

Die Effekte von Schule auf Leistungsentwicklung und Leistungsunterschiede nach sozialer Herkunft: Eine Längsschnittstudie zu Ursachen von und Maßnahmen gegen Bildungsungleichheiten

Nachbauer, Max

Veröffentlichungsversion / Published Version

Dissertation / phd thesis

Empfohlene Zitierung / Suggested Citation:

Nachbauer, M. (2023). *Die Effekte von Schule auf Leistungsentwicklung und Leistungsunterschiede nach sozialer Herkunft: Eine Längsschnittstudie zu Ursachen von und Maßnahmen gegen Bildungsungleichheiten*. (Empirische Erziehungswissenschaft, 83). Münster: Waxmann. <https://doi.org/10.31244/9783830997320>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see: <https://creativecommons.org/licenses/by/4.0>



Max Nachbauer

Die Effekte von Schule auf Leistungsentwicklung und Leistungsunterschiede nach sozialer Herkunft

**Eine Längsschnittstudie zu
Ursachen von und Maßnahmen gegen
Bildungsungleichheiten**

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 83

Max Nachbauer

Die Effekte von Schule auf Leistungsentwicklung und Leistungsunterschiede nach sozialer Herkunft

Eine Längsschnittstudie zu Ursachen von
und Maßnahmen gegen Bildungsungleichheiten



Waxmann 2023

Münster • New York

Das vorliegende Werk wurde als Dissertation an der Fakultät Humanwissenschaften der Otto-Friedrich-Universität Bamberg eingereicht.

Diese Arbeit nutzt Daten des Nationalen Bildungspanels (NEPS; Blossfeld & Roßbach, 2019). Das NEPS wird vom Leibniz-Institut für Bildungsverläufe (LifBi, Bamberg) in Kooperation mit einem deutschlandweiten Netzwerk durchgeführt.

Ich danke für die Unterstützung durch den Open-Access-Publikationsfonds der Otto-Friedrich-Universität Bamberg.

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.

Empirische Erziehungswissenschaft, Bd. 83

ISSN 1862-2127

Print-ISBN 978-3-8309-4732-5

E-Book-ISBN 978-3-8309-9732-0

<https://doi.org/10.31244/9783830997320>

Waxmann Verlag GmbH, 2023
Steinfurter Straße 555, 48159 Münster

www.waxmann.com

info@waxmann.com

Umschlaggestaltung: Gregor Pleßmann, Ascheberg

Umschlagbild: © FatCamera – istock

Satz: MTS. Satz & Layout, Münster



Dieses Werk ist lizenziert unter der Creative-Commons-Lizenz CC BY 4.0. Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Zitate etc.) sind von der CC-Lizenz ausgenommen und für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen.

Vorwort

Beim vorliegenden Werk handelt es sich um eine geringfügig überarbeitete Fassung meiner Dissertation, welche ich im Jahr 2021 an der Fakultät Humanwissenschaften der Otto-Friedrich-Universität Bamberg eingereicht habe. Es ist mir eine große Freude, dass die Erkenntnisse meiner mehrjährigen Forschungsarbeiten nun einer größeren Leserschaft zugänglich gemacht werden. Die Wirksamkeit von Schule und Unterricht sowie Bildungsungleichheiten nach sozialer Herkunft zählen gegenwärtig zu den wichtigsten Themen der empirischen Bildungsforschung. Mit meinen Analysen zu den Ursachen von Kompetenzunterschieden nach sozialer Herkunft und möglichen schulischen Maßnahmen zur Verringerung dieser Kompetenzunterschiede möchte ich einen Beitrag zur Debatte um gute Schulen und Bildungsgerechtigkeit leisten. Es steht zu hoffen, dass die gewonnenen Erkenntnisse für die pädagogische Praxis und Bildungspolitik von Nutzen sind.

Ich möchte mich bei einer Reihe von Menschen bedanken, die auf unterschiedliche Weise zur Entstehung des vorliegenden Werks beigetragen haben. Dank gebührt zunächst einmal meinem Doktorvater Hans-Günther Roßbach. Der Grundstein für mein Interesse an der empirischen Bildungsforschung wurde maßgeblich durch die Lehrveranstaltungen von Herrn Roßbach gelegt, die ich während meines Studiums an der Otto-Friedrich-Universität Bamberg besucht habe. Im Laufe der Promotion hat mir Herr Roßbach in zahlreichen Beratungsgesprächen viele hilfreiche Rückmeldungen und Hinweise für die Schärfung des inhaltlichen Profils der Arbeit gegeben. Weiterhin möchte ich mich bei meinem Zweitbetreuer Hartmut Ditton bedanken. Aufgrund seiner methodischen Expertise war er mir insbesondere bei der Durchführung der empirischen Analysen eine große Hilfe. Dank gebührt außerdem der Bamberg Graduate School of Social Sciences, die mir in den ersten Jahren der Promotion eine anregende Lernumwelt geboten hat. Auf wissenschaftlicher Seite möchte ich mich insbesondere bei Claus Carstensen, Cordula Artelt, Sabine Weinert, Steffen Schindler und Ulrich Schroeders bedanken. Auf Seite der Verwaltung bedanke ich mich bei Thomas Saalfeld, Miriam Schneider und Marc Scheibner. Ebenfalls bedanken möchte ich mich bei den Mitarbeiter/-innen des Leibniz-Instituts für Bildungsverläufe, und hier insbesondere bei Daniel Fuß, für zahlreiche Hilfestellungen bei der Arbeit mit den Daten des Nationalen Bildungspanels. Für Beratungsgespräche im Rahmen von Konferenzen und Summer

Schools bedanke ich mich bei Knut Schwippert, Jürgen Baumert und Leonidas Kyriakides.

Bedanken möchte ich mich weiterhin bei meiner Familie. Die schönen und ausgelassenen Stunden mit meinem Vater Dieter, meiner Mutter Sibylla, meiner Schwester Kathrin, meinem Schwager Willi und meinen Neffen Kalle und Matti waren ein wichtiger Ausgleich zu der hohen Arbeitsbelastung. Bei Friederike möchte ich mich für das große Verständnis und die Rücksichtnahme bedanken. Besonderer Dank gebührt schließlich meinen Freund/-innen Riccardo, Michi, Valentina und Christopher für die vielen Stunden des Korrekturlesens. Ihre Rückmeldungen waren von großem Wert für mich und haben die Qualität der Arbeit erheblich verbessert.

Berlin, im Juli 2023
Max Nachbauer

Zusammenfassung

Das vorliegende Werk befasst sich mit dem Lernerfolg von Schüler/-innen und der Rolle, die die Bildungseinrichtung Schule einerseits und der familiäre Hintergrund der Schüler/-innen andererseits hierfür spielen. Es verbindet den Forschungsbereich der Wirksamkeit von Unterricht und Schule mit dem Forschungsbereich der Bildungsungleichheiten. Ausgangspunkt sind die wiederholt festgestellten Kompetenzunterschiede zwischen sozial privilegierten und sozial benachteiligten Schüler/-innen. Ungleiche Bildungschancen stellen eine anhaltende gesellschaftliche Herausforderung dar, deren Ursachen bisher allerdings noch nicht ausreichend geklärt sind. Ebenso fehlt es an Evidenzen über schulische Maßnahmen, mit denen Kompetenzunterschiede nach sozialer Herkunft verringert werden können. Diese Problemstellungen werden im vorliegenden Werk einer empirischen Untersuchung unterzogen.

Im ersten Teil werden *theoretische Grundlagen* erörtert und es wird ein *Überblick über den vorhandenen Forschungsstand* gegeben. In theoretischer Hinsicht wird zunächst das zentrale Konzept der pädagogischen Wirksamkeit erläutert. Weiterhin werden zwei theoretische Modelle entwickelt. Erstens wird ein Modell zu den vermittelnden Mechanismen zwischen der sozialen Herkunft von Schüler/-innen und ihrem Lernerfolg entwickelt. Dieses Modell integriert bildungssoziologische und pädagogisch-psychologische Theorieansätze. Zweitens wird ein Modell der Wirksamkeit von Schule entwickelt, welches insbesondere das Zusammenwirken von schulischen und familiären Einflüssen betrachtet. Hierauf aufbauend wird ein Überblick über vorhandene empirische Befunde gegeben. Der Überblick zeigt, dass in mehreren relevanten Forschungsfeldern nur eine begrenzte Anzahl an aussagekräftigen Studien zur Verfügung steht.

Den zweiten Teil bildet eine *empirische Analyse*. Es handelt sich um eine multivariate Längsschnittstudie auf der Grundlage des Nationalen Bildungspanels. Untersucht werden Gymnasien in ganz Deutschland im Zeitraum zwischen der fünften und der siebten Jahrgangsstufe (1523 Schüler/-innen in 120 Klassen). Als abhängige Variable wird der Erwerb von Mathematikkompetenz betrachtet.

Die Analysen umfassen drei Bereiche von Fragestellungen. Der erste Bereich bezieht sich auf *vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg*. Hierbei werden die Ursachen von Kompetenzunterschieden nach sozialer Herkunft untersucht. Die Analysen zielen darauf ab, Mediationsvariablen zu identifizieren, also Variablen, die von der sozialen Herkunft bedingt werden und ihrerseits den Lernerfolg von Schüler/-innen bedingen. Die

empirischen Befunde sprechen dafür, dass die kognitiven Lernvoraussetzungen von Schüler/-innen, ihre Freizeitaktivitäten und die Zusammensetzung ihres Freundeskreises bedeutsame Mediatoren darstellen. Demnach spielen diese Faktoren eine Rolle für die Entstehung von Kompetenzunterschieden nach sozialer Herkunft.

Der zweite Bereich von Fragestellungen behandelt die *Effekte von Klassen*. Klasseneffekte beziehen sich auf die Frage, inwieweit der Lernerfolg von Schüler/-innen in verschiedenen Klassen unterschiedlich ausfällt. Es zeigt sich klar, dass der Kompetenzzuwachs der Schüler/-innen bedeutsam zwischen Klassen variiert. Bezüglich der Kompetenzunterschiede nach sozialer Herkunft sind die Ergebnisse nicht ganz konsistent, aber in der Tendenz sprechen die Ergebnisse dafür, dass auch Kompetenzunterschiede nach sozialer Herkunft bedeutsam zwischen Klassen variieren. Analysen zum Verhältnis zwischen den beiden Zielkriterien ergeben, dass ein hoher Kompetenzzuwachs und geringe Kompetenzunterschiede miteinander vereinbar sind (sich also nicht gegenseitig ausschließen).

Der dritte Bereich von Fragestellungen bezieht sich auf die *Effekte von pädagogischen Merkmalen*. Hierbei steht die Frage nach wirksamen schulischen Maßnahmen im Vordergrund. Es wird untersucht, inwieweit Unterrichtsmerkmale, Ganztagsangebote und Klassenmerkmale den Lernerfolg der Schüler/-innen beeinflussen. Die empirischen Befunde legen nahe, dass ein hoher Kompetenzzuwachs der Schüler/-innen durch eine effiziente Klassenführung, ein hohes Ausmaß an kognitiver Aktivierung, ein mittleres Maß an Klassenunterricht und Partner-/Gruppenarbeiten sowie eine leistungsstarke Klassenkomposition begünstigt wird. Individualisierung im Unterricht und die Häufigkeit von Förderunterricht für leistungsstarke Schüler/-innen haben dagegen negative Effekte. Bezüglich der Kompetenzunterschiede nach sozialer Herkunft zeigt sich, dass geringe Kompetenzunterschiede durch häufige Partner-/Gruppenarbeiten (insbesondere bei leistungsheterogenen Gruppen), viel Unterstützung durch die Lehrkraft, regelmäßigen Förderunterricht in Mathematik und eine leistungsstarke Klassenkomposition begünstigt werden. Demgegenüber führen Förderunterricht für leistungsstarke Schüler/-innen und ein hoher Anteil an Schüler/-innen mit Migrationshintergrund zu einer Verstärkung von Kompetenzunterschieden.

Im letzten Kapitel wird die methodische Vorgehensweise der Studie diskutiert. Unter Rückbezug auf die empirischen Befunde werden Empfehlungen formuliert, wie Schule und Unterricht gestaltet werden sollten, um einen hohen Kompetenzzuwachs und geringe Kompetenzunterschiede nach sozialer Herkunft zu erreichen. Abschließend werden Implikationen für zukünftige Studien der empirischen Bildungsforschung erörtert.

Abstract

This study examines how students' learning success is affected by schools on the one hand and family background on the other hand. It brings together research on educational effectiveness and research on educational inequality. Numerous studies have shown that there is a considerable achievement gap between students from high socioeconomic status families and students from low socioeconomic status families. These educational inequalities remain a major societal challenge. However, the causes of the socioeconomic achievement gap are still not fully understood. Further, evidence on school-based interventions to reduce the achievement gap is lacking. These problems are addressed by the empirical investigations presented in this study.

The first part comprises the *theoretical framework* and a *review of research*. As a theoretical foundation, the concept of educational effectiveness is elaborated on. Moreover, two theoretical models are developed. The first model describes mediating mechanisms between students' social background and their learning success. It integrates theoretical approaches from sociology of education and educational psychology. The second model focuses on factors related to effective schooling and teaching. Particularly, it considers how school and family factors work together. Building on this theoretical framework, a review of empirical studies is provided. The review reveals that there is a lack of high-quality studies in several fields of research.

In the second part, an *empirical analysis* is presented. The multivariate and longitudinal analysis is based on the National Educational Panel Study. The sample covers high-track secondary schools all over Germany (1523 students in 120 classes). The dependant variable is achievement gain in mathematics from fifth to seventh grade.

Three domains of research questions are examined. The first domain refers to *mediating mechanisms between social background and learning success*. This means that the causes of the socioeconomic achievement gap are investigated. Analyses aim at identifying mediators, thus variables that are affected by social background and in turn affect learning success. Empirical results suggest that students' cognitive learning preconditions, their recreational activities and the composition of their peer group are relevant mediators. Therefore, these factors contribute to the emergence of the socioeconomic achievement gap.

The second domain of research questions focuses on *class effects*. Class effects refer to the question whether school classes differ in regard to students' learning success. Analyses clearly show that students' achievement gains vary substantially between school classes. Results on socioeconomic achievement gaps are not quite consistent, but there is a tendency that socioeconomic achievement gaps also vary between school classes. Analyses on the relation between the two criteria reveal that a high achievement gain and a small achievement gap are compatible goals (thus do not exclude each other).

The third domain of research questions addresses the *effectiveness of school and teaching factors*. These analyses aim at identifying school-based interventions that promote student learning. Empirical investigations focus on teaching characteristics, extended educational offerings (provided in the afternoon) and class characteristics. Empirical results suggest that high achievement gains are facilitated by an efficient classroom management, a high degree of cognitive activation, a moderate degree of whole-class instruction and partner/group work as well as a high-ability class composition. In contrast, differentiated instruction and advanced courses for high-achieving students both have negative effects. In terms of socioeconomic achievement gaps, it is revealed that small achievement gaps are facilitated by frequent partner/group work (especially with mixed-ability groups), intense teacher support, frequent remedial courses in mathematics and a high-ability class composition. On the other hand, advanced courses for high-achieving students and a class composition with many migrant students increase socioeconomic achievement gaps.

In the last chapter, the methodological approach of the study is discussed. Referring to the empirical results, recommendations for schooling and teaching are made. Interventions that promote high achievement gains and small socioeconomic achievement gaps are illustrated. Finally, implications for future studies in the field of educational research are drawn.

Inhalt

1.	Einleitung	17
2.	Pädagogische Wirksamkeit	23
2.1	Zielsetzungen von Schule und Unterricht	23
2.1.1	Lernziele nach Persönlichkeitsbereichen	23
2.1.2	Durchschnittlicher Lernerfolg und Unterschiede im Lernerfolg	25
2.2	Konzeptualisierung pädagogischer Wirksamkeit	27
2.2.1	Definition pädagogischer Wirksamkeit	27
2.2.2	Pädagogische Faktoren	28
2.2.3	Die Rolle der Schüler/-innen	30
2.2.4	Lernaktivitäten und Lernvoraussetzungen	31
2.2.5	Lernwirkung, Lernerfolg und pädagogischer Effekt	33
2.2.6	Zusammenwirken pädagogischer Merkmale	35
2.2.7	Dimensionen pädagogischer Wirksamkeit	36
2.3	Pädagogische Wirksamkeitsforschung	38
2.3.1	Ziele und Fragestellungen der pädagogischen Wirksamkeitsforschung ..	39
2.3.2	Forschungsmethoden und -designs der pädagogischen Wirksamkeitsforschung	41
2.3.3	Standards für Wirksamkeitsstudien	44
2.4	Zusammenfassung des zweiten Kapitels	46
3.	Effekte der sozialen Herkunft	47
3.1	Bildungsungleichheiten nach sozialer Herkunft	48
3.1.1	Konzeptualisierung und Operationalisierung sozialer Herkunft	49
3.1.2	Soziale Herkunft und Bildungserfolg	50
3.2	Theoretisches Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	51
3.2.1	Theoretische Grundlagen des Modells zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	52
3.2.2	Beschreibung des Modells zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	53
3.3	Forschungsansätze zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	57
3.4	Unterschiede bei familiären Ressourcen	59
3.5	Unterschiede bei familiären Lernumwelten	61
3.5.1	Unterschiede beim Erziehungsstil der Eltern	61
3.5.2	Unterschiede bei der kognitiven Anregung durch die Eltern	63
3.5.3	Unterschiede bei der schulischen Beteiligung der Eltern	64
3.6	Unterschiede bei frühpädagogischen Einrichtungen	66

3.7	Unterschiede bei Schulen	68
3.7.1	Unterschiede in den besuchten Schulen	68
3.7.2	Unterschiede in schulischen Lernaktivitäten	70
3.8	Unterschiede bei schulbezogenen Einrichtungen	72
3.9	Unterschiede bei Freizeiteinrichtungen und Freizeitaktivitäten	73
3.10	Unterschiede bei Gleichaltrigengruppen	74
3.11	Unterschiede bei Lernvoraussetzungen	75
3.11.1	Unterschiede bei kognitiven Merkmalen der Heranwachsenden	76
3.11.2	Unterschiede bei motivationalen Merkmalen der Heranwachsenden	78
3.11.3	Unterschiede bei psychischen Dispositionen der Heranwachsenden	79
3.12	Zusammenfassung des dritten Kapitels	81
4.	Effekte von Klassen und Schulen	83
4.1	Theoretische Positionen zur Rolle der Schule	84
4.2	Schulische Effekte auf den durchschnittlichen Lernerfolg	86
4.2.1	Absolute schulische Effekte auf den durchschnittlichen Lernerfolg	86
4.2.2	Relative schulische Effekte auf den durchschnittlichen Lernerfolg	87
4.2.3	Relative schulische Effekte auf den durchschnittlichen Lernerfolg bei verschiedenen Schülergruppen	90
4.3	Schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg	92
4.3.1	Absolute schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg	92
4.3.2	Relative schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg	95
4.4	Die Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg	97
4.5	Zusammenfassung des vierten Kapitels	99
5.	Effekte von pädagogischen Merkmalen	102
5.1	Theoretisches Modell der Wirksamkeit von Schule	104
5.1.1	Theoretische Grundlagen des Modells der Wirksamkeit von Schule	104
5.1.2	Beschreibung des Modells der Wirksamkeit von Schule	105
5.2	Mechanismen der Effekte von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg	109
5.2.1	Förderung verschiedener Schülergruppen	110
5.2.2	Differenziell wirksame Lerngelegenheiten und unterschiedliche Lerngelegenheiten	111
5.2.3	Vertiefte Betrachtung differenziell wirksamer Lerngelegenheiten	115
5.2.4	Vertiefte Betrachtung unterschiedlicher Lerngelegenheiten	117
5.3	Forschungsansätze zu Effekten von pädagogischen Merkmalen	119
5.4	Effekte von Unterrichtsmerkmalen	122
5.4.1	Effekte des Unterrichtsumfangs	122
5.4.2	Effekte von Sozialformen	123

5.4.3	Effekte von instruktionalen und konstruktivistischen Unterrichtsmethoden	126
5.4.4	Effekte von Klassenführung	130
5.4.5	Effekte von Klarheit und Strukturiertheit	131
5.4.6	Effekte von kognitiver Aktivierung	133
5.4.7	Effekte von Übung und Wiederholung	135
5.4.8	Effekte von Unterstützung und Feedback	137
5.4.9	Effekte von Individualisierung	140
5.5	Effekte von Ganztagsangeboten	142
5.5.1	Effekte der schulischen Organisationsform	144
5.5.2	Effekte von spezifischen Programmen	145
5.6	Effekte von Merkmalen der Lerngruppe	147
5.6.1	Effekte der Klassengröße	147
5.6.2	Effekte der Schülerkomposition	148
5.7	Zusammenfassung des fünften Kapitels	151
6.	Fragestellungen	153
6.1	Herleitung der Fragestellungen	153
6.2	Formulierung der Fragestellungen	155
6.3	Auswahl des pädagogischen Kontexts	158
7.	Forschungsdesign und Forschungsmethoden	162
7.1	Datengrundlage	162
7.1.1	Das Nationale Bildungspanel	163
7.1.2	Beschreibung der Stichprobe	164
7.2	Variablen	166
7.2.1	Auswahl der Untersuchungsvariablen	166
7.2.2	Variablen auf der Schülerebene	168
7.2.3	Variablen auf der Klassen- und Schulebene	172
7.3	Analysen	174
7.3.1	Überblick über die verwendeten Analysemethoden	175
7.3.2	Analysestrategie für vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg	177
7.3.3	Analysestrategie für Effekte von Klassen	179
7.3.4	Analysestrategie für Effekte von pädagogischen Merkmalen	186
7.3.5	Umsetzung der Analysen	190
7.3.6	Robustheitsanalysen	191
7.4	Zusammenfassung des siebten Kapitels	192
8.	Ergebnisse	193
8.1	Einleitende Bemerkungen	193
8.2	Deskriptive Ergebnisse	195

8.3	Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	197
8.3.1	Ergebnisse zur Rolle der elterlichen Lernbegleitung	198
8.3.2	Ergebnisse zur Rolle der Nutzung von Ganztagsangeboten	200
8.3.3	Ergebnisse zur Rolle von Nachhilfe	203
8.3.4	Ergebnisse zur Rolle von Freizeitaktivitäten	205
8.3.5	Ergebnisse zur Rolle der Gleichaltrigengruppe	207
8.3.5	Ergebnisse zur Rolle von Lernvoraussetzungen	209
8.3.6	Ergebnisse des Gesamtmodells zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	213
8.3.7	Diskussion der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	215
8.4	Ergebnisse zu Effekten von Klassen	218
8.4.1	Ergebnisse zu relativen Klasseneffekten auf den durchschnittlichen Lernerfolg	218
8.4.2	Ergebnisse zu relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg	221
8.4.3	Ergebnisse zum Verhältnis zwischen durchschnittlichem Lernerfolg und herkunftsbedingten Unterschieden im Lernerfolg	225
8.4.4	Diskussion der Ergebnisse zu Effekten von Klassen	225
8.5	Ergebnisse zu Effekten von pädagogischen Merkmalen	227
8.5.1	Ergebnisse zu Effekten von Unterrichtsmerkmalen	229
8.5.2	Diskussion der Ergebnisse zu Effekten von Unterrichtsmerkmalen	234
8.5.3	Ergebnisse zu Effekten von Ganztagsangeboten	236
8.5.4	Diskussion der Ergebnisse zu Effekten von Ganztagsangeboten	239
8.5.5	Ergebnisse zu Effekten von Merkmalen der Lerngruppe	242
8.5.6	Diskussion der Ergebnisse zu Effekten von Merkmalen der Lerngruppe	245
8.5.7	Ergebnisse der Gesamtmodelle zu Effekten von pädagogischen Merkmalen	246
9.	Abschließende Zusammenfassung und Diskussion	251
9.1	Bewertung der Forschungsstrategie	251
9.2	Diskussion der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	253
9.2.1	Zusammenfassung der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	254
9.2.2	Einordnung der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg in den Forschungsstand	255
9.2.3	Bewertung der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	258
9.3	Diskussion der Ergebnisse zu Effekten von Klassen	259
9.3.1	Zusammenfassung der Ergebnisse zu Effekten von Klassen	260
9.3.2	Einordnung der Ergebnisse zu Effekten von Klassen in den Forschungsstand	261

9.3.3	Bewertung der Ergebnisse zu Effekten von Klassen	262
9.4	Diskussion der Ergebnisse zu Effekten von pädagogischen Merkmalen ..	263
9.4.1	Zusammenfassung der Ergebnisse zu Effekten von pädagogischen Merkmalen	263
9.4.2	Einordnung der Ergebnisse zu Effekten von pädagogischen Merkmalen in den Forschungsstand	266
9.4.3	Bewertung der Ergebnisse zu Effekten von pädagogischen Merkmalen .	268
9.5	Vergleichende Gegenüberstellung ausgewählter Befunde	270
9.6	Implikationen für die pädagogische Praxis und Bildungspolitik	271
9.7	Implikationen für die empirische Bildungsforschung	276
9.7.1	Implikationen für Studien zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg	276
9.7.2	Implikationen für Wirksamkeitsstudien zu Unterschieden im Lernerfolg	277
Literatur		280
Anhang A: Erhebungsinstrumente		315
Anhang B: Ergebnisse der konfirmatorischen Faktorenanalysen		322
Anhang C: Abbildungen von Regressionslinien		325

1. Einleitung

In modernen Gesellschaften ist Bildung ein entscheidender Faktor für eine erfolgreiche Lebensgestaltung. Moderne Gesellschaften sind unter anderem durch ein hohes Maß an individuellen Freiheiten, vielfältigen Möglichkeiten demokratischer Mitbestimmung und soziale Mobilität beim Zugang zu gesellschaftlichen Positionen gekennzeichnet (Degele, 2014, S. 326). Wie Individuen ihre Freiheiten und Mitbestimmungsmöglichkeiten nutzen, aber auch auf welche gesellschaftlichen Positionen sie gelangen, hängt wesentlich von ihrer Bildung ab. So haben höher gebildete Personen durchschnittlich eine bessere Gesundheit, eine höhere Wahlbeteiligung, engagieren sich häufiger ehrenamtlich, werden seltener kriminell, erhalten höhere Einkommen und sind weniger von Arbeitslosigkeit betroffen (Autorengruppe Bildungsberichterstattung, 2012, S. 9, 207 ff.; Baumert et al., 2001b, S. 31; Lochner & Moretti, 2004; Schuller & Desjardins, 2007). An diesen Befunden wird deutlich, dass Bildung nicht nur einen Nutzen für das Individuum selbst hat, sondern auch für die Gesellschaft, in der das Individuum lebt. Vor diesem Hintergrund kann das Gelingen von Bildungsprozessen als ein Ziel von höchster gesellschaftlicher Relevanz angesehen werden.

Bildungsprozesse erfolgen zu einem großen Teil in öffentlichen Bildungseinrichtungen wie Kindergärten, Schulen und Hochschulen. Der Schule kommt hierbei aufgrund der gesetzlichen Schulpflicht und der hohen Dauer des Schulbesuchs eine besondere Rolle zu. Die Ausgestaltung von Schule und Unterricht wird durch ein komplexes Gefüge von *Entscheidungen auf unterschiedlichen Ebenen* bedingt. Mehrere Gruppen von Entscheidungsträger/-innen sind daran beteiligt. Lehrkräfte entscheiden, welche Unterrichtsmethoden sie verwenden oder wie sie mit problematischen Verhaltensweisen von Schüler/-innen umgehen. Schulleiter/-innen entscheiden, welche neuen Lehrkräfte sie einstellen oder nach welchem Prinzip neue Klassen gebildet werden. Bildungspolitiker/-innen und Mitarbeiter/-innen der Bildungsverwaltung entscheiden über Vorgaben zur Aus- und Weiterbildung von Lehrkräften oder über den Ausbau von Ganztagschulen. Angesichts der dargelegten Bedeutung von Bildung können diese Entscheidungen weitreichende Konsequenzen haben, sowohl für die einzelnen Schüler/-innen als auch für die Gesellschaft als Ganzes.

Dies wirft die Frage auf, auf welcher Grundlage die Entscheidungsträger/-innen des Schulsystems ihre Entscheidungen fällen. Ein zentrales Kriterium

bei der Entscheidungsfindung sind zunächst einmal die *Zielsetzungen*, welche durch die Schule als Bildungseinrichtung erreicht werden sollen. Grundsätzlich ist die Ausgestaltung von Schule und Unterricht auf eine Vielzahl von unterschiedlichen Zielen ausgerichtet. Neben fachlichen Lernzielen (z. B. Mathematikkompetenz, Lesekompetenz) werden auch übergreifende Lernziele (z. B. soziale Kompetenzen, Wertorientierungen) als relevant angesehen (Baumert et al., 2001b, S. 22; Weinert, 2000, S. 8 ff.). Hinsichtlich der genannten Lernziele sollen die Schüler/-innen einen möglichst hohen Lernerfolg erreichen. Zusätzlich sollen Unterschiede zwischen Schüler/-innen bezüglich ihres Lernerfolgs gering gehalten werden (Helmke, 2014, S. 38). Unterschiede im Lernerfolg beziehen sich beispielsweise darauf, dass Jungen und Mädchen, Schüler/-innen mit und ohne Migrationshintergrund oder Schüler/-innen verschiedener sozialer Herkunft unterschiedlich gute Fachkompetenzen erwerben. Dieser Themenbereich wurde durch die erste PISA-Studie im Jahr 2000 (Baumert et al., 2001a) wieder verstärkt in die öffentliche Diskussion gebracht. Eines der zentralen Ergebnisse des internationalen Vergleichs waren die in Deutschland ausgesprochen starken Kompetenzunterschiede zwischen Kindern von Eltern mit hohen Berufspositionen (z. B. Ärzt/-innen, Anwälte/-innen) und Kindern von Eltern mit niedrigen Berufspositionen (z. B. Facharbeiter/-innen, ungelernete Arbeiter/-innen) (Baumert & Schümer, 2001). Zahlreiche weitere Studien bestätigen, dass der Lernerfolg von Schüler/-innen bedeutsam von ihrem familiären Hintergrund abhängt (Sirin, 2005). Ungleiche Bildungschancen stellen eine anhaltende Herausforderung dar, deren Ursachen bisher allerdings noch nicht ausreichend geklärt sind (Neumann et al., 2014, S. 193). Obwohl also weiterer Forschungsbedarf besteht, haben die vorhandenen Befunde bereits ein Problembewusstsein erzeugt. So wird der Förderung sozial benachteiligter Kinder auf Seiten der Bildungspolitik eine hohe Priorität beigemessen (Kultusministerkonferenz, 2002).

Inwieweit es gelingt, den Lernerfolg der Schüler/-innen zu verbessern und Unterschiede im Lernerfolg zu verringern, hängt maßgeblich davon ab, wie Schule und Unterricht konkret ausgestaltet werden. Wesentlich ist hierbei, ob die Entscheidungsträger/-innen des Schulsystems wirksame Maßnahmen zur Erreichung dieser Zielsetzungen wählen. Diesbezüglich wird in der Erziehungswissenschaft seit der Jahrtausendwende das Potential einer *evidenzbasierten Vorgehensweise* diskutiert (Tippelt & Reich-Claassen, 2010, S. 22). Damit ist gemeint, dass Entscheidungen über zu ergreifende Maßnahmen auf der Grundlage von empirischen Befunden gefällt werden. Das für eine evidenzbasierte pädagogische Praxis und Bildungspolitik benötigte empirische Orientierungswissen wird insbesondere durch die *empirische Bildungsforschung* (Rein-

ders et al., 2015) zur Verfügung gestellt. Innerhalb dieses Forschungsfeldes ist hierbei die *pädagogische Wirksamkeitsforschung* hervorzuheben (Chapman et al., 2016a). Dieser Forschungsansatz untersucht die Wirksamkeit von pädagogischen Einrichtungen und relevanter Einflussfaktoren innerhalb dieser Einrichtungen (z. B. Unterrichtsmethoden, Klassengröße) für die Förderung des Lernerfolgs.

Umfassende Forschungsaktivitäten zur Wirksamkeit von Schule und Unterricht begannen in den 1970er Jahren, insbesondere in den USA und England (Chapman et al., 2016b, S. 1f.; Creemers et al., 2010, S. 5). In den darauffolgenden Jahrzehnten gewann das Forschungsfeld zunehmend an Bedeutung und in zahlreichen Ländern wurden Studien zur Wirksamkeit von Schule und Unterricht durchgeführt. Auch in Deutschland liegt inzwischen eine größere Anzahl von Untersuchungen vor (z. B. Baumert & Köller, 1998; Kunter et al., 2011; Weinert & Helmke, 1997). Dementsprechend steht eine fundierte Wissensbasis über schulische Einflussfaktoren auf den Lernerfolg von Schüler/-innen zur Verfügung (Hattie, 2012; Lipowsky, 2006; Muijs et al., 2014; Reynolds et al., 2014). Allerdings berücksichtigt der weitaus größte Teil der vorhandenen Studien lediglich den *durchschnittlichen Lernerfolg* als Kriterium für Wirksamkeit, nicht aber *Unterschiede im Lernerfolg* zwischen Schüler/-innen (Kyriakides, 2007, S. 41). Zu schulischen Einflussfaktoren auf herkunftsbedingte Unterschiede im Lernerfolg gibt es bisher sowohl international (z. B. Lee & Bryk, 1989; Gustafsson et al., 2018) als auch in Deutschland (z. B. Pfeifer, 2011; Schwippert, 2001) nur relativ wenige Studien. Folglich bezieht sich der größte Teil der verfügbaren Evidenzen auf Maßnahmen zur Förderung des schulischen Lernens im Allgemeinen. Empirisch fundiertes Orientierungswissen über Maßnahmen, mit denen Kompetenzunterschiede nach sozialer Herkunft verringert werden können, ist dagegen nur in begrenztem Maße vorhanden. Demnach fehlt den Entscheidungsträger/-innen des Schulsystems gerade in einem Bereich mit hohem Handlungsbedarf (Kultusministerkonferenz, 2002) eine solide empirische Grundlage.

An dieser Stelle setzt die vorliegende Forschungsarbeit an. Sie befasst sich mit dem Lernerfolg von Schüler/-innen unterschiedlicher sozialer Herkunft und der Rolle, welche die Bildungseinrichtung Schule dabei spielt. Drei Bereiche von Fragestellungen werden durch empirische Analysen bearbeitet. Erstens werden *vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg* untersucht. Im Vordergrund steht die Frage, aufgrund welcher Faktoren sozial privilegierte Schüler/-innen bessere Kompetenzen erwerben als sozial benachteiligte Schüler/-innen. Vermittelnde Mechanismen beziehen sich auf Mediationsvariablen, also auf Variablen, die von der sozialen Herkunft bedingt

werden und ihrerseits den Lernerfolg von Schüler/-innen bedingen. Konkret wird untersucht, inwieweit Kompetenzunterschiede nach sozialer Herkunft darauf zurückzuführen sind, dass sozial privilegierte Schüler/-innen eine bessere Lernbegleitung von ihren Eltern erhalten, häufiger schulische Ganztagsangebote nutzen, häufiger Nachhilfe nutzen, anregendere Freizeitaktivitäten ausüben, mehr schulisch engagierte Freund/-innen haben oder über bessere kognitive und motivationale Lernvoraussetzungen verfügen als sozial benachteiligte Schüler/-innen.

Der zweite Fragenbereich bezieht sich auf die *Effekte von Klassen*. Spezifisch werden relative Klasseneffekte untersucht. Diese verweisen auf die Frage, welche Bedeutung es für den Lernerfolg von Schüler/-innen hat, welche Klasse sie besuchen. Hierfür wird analysiert, inwieweit sich Klassen darin unterscheiden, wie hoch der durchschnittliche Lernerfolg der Schüler/-innen ausfällt. Unter dem Gesichtspunkt von Herkunftseffekten wird außerdem analysiert, inwieweit sich Klassen darin unterscheiden, wie stark Unterschiede im Lernerfolg zwischen Schüler/-innen verschiedener sozialer Herkunft ausfallen. Weiterhin wird überprüft, in welchem Verhältnis der durchschnittliche Lernerfolg und herkunftsbedingte Unterschiede im Lernerfolg zueinanderstehen. Hierbei ist von Interesse, ob es möglich ist, gleichzeitig einen hohen durchschnittlichen Lernerfolg und geringe Unterschiede im Lernerfolg zu erreichen.

Der letzte Fragenbereich betrifft die *Effekte von pädagogischen Merkmalen*. Effekte von pädagogischen Merkmalen beziehen sich auf die Frage, welche Merkmale von Schule und Unterricht Zusammenhänge zum Lernerfolg der Schüler/-innen aufweisen (z. B. bei häufigem Einsatz von Gruppenarbeiten fällt der Kompetenzzuwachs höher aus als bei seltenem Einsatz von Gruppenarbeiten). Auch bei diesen Analysen werden herkunftsbedingte Unterschiede im Lernerfolg berücksichtigt. Es wird also zusätzlich betrachtet, inwieweit Zusammenhänge zwischen pädagogischen Merkmalen und herkunftsbedingten Unterschieden im Lernerfolg bestehen (z. B. bei häufigem Einsatz von Gruppenarbeiten fallen Kompetenzunterschiede geringer aus als bei seltenem Einsatz von Gruppenarbeiten). Untersucht werden pädagogische Merkmale in den drei Bereichen Unterrichtsmerkmale (z. B. Klassenführung, kognitive Aktivierung), Ganztagsangebote (z. B. Hausaufgabenbetreuung, Förderunterricht) und Merkmale der Lerngruppe (z. B. Klassengröße, Klassenkomposition).

Die vorliegende Arbeit umfasst zwei Teile. Im *ersten Teil* werden theoretische Grundlagen erarbeitet und es wird ein Überblick über den vorhandenen Forschungsstand gegeben. Im nächsten Kapitel erfolgt zunächst eine theoretische Konzeptualisierung *pädagogischer Wirksamkeit*. Hierdurch soll ein grundlegendes Verständnis über den untersuchten Gegenstandsbereich

vermittelt werden. Weiterhin werden Ansätze zur Erforschung pädagogischer Wirksamkeit erläutert. Es wird aufgezeigt, welche Möglichkeiten der empirischen Untersuchung von pädagogischer Wirksamkeit bestehen und welche Schwierigkeiten beachtet werden müssen.

Das dritte Kapitel befasst sich schwerpunktmäßig mit den *Effekten der sozialen Herkunft* auf den Lernerfolg. Das Kapitel bildet die Grundlage für den ersten Bereich von Fragestellungen. Es wird ein theoretisches Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg entwickelt, welches bestehende Erklärungsansätze in einem interdisziplinären Ansatz zusammenführt. Basierend auf dem theoretischen Modell wird ein Überblick über den empirischen Forschungsstand zu vermittelnden Mechanismen gegeben. Resümierend wird herausgearbeitet, dass die Aussagekraft vieler Studien als begrenzt einzuschätzen ist und daher weiterer Forschungsbedarf besteht.

Das vierte Kapitel fokussiert auf die Bildungseinrichtung Schule, und zwar spezifisch auf die *Effekte von Klassen und Schulen*. In diesem Kapitel werden die Grundlagen für den zweiten Bereich von Fragestellungen erarbeitet. Im Vordergrund steht die Frage, welche Effekte bei den Schüler/-innen durch den Besuch von Klassen und Schulen entstehen. Konzeptuell wird eine Unterscheidung zwischen schulischen Effekten auf den durchschnittlichen Lernerfolg und schulischen Effekten auf herkunftsbedingte Unterschiede im Lernerfolg vorgenommen. Es wird ein Forschungsüberblick gegeben, der verdeutlicht, dass schulische Effekte auf den durchschnittlichen Lernerfolg gut untersucht sind, während bei schulischen Effekten auf herkunftsbedingte Unterschiede im Lernerfolg nur eine überschaubare Anzahl an Studien vorliegt.

Im fünften Kapitel stehen die *Effekte von pädagogischen Merkmalen* im Vordergrund. Das Kapitel bildet die Grundlage für den dritten Bereich von Fragestellungen und fokussiert somit Zusammenhänge zwischen pädagogischen Merkmalen und dem Lernerfolg der Schüler/-innen. Es wird ein Modell der Wirksamkeit von Schule entwickelt, welches eine theoretische Ordnung der Lernumwelt Schule leistet und hierbei auch das Zusammenwirken von schulischen und familiären Einflüssen berücksichtigt. Basierend auf diesen Überlegungen wird ein Überblick über vorhandene empirische Befunde zu den Effekten von pädagogischen Merkmalen gegeben, welcher sich auf die Bereiche Unterrichtsmerkmale, Ganztagsangebote und Merkmale von Lerngruppen konzentriert. Die im vorausgehenden Kapitel eingeführte Unterscheidung zwischen Effekten auf den durchschnittlichen Lernerfolg und Effekten auf herkunftsbedingte Unterschiede im Lernerfolg wird aufgegriffen. Auch in diesem Forschungsfeld ist festzustellen, dass Effekte auf den durchschnittlichen

Lernerfolg deutlich besser untersucht sind als Effekte auf herkunftsbedingte Unterschiede im Lernerfolg.

Der *zweite Teil* der Arbeit beschreibt eine empirische Untersuchung. Unter Rückgriff auf den zuvor gegebenen Forschungsüberblick werden im sechsten Kapitel die *Fragestellungen* hergeleitet und ausformuliert. Hierbei wird auch dargelegt, warum als Kontext für die empirische Untersuchung die Schulform Gymnasium und das Unterrichtsfach Mathematik gewählt wurde.

Im siebten Kapitel werden *Forschungsdesign und Forschungsmethoden* erläutert. Das Forschungsdesign ist eine multivariate Längsschnittstudie. Die Analysen basieren auf Daten des Nationalen Bildungspanels (Blossfeld & Roßbach, 2019). Die Erhebung der Daten erfolgt mit Tests, Fragebögen für Schüler/-innen, Lehrkräften und Schulleiter/-innen sowie Telefoninterviews mit Eltern. Für die Auswertung der Daten werden Mehrebenenmodelle in Kombination mit Pfadmodellen und Strukturgleichungsmodellen verwendet.

Im achten Kapitel werden die *Ergebnisse* präsentiert. Die Analysen zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg ergeben, dass insbesondere die kognitiven Lernvoraussetzungen von Schüler/-innen und ihre Freizeitaktivitäten als Mediatoren in Betracht kommen. Die Analysen zu relativen Klasseneffekten zeigen, dass sich Klassen bedeutsam bezüglich des durchschnittlichen Lernerfolgs und in der Tendenz auch bezüglich herkunftsbedingter Unterschiede im Lernerfolg unterscheiden. Ein hoher durchschnittlicher Lernerfolg geht hierbei mit geringen Unterschieden im Lernerfolg einher. In den Analysen zu den Effekten von pädagogischen Merkmalen wird festgestellt, dass die Effizienz der Klassenführung, das Ausmaß an kognitiver Aktivierung und das mathematische Ausgangsniveau der Klasse den durchschnittlichen Lernerfolg fördern, während Aufgaben mit Hilfe der Lehrkraft, Förderunterricht in Mathematik und das mathematische Ausgangsniveau der Klasse herkunftsbedingte Unterschiede im Lernerfolg gering halten.

Den Abschluss bildet das neunte Kapitel mit einer *Zusammenfassung und Diskussion*. Die Vorgehensweise und Ergebnisse der Untersuchung werden kritisch bewertet. Weiterhin werden Implikationen sowohl für die pädagogische Praxis und Bildungspolitik als auch für die empirische Bildungsforschung erörtert.

2. Pädagogische Wirksamkeit

In diesem Kapitel erfolgt eine grundlegende Einführung in den Gegenstandsbereich der pädagogischen Wirksamkeit. Die Funktion des Kapitels liegt darin, pädagogische Wirksamkeit begrifflich zu bestimmen, wesentliche Charakteristika pädagogischer Wirksamkeit aufzuzeigen sowie zu erörtern, wie pädagogische Wirksamkeit wissenschaftlich untersucht werden kann. Das Kapitel unterteilt sich in drei Abschnitte. Zunächst wird auf relevante Zielsetzungen von Schule und Unterricht eingegangen (Kapitel 2.1). Im Weiteren wird der Begriff der pädagogischen Wirksamkeit theoretisch konzeptualisiert (Kapitel 2.2). Anschließend werden Ansätze zur Erforschung pädagogischer Wirksamkeit erläutert (Kapitel 2.3).

2.1 Zielsetzungen von Schule und Unterricht

Die Zielsetzung von Schule und Unterricht sind von wesentlicher Bedeutung für die Betrachtung von pädagogischer Wirksamkeit. Zielsetzungen bilden die normative Grundlage, indem sie wünschenswerte Ergebnisse benennen, anhand derer die Wirksamkeit von Schule und Unterricht beurteilt werden kann (Ditton, 2000, S. 77; Scheerens, 2000, S. 19). Das bedeutet, eine Schule oder eine Lehrkraft ist umso wirksamer, je mehr die bei den Schüler/-innen entstandenen Lernwirkungen den angestrebten Zielen entsprechen. Im Weiteren erfolgt eine überblicksartige Darstellung von pädagogischen Zielen, die in der Literatur als relevant angesehen werden (Kunter, 2005; Weinert, 2000). Neben der begrifflichen Klärung von Zielsetzungen werden auch knapp Begründungen der Zielsetzungen behandelt. Die Darstellung der Ziele folgt einer Klassifikation anhand von zwei Unterscheidungsdimensionen (Helmke, 2014, S. 38; Kyriakides et al., 2018a, S. 5 ff.). Die erste Dimension liegt auf der Ebene einzelner Schüler/-innen und unterscheidet Lernziele nach inhaltlichen Gesichtspunkten. Die zweite Dimension liegt auf der Ebene von Schülergruppen und unterscheidet den durchschnittlichen Lernerfolg der Schüler/-innen von Unterschieden im Lernerfolg zwischen den Schüler/-innen.

2.1.1 Lernziele nach Persönlichkeitsbereichen

Ein *Lernziel* benennt ein als wünschenswert angesehenes Persönlichkeitsmerkmal (in Anlehnung an Klauer & Leutner, 2012, S. 28). Hierbei wird im Weiteren

ein inhaltlich weit gefasster Begriff von Persönlichkeit verwendet, welcher auch kognitive Merkmale einschließt. Somit werden beispielsweise Wissen, Mathematikkompetenz oder Intelligenz als Persönlichkeitsmerkmale angesehen (Asendorpf, 2007, S. 195 ff.). Lernziele können sich auf verschiedene Bereiche der Persönlichkeit beziehen (z. B. Fachkompetenzen, soziale Kompetenzen, Wertorientierungen). Traditionell steht im schulischen Unterricht die Vermittlung von Fachwissen und Fachkompetenzen im Vordergrund. Die angestrebten Fachkompetenzen umfassen Kompetenzen in mehreren Inhaltsbereichen (z. B. Mathematikkompetenz, Lesekompetenz). Die vorliegende Arbeit befasst sich schwerpunktmäßig mit Mathematikkompetenz, welche daher im Weiteren tiefergehend betrachtet wird.

Mathematisches Wissen bezieht sich auf die Verfügbarkeit von mathematischen Inhalten im deklarativen Langzeitgedächtnis. Hierauf aufbauend beschreibt *Mathematikkompetenz* das Vermögen, mathematisches Wissen flexibel zur Bearbeitung von vielfältigen Problemen anzuwenden (Klieme et al., 2001a, S. 139; Weinert et al., 2011, S. 77). Insofern betont der Kompetenzbegriff (Synonym: Fähigkeit) die Verbindung zwischen Wissen und Können sowie die Befähigung zur Bewältigung unterschiedlicher Lebenssituationen (Klieme, 2004, S. 3). Als Teilaspekte umfasst Mathematikkompetenz unter anderem die Fähigkeit, die mathematische Struktur eines Problems zu erkennen, die Fähigkeit, mathematische Konzepte und Verfahren als Werkzeuge anzuwenden und die Fähigkeit, begründete mathematische Urteile abzugeben (in Anlehnung an Klieme et al., 2001a, S. 141 f.). Mathematikkompetenz kann hierbei in mehrere Inhaltsbereiche unterteilt werden, z. B. Quantitäten, Veränderungen und Verhältnisse, Raum und Formen sowie Daten und Wahrscheinlichkeiten (Weinert et al., 2011, S. 78).

Die *Relevanz des Lernziels Mathematikkompetenz* ergibt sich aus den Anforderungen moderner Gesellschaften. Das Konzept der Grundbildung (Englisch: literacy) betont hierbei die Funktionalität der durch Schulen vermittelten Kompetenzen. Demnach ermöglicht Mathematikkompetenz den Schüler/-innen eine erfolgreiche Lebensführung und eine aktive Teilnahme an der Gesellschaft (Baumert et al., 2001b, S. 16; Weinert et al., 2011, S. 72 f.). Mathematische Probleme stellen sich sowohl im Alltag als auch in vielen Berufsfeldern. Beispiele aus dem Alltag sind Finanzen, Mobilität und Hauswirtschaft. Im beruflichen Kontext spielt Mathematik z. B. in den Naturwissenschaften, in Ingenieursberufen, in der Betriebswirtschaft, im Handwerk, in Gesundheitsberufen und in sozialwissenschaftlichen Berufen eine zentrale Rolle. Personen mit guten Mathematikkompetenzen gelingt die Bewältigung derartiger Anforderungen besser als Personen mit schlechten Mathematikkompetenzen. Vor

diesem Hintergrund ist die Förderung von Mathematikkompetenz schon lange ein bildungspolitisch fest verankertes Ziel von Schule und Unterricht und wurde in jüngerer Zeit durch die Formulierung von Bildungsstandards noch einmal konkretisiert (Kultusministerkonferenz, 2004, S. 13 f.).

Neben der Förderung der Mathematikkompetenz sind weitere Lernziele von Bedeutung für den Mathematikunterricht. Da diese für die vorliegende Arbeit nur eine untergeordnete Rolle spielen, werden sie nachfolgend nur überblicksartig benannt. Weitere fachspezifische Lernziele sind die Förderung von Interesse an Mathematik und die Verhinderung von Angst vor Mathematik (Kunter, 2005; Kunter & Voss, 2011, S. 101). Daneben sind auch fächerübergreifende Lernziele relevant. Zu nennen sind hierbei die Förderung sozialer Kompetenzen und die Vermittlung von Werten (Weinert, 2000). Schon an diesem knappen Überblick wird deutlich, dass eine große Anzahl an inhaltlich unterschiedlichen Lernzielen für die Einschätzung pädagogischer Wirksamkeit herangezogen werden kann.

2.1.2 Durchschnittlicher Lernerfolg und Unterschiede im Lernerfolg

Die Vielfalt potenzieller Zielkriterien wird durch die zweite Unterscheidungsdimension sogar noch erweitert. Für jedes der genannten Lernziele können bei der Betrachtung einer Gruppe von Schüler/-innen (z.B. eine Klasse, eine Schule) wiederum mehrere Zielsetzungen formuliert werden. Dies wird im Weiteren anhand der Mathematikkompetenz veranschaulicht. Eine naheliegende Zielsetzung ist, dass die Schüler/-innen einen hohen *durchschnittlichen Lernerfolg* erreichen, also im Durchschnitt einen hohen Kompetenzzuwachs aufweisen. Dies wird in der deutschsprachigen Literatur als „Qualifizierung“ bezeichnet (Helmke, 1988, S. 45; Schwippert, 2001, S. 34), in der englischsprachigen Literatur als „quality dimension“ (Kyriakides & Creemers, 2011, S. 238 f.; Reynolds et al., 2014, S. 205). Die Betrachtung lediglich des Durchschnitts ist allerdings nicht ausreichend, denn hierbei bleibt unberücksichtigt, wie sich der Lernerfolg zwischen den Schüler/-innen verteilt. Im Extremfall könnte es sein, dass einige Schüler/-innen sehr viel lernen, während andere Schüler/-innen überhaupt nichts lernen.

Vor diesem Hintergrund erscheint es sinnvoll, *Unterschiede im Lernerfolg* zwischen den Schüler/-innen bei der Zielformulierung zu berücksichtigen (Helmke, 2014, S. 38; Kyriakides et al., 2018a, S. 9). Dies wird in der englischsprachigen Literatur als „equity dimension“ bezeichnet (Kyriakides & Creemers, 2011, S. 238 f.; Reynolds et al., 2014, S. 205). In Bezug auf geringe

Unterschiede im Lernerfolg können verschiedene Zielsetzungen angestrebt werden. Eine in der Literatur mehrfach formulierte Zielsetzung bezieht sich darauf, dass Kompetenzunterschiede zwischen Schüler/-innen im Zeitverlauf reduziert werden sollen. Dies wird in der deutschsprachigen Literatur als „Egalisierung“ bezeichnet (Helmke, 1988, S. 45; Schwippert, 2001, S. 34). Diese Zielsetzung bezieht sich auf unsystematische Unterschiede, d. h., es werden Unterschiede zwischen Schüler/-innen jeglicher Art betrachtet. Alternativ können bei der Zielformulierung systematische Unterschiede zwischen Schüler/-innen herangezogen werden, also Unterschiede zwischen bestimmten Gruppen von Schüler/-innen. Diese Gruppen werden in der Regel anhand von äußeren Merkmalen der Schüler/-innen gebildet (z. B. soziale Herkunft, Migrationshintergrund, Geschlecht). Als Ziel lässt sich beispielsweise formulieren, dass Unterschiede im Lernerfolg zwischen Schüler/-innen verschiedener sozialer Herkunft möglichst gering ausfallen sollen. Hierfür wird in der deutschsprachigen Literatur der Begriff „Entkoppelung“ verwendet (Schwippert, 2001, S. 106). Analog können geringe Unterschiede zwischen Mädchen und Jungen sowie geringe Unterschiede zwischen Schüler/-innen mit und ohne Migrationshintergrund als weitere relevante Zielsetzungen angesehen werden.

Eine wesentliche Begründung für Zielsetzungen, die sich auf Unterschiede im Lernerfolg beziehen, liegt in den weitreichenden Konsequenzen erfolgreicher Lernprozesse für Individuen. Damit angesprochen ist der Themenbereich der *Bildungschancen*. Bildungschancen beziehen sich auf individuelle Wahrscheinlichkeiten, Bildungsprozesse erfolgreich abzuschließen (z. B. eine Kompetenz erlernen, einen Schulabschluss erwerben). Eine faire Verteilung von Bildungschancen kann als ein zentraler Aspekt von Bildungsgerechtigkeit angesehen werden (OECD, 2013, S. 27). Eine tiefergehende Erörterung der Frage, was unter fairen Chancen zu verstehen ist, kann im Rahmen der vorliegenden Arbeit nicht geleistet werden. An dieser Stelle genügt es festzuhalten, dass systematische Unterschiede im Lernerfolg der Idealvorstellung gerechter Bildungschancen widersprechen. Ungleiche Bildungschancen nach sozialer Herkunft, Migrationshintergrund oder Geschlecht sind als problematisch anzusehen, weil Schüler/-innen keine Möglichkeit haben, diese äußeren Merkmale zu beeinflussen (Nachbauer & Kyriakides, 2020, S. 308). Dies gilt insbesondere vor dem Hintergrund der Auswirkungen von Bildung auf Gesundheit, Einkommen und Arbeitslosigkeitsrisiko (Autorengruppe Bildungsberichterstattung, 2012, S. 9, 207 ff.; Baumert et al., 2001b, S. 31). In diesem Sinne führen ungleiche Bildungschancen zu ungleichen Lebenschancen.

Im Einklang mit den angebrachten Argumenten findet das Ziel der Entkoppelung des Lernerfolgs von der sozialen Herkunft breite Zustimmung sowohl

in der Bevölkerung (Woessmann et al., 2019, S. 6) als auch in der Politik (Kultusministerkonferenz, 2002, S. 7; Kultusministerkonferenz, 2015, S. 4; Sammons, 2010, S. 56) und in der Wissenschaft (Autorengruppe Bildungsbericht-erstattung, 2012, S. 210; Ditton, 2002, S. 206; Helmke & Schrader, 2010, S. 93).

2.2 Konzeptualisierung pädagogischer Wirksamkeit

Nachdem im letzten Abschnitt ein Verständnis für relevante Zielsetzungen von Schule und Unterricht erarbeitet wurde, erfolgt im Weiteren eine Begriffsbestimmung pädagogischer Wirksamkeit. Zunächst wird eine Definition von pädagogischer Wirksamkeit erarbeitet. Anschließend werden zwei zentrale Perspektiven auf pädagogische Faktoren beleuchtet (pädagogische Einheiten vs. pädagogische Merkmale). Weiterhin wird auf die Rolle der Schüler/-innen eingegangen, mit einem Fokus auf Lernaktivitäten und Lernvoraussetzungen. Dann werden die Begriffe Lernwirkung, Lernerfolg und pädagogischer Effekt geklärt. Abschließend werden das Zusammenwirken pädagogischer Merkmale und verschiedene Dimensionen pädagogischer Wirksamkeit behandelt.

2.2.1 Definition pädagogischer Wirksamkeit

Der Begriff Wirksamkeit beschreibt im Allgemeinen das Ausmaß, in dem eine bestimmte Maßnahme eine erwünschte Wirkung erzeugt. Als synonyme Begriff kann Effektivität angesehen werden. Im pädagogischen Bereich bezieht sich Wirksamkeit auf Lehr-Lern-Prozesse und die Kontexte, in denen Lehr-Lern-Prozesse erfolgen. Die Wirksamkeit von Lehrmethoden kann beispielsweise folgendermaßen definiert werden: „Eine Lehrmethode ist effektiv, wenn sie einen Lernzuwachs, eine Verbesserung der fraglichen Leistung bringt.“ (Klauer & Leutner, 2012, S. 99). Pädagogische Wirksamkeit ist als Begriff allerdings weiter gefasst als die Wirksamkeit von Lehrmethoden, da sich Wirksamkeit auch auf andere pädagogische Faktoren beziehen kann (z. B. strukturelle Bedingungen wie die Klassengröße). Im Rahmen dieser Arbeit wird daher in Anlehnung an die Definition von Klauer und Leutner folgende Definition verwendet: *pädagogische Wirksamkeit liegt vor, wenn ein pädagogischer Faktor bei einem Individuum eine Lernaktivität anregt und dadurch eine Lernwirkung erzeugt*. Ein einfaches grafisches Modell pädagogischer Wirksamkeit liefert Abbildung 1.

Die Definition verdeutlicht, dass pädagogische Wirksamkeit einen *Kausalzusammenhang* beschreibt: ein pädagogischer Faktor ist die Ursache für eine Lernwirkung (Creemers et al., 2010, S. 19; Kyriakides, 2007, S. 46; Scheerens &

Bosker, 1997, S. 26). Hierbei ist hervorzuheben, dass es sich im Falle von pädagogischer Wirksamkeit nicht um deterministische, sondern um *probabilistische Ursache-Wirkungs-Zusammenhänge* handelt (zur Unterscheidung siehe Döring & Bortz, 2016, S. 49; Helmke & Schrader, 2008, S. 28). Demnach erzeugt ein pädagogischer Faktor die entsprechende Lernwirkung nicht immer (d. h. in allen Fällen), sondern in der Regel (d. h. in vielen Fällen). Weiterhin verweist die Definition darauf, dass pädagogische Wirksamkeit einen *Mediationsprozess* impliziert. Pädagogische Faktoren erzeugen Lernwirkungen also nicht direkt, sondern vermittelt über das Anregen von Lernaktivitäten (Helmke, 2014, S. 71; Klieme et al., 2006, S. 131).

Inhaltlich verwandt mit dem Begriff der Wirksamkeit ist der Begriff der Effizienz. In pädagogischen Kontexten bezieht sich *Effizienz* darauf, dass die Wirksamkeit von Lehr-Lern-Prozessen ins Verhältnis zu den aufgewendeten Ressourcen gesetzt wird. Ressourcen beziehen sich auf die benötigte Zeit und die entstehenden Kosten für Personal, Material und Räumlichkeiten (Klauer & Leutner, 2012, S. 99; Reynolds et al., 2014, S. 219). Es wird also betrachtet, welche Lernwirkungen mit welchem Ressourcenaufwand erzeugt werden. Die Berücksichtigung der aufgewendeten Ressourcen erscheint für eine vergleichende Bewertung verschiedener Vorgehensweisen unerlässlich. Werden beispielsweise zwei Unterrichtsmethoden bezüglich des erreichten Lernerfolgs der Schüler/-innen miteinander verglichen, sollte der Lernerfolg ins Verhältnis zur benötigten Zeit gesetzt werden (Klauer & Leutner, 2012, S. 99; Hattie, 2013, S. 11). Entsprechendes gilt für die Aufwendung finanzieller Ressourcen.

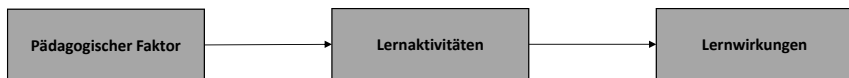


Abb. 1: Theoretisches Modell pädagogischer Wirksamkeit

2.2.2 Pädagogische Faktoren

Im Schulbereich kann der umfassende Begriff der pädagogischen Wirksamkeit in mehrere *Kategorien* unterteilt werden. Zentrale Unterkategorien sind Unterrichtswirksamkeit, Lehrkraftwirksamkeit, Schulwirksamkeit und Systemwirksamkeit (in Anlehnung an Scheerens, 2014, S. 282). Diese Konzepte beziehen sich auf die verschiedenen Ebenen von Schule (Schülerebene, Klassenebene, Schulebene, Systemebene) und spiegeln somit die hierarchische Struktur von Schule wider (d. h., mehrere Schüler/-innen bilden eine Klasse, mehrere Klassen bilden eine Schule, mehrere Schulen bilden ein Schulsystem). Unter-

richtswirksamkeit bezieht sich auf die Wirksamkeit von Unterrichtsprozessen. Ein verwandter Begriff ist Lehrkraftwirksamkeit, welcher sich sowohl auf das Handeln der Lehrkraft im Unterricht (in diesem Fall ist Lehrkraftwirksamkeit gleichbedeutend mit Unterrichtswirksamkeit) als auch auf die Person der Lehrkraft beziehen kann (z. B. pädagogische Überzeugungen). Schulwirksamkeit bezeichnet die Wirksamkeit von einzelnen Schulen und Systemwirksamkeit entsprechend die Wirksamkeit von ganzen Schulsystemen. Im Weiteren wird primär der Begriff „pädagogische Wirksamkeit“ verwendet. Dieses umfassende Konzept hat den Vorteil, dass es eine Theoriebildung ermöglicht, welche die verschiedenen Teilaspekte integriert und ihr komplexes Zusammenwirken berücksichtigt (Creemers & Kyriakides, 2008, S. 3 f.).

Es können zwei Perspektiven auf wirksame pädagogische Faktoren unterschieden werden (in Anlehnung an Helmke, 2014, S. 24 und Helmke & Weichert, 1997a, S. 128): die Wirksamkeit von pädagogischen Einheiten und die Wirksamkeit von pädagogischen Merkmalen. Im Schulkontext sind *pädagogische Einheiten* zusammenhängende Gruppen von Schüler/-innen und pädagogischen Fachkräften wie beispielsweise Klassen, Schulen oder Schulsysteme. Wirksamkeit bezieht sich in diesem Fall darauf, dass bei den Schüler/-innen durch den Besuch einer pädagogischen Einheit Lernwirkungen entstehen. Das Ausmaß der entstandenen Lernwirkungen kann hierbei zwischen verschiedenen pädagogischen Einheiten variieren. Diese Perspektive auf Wirksamkeit findet sich z. B. in den englischen „School Performance Tables“ (Department for Education, 2019), in denen Schulen anhand der Testergebnisse ihrer Schüler/-innen geordnet werden, oder in internationalen Schulleistungsstudien, in denen Rangreihen von nationalen Schulsystemen anhand der erreichten Kompetenzstände der Schüler/-innen gebildet werden (OECD, 2016a, S. 5).

In Bezug auf die Bereitstellung eines empirisch fundierten Orientierungswissens ist diese Perspektive allerdings nicht ausreichend, weil unklar bleibt, aus welchen Gründen bestimmte Einheiten wirksamer sind als andere Einheiten. Hierfür ist die Betrachtung der Wirksamkeit von *pädagogischen Merkmalen* sinnvoll. Ein pädagogisches Merkmal ist eine Eigenschaft von Klassen, Schulen oder Schulsystemen. Es handelt sich um eine Variable, die bei verschiedenen Einheiten unterschiedliche Ausprägungen annehmen kann. Beispiele für pädagogische Merkmale im Schulbereich sind die Häufigkeit von Gruppenarbeiten, die Klassengröße oder das Schulklima. Die Wirksamkeit eines pädagogischen Merkmals bezieht sich darauf, ob dieses Merkmal in einem kausalen Zusammenhang mit den Lernwirkungen steht. Das bedeutet, dass Veränderungen in den Ausprägungen des pädagogischen Merkmals zu Veränderungen in den Lernwirkungen führen. Hierbei sind verschiedene Arten von

Zusammenhängen zwischen pädagogischen Merkmalen und Lernwirkungen zu unterscheiden (Creemers et al., 2010, S. 22; Ditton, 2002, S. 205 f.). Ein linearer Zusammenhang liegt vor, wenn der Lernfortschritt umso größer ausfällt, je stärker das pädagogische Merkmal ausgeprägt ist (z. B. je mehr Gruppenarbeiten, desto höher der Lernfortschritt). Daneben sind auch kurvilineare Zusammenhänge möglich, d. h., der Lernfortschritt fällt bei einer mittleren Ausprägung höher aus als bei einer hohen oder niedrigen Ausprägung (z. B. höchster Lernfortschritt bei mittlerem Anspruchsniveau des Unterrichts, also keine Überforderung und keine Unterforderung). Die Betrachtung pädagogischer Wirksamkeit anhand von pädagogischen Merkmalen findet sich beispielsweise in Zusammenstellungen von Merkmalen wirksamen Unterrichts (Brophy, 2000; Helmke, 2007b) oder Merkmalen wirksamer Schulen (Reynolds & Teddlie, 2000; Sammons et al., 1995). Die Frage nach der Zulässigkeit von Kausalschlussfolgerungen wird in Kapitel 2.3.2 „Forschungsmethoden und -designs der pädagogischen Wirksamkeitsforschung“ ausführlicher behandelt.

2.2.3 Die Rolle der Schüler/-innen

Bei der Betrachtung pädagogischer Faktoren stehen äußere Einflüsse auf die Lernprozesse von Schüler/-innen im Vordergrund. Im Weiteren wird die Rolle der Schüler/-innen erörtert, denen bei schulischen Bildungsprozessen naturgemäß eine wesentliche Bedeutung zukommt. Zunächst ist auf den *aktiven Beitrag von Schüler/-innen beim Lernen* hinzuweisen, welcher insbesondere von kognitiven und konstruktivistischen Lerntheorien betont wird (Mietzel, 2007, S. 42, 212 ff.). Der Erwerb von Fachkompetenzen im Rahmen schulischen Unterrichts hängt wesentlich davon ab, welche Lernaktivitäten ein Schüler/eine Schülerin im Unterricht ausführt, also z. B. ob er/sie bei der Besprechung neuer Inhalte aktiv mitdenkt oder gedanklich abschweift. Dieser Sachverhalt wird durch sogenannte Angebots-Nutzungs-Modelle von Unterricht beschrieben (Helmke, 2014, S. 71; Klieme et al., 2006, S. 131). Demzufolge stellt das unterrichtliche Handeln der Lehrkraft ein Lernangebot dar, welches von den Schüler/-innen durch entsprechende Lernaktivitäten genutzt werden muss, damit Lernwirkungen entstehen. Hieran wird deutlich, dass die Lehrkraft Lernwirkungen nicht eigenständig erzeugen kann. Vielmehr müssen sowohl die Lehrkraft als auch die Schüler/-innen jeweils eigenständige Beiträge leisten. In diesem Sinn stellen Wirkungen des Unterrichts eine „Koproduktion“ von Lehrkraft und Schüler/-innen dar (Fend, 2000, S. 58; Helmke 2007, S. 3). Das Konzept der pädagogischen Wirksamkeit fokussiert hierbei auf den Beitrag der

Lehrkraft, ohne dass die Beiträge der Schüler/-innen dadurch in Frage gestellt werden.

Weiterhin hervorzuheben ist, dass Wirkungszusammenhänge zwischen Lehrkräften und Schüler/-innen nicht nur in eine Richtung, sondern in beide Richtungen bestehen. Das bedeutet, dass Lehrkräfte zwar Schüler/-innen beeinflussen, dass aber gleichzeitig auch *Schüler/-innen Lehrkräfte beeinflussen*. Dieser Sachverhalt wird als reziproke Effekte bezeichnet (Helmke & Schrader, 2010, S. 98). Dies zeigt sich im Rahmen von Lehr-Lern-Prozessen in der Form, dass Schüler/-innen aktiv auf das Handeln von Lehrkräften Einfluss nehmen. Beispielsweise fordern Schüler/-innen bei Verständnisschwierigkeiten gezielt Unterstützung durch die Lehrkraft ein oder bitten die Lehrkraft nach der Bearbeitung einer Aufgabe um Feedback. Neben einer direkten Einflussnahme auf das unterrichtliche Handeln bestehen weitere Einflüsse der Schüler/-innen auf das Erleben der Lehrkraft. Beispielsweise führen massive Unterrichtsstörungen durch die Schüler/-innen zu emotionalen Belastungen bei der Lehrkraft (Thiel, 2016, S. 31). Emotionale Belastungen können sich ihrerseits auf das Unterrichtshandeln der Lehrkraft auswirken.

Schließlich ist darauf hinzuweisen, dass sich Schüler/-innen im Unterricht nicht nur die Rolle der Lernenden einnehmen, sondern auch *Einfluss auf die Lernaktivitäten ihrer Mitschüler/-innen* nehmen. Diese Einflussnahme kann sowohl absichtsvoll erfolgen (z. B. Unterstützung bei Lernproblemen) als auch unbeabsichtigt (z. B. Vorbildverhalten durch aktive Teilnahme am Unterricht). In diesem Sinne umfassen die unterrichtlichen Lernangebote für einen bestimmten Schüler/eine bestimmte Schülerin sowohl das Verhalten der Lehrkraft als auch das Verhalten der Mitschüler/-innen.

2.2.4 Lernaktivitäten und Lernvoraussetzungen

Da die Lernaktivitäten von Schüler/-innen den vermittelnden Prozess zwischen pädagogischen Faktoren und Lernwirkungen darstellen, verdienen diese eine tiefgehende Betrachtung. *Lernen* bezieht sich auf die Veränderung von Persönlichkeitsmerkmalen (z. B. Verhaltensweisen, Wissen, Einstellungen) durch Umwelterfahrungen oder Übung (in Anlehnung an Mietzel, 2007, S. 33 f.). Darauf aufbauend bezeichnen *Lernaktivitäten* kognitive und verhaltensbezogene Aktivitäten eines Individuums, welche mit Lernprozessen einhergehen. Im Kontext pädagogischer Wirksamkeit ist die Betrachtung von Lernen relevant, weil hierdurch erklärt werden kann, aufgrund welcher Mechanismen pädagogische Merkmale wirksam sind. Solche Überlegungen ermöglichen es, theoriegeleitet pädagogische Merkmale zu ermitteln, bei denen lernförderliche

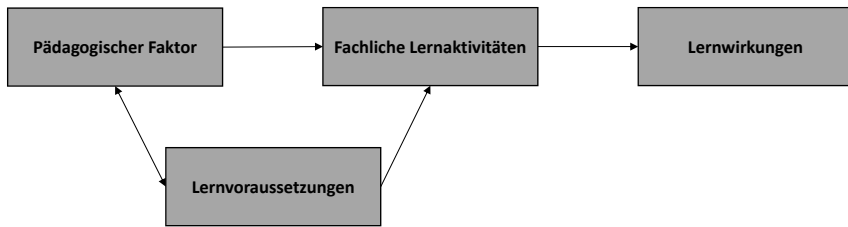


Abb. 2: Theoretisches Modell der Mechanismen pädagogischer Wirksamkeit

Effekte zu erwarten sind (a priori). Weiterhin kann überprüft werden, ob bei einem empirisch vorfindbaren Zusammenhang zwischen einem pädagogischen Merkmal und einer Lernwirkung ein Kausaleffekt plausibel erscheint (a posteriori). Insofern können theoretische Erklärungen Kausalinterpretationen fundieren und sind somit auch relevant für die Beurteilung von Evidenz.

Für die vorliegende Arbeit ist primär von Interesse, durch welche Mechanismen der *Erwerb von Fachkompetenzen* gefördert wird. Grundlegend ist die Gegenüberstellung von direkten Mechanismen, d. h., ein pädagogisches Merkmal wirkt sich direkt auf die fachlichen Lernaktivitäten aus, und indirekten Mechanismen, d. h., ein pädagogisches Merkmal wirkt sich auf Lernvoraussetzungen des Schülers/der Schülerin aus, welche ihrerseits die fachlichen Lernaktivitäten beeinflussen (Helmke, 2014, S. 169). Abbildung 2 gibt einen Überblick über direkte und indirekte Mechanismen beim Erwerb von Fachkompetenzen.

Die *direkten Mechanismen* unterteilen sich in die Förderung der Quantität fachlicher Lernaktivitäten und die Förderung der Qualität fachlicher Lernaktivitäten (Helmke, 2007b, S. 7). Die *Quantität fachlicher Lernaktivitäten* betrifft den zeitlichen Umfang der gedanklichen Beschäftigung mit Lerninhalten. Ein hierauf basierender Mechanismus wäre zum Beispiel, dass eine effiziente Klassenführung bewirkt, dass die Schüler/-innen sich insgesamt mehr Zeit mit den Lerninhalten beschäftigen, was wiederum zu einem höheren Lernerfolg führt (Klieme et al., 2006, S. 131). Für inhaltliche Überlegungen, welche pädagogischen Merkmale die Quantität fachlicher Lernaktivitäten beeinflussen, bieten sich insbesondere Modelle aktiver Lernzeit an (Helmke, 2014, S. 78). Der Quantität kann die *Qualität fachlicher Lernaktivitäten* gegenübergestellt werden, welche sich auf die Art der erfolgenden Denkprozesse bezieht.¹ Beispielsweise kann die Wirksamkeit von kognitiv aktivierenden Aufgabenstellungen damit erklärt werden, dass diese Aufgaben zu einer tiefen Verarbeitung der

1 Der Begriff Qualität wird sowohl deskriptiv, im Sinne von Art und Weise bzw. Beschaffenheit, als auch normativ, im Sinne von Güte verwendet (Helmke, 2014, S. 18). An dieser Stelle ist Qualität im deskriptiven Sinne gemeint.

Lerninhalte führen, was wiederum den Lernerfolg fördert (Klieme et al., 2006, S. 131). Welche pädagogischen Merkmale die Qualität fachlicher Lernaktivitäten beeinflussen, kann insbesondere aus kognitiven und konstruktivistischen Lerntheorien abgeleitet werden (z. B. Craik & Lockhart, 1972; Sweller, 1988).

Die Gruppe der *indirekten Mechanismen* bezieht sich auf die Förderung verschiedener Lernvoraussetzungen der Schüler/-innen. *Lernvoraussetzungen* bezeichnen psychologische Merkmale von Schüler/-innen, welche einen Einfluss auf ihren Lernerfolg haben. Von zentraler Bedeutung ist hierbei die *Lernmotivation* der Schüler/-innen, also die Bereitschaft zum Ausführen von Lernaktivitäten. Die Wirksamkeit eines pädagogischen Merkmals für fachliches Lernen kann demnach darauf zurückzuführen sein, dass es zu einer höheren Lernmotivation der Schüler/-innen führt, welche ihrerseits zu einer höheren Quantität und/oder Qualität fachlicher Lernaktivitäten führt. Wenn Schüler/-innen im Unterricht beispielsweise die Möglichkeit erhalten, die Nützlichkeit der Lerninhalte zu verstehen, fördert dies das Interesse an den Lerninhalten und dadurch auch den Kompetenzerwerb (Hulleman & Harackiewicz, 2009). Die Bestimmung von pädagogischen Merkmalen, welche über den Mechanismus der Förderung von Lernmotivation wirken, ist mithilfe von Theorien der Lernmotivation möglich (Spinath, 2015). In ähnlicher Weise können sich pädagogische Merkmale auf den Erwerb von Fachkompetenzen auswirken, weil sie andere relevante Lernvoraussetzungen der Schüler/-innen fördern, beispielsweise *Sprachkompetenzen* oder *metakognitive Kompetenzen* (diese Lernvoraussetzungen werden in Kapitel 3.11 „Unterschiede bei Lernvoraussetzungen“ näher erläutert).

2.2.5 Lernwirkung, Lernerfolg und pädagogischer Effekt

Im Weiteren werden die Begriffe Lernwirkung, Lernerfolg und pädagogischer Effekt näher erläutert. Eine *Lernwirkung* ist das Resultat eines Lernprozesses. Entsprechend bezeichnet der Begriff Lernwirkung (Synonyme: Lerneffekt, Lernergebnis, Lernfortschritt) den Unterschied zwischen der ursprünglichen Ausprägung eines Persönlichkeitsmerkmals und der durch das Lernen entstandenen neuen Ausprägung dieses Persönlichkeitsmerkmals. Ein verwandter Begriff ist *Lernerfolg*, welcher allerdings eine Bewertung impliziert und sich somit auf eine als wünschenswert angesehene Lernwirkung bezieht. Eine Lernwirkung, die durch einen bestimmten pädagogischen Faktor erzeugt wurde, kann als *pädagogischer Effekt* (Synonym: pädagogische Wirkung) bezeichnet werden (in Anlehnung an Brezinka, 1995, S. 197 ff.).

Bei der Betrachtung von pädagogischen Effekten ist wiederum die Unterscheidung von pädagogischen Einheiten und pädagogischen Merkmalen sinnvoll. Der *Effekt einer pädagogischen Einheit* besteht in den Lernwirkungen, die bei den Lernenden durch den Besuch dieser Einheit entstehen. Beispielsweise ist ein Effekt einer Schule der bei den Schüler/-innen angeregte Zuwachs der Mathematikkompetenzen im Laufe eines Schuljahres. Hierbei können absolute Effekte und relative Effekte unterschieden werden (Reynolds et al., 2014, S. 206). Im genannten Beispiel wäre der *absolute Effekt* der Schule der Unterschied zwischen dem durch die Schule angeregten Zuwachs der Mathematikkompetenzen und dem Zuwachs der Mathematikkompetenzen, der erfolgt wäre, wenn die Schüler/-innen keine Schule besucht hätten. Absolute Schuleffekte beziehen sich also darauf, welche Bedeutung es für den Lernerfolg von Schüler/-innen macht, *ob sie eine Schule besuchen oder nicht*. Absolute Effekte sind im Schulbereich aufgrund der gesetzlichen Schulpflicht häufig nicht ermittelbar. Daher wird bei Schulen in der Regel der *relative Effekt* betrachtet (Reynolds et al., 2014, S. 206). Im obigen Beispiel wäre der relative Effekt der Schule der Unterschied zwischen dem durch die Schule angeregten Zuwachs der Mathematikkompetenzen und dem Zuwachs der Mathematikkompetenzen, der erfolgt wäre, wenn die Schüler/-innen eine andere Schule besucht hätten. Relative Schuleffekte beziehen sich also darauf, welche Bedeutung es für den Lernerfolg von Schüler/-innen macht, *welche Schule sie besuchen* (Creemers et al., 2010, S. 7). Schuleffekte können auch in Bezug auf Unterschiede im Lernerfolg betrachtet werden. Bei der Betrachtung von Schüler/-innen verschiedener sozialer Herkunft beziehen sich relative Schuleffekte beispielsweise darauf, inwieweit Kompetenzunterschiede nach sozialer Herkunft in verschiedenen Schulen unterschiedlich stark ausfallen.

Bei der Betrachtung von pädagogischen Merkmalen sind Effekte inhaltlich anders zu bestimmen. Der *Effekt eines pädagogischen Merkmals* besteht im Unterschied zwischen den Lernwirkungen, die sich bei verschiedenen Ausprägungen dieses Merkmals ergeben. Beispielsweise kann der Zuwachs der Mathematikkompetenzen der Schüler/-innen bei häufiger Verwendung von Gruppenarbeiten und bei seltener Verwendung von Gruppenarbeiten miteinander verglichen werden. Neben dem Effekt auf den durchschnittlichen Lernerfolg kann auch der Effekt auf Unterschiede im Lernerfolg betrachtet werden. Im genannten Beispiel wäre ein Effekt von Gruppenarbeiten gegeben, wenn Kompetenzunterschiede nach sozialer Herkunft bei häufiger Verwendung von Gruppenarbeiten geringer ausfallen als bei seltener Verwendung von Gruppenarbeiten.

2.2.6 Zusammenwirken pädagogischer Merkmale

Da pädagogische Merkmale Eigenschaften von Klassen/Schulen darstellen und jede Klasse/Schule über mehrere Eigenschaften verfügt, wirken pädagogische Merkmale nicht unabhängig voneinander, sondern in komplexer Weise zusammen. Dies wird im Weiteren ausgeführt. Relevant ist zunächst die Unterscheidung zwischen Prozessmerkmalen und Strukturmerkmalen (Bäumer et al., 2011, S. 94 f.; Ditton, 2000, S. 76; Scheerens et al., 2007, S. 27 f.). Inhaltlich verwandt ist die Gegenüberstellung von proximalen Merkmalen und distalen Merkmalen (Helmke, 2014, S. 29). *Prozessmerkmale* beziehen sich auf das Verhalten von pädagogischen Fachkräften. Sie sind zeitlich relativ variabel, d. h., sie können sich in relativ kurzen Zeitperioden bedeutsam verändern. Typischerweise haben Prozessmerkmale *direkte Effekte*, d. h., sie beeinflussen die Lernprozesse von Schüler/-innen unmittelbar. Als Beispiele können Unterrichtsmerkmale wie Klassenführung oder kognitive Aktivierung genannt werden. Demgegenüber handelt es sich bei *Strukturmerkmalen* um zeitlich relativ stabile Merkmale. Hierzu zählen Merkmale von pädagogischen Fachkräften (z. B. pädagogische Überzeugungen), Merkmale von Lerngruppen (z. B. Klassengröße) und räumlich-materielle Bedingungen (z. B. Lernmaterial). Strukturmerkmale sind durch *indirekte Effekte* gekennzeichnet, d. h., sie haben keinen direkten Einfluss auf die Lernprozesse der Schüler/-innen, sondern sie wirken vermittelt über Prozessmerkmale (z. B. beeinflussen die pädagogischen Überzeugungen der Lehrkraft ihre Unterrichtsgestaltung). Im statistischen Sinne liegt eine Mediation vor (Baron & Kenny, 1986). Abbildung 3 veranschaulicht das Zusammenwirken von Struktur- und Prozessmerkmalen.

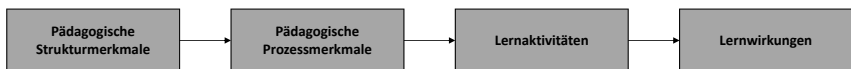


Abb. 3: Theoretisches Modell des Zusammenwirkens von Struktur- und Prozessmerkmalen

Indirekte Effekte stellen somit eine zweite Wirkweise pädagogischer Merkmale dar: Strukturmerkmale sind wirksam, weil sie einen günstigen Effekt auf Prozessmerkmale haben. Bezüglich der Art und Weise, wie Strukturmerkmale das pädagogische Handeln von Fachkräften bedingen, sind insbesondere zwei Mechanismen zu unterscheiden. Der erste Mechanismus besteht darin, dass Strukturmerkmale *Handlungsmöglichkeiten* beeinflussen, d. h., es werden neue Handlungsoptionen eröffnet. Dieser Mechanismus bezieht sich auf den Aspekt des „Könnens“. Beispielsweise führt das Wissen der Lehrkraft über Klassenführung dazu, dass sie bestimmte Techniken einsetzen kann (Piowar et al., 2013, S. 7). Dass Handlungsoptionen zur Verfügung stehen, heißt allerdings

nicht zwangsläufig, dass diese auch genutzt werden. Hierfür spielt der zweite Mechanismus eine wesentliche Rolle. Dieser bezieht sich darauf, dass Strukturmerkmale *Handlungsbereitschaften* beeinflussen, d.h., Präferenzen für bestimmte Handlungsoptionen verstärken oder abschwächen. Dies betrifft den Aspekt des „Wollens“. Ob eine Lehrkraft beispielsweise die ihr zur Verfügung stehenden Techniken der Klassenführung tatsächlich anwendet, hängt davon ab, welche lerntheoretischen Überzeugungen sie hat (Kunter et al., 2013, S. 814).

Bezüglich des Zusammenwirkens pädagogischer Merkmale ist weiterhin zwischen Haupteffekten und Interaktionseffekten zu unterscheiden (Ditton, 2002, S. 205 f.; Helmke, 2014, S. 31). *Haupteffekte* liegen vor, wenn zwei Merkmale jeweils eigenständige Effekte haben und somit jeweils unabhängig vom anderen Merkmal wirken. Demgegenüber sind *Interaktionseffekte* gegeben, wenn sich die Effekte zweier Merkmale gegenseitig bedingen (Caro et al., 2016, S. 33). Im Sinne einer statistischen Moderation (Baron & Kenny, 1986) hängt der Effekt eines Merkmals von der Ausprägung des anderen Merkmals ab. Ein Interaktionseffekt zeigt sich beispielsweise im folgenden Befund: Strukturierungshilfen der Lehrkraft, also unterstützende Maßnahmen wie Hinweise und Vorschläge, haben nur dann einen förderlichen Effekt auf den Lernerfolg der Schüler/-innen, wenn die Lehrkraft über eine hohe diagnostische Kompetenz verfügt (Weinert & Helmke 1987, S. 27 f.).

2.2.7 Dimensionen pädagogischer Wirksamkeit

Als ein zentrales Charakteristikum pädagogischer Wirksamkeit ist die Komplexität dieses Konzepts zu nennen. Mit Komplexität ist konkret gemeint, dass sich Wirksamkeit auf unterschiedliche Aspekte beziehen und zwischen diesen Aspekten variieren kann (Creemers et al., 2010, S. 24 ff.; Helmke, 2014, S. 83; Muijs et al., 2005, S. 52 f.; Sammons, 2010, S. 55; Scheerens & Bosker, 1997, S. 80 ff.). Zentrale Dimensionen pädagogischer Wirksamkeit sind Zielsetzungen, Zeiträume, Gruppen von Schüler/-innen und institutionelle Kontexte. Diese Dimensionen gelten sowohl für pädagogische Einheiten als auch für pädagogische Merkmale, werden im Weiteren aber anhand von pädagogischen Merkmalen veranschaulicht.

Es wurde bereits ausgeführt, dass die *Zielsetzungen von Schule und Unterricht* den normativen Maßstab für die Einschätzung von Wirksamkeit darstellen. Die Wirksamkeit eines pädagogischen Faktors kann in Abhängigkeit von der betrachteten Zielsetzung variieren. Beispielsweise fördert kognitive Aktivierung zwar die Entwicklung der Mathematikkompetenz, nicht aber die Entwicklung der Lernfreude (Kunter & Voss, 2011, S. 104 f.).

Zweitens kann pädagogische Wirksamkeit vom betrachteten *Zeitraum* abhängen. Hierbei ist der Unterschied zwischen kurzfristigen und langfristigen Lernwirkungen zentral. Beispielsweise zeigen Ziegler und Stern (2014), dass eine geordnete Vorgehensweise im Mathematikunterricht (zuerst Additionsaufgaben, dann Multiplikationsaufgaben) kurzfristig zu besseren Kompetenzen führt als eine verschachtelte Vorgehensweise (abwechselnd Additionsaufgaben und Multiplikationsaufgaben), die verschachtelte Vorgehensweise aber langfristig zu besseren Kompetenzen führt (S. 142 f.).

Drittens kann die Wirksamkeit eines pädagogischen Faktors bei verschiedenen *Gruppen von Schüler/-innen* unterschiedlich ausfallen, was auch als differenzielle Wirksamkeit bezeichnet wird (Kyriakides, 2004). Die Gegenüberstellung von Gruppen kann sich entweder auf Schüler/-innen mit verschiedenen Lernvoraussetzungen (z. B. Intelligenz, Fachinteresse) oder auf Schüler/-innen mit verschiedenen äußeren Merkmalen (z. B. Geschlecht, soziale Herkunft, Migrationshintergrund) beziehen. Beispielsweise hat beim entdeckenden Lernen die Lernsteuerung durch die Lehrkraft einen förderlichen Effekt auf den Lernerfolg von Schüler/-innen mit ungünstigen Lernvoraussetzungen, nicht aber auf den Lernerfolg von Schüler/-innen mit günstigen Lernvoraussetzungen (Möller et al., 2002, S. 183). Differenzielle Wirksamkeit ist inhaltlich eng verbunden mit Wirksamkeit bezüglich geringer Unterschiede im Lernerfolg (Kyriakides, 2004, S. 144 f.), weil sich beide Sachverhalte auf die Gegenüberstellung von Gruppen von Schüler/-innen beziehen. Dies wird im fünften Kapitel genauer beleuchtet (siehe Kapitel 5.2.2 „Differenziell wirksame Lerngelegenheiten und unterschiedliche Lerngelegenheiten“).

Schließlich kann pädagogische Wirksamkeit zwischen verschiedenen *Kontexten* variieren. Kontexte beziehen sich auf den äußeren Rahmen von Schulen, wozu die Schulform, die regionale Lage (z. B. Stadt vs. Land) oder das Land zählen. Ein Befund ist beispielsweise, dass persönliche Beziehungen zwischen Lehrkräften und Schüler/-innen in einigen Ländern positive Effekte auf den Lernerfolg haben, in anderen Ländern dagegen negative Effekte (Caro et al., 2016, S. 35).

An den dargestellten Befunden wird deutlich, dass die Einschätzung der Wirksamkeit eines pädagogischen Faktors sehr unterschiedlich ausfallen kann in Abhängigkeit davon, welche Aspekte betrachtet werden. Insofern erscheint es als eine angemessene Perspektive, einen pädagogischen Faktor nicht durch *eine Wirksamkeit* zu charakterisieren, sondern durch *mehrere Wirksamkeiten* (z. B. Wirksamkeit für Lernziel A vs. Lernziel B, Wirksamkeit für Schülergruppe X vs. Schülergruppe Y). Bei pädagogischen Merkmalen kann es hierbei auch zu Konflikten zwischen verschiedenen Aspekten kommen. Das ist der Fall,

wenn eine bestimmte Ausprägung für eine Zielsetzung oder eine Gruppe von Schüler/-innen günstig, aber für eine andere Zielsetzung oder eine Gruppe von Schüler/-innen ungünstig ist. Beispielsweise berichten Weinert und Helmke (1987), dass eine intensive Unterstützung durch die Lehrkräfte die Kompetenzentwicklung der Schüler/-innen fördert, aber die Entwicklung der Lernfreude beeinträchtigt (S. 26 ff.).

2.3 Pädagogische Wirksamkeitsforschung

Nachdem im vorherigen Abschnitt pädagogische Wirksamkeit begrifflich bestimmt wurde, wird in diesem Abschnitt ein Überblick über Ansätze zur Erforschung pädagogischer Wirksamkeit gegeben. Im Vordergrund steht die Frage, wie pädagogische Wirksamkeit empirisch untersucht wird. Dies ermöglicht die Einordnung der vorliegenden Arbeit in einen größeren wissenschaftlichen Kontext. Weiterhin wird ein Verständnis dafür vermittelt, wie die in den nachfolgenden Kapiteln berichteten Befunde durch empirische Untersuchungen gewonnen wurden. Außerdem stellen die Ausführungen die Grundlage für die empirischen Analysen der vorliegenden Arbeit dar.

Im Schulbereich sind zwei Entwicklungslinien der Erforschung pädagogischer Wirksamkeit auszumachen, nämlich die Unterrichtswirksamkeitsforschung und die Schulwirksamkeitsforschung (Chapman et al., 2016b, S. 1 ff.). In beiden Feldern begann die systematische Forschung in den 1970er Jahren, insbesondere in Nordamerika und Großbritannien (Brookover et al., 1979; Rutter et al., 1979). In den folgenden Jahrzehnten kam es in zahlreichen Ländern zu umfangreichen Forschungsaktivitäten. Hierdurch wurden sowohl eine breite Wissensbasis erarbeitet als auch vielfältige Entwicklungen bezüglich Theorie und Methodologie angestoßen (Reynolds et al., 2014, S. 198). Wenngleich historisch Unterschiede in Fragestellungen und Forschungsmethoden bestanden, sind die Unterrichts- und Schulwirksamkeitsforschung inzwischen so eng miteinander verbunden, dass sie auch integrativ unter dem Begriff pädagogische Wirksamkeitsforschung (Englisch: educational effectiveness research) zusammengefasst werden. Die pädagogische Wirksamkeitsforschung kann ihrerseits als ein Teilgebiet der empirischen Bildungsforschung angesehen werden (Reinders et al., 2015).

Nachfolgend werden zunächst die zentralen Ziele und Fragestellungen der pädagogischen Wirksamkeitsforschung dargestellt. Danach werden typische Forschungsmethoden und -designs behandelt. Abschließend wird auf Standards für Wirksamkeitsstudien eingegangen.

2.3.1 Ziele und Fragestellungen der pädagogischen Wirksamkeitsforschung

Das primäre Ziel der pädagogischen Wirksamkeitsforschung, wie auch der empirischen Bildungsforschung im Allgemeinen, liegt darin, wissenschaftliche Erkenntnisse zu gewinnen, durch die das Bildungssystem verbessert werden kann (Gräsel, 2015, S. 16; Chapman et al., 2016b, S. 3). Die gewonnenen Erkenntnisse werden den Entscheidungsträger/-innen im Bildungswesen (z. B. Lehrkräfte, Schulleiter/-innen, Bildungspolitiker/-innen) in der Form eines *empirisch fundierten Orientierungswissens* zur Verfügung gestellt. Hierdurch werden pädagogische Fachkräfte und Bildungspolitiker/-innen in die Lage versetzt, relevante Entscheidungen auf der Grundlage empirischer Befunde über die Effekte von Maßnahmen zu fällen. Diese Vorgehensweise ist ein zentrales Merkmal einer *evidenzbasierten pädagogischen Praxis und Bildungspolitik* (Bromme et al., 2016, S. 131f.; Tippelt & Reich-Claassen, 2010, S. 22f.). Eine evidenzbasierte Entscheidungsfindung kann von einer intuitiven Entscheidungsfindung (z. B. Bauchgefühl) und einer theoretischen Entscheidungsfindung (z. B. Abwägen von theoretischen Argumenten) abgegrenzt werden. Der Befürwortung einer evidenzbasierten Entscheidungsfindung liegt die Annahme zugrunde, dass diese Vorgehensweise zu einem größeren Lernerfolg der Schüler/-innen führt als andere Arten der Entscheidungsfindung. Gestützt wird diese Annahme durch Schulentwicklungsansätze, welche Erkenntnisse der pädagogischen Wirksamkeitsforschung umsetzen und nachweislich den Lernerfolg der Schüler/-innen erhöhen (Creemers & Kyriakides, 2012; Slavin et al., 2009).

Eine notwendige Voraussetzung für eine evidenzbasierte Entscheidungsfindung ist, dass für die Wirksamkeit der in Frage kommenden Maßnahmen ausreichende Evidenzen vorliegen. Im Kontext pädagogischer Wirksamkeit bezeichnet *Evidenz* das Ausmaß, in dem die Annahme, dass ein pädagogischer Faktor bestimmte Wirkungen erzeugt, durch empirische Befunde belegt werden kann (in Anlehnung an Bromme et al., 2016, S. 132). Evidenz verweist auf die interne Validität eines Befunds, also auf die Zulässigkeit der Schlussfolgerung eines Ursache-Wirkungs-Zusammenhangs auf der Grundlage einer Korrelation. Die vorliegende Evidenz hängt sowohl von der Anzahl der verfügbaren Studien als auch von der Qualität der verfügbaren Studien ab (siehe Kapitel 2.3.3 „Standards für Wirksamkeitsstudien“). Einschätzungen von vorliegenden Evidenzen sind anspruchsvolle Aufgaben, die eine fundierte Kenntnis des Forschungsstandes und vertieftes forschungsmethodologisches Wissen erfordern. In diesem Sinne können Einschätzungen von Evidenz als unver-

zichtbarer Bestandteil des durch die pädagogische Wirksamkeitsforschung zur Verfügung gestellten Orientierungswissens angesehen werden.

Zur Gewinnung relevanten Orientierungswissens fokussiert die schulbezogene Wirksamkeitsforschung vorrangig auf Fragestellungen in drei Bereichen. Von grundlegendem Interesse ist zunächst einmal die *Größe der Effekte von pädagogischen Einheiten*. Im Vordergrund steht die rein beschreibende Frage, wie groß Unterschiede in Lernwirkungen zwischen verschiedenen Klassen oder Schulen ausfallen (Ditton, 2011, S. 247; Scheerens, 2000, S. 18). Von Interesse ist hierbei auch die direkte Gegenüberstellung von Unterschieden zwischen Klassen und Unterschieden zwischen Schulen, da hieran erkennbar wird, welche relative Bedeutung die verschiedenen Ebenen jeweils für den Lernerfolg von Schüler/-innen haben.

Der zweite Bereich stellt den eigentlichen Kern der pädagogischen Wirksamkeitsforschung dar. Hierbei geht es um Erklärungen, warum der Lernerfolg der Schüler/-innen in bestimmten Klassen/Schulen höher ausfällt als in anderen Klassen/Schulen (Creemers & Kyriakides, 2008, S. 13; Scheerens, 2000, S. 19). Diese Perspektive fokussiert auf die *Wirksamkeit pädagogischer Merkmale*. Grundlegend ist die Frage, hinsichtlich welcher pädagogischer Merkmale sich hoch wirksame Klassen/Schulen von niedrig wirksamen Klassen/Schulen unterscheiden (Ditton, 2011, S. 247). Die Bearbeitung dieser Frage ist relativ einfach möglich, da sie sich inhaltlich auf Korrelationen bezieht. Die interessantere und unter dem Gesichtspunkt des Orientierungswissens relevantere Frage lautet dagegen, welche Merkmale von Klassen/Schulen wirksam sind, d.h., welche Merkmale die Ursachen für die entstandenen Lernwirkungen sind. Die Bearbeitung der letztgenannten Frage ist methodologisch allerdings deutlich herausfordernder, da eine Überprüfung von Kausalbeziehungen notwendig ist.

Ein dritter Fragenkomplex behandelt die *Konsistenz pädagogischer Effekte* über verschiedene Zielkriterien hinweg. Im Vordergrund steht die Frage, ob Klassen/Schulen, die hinsichtlich Zielsetzung A günstige Lernwirkungen erzeugen, auch hinsichtlich Zielsetzung B günstige Lernwirkungen erzeugen. Dies erlaubt Schlussfolgerungen über das Verhältnis zwischen verschiedenen Zielsetzungen, also darüber, ob verschiedene Zielsetzungen vereinbar, unabhängig oder konfligierend sind (Reynolds et al., 2014, S. 207). Eine positive Korrelation zwischen Lernwirkungen spricht für vereinbare Ziele (d.h., die Erreichung von Ziel A begünstigt die Erreichung von Ziel B), eine Nullkorrelation für unabhängige Ziele (d.h., die Erreichung von Ziel A hat keinen Einfluss auf die Erreichung von Ziel B) und eine negative Korrelation für konfligierende Ziele (d.h., die Erreichung von Ziel A beeinträchtigt die Erreichung von

Ziel B). Untersucht werden sowohl die Vereinbarkeit verschiedener Lernziele wie Fachkompetenzen und Lernfreude (z. B. Gruehn, 1995) als auch die Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen Unterschieden im Lernerfolg (z. B. Kyriakides & Creemers, 2011).

2.3.2 Forschungsmethoden und -designs der pädagogischen Wirksamkeitsforschung

Die pädagogische Wirksamkeitsforschung ist ein empirisch orientierter Forschungsansatz und basiert somit auf der systematischen und an wissenschaftlichen Gütekriterien orientierten Erhebung und Auswertung von empirischen Informationen. Erhebungen im Rahmen von Wirksamkeitsstudien beziehen sich einerseits auf den Bereich der pädagogischen Merkmale und andererseits auf den Bereich des Lernens. Bezüglich der *Erhebung von Lernen* gilt, dass Lernprozesse nicht direkt gemessen werden können, sondern nur anhand von Lernwirkungen nachvollzogen werden können. Aussagekräftig sind insbesondere längsschnittliche Erhebungen, mit denen Veränderungen von Schülermerkmalen erfasst werden können (z. B. Mathematikkompetenz am Anfang und am Ende des Schuljahres). Schülermerkmale im kognitiven Bereich werden in der Regel mit standardisierten Leistungstests erfasst, Schülermerkmale in anderen Bereichen (z. B. Lernmotivation, Sozialverhalten) typischerweise durch Befragungen von Schüler/-innen (Baumert & Köller, 1998, S. 14 ff.).

Für die *Erhebung von pädagogischen Merkmalen* kommen neben Befragungen von Lehrkräften, Schulleiter/-innen oder Schüler/-innen auch Beobachtungen und Dokumentenanalysen zum Einsatz (Gärtner et al., 2022, S. 148; Helmke, 2014, S. 272 ff.; Kunter & Voss, 2011, S. 90 ff.). Diese Erhebungsmethoden haben jeweils spezifische Vor- und Nachteile, wie am Beispiel der Messung von Unterrichtsmerkmalen überblicksartig veranschaulicht werden soll (Göllner et al., 2016, S. 63 ff.; Lüdtke et al., 2009, S. 120; Praetorius, 2013, S. 176 ff.). Die Stärken der Lehrkräftebefragung liegen darin, dass große Verhaltensstichproben erfasst werden können (Trends über mehrere Unterrichtsstunden hinweg) und dass ein geringer Aufwand notwendig ist. Allerdings besteht die Gefahr, dass Lehrkräfte ihre Antworten (bewusst oder unbewusst) verzerren, beispielsweise als Selbstwertschutz oder aufgrund von Verzerrung nach sozialer Erwünschtheit. Befragungen von Schüler/-innen können ebenfalls große Verhaltensstichproben erfassen und sind mit geringem Aufwand verbunden. Zusätzlich wird eine große Informationsmenge erzeugt, da in der Regel jeweils die gesamte Klasse den Unterricht einschätzt. Allerdings besteht bei Schüler/-innen die Gefahr einer Verzerrung durch Sympathie bzw. Antipathie zur

Lehrkraft. Beobachtungen haben die Vorteile, dass Beobachter/-innen eine neutrale Außenperspektive einnehmen und alle verfügbaren kognitiven Ressourcen für die Einschätzung des Unterrichts aufwenden können. Allerdings kann nur eine kleine Verhaltensstichprobe (eine Unterrichtsstunde) erfasst werden und der Aufwand ist sehr hoch.

Innerhalb der pädagogischen Wirksamkeitsforschung dominieren *quantitative Forschungsansätze* (Chapman et al., 2016b, S. 2). Das bedeutet, die Auswertung empirischer Informationen erfolgt vorrangig mithilfe statistischer Verfahren. Die statistischen Analysemethoden wurden in der noch jungen Geschichte des Forschungsfeldes bereits bedeutsam weiterentwickelt und gegenwärtig steht eine ganze Reihe unterschiedlicher Verfahren zur Verfügung (Creemers et al., 2010). Von besonderer Bedeutung sind hierbei Mehrebenenmodelle und latente Variablenmodelle (Muijs, 2012, S. 61). Eine Darstellung ausgewählter Verfahren erfolgt in Kapitel 7.3 „Analysen“.

Bezüglich der verwendeten Forschungsdesigns ist festzuhalten, dass für Wirksamkeitsstudien sowohl korrelative Designs (z. B. Klieme et al., 2006; Kunter & Voss, 2011) als auch experimentelle Designs (z. B. Hulleman & Harackiewicz, 2009; Tuovinen & Sweller, 1999) verwendet werden. Beide Designs weisen spezifische Vor- und Nachteile auf. Bei *korrelativen Studien* werden die natürlicherweise vorkommenden Unterschiede in den Merkmalen von Klassen/Schulen und dem Lernerfolg der Schüler/-innen analysiert. Eine zentrale Herausforderung ist die *Ermittlung der pädagogischen Effekte* von Klassen/Schulen, da sich Klassen/Schulen in den Ausgangsbedingungen ihrer Schüler/-innen unterscheiden. *Ausgangsbedingungen* von Schüler/-innen beziehen sich auf individuelle Einflussfaktoren auf den Lernerfolg, die nicht zur Schule gehören (z. B. Intelligenz der Schüler/-innen, Lernunterstützung durch die Eltern). Da eine Angleichung der Ausgangsbedingungen nicht möglich ist, erfolgt in Korrelationsstudien eine Adjustierung mithilfe statistischer Verfahren. Hierfür haben sich sogenannte *Mehrwertmodelle* (Englisch: value added model) etabliert, durch die der spezifisch durch Klassen/Schulen geleistete Beitrag zum Lernerfolg der Schüler/-innen geschätzt wird (Sammons, 1999, 26 f.). Bei einem Mehrwertmodell werden relevante Ausgangsbedingungen der Schüler/-innen erfasst und als Kontrollvariablen berücksichtigt. Typische Kontrollvariablen sind der Ausgangswert des betrachteten Lernziels (z. B. Mathematikkompetenz am Anfang des Schuljahres), Lernvoraussetzungen (z. B. Intelligenz) und familiäre Merkmale (z. B. sozialer Status der Eltern) (Scheerens & Bosker, 1997, S. 53 ff.).

Die *Ermittlung wirksamer pädagogischer Einheiten* erfolgt anhand eines sozialen Maßstabs, d. h., Klassen/Schulen werden hinsichtlich der bei ihren

Schüler/-innen entstandenen Lernwirkungen miteinander verglichen (Reynolds et al., 2014, S. 205; Sammons et al., 1995, S. 7). Dieser Ansatz basiert somit auf der Betrachtung relativer pädagogischer Effekte. Durch Mehrwertmodelle werden hierbei faire Vergleiche gewährleistet, d. h., Ähnliches wird mit Ähnlichem verglichen (Sammons, 1999, S. 26; Scheerens, 2000, S. 18). Demnach werden Klassen/Schulen als *hoch wirksam* angesehen, wenn die Schüler/-innen einen höheren Lernerfolg aufweisen als Schüler/-innen in Klassen/Schulen mit ähnlichen Ausgangsbedingungen. *Niedrig wirksame* Klassen/Schulen sind entsprechend dadurch gekennzeichnet, dass die Schüler/-innen einen niedrigeren Lernerfolg aufweisen als Schüler/-innen in Klassen/Schulen mit ähnlichen Ausgangsbedingungen (in Anlehnung an Sammons et al., 1995, S. 7).

Basierend auf den ermittelten Effekten von Klassen/Schulen kann untersucht werden, hinsichtlich welcher *Merkmale* sich hoch wirksame Klassen/Schulen und niedrig wirksame Klassen/Schulen unterscheiden. Im Vordergrund stehen hierbei statistische Zusammenhänge zwischen pädagogischen Merkmalen und Lernwirkungen. Schlussfolgerungen von Merkmalen wirksamer Klassen/Schulen auf wirksame Merkmale (in dem Sinne, dass diese Merkmale die Ursache für den hohen Lernerfolg der Schüler/-innen darstellen) sind allerdings nicht ohne Weiteres zulässig. Das liegt daran, dass bei korrelativen Studien eine große Vielzahl an Wirkzusammenhängen möglich ist. Beispielsweise unterscheiden sich hoch und niedrig wirksame Klassen/Schulen in der Regel hinsichtlich vieler Merkmale, von denen einige tatsächlich kausal wirken, andere dagegen nur korrelieren (Creemers et al., 2010, S. 40 f.). Weiterhin können umgekehrte Wirkrichtungen vorliegen, d. h., Merkmale von Schüler/-innen beeinflussen pädagogische Merkmale (Scheerens & Bosker, 1997, S. 28). Mithilfe angemessener statistischer Verfahren können komplexe Wirkzusammenhänge dieser Art überprüft werden und dadurch die interne Validität erhöht werden. Außerdem spielt bei Kausalinterpretationen im Rahmen von korrelativen Studien Theorie eine bedeutende Rolle (Creemers et al., 2010, S. 53; Goldstein, 1997, S. 376). Von Bedeutung ist insbesondere eine theoretische Erklärung für die Wirksamkeit eines pädagogischen Merkmals. Nur wenn sich plausible Mechanismen angeben lassen, auf welche Art und Weise ein pädagogisches Merkmal das Zielkriterium beeinflusst, kann in Betracht gezogen werden, dass es sich tatsächlich um einen Kausalzusammenhang handelt. Während theoretisch und methodisch angemessene Korrelationsstudien zumindest fundierte Hinweise auf Kausaleffekte liefern können, ist ein eindeutiger Nachweis von Ursache-Wirkungs-Beziehungen mit diesem Forschungsdesign nicht möglich (Scheerens & Bosker, 1997, S. 28).

Hierfür werden *experimentelle Studien* benötigt. Bei diesem Forschungsdesign werden Schüler/-innen nach dem Zufallsprinzip auf zwei (oder mehr) Gruppen aufgeteilt. Anschließend wird ein pädagogisches Merkmal so manipuliert, dass das Merkmal bei der einen Gruppe eine bestimmte Ausprägung aufweist und bei der anderen Gruppe eine andere Ausprägung (z. B. in einer Gruppe bearbeiten die Schüler/-innen Aufgaben in Einzelarbeit, in der anderen Gruppe bearbeiten die Schüler/-innen Aufgaben in Partnerarbeit). Durch die randomisierte Gruppenbildung sind die beiden Gruppen in ihren Ausgangsbedingungen miteinander vergleichbar (z. B. Intelligenz, soziale Herkunft). Folglich können Unterschiede im Lernerfolg der Schüler/-innen eindeutig auf das manipulierte pädagogische Merkmal zurückgeführt werden. Der zentrale Vorteil experimenteller Studien ist demnach ihre hohe interne Validität, d. h., es kann mit hoher Sicherheit auf Ursache-Wirkungs-Zusammenhänge geschlossen werden (Döring & Bortz, 2016, S. 94). Ein grundsätzliches Problem liegt darin, dass eine randomisierte Gruppenbildung im Schulbereich nur selten möglich ist. Weiterhin wird die Durchführbarkeit von Experimenten dadurch eingeschränkt, dass bestimmte pädagogische Merkmale nicht oder nur im begrenzten Maße manipuliert werden können (z. B. Überzeugungen der Lehrkraft, Schulklima). Schließlich sind experimentelle Studien vergleichsweise aufwendig, da pro Experiment jeweils nur ein pädagogisches Merkmal untersucht werden kann. Dies mögen einige der Gründe sein, warum innerhalb der pädagogischen Wirksamkeitsforschung weniger experimentelle Studien als korrelative Studien durchgeführt werden (Muijs, 2012, S. 61).

2.3.3 Standards für Wirksamkeitsstudien

Nachdem im vorherigen Absatz beschreibend dargestellt wurde, wie pädagogische Wirksamkeit untersucht werden kann, wird im Weiteren beleuchtet, was eine hochwertige Wirksamkeitsstudie auszeichnet. In der Literatur finden sich mehrfach Überlegungen zu wichtigen Standards für Wirksamkeitsstudien. Bei diesen Standards handelt es sich um Gütekriterien, die wissenschaftliche Studien erfüllen sollten, damit sie belastbare Evidenzen für pädagogische Wirksamkeit liefern. Die in der Literatur genannten Empfehlungen können folgendermaßen zusammengefasst werden (Ditton, 2000, S. 75 ff.; Ditton, 2002, S. 205 ff.; Goldstein, 1997, S. 376; Helmke, 2014, S. 53 ff.; Reynolds et al., 2014, S. 203 ff.):

- 1) Wirksamkeitsstudien sollten eine *fundierte theoretische Grundlage* aufweisen. Die Theorie sollte eine begründete Auswahl von Untersuchungsvariablen ermöglichen, eine klare Konzeptualisierung und Ordnung der Vari-

ablen leisten sowie Erklärungen für die Wirksamkeit von pädagogischen Merkmalen bereitstellen.

- 2) Angesichts der Komplexität des Gegenstandsbereichs erscheint die *Berücksichtigung mehrerer Dimensionen pädagogischer Wirksamkeit* angezeigt. Differenzierte Analysen unterscheiden zwischen verschiedenen Lernzielen, durchschnittlichem Lernerfolg und Unterschieden im Lernerfolg, kurzfristigen und langfristigen Lernwirkungen, Wirksamkeit bei verschiedenen Schülergruppen sowie Wirksamkeit in verschiedenen institutionellen Kontexten.
- 3) *Geeignete Forschungsdesigns* für Wirksamkeitsstudien sind einerseits experimentelle Designs und andererseits korrelative Designs mit multivariater und längsschnittlicher Anlage. Ein multivariater Erhebungsplan berücksichtigt neben pädagogischen Merkmalen auch die Ausgangsbedingungen der Schüler/-innen (z.B. Intelligenz, soziale Herkunft). Der Untersuchungszeitraum sollte mindestens ein Schuljahr und zwei Messzeitpunkte umfassen. Eine höhere Anzahl an Schuljahren und Messzeitpunkten ist wünschenswert.
- 4) *Angemessene Erhebungsmethoden* sind dadurch gekennzeichnet, dass für die Erfassung der einzelnen Untersuchungsvariablen jeweils geeignete Datenquellen und Messinstrumente verwendet werden. Bei Variablen, bei denen mehrere Perspektiven ihre Berechtigung haben (z.B. Unterrichtsmerkmale), ist die Kombination mehrerer Datenquellen sinnvoll.
- 5) *Angemessene Auswertungsmethoden* berücksichtigen die spezifischen Anforderungen der untersuchten Daten. Für hierarchische Daten sind Mehrebenenmodelle geeignet, für latente Variablen Strukturgleichungsmodelle und Item-Response-Theorie-Modelle.
- 6) Weiterhin sollten sich die Analysen nicht auf lineare Haupteffekte von pädagogischen Merkmalen beschränken, sondern es sollten *komplexe Wirkzusammenhänge überprüft werden*. Hierzu zählen kurvilineare Effekte, indirekte Effekte, Interaktionseffekte und reziproke Effekte.

Es ist anzumerken, dass es sich bei den genannten Standards um wünschenswerte Idealvorstellungen handelt. Allein schon aus pragmatischen Gründen wird es in der Forschungspraxis nur selten möglich sein, alle Standards zu verwirklichen. Die Standards geben aber in dem Sinne eine Orientierung, als dass die von einer Studie gelieferten Evidenzen umso belastbarer ausfallen, je stärker diese Standards berücksichtigt werden.

2.4 Zusammenfassung des zweiten Kapitels

Zