliefs and Practices of University Academics. *Review of Educational Research*, 72 (2), 177–228. <https://doi.org/10.3102/00346543072002177>
- Kielblock, S. (2018). *Inclusive Education for All: Development of an Instrument to Measure the Teacher's Attitudes*. Dissertation, Justus Liebig Universität. Giessen. <https://d-nb.info/1162053941/34>
- Kiely, M. T., Brownell, M. T., Lauterbach, A. A. & Benedict, A. E. (2015). Teachers' Beliefs about Students with Special Needs and Inclusion. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 475–491). New York: Routledge.
- Klaghofer, R. & Krüger-Haenisch, E.-M. (1979). Die curriculare Analyse der Leistungstests. In H. Haenisch, H. Lukesch, R. Klaghofer & E.-M. Krüger-Haenisch (Hrsg.), *Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik* (Schule und Weiterbildung, Bd. 8, S. 363–476). Paderborn: Schöningh.
- Kleining, G. & Moore, H. (1968). Soziale Selbsteinstufung (SSE). Ein Instrument zur Messung sozialer Schichten. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 20 (3), 502–552.
- Klieme, E. (2016). Schulqualität, Schuleffektivität und Schulentwicklung – Welche Erkenntnis eröffnet empirische Forschung? In U. Steffens & T. Bargel (Hrsg.),

- Schulqualität – Bilanz und Perspektiven. Grundlagen der Qualität von Schule 1* (Beiträge zur Schulentwicklung, Bd. 1, S. 45–64). Münster: Waxmann.
- Klieme, E. (2019). Unterrichtsqualität. In M. Haring, C. Rohlf's & M. Gläser-Zikuda (Hrsg.), *Handbuch Schulpädagogik* (S. 393–408). Münster: Waxmann.
- Klieme, E. & Warwas, J. (2011). Konzepte der Individuellen Förderung. *Zeitschrift für Pädagogik*, 56 (6), 805–818.
- Knigge, M. & Rotter, C. (2015). Unterrichtsplanung bei Lehramtsstudierenden im Falle der Wahrnehmung von vermeintlich „besonderen“ Schülerinnen und Schülern und ihr Zusammenhang mit Selbstwirksamkeitsüberzeugungen und Einstellungen in Bezug zu Inklusion – beispielhafte Mixed-Method-Analysen aus der EiLink-Studie. *Empirische Sonderpädagogik*, 7 (3), 223–240.
- Koch, J.-J. (1972). *Lehrer – Studium und Beruf. Einstellungswandel in den beiden Phasen der Ausbildung*. Ulm: Süddeutsche Verlagsgesellschaft.
- König, J. (2012). Teachers' Pedagogical Beliefs: Current and Future Research. In J. König (Hrsg.), *Teachers' Pedagogical Beliefs. Definition and Operationalisation – Connections to Knowledge and Performance – Development and Change* (S. 7–11). Münster: Waxmann.
- König, J. (2014). Kompetenz und Profession von Lehrerinnen und Lehrern. In F. Kessl (Hrsg.), *Prekarisierung der Pädagogik – Pädagogische Prekarisierung? Erziehungswissenschaftliche Vergewisserungen* (S. 91–106). Weinheim: Beltz Juventa.
- Krüger-Haenisch, E.-M. (o. J.). *Erhebungsinstrumente der Leistungsstudie 1978 (außer Leistungstests)*. Projekt: „Vergleichsuntersuchung zwischen Gesamtschulen und Schulen des traditionellen Schulsystems“. Konstanz: Universität Konstanz, Zentrum I Bildungsforschung.
- Krull, J. L. & MacKinnon, D. P. (2001). Multilevel Modeling of Individual and Group Level Mediated Effects. *Multivariate Behavioral Research*, 36 (2), 249–277. https://doi.org/10.1207/S15327906MBR3602_06
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M. & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (Dortmunder Symposium der Empirischen Bildungsforschung, Bd. 1, S. 9–31). Münster: Waxmann.
- Kunter, M. & Pohlmann, B. (2015). Lehrer. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (Springer-Lehrbuch, 2., vollständig überarb. und aktualisierte Aufl., S. 261–281). Berlin: Springer. https://doi.org/10.1007/978-3-642-41291-2_11
- Läge, D. & McCombie, G. (2015). Berufsbezogene Lehrerüberzeugungen als pädagogische Bezugssysteme erfassen. Ein Vergleich von angehenden und berufstätigen

- Lehrpersonen der verschiedenen Schulstufen in der Schweiz. *Zeitschrift für Pädagogik*, 61 (1), 118–143.
- Langer, W. (2010). Mehrebenenanalyse mit Querschnittsdaten. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 741–774). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92038-2_28
- Leistungsmessung in Gesamtschulen. (1974). *Englisch-Einstufungs-Test (EET 9+)* (Werkstattbericht „In eigener Sache“). Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Linacre, J. M. (2002). What Do Infit and Outfit, Mean-square and Standardized mean? *Rasch Measurement Transactions*, 16 (2), 878. <https://www.rasch.org/rmt/rmt162f.htm>
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. *Zeitschrift für Pädagogik* (51), 47–70.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58 (2), 103–117. <https://doi.org/10.1026/0033-3042.58.2.103>
- 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 (2), 120–131. <https://doi.org/10.1016/j.cedpsych.2008.12.001>
- Lukesch, H. (1979a). Leistungsvergleich zwischen Gesamtschulen und herkömmlichen Schulen am Ende der Pflichtschulzeit in Nordrhein-Westfalen. In H. Haenisch, H. Lukesch, R. Klaghofer & E.-M. Krüger-Haenisch (Hrsg.), *Gesamtschule und dreigliedriges Schulsystem in Nordrhein-Westfalen – Schulleistungsvergleiche in Deutsch, Mathematik, Englisch und Physik* (Schule und Weiterbildung, Bd. 8, S. 227–364). Paderborn: Schöningh. <https://epub.uni-regensburg.de/2771/1/Leistungsvergleich.pdf> [10.10.2019].
- Lukesch, H. (1979b). *Schulleistungsvergleiche zwischen Gesamtschulen in Hessen und Schulen des gegliederten Schulsystems am Ende der Pflichtschulzeit* (Arbeitsbericht Nr. 13). Konstanz: Universität Konstanz, Zentrum I Bildungsforschung.
- Maas, C. J. M. & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1 (3), 86–92. <https://doi.org/10.1027/1614-2241.1.3.86>
- MacKinnon, D. P. (2008). *Introduction to Statistical Mediation Analysis*. New York: Lawrence Erlbaum Associates.
- Martínez, J. F. (2012). Consequences of Omitting the Classroom in Multilevel Models of Schooling. An Illustration Using Opportunity to Learn and Reading Achievement. *School Effectiveness and School Improvement*, 23 (3), 305–326. <https://doi.org/10.1080/09243453.2012.678864>

- Mattes, M. (2018). Gesamtschule im Flächenversuch. Erfahrungen mit einem bildungspolitischen Experiment im hessischen Wetzlar 1965-1990. *Jahrbuch für Historische Bildungsforschung*, 23, 29–58.
- Mattes, M. & Reh, S. (2019). Entstehung und Durchführung der Gesamtschul-Studien in den 1970er Jahren. Monika Mattes und Sabine Reh im Zeitzeugengespräch mit Helmut Fend. *Zeitschrift für Pädagogik*, 65 (1), 6–23.
- McLeod, D. B. & McLeod, S. (2002). Synthesis – Beliefs and Mathematics Education: Implications for Learning, Teaching and Research. In G. C. Leder, E. Pehkonen & G. Törner (Hrsg.), *Beliefs: A Hidden Variable in Mathematics Education?* (Mathematics Education Library, Bd. 31, S. 115–123). Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47958-3_7
- Messner, R. (2016). Die Einzelschule als pädagogische Handlungseinheit und das Zusammenspiel der Handlungsebenen und institutionellen Akteure. In U. Steffens & T. Bargel (Hrsg.), *Schulqualität – Bilanz und Perspektiven. Grundlagen der Qualität von Schule 1* (Beiträge zur Schulentwicklung, Bd. 1, S. 95–114). Münster: Waxmann.
- Möller, K., Hardy, I., Jonen, A., Kleickmann, T. & Blumberg, E. (2006). Naturwissenschaften in der Primarstufe. Zur Förderung konzeptuellen Verständnisses durch Unterricht und zur Wirksamkeit von Lehrerfortbildungen. In M. Prenzel & L. Allolio-Näcke (Hrsg.), *Unterrichtsforschung zur Bildungsqualität von Schule* (S. 161–193). Münster: Waxmann.
- Moosbrugger, H. (2012). Item-Response-Theorie (IRT). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2., aktualisierte und überarbeitete Auflage, S. 227–274). Berlin: Springer. https://doi.org/10.1007/978-3-642-20072-4_10
- Muijs, D. & Reynolds, D. (2002). Teachers' Beliefs and Behaviors: What Really Matters? *Journal of Classroom Interaction*, 37 (2), 3–15.
- Müller, C., Felbrich, A. & Blömeke, S. (2008). Überzeugungen zum Lehren und Lernen von Mathematik. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare. Erste Ergebnisse zur Wirksamkeit der Lehrerbildung* (S. 247–276). Münster: Waxmann.
- Murphy, N. & Messer, D. (2010). Differential Benefits from Scaffolding and Children Working Alone. *Educational Psychology*, 20 (1), 17–31. <https://doi.org/10.1080/014434100110353>
- Muthén, L. K. & Muthén, B. O. (1998-2013) Mplus (Version 7.11) [Computer software]. Los Angeles, CA.
- Muthén, L. K. & Muthén, B. O. (2015). *Mplus User's Guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Nespor, J. (1987). The Role of Beliefs in the Practice of Teaching. *Journal of Curriculum Studies*, 19 (4), 317–328. <https://doi.org/10.1080/0022027870190403>

- OECD (Hrsg.). (2009). *Creating Effective Teaching and Learning Environments. First Results from TALIS (Teaching and Learning International Survey)*. Paris: OECD.
- Opdenakker, M.-C. & Van Damme, J. (2000). The Importance of Identifying Levels in Multilevel Analysis: An Illustration of the Effects of Ignoring the Top or Intermediate Levels in School Effectiveness Research. *School Effectiveness and School Improvement*, 11 (1), 103–130.
- Oser, F. (1998a). Das Ethos der Lehrperson: Theoretische Grundlagen in der Übersicht. In F. Oser (Hrsg.), *Ethos – Die Vermenschlichung des Erfolgs. Zur Psychologie der Berufsmoral von Lehrpersonen* (Schule und Gesellschaft, Bd. 16, S. 35–42). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-97398-6_2
- Oser, F. (Hrsg.). (1998b). *Ethos – Die Vermenschlichung des Erfolgs. Zur Psychologie der Berufsmoral von Lehrpersonen* (Schule und Gesellschaft, Bd. 16). Opladen: Leske + Budrich. <https://doi.org/10.1007/978-3-322-97398-6>
- Oser, F. & Blömeke, S. (2012). Überzeugungen von Lehrpersonen. Einführung in den Thementeil. *Zeitschrift für Pädagogik*, 58 (4), 415–421.
- Pajares, F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, 62 (3), 307–332. <https://doi.org/10.3102/00346543062003307>
- Pant, H. A., Vock, M., Pöhlmann, C. & Köller, O. (2008). Offenheit für Innovationen. Befunde aus einer Studie zur Rezeption der Bildungsstandards bei Lehrkräften und Zusammenhänge mit Schülerleistungen. *Zeitschrift für Pädagogik*, 54 (6), 827–845.
- Pauli, C., Reusser, K. & Grob, U. (2007). Teaching for Understanding and/or Self-regulated Learning? A Video-based Analysis of Reform-oriented Mathematics Instruction in Switzerland. *International Journal of Educational Research*, 46 (5), 294–305. <https://doi.org/10.1016/j.ijer.2007.10.004>
- Philipp, R. (2007). Mathematics Teachers' Beliefs and Affect. In F. K. Lester (Hrsg.), *Second Handbook of Research on Mathematics Teaching and Learning. A Project of the National Council of Teachers of Mathematics* (S. 257–315). Charlotte, NC: Information Age Publishing.
- Phipps, S. & Borg, S. (2009). Exploring Tensions between Teachers' Grammar Teaching Beliefs and Practices. *System*, 37 (3), 380–390. <https://doi.org/10.1016/j.system.2009.03.002>
- R Core Team. (2017) R: A Language and Environment for Statistical Computing [Computer software]. Vienna, Austria. <https://www.R-project.org/>
- Rakoczy, K. (2006). Motivationsunterstützung im Mathematikunterricht. Zur Bedeutung von Unterrichtsmerkmalen für die Wahrnehmung von Schülerinnen und Schüler. *Zeitschrift für Pädagogik*, 52 (6), 822–843.
- Rammstedt, B. (2010). Reliabilität, Validität, Objektivität. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 239–258). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92038-2_11

- Raths, L. E., Harmin, M. & Simon, S. B. (1987). Selections from Values and Teaching. In P. F. Carbone (Hrsg.), *Value Theory and Education* (S. 198–214). Malabar: Krieger.
- Rauch, D. & Hartig, J. (2012). Interpretation von Testwerten in der IRT. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2., aktualisierte und überarbeitete Auflage, S. 253–274). Berlin: Springer.
- Reusser, K., Pauli, C. & Elmer, A. (2011). Berufsbezogene Überzeugungen von Lehrerinnen und Lehrern. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 478–495). Münster: Waxmann.
- Richardson, V. (1996). The Role of Attitudes and Beliefs in Learning to Teach. In J. Sikula (Hrsg.), *Handbook of Research on Teacher Education* (2. Aufl., S. 102–119). New York: Macmillan.
- Robitzsch, A., Kiefer, T. & Wu, M. (2017) TAM: Test Analysis Modules. R Package Version 2.8-21 [Computer software]. <https://CRAN.R-project.org/package=TAM>
- Rokeach, M. (1960). *The Open and Closed Mind. Investigations into the Nature of Belief Systems and Personality Systems*. New York: Basic Books.
- Rokeach, M. (1969). *Beliefs, Attitudes, and Values. A Theory of Organization and Change* (2. Aufl.). San Francisco: Jossey-Bass.
- Rost, J. (1996). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2., vollständig überarbeitete und erweiterte Auflage). Bern: Huber.
- Rothland, M. (2012). The Professional Motivation, Job-related Beliefs and Patterns of Work-related Coping Behaviour of Teacher Training Students. In J. König (Hrsg.), *Teachers' Pedagogical Beliefs. Definition and Operationalisation – Connections to Knowledge and Performance – Development and Change* (S. 71–90). Münster: Waxmann.
- Rubie-Davies, C. (2015). Teachers' Instrucional Beliefs and the Classroom Climate. In H. Fives & M. G. Gill (Hrsg.), *International Handbook of Research on Teachers' Beliefs* (Educational Psychology Handbook, S. 266–283). New York: Routledge.
- Schafer, J. L. & Graham, J. W. (2002). Missing Data. Our View of the State of the Art. *Psychological Methods*, 7 (2), 147–177. <https://doi.org/10.1037//1082-989X.7.2.147>
- Scheunpflug, A. (2005). Job oder Berufung? Gibt es eine Berufsethik für Lehrkräfte? In G.-B. von Carlsburg & M. Heitger (Hrsg.), *Der Lehrer – ein (un)möglicher Beruf* (Erziehungskonzeptionen und Praxis, Bd. 61, S. 57–67). Frankfurt am Main: Lang.
- Schlee, J. (2017). Die Menschenbildannahmen des Forschungsprogramms „Subjektive Theorien in ihrer Bedeutung für Haltungen und Sichtweisen von Lehrkräften“. In J. Standop, E. D. Röhrig & R. Winkels (Hrsg.), *Menschenbilder in Schule und Unterricht* (S. 213–227). Weinheim: Beltz Juventa.
- Schmid, C., Dohrmann, J., Feldhoff, T., Steinert, B., Wurster, S. & Klieme, E. (2019). *Skalenhandbuch Drei-Länder-Studie (DLS) 1978/79 von Helmut Fend. Dokumenta-*

- tion der Erhebungsinstrumente für Schülerinnen und Schüler sowie für Lehrkräfte. <http://nbn-resolving.de/urn:nbn:de:0111-pedocs-172975>
- Schmotz, C., Felbrich, A. & Kaiser, G. (2010). Überzeugungen angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *TEDS-M 2008 – Professionelle Kompetenz und Lernangelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich* (S. 279–305). Münster: Waxmann.
- Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. San Diego: Academic Press. <https://doi.org/10.1016/C2013-0-05012-8>
- Schoenfeld, A. H. (2002). How Can We Examine the Connection between Teachers' World Views and Their Educational Practices? *Issues in Education*, 8 (2), 217–227.
- Scholz, E. (1909). Takt, pädagogischer. In W. Rein (Hrsg.), *Encyklopädisches Handbuch der Pädagogik* (Bd. 9, 2. Aufl., S. 102–108). Langensalza: Beyer.
- Schultze, W. (1974). *Die Leistungen im naturwissenschaftlichen Unterricht in der Bundesrepublik im internationalen Vergleich* (Mitteilungen und Nachrichten). Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Schwer, C. & Solzbacher, C. (Hrsg.). (2014). *Professionelle pädagogische Haltung. Historische, theoretische und empirische Zugänge zu einem viel strapazierten Begriff*. Bad Heilbrunn: Klinkhardt.
- Seidel, T. & Shavelson, R. J. (2007). Teaching Effectiveness Research in the Past Decade. The Role of Theory and Research Design in Disentangling Meta-Analysis Results. *Review of Educational Research*, 77 (4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Seifried, J. (2009). *Unterricht aus der Sicht von Handelslehrern* (Konzepte des Lehrens und Lernens, Bd. 16). Frankfurt am Main: Lang.
- Sembill, D. & Seifried, J. (2009). Konzeptionen, Funktionen und internationale Veränderungen von Sichtweisen. In O. Zlatkin-Troitschanskaia (Hrsg.), *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (Beltz-Bibliothek, S. 345–355). Weinheim: Beltz.
- Shulman, L. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, 15 (2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- Snijders, T. A. B. & Bosker, R. J. (1999). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. London: Sage.
- Standop, J. (2017). Die Bedeutung von Menschenbildern für das Unterrichts- und Erziehungshandeln von Lehrerinnen und Lehrern. In J. Standop, E. D. Röhrig & R. Winkels (Hrsg.), *Menschenbilder in Schule und Unterricht* (S. 259–274). Weinheim: Beltz Juventa.
- Staub, F. C. & Stern, E. (2002). The Nature of Teachers' Pedagogical Content Beliefs Matters for Students' Achievement Gains: Quasi-experimental Evidence from Elementary Mathematics. *Journal of Educational Psychology*, 94 (2), 344–355. <https://doi.org/10.1037//0022-0663.94.2.344>

- Steffens, U. (2012). *Schulqualitätsdiskussion in Deutschland – ihre Entwicklung im Überblick* (IQ-Kompakt Bd. 10). Wiesbaden: Hessisches Kultusministerium, Institut für Qualitätsentwicklung (IQ).
- Steffens, U. & Bargel, T. (2016). Die Diskussion um Schulqualität – Anfänge, Wege und Erträge des „Arbeitskreises Qualität von Schule“. In U. Steffens & T. Bargel (Hrsg.), *Schulqualität – Bilanz und Perspektiven. Grundlagen der Qualität von Schule 1* (Beiträge zur Schulentwicklung, Bd. 1). Münster: Waxmann.
- Taibi, M. (2012). The Development of Professional Beliefs during Teacher Education at University. In J. König (Hrsg.), *Teachers’ Pedagogical Beliefs. Definition and Operationalisation – Connections to Knowledge and Performance – Development and Change* (S. 53–70). Münster: Waxmann.
- Terhart, E. (2000). Qualität und Qualitätssicherung im Schulsystem. Hintergründe – Konzepte – Probleme. *Zeitschrift für Pädagogik*, 46 (6), 809–829.
- Thompson, A. G. (1992). Teachers’ Beliefs and Conceptions: A Synthesis of the Research. In D. A. Grouws (Hrsg.), *Handbook of research on mathematics teaching and learning. A Project of the National Council of Teachers of Mathematics* (S. 127–146). New York: Macmillan.
- Trautmann, M. & Wischer, B. (2008). Das Konzept der Inneren Differenzierung — eine vergleichende Analyse der Diskussion der 1970er Jahre mit dem aktuellen Heterogenitätsdiskurs. In M. A. Meyer, M. Prenzel & S. Hellekamps (Hrsg.), *Perspektiven der Didaktik* (Zeitschrift für Erziehungswissenschaft, Sonderheft, Bd. 9, S. 159–172). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91775-7_11
- Trautwein, C. (2013). Lehrbezogene Überzeugungen und Konzeptionen – eine konzeptuelle Landkarte. *Zeitschrift für Hochschulentwicklung*, 8 (3), 1–14. <https://doi.org/10.3217/zfhe-8-03/02>
- Turner, J. C., Christensen, A. & Meyer, D. K. (2009). Teachers’ Beliefs about Student Learning and Motivation. In A. G. Dworkin & L. J. Saha (Hrsg.), *International Handbook of Research on Teachers and Teaching* (Springer International Handbooks of Education, Bd. 21, S. 361–371). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-73317-3_23
- Tyack, D. & Tobin, W. (1994). The “Grammar” of Schooling: Why Has it Been so Hard to Change? *American Educational Research Journal*, 31 (3), 453–479. <https://doi.org/10.3102/00028312031003453>
- Urban, D. & Mayerl, J. (2018). *Angewandte Regressionsanalyse. Theorie, Technik und Praxis* (5. überarb. Aufl.). Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-658-01915-0>

- Van der Linden, W. J. & Hambleton, R. K. (Hrsg.). (1997). *Handbook of Modern Item Response Theory*. New York: Springer. <https://doi.org/10.1007/978-1-4757-2691-6>
- Voss, T., Kleickmann, T., Kunter, M. & Hachfeld, A. (2011). Überzeugungen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 235–257). Münster: Waxmann.
- Wahl, D. (2002). Mit Training vom trägen Wissen zum kompetenten Handeln? *Zeitschrift für Pädagogik*, 48 (2), 227–241.
- Wahl, D. (2006). *Lernumgebungen erfolgreich gestalten. Vom trägen Wissen zum kompetenten Handeln* (2., erw. Aufl.). Bad Heilbrunn: Klinkhardt.
- Warm, T. (1989). Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*, 54 (3), 427–450. <https://doi.org/10.1007/BF02294627>
- Weiß, R. H. (1972). *Grundintelligenztest CFT 2 Skala 2: Handanweisung für die Durchführung, Auswertung und Interpretation*. Braunschweig: Westermann.
- Wildfeuer, A. G. (2009). Praxis. In P. Kolmer & A. G. Wildfeuer (Hrsg.), *Neues Handbuch philosophischer Grundbegriffe* (S. 1774–1804). Freiburg im Breisgau: Alber.
- Wolff, H.-G. & Bacher, J. (2010). Hauptkomponentenanalyse und explorative Faktorenanalyse. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 333–365). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92038-2_15
- Woolfolk Hoy, A., Davis, H. & Pape, S. J. (2006). Teacher Knowledge and Beliefs. In P. A. Alexander & P. H. Winne (Hrsg.), *Handbook of Educational Psychology* (2. Aufl., S. 715–737). Mahwah, N.J.: Erlbaum.
- Wright, B. D. & Linacre, J. M. (1994). Reasonable Mean-Square Fit Values. *Rasch Measurement Transactions*, 8 (3), 370. <https://www.rasch.org/rmt/rmt83b.htm>
- Wu, M. & Adams, R. (2007). *Applying the Rasch Model to Psycho-social Measurement. A Practical Approach*. Melbourne: Educational Measurement Solutions.
- Wurster, S. & Feldhoff, T. (2019). Schul- und Unterrichtsqualität aus der Mehrebenenperspektive: Ist die Schule oder die Klasse die relevante pädagogische Handlungseinheit? *Zeitschrift für Pädagogik*, 65 (1), 24–39

Abbildungsverzeichnis

Abbildung 2.1:	Verhältnis von Überzeugungen zu benachbarten Begriffen	19
Abbildung 4.1:	Die drei Funktionen von Überzeugungen	45
Abbildung 5.1:	Schematische Darstellung der Forschungsfragen und Hypothesen..	64
Abbildung 6.1:	Erhebungsschwerpunkte der Instrumente für Schülerinnen und Schüler der Drei-Länder-Studie auf einen Blick.....	73
Abbildung 6.2:	Abfolge der Fragebögen und Tests in der Drei-Länder-Studie 1978/79.....	74
Abbildung 6.3:	Ausschnitt eines Rohdatensatzes der Drei-Länder-Studie in der SPSS-Variablenansicht.....	76
Abbildung 6.4:	Wright Map für die 24 ausgewählten Items in Mathematik.....	115
Abbildung 6.5:	Itemcharakteristische Funktionen und empirische Antwortkurven für zwei Items des Englischtests.....	117
Abbildung 6.6:	Wright Map für die 34 ausgewählten Items in Englisch.....	118
Abbildung 7.1:	Beispielitem für die Skala „Adaptivität“ aus dem Fragebogen „Schülerheft 4“ der Drei-Länder-Studie	132
Abbildung 7.2:	Instruktion und Erläuterung des Antwortformats aus dem Fragebogen „Schülerheft 4“ der Drei-Länder-Studie für das Fach Englisch	132
Abbildung 7.3:	Mediationsmodell	141
Abbildung 8.1:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Adaptivität als Mediator	182
Abbildung 8.2:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Adaptivität als Mediator.....	185
Abbildung 8.3:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator	185
Abbildung 8.4:	Grafisches Mediationsmodell zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Adaptivität als Mediator...	186
Abbildung 8.5:	Grafisches Mediationsmodell zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Adaptivität als Mediator	187

Abbildung 8.6:	Grafisches Mediationsmodell zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator	188
Abbildung 8.7:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Adaptivität als Mediator	188
Abbildung 8.8:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Adaptivität als Mediator.....	189
Abbildung 8.9:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator	189
Abbildung 8.10:	Grafisches Mediationsmodell zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Adaptivität als Mediator.....	191
Abbildung 8.11:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Disziplin als Mediator	192
Abbildung 8.12:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Disziplin als Mediator	193
Abbildung 8.13:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Förderung mit Disziplin als Mediator.....	193
Abbildung 8.14:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Disziplin als Mediator	194
Abbildung 8.15:	Grafisches Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Disziplin als Mediator	195
Abbildung 8.16:	Grafisches Mediationsmodell zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von der Reformbereitschaft der Lehrkräfte mit Disziplin als Mediator	196

Abbildung 8.17:	Grafisches Mediationsmodell zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von der Einstellung der Lehrkräfte zur Notwendigkeit von Druck mit Disziplin als Mediator	196
-----------------	---	-----

Tabellenverzeichnis

Tabelle 2.1:	Unterschiede zwischen Wissen und Überzeugungen	25
Tabelle 5.1:	Erwartete Zusammenhänge zwischen den Überzeugungen von Lehrkräften und den Unterrichtsmerkmalen	62
Tabelle 6.1:	Anzahl der Schulen in der Drei-Länder-Studie nach Schulform und Bundesland	68
Tabelle 6.2:	Veränderungen an den Skalen der Drei-Länder-Studie in der Übersicht	81
Tabelle 6.3:	Skalen der Drei-Länder-Studie für Schülerinnen und Schüler, originale und neue Fassungen – 1. Teil: affektive Skalen und elterliches Verhalten	83
Tabelle 6.4:	Skalen der Drei-Länder-Studie für Schülerinnen und Schüler, originale und neue Fassungen – 2. Teil: Unterrichtsskalen	90
Tabelle 6.5:	Skalen der Drei-Länder-Studie für Lehrkräfte, originale und neue Fassungen	100
Tabelle 6.6:	Kennwerte der verschiedenen Versionen aller Tests der Drei-Länder-Studie aus der IRT-Skalierung	114
Tabelle 7.1:	Anzahl der Schulen in der Sekundäranalyse nach Schulform und Bundesland	120
Tabelle 7.2:	Anzahl und Anteile der Schülerinnen und Schüler nach Schulform	122
Tabelle 7.3:	Anzahl und Anteile der Lehrkräfte nach Schulform	123
Tabelle 7.4:	Itemformulierungen und Itemtrennschärfen der Skala „Reformbereitschaft“	125
Tabelle 7.5:	Itemformulierungen und Itemtrennschärfen der Skala „Notwendigkeit von Druck bei Lernprozessen“	127
Tabelle 7.6:	Itemformulierungen und Itemtrennschärfen der Skala „Einstellung zur Förderung von schwachen Schülerinnen und Schülern“	128
Tabelle 7.7:	Deskriptive Statistiken der Überzeugungen von Lehrkräften in der Sekundäranalyse	129
Tabelle 7.8:	Verwendete Unterrichtsmerkmale der Sekundäranalyse	130
Tabelle 7.9:	Itemformulierungen und Itemtrennschärfen der Skala „Adaptivität des Englisch- bzw. Mathematikunterrichts“	131
Tabelle 7.10:	Itemformulierungen und Itemtrennschärfen der Skala „Disziplin im Englisch- bzw. Mathematikunterricht“	134
Tabelle 7.11:	Itemformulierungen und Itemtrennschärfen der Skala „Wohlfühlen im Englisch- bzw. Mathematikunterricht“	135
Tabelle 7.12:	Anteil fehlender Angaben bei Variablen der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie ..	145

Tabelle 7.13:	Anteil fehlender Angaben bei Überzeugungen der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie.....	146
Tabelle 8.1:	Deskriptive Verteilung der Überzeugungen von Lehrkräften, der Unterrichtsmerkmale und der schulischen Ergebnisse von Schülerinnen und Schülern im Englischunterricht	149
Tabelle 8.2:	Deskriptive Verteilung der Überzeugungen von Lehrkräften, der Unterrichtsmerkmale und der schulischen Ergebnisse von Schülerinnen und Schülern im Mathematikunterricht	151
Tabelle 8.3:	Korrelationen der Überzeugungen von Lehrkräften	152
Tabelle 8.4:	Korrelationen der Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen und schulischen Ergebnissen der Schülerinnen und Schüler im Englischunterricht	154
Tabelle 8.5:	Korrelationen der Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen und den schulischen Ergebnissen der Schülerinnen und Schüler im Mathematikunterricht	156
Tabelle 8.6:	Mehrebenenanalysen zur Adaptivität des Englischunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	159
Tabelle 8.7:	Mehrebenenanalysen zur Adaptivität des Mathematikunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	160
Tabelle 8.8:	Mehrebenenanalysen zur Disziplin im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	162
Tabelle 8.9:	Mehrebenenanalysen zur Disziplin im Mathematikunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	164
Tabelle 8.10:	Übersicht über die Ergebnisse aus Mehrebenenanalysen zu Unterrichtsmerkmalen des Englischunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	166
Tabelle 8.11:	Übersicht über die Ergebnisse aus Mehrebenenanalysen zu Unterrichtsmerkmalen des Mathematikunterrichts in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	169
Tabelle 8.12:	Mehrebenenanalysen zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte und der Adaptivität des Englischunterrichts.....	173
Tabelle 8.13:	Mehrebenenanalysen zu den Englischleistungen der Schülerinnen und Schüler in Abhängigkeit von den Überzeugungen der Lehrkräfte und der Adaptivität des Englischunterrichts.....	175
Tabelle 8.14:	Mehrebenenanalysen zum Wohlfühlen der Schülerinnen und Schüler im Mathematikunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte und der Adaptivität des Mathematikunterrichts	177

Tabelle 8.15:	Mehrebenenanalysen zu den Mathematikleistungen der Schülerinnen und Schüler in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	179
Tabelle 8.16:	Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte mit Adaptivität als Mediator..	183
Tabelle 8.17:	Totaler, indirekter und direkter Effekt der Adaptivität beim Mediationsmodell zum Wohlfühlen der Schülerinnen und Schüler im Englischunterricht in Abhängigkeit von den Überzeugungen der Lehrkräfte.....	184
Tabelle 8.18:	Übersicht über die Ergebnisse der Mediationsanalysen zu Adaptivität im Englischunterricht	186
Tabelle 8.19:	Übersicht über die Ergebnisse der Mediationsanalysen zur Adaptivität im Mathematikunterricht	190
Tabelle 8.20:	Übersicht über die Ergebnisse der Mediationsanalysen zu Disziplin im Englischunterricht.....	192
Tabelle 8.21:	Übersicht über die Ergebnisse der Mediationsanalysen zur Disziplin im Mathematikunterricht	194
Tabelle 9.1:	Zusammenhang von Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen im Englischunterricht: Ergebnisse der Mehrebenen-Regressionsanalysen im Überblick.....	198
Tabelle 9.2:	Zusammenhang von Überzeugungen der Lehrkräfte mit Unterrichtsmerkmalen im Mathematikunterricht: Ergebnisse der Mehrebenen-Regressionsanalysen im Überblick.....	199
Tabelle 9.3:	Adaptivität im Englischunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick.....	203
Tabelle 9.4:	Adaptivität im Mathematikunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick.....	204
Tabelle 9.5:	Disziplin im Englischunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick.....	205
Tabelle 9.6:	Disziplin im Mathematikunterricht als Mediator: Ergebnisse der Mediationsanalysen zur Erklärung der Lernergebnisse von Schülerinnen und Schülern durch Überzeugungen von Lehrkräften im Überblick.....	206

Tabelle A.1:	Geschlechterverteilung der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie.....	237
Tabelle A.2:	Migrationshintergrund der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie.....	237
Tabelle A.3:	Sozialstatus der Familie der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie.....	238
Tabelle A.4:	Geschlechterverteilung der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie.....	239
Tabelle A.5:	Altersverteilung der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie.....	239
Tabelle A.6:	Berufserfahrung der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie.....	240
Tabelle A.7:	Berufserfahrung der Lehrkräfte an dieser Schule in der Sekundäranalyse der Drei-Länder-Studie.....	240

Anhang

Tabelle A.1: Geschlechterverteilung der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie

s106r Zuerst kommen jetzt einige Fragen zu Deiner Person. Geschlecht						
	Englischunterricht			Mathematikunterricht		
	Häufigkeit	Prozent	Gültige Prozente	Häufigkeit	Prozent	Gültige Prozente
männlich	472	45,3	48,8	555	46,7	50,4
weiblich	495	47,5	51,2	546	46,0	49,6
Gesamt ohne fehlende Werte	967	92,7	100	1101	92,7	100
Fehlend	76	7,3		87	7,3	
Gesamt	1043	100		1188	100	

Tabelle A.2: Migrationshintergrund der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie

s111 Welche Sprache wird bei Euch zu Hause überwiegend gesprochen? (rekodiert)						
	Englischunterricht			Mathematikunterricht		
	Häufigkeit	Prozent	Gültige Prozente	Häufigkeit	Prozent	Gültige Prozente
nicht Deutsch	32	3,1	3,4	38	3,2	3,5
Deutsch	916	87,8	96,6	1062	89,4	96,5
Gesamt ohne fehlende Werte	948	90,9	100	1100	92,6	100
Fehlend	95	9,1		88	7,4	
Gesamt	1043	100		1188	100	

Tabelle A.3: Sozialstatus der Familie der Schülerinnen und Schüler in der Sekundäranalyse der Drei-Länder-Studie

s123r Welchem Kästchen würdest Du den Beruf Deines Vaters (bzw. Erziehungsberechtigten) zuordnen? Bitte streiche hier dasjenige Kästchen an, zu dem der Beruf Deines Vaters am ehesten passt.						
	Englischunterricht			Mathematikunterricht		
	Häufig- keit	Pro- zent	Gültige Prozente	Häufig- keit	Pro- zent	Gültige Prozente
1 = Parkwächter, Müllabfuhr- arbeiter, Gepäckträger, Kanalisa- tionsarbeiter, Zeitungsbotin, Garderobefrau, Gelegenheits- arbeiterin, Waschfrau	10	1,0	1,1	10	0,8	0,9
2 = Bauarbeiter, Asphaltierer, Lagerarbeiter, Gärtnergehilfe, Textilarbeiterin, Ladenhilfe, Fließbandarbeiterin, Büglerin	91	8,7	9,8	104	8,8	9,8
3 = Maschinenschlosser, Möbel- schreiner, Heizungsmonteur, Bauer (kleiner Hof), Vorarbeiterin in einem Industriebetrieb, Friseu- se, Maschinenerste, Verkäuferin	250	24,0	26,8	327	27,5	30,8
4 = Montageführer, Buchhalter, Malermeister, Bauer (mittlerer Hof), Technische Zeichnerin, Meisterin in der Bekleidungs- industrie, Laborantin, Besitzerin einer kleinen Wäscherei	196	18,8	21,0	213	17,9	20,1
5 = Elektroingenieur, Amtsvor- steher, Polizeiinspektor, Versi- cherungsinspektor, Volksschul- lehrerin, Selbständige Buchhänd- lerin, Hauptbuchhalterin, Beam- tin im mittleren Dienst	240	23,0	25,7	266	22,4	25,0
6 = Gymnasialdirektor, Oberarzt, Rechtsanwalt, Technischer Direk- tor, Studienrätin, Ärztin, Innen- architektin, Prokuristin in einer größerer Firma	116	11,1	12,4	113	9,5	10,6
7 = Oberlandesgerichtspräsident, Generalkonsul, Bankpräsident, Generaldirektor, Bundes- verfassungsrichterin, Universi- tätsprofessorin, Senatspräside- ntin, Diplomatin	30	2,9	3,2	29	2,4	2,7
Gesamt ohne fehlende Werte	933	89,5	100	1062	89,4	100
Fehlend	110	10,5		126	10,6	
Gesamt	1043	100		1188	100	

Tabelle A.4: Geschlechterverteilung der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie

L242 Wir bitten Sie noch um einige Angaben zu Ihrer Person. Geschlecht						
	Englischlehrkräfte			Mathematiklehrkräfte		
	Häufigkeit	Prozent	Gültige Prozente	Häufigkeit	Prozent	Gültige Prozente
männlich	60	40,8	43,8	95	57,9	71,4
weiblich	77	52,4	56,2	38	23,2	28,6
Gesamt ohne fehlende Werte	137	93,2	100	133	81,1	100
Fehlend	10	6,8		31	18,9	
Gesamt	147	100		164	100	

Tabelle A.5: Altersverteilung der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie

L243 Wir bitten Sie noch um einige Angaben zu Ihrer Person. Alter						
	Englischlehrkräfte			Mathematiklehrkräfte		
	Häufigkeit	Prozent	Gültige Prozente	Häufigkeit	Prozent	Gültige Prozente
bis 25 Jahre	5	3,4	3,6	6	3,7	4,3
26 bis 35 Jahre	93	63,3	66,4	77	47,0	54,6
36 bis 45 Jahre	23	15,6	16,4	32	19,5	22,7
46 bis 55 Jahre	15	10,2	10,7	15	9,1	10,6
56 bis 65 Jahre	3	2	2,1	11	6,7	7,8
über 65 Jahre	1	0,7	0,7	–	–	–
Gesamt ohne fehlende Werte	140	95,2	100	141	86,0	100
Fehlend	7	4,8		23	14,0	
Gesamt	147	100		164	100	

Tabelle A.6: Berufserfahrung der Lehrkräfte in der Sekundäranalyse der Drei-Länder-Studie

L103 Seit wieviel Jahren sind Sie Lehrer? (Ab erster Staatsprüfung, Ausfalljahre bitte abziehen)						
	Englischlehrkräfte			Mathematiklehrkräfte		
	Häufigkeit	Prozent	Gültige Prozente	Häufigkeit	Prozent	Gültige Prozente
bis 2 Jahre	16	10,9	11,4	13	7,9	8,7
3 bis 5 Jahre	38	25,9	27,1	38	23,2	25,5
6 bis 10 Jahre	41	27,9	29,3	45	27,4	30,2
11 und mehr Jahre	45	30,6	32,1	53	32,3	35,6
Gesamt ohne fehlende Werte	140	95,2	100	149	90,9	100
Fehlend	7	4,8		15	9,1	
Gesamt	147	100		164	100	

Tabelle A.7: Berufserfahrung der Lehrkräfte an dieser Schule in der Sekundäranalyse der Drei-Länder-Studie

L104 Wie lange unterrichten Sie schon an dieser Schule?						
	Englischlehrkräfte			Mathematiklehrkräfte		
	Häufigkeit	Prozent	Gültige Prozente	Häufigkeit	Prozent	Gültige Prozente
bis zu 2 Jahren	43	29,3	30,7	42	25,6	28,0
3 bis 5 Jahre	45	30,6	32,1	38	23,2	25,3
6 bis 10 Jahre	43	29,3	30,7	44	26,8	29,3
über 10 Jahre	9	6,1	6,4	26	15,9	17,3
Gesamt ohne fehlende Werte	140	95,2	100	150	91,5	100
Fehlend	7	4,8		14	8,5	
Gesamt	147	100		164	100	

Abstract

This study analyses the pedagogical beliefs of teachers as a central aspect of their professional competence which is significant for the quality of school and instruction. It examines which instructional characteristics general pedagogical beliefs of teachers are related to and how these beliefs, mediated by instruction, are related to student outcomes. For this a secondary analysis of Helmut Fend's "Drei-Länder-Studie" in Germany from 1978 and 1979 is carried out.

The aim of this study is to analyse the relationships between general pedagogical beliefs, instruction and student outcomes, taking into account the current state of theoretical and methodological research. Such general pedagogical beliefs are relevant to all school subjects; in the context of this study, general pedagogical beliefs of English and mathematics teachers are considered. Three teachers' beliefs are in focus: Teachers' attitudes towards the support of weak students and teachers' attitudes towards the necessity of pressure in learning processes relate to a comprehensive understanding of instruction that enhances learning. Thirdly, the willingness of teachers to reform is included as a general pedagogical belief which indicates how resistantly or flexibly schools should react to societal demands.

The first research question examines the relation of teachers' beliefs and instruction. Then relations between teachers' beliefs and student outcomes are addressed: The second research question investigates the relations between teachers' beliefs and student well-being. The third research question analyses the relations between teachers' beliefs and student achievement in mathematics or English.

The analyses of the first research question, the relation of teachers' beliefs to their instruction, examine various aspects of instruction. Seven constructs are considered, which depict a supportive teaching climate and four constructs, which illustrate classroom management. Overall, the connection between the beliefs investigated here and the instructional characteristics can be empirically confirmed primarily for characteristics of a supportive climate in English lessons. Only here correlations between pedagogical beliefs of teachers and the respective instructional characteristics emerge consistently. Such correlations cannot be

systematically found for the four constructs of classroom management or for mathematics classes.

The scale “adaptivity” as an aspect of a supportive climate and the scale “discipline” as an aspect of classroom management are of particular importance in the analyses of the second and third research question, since the mediating function of teachers’ beliefs on student outcomes is examined for these two instructional characteristics.

The empirical analyses of the second question largely support the assumption of a connection between the teachers’ pedagogical beliefs and the students’ well-being in class. Students taught by teachers who showed a high willingness to reform in English or mathematics lessons or a positive attitude towards supporting weak students in English lessons felt more comfortable in class than their classmates. Students whose teacher considered pressure in learning to be necessary felt less comfortable in both English and mathematics classes. Regarding the willingness to reform and the necessity of pressure in learning processes, it can still be assumed that this relation is mediated by adaptivity, at least in English lessons. The assumption of discipline as a mediator of teachers’ pedagogical beliefs on the well-being of students cannot be empirically confirmed. In mathematics lessons, no mediation processes of pedagogical beliefs on students’ well-being could be determined.

The analyses for the third research question show no mediation of teachers’ pedagogical beliefs by the adaptivity or discipline of their instruction for students’ English and mathematics achievement.

In summary, the results of this secondary analysis show that in English classes, teachers’ pedagogical beliefs in the late 1970s in Germany – mediated by adaptive instruction – were related to students’ learning outcomes, especially in the affective domain. They further lead to the conclusion that general pedagogical beliefs of teachers were related to a supportive climate in English classes. Moreover, these pedagogical beliefs were significant for students’ well-being and, to some extent, for their achievement in English and mathematics. These findings are also of fundamental theoretical importance beyond the historical context.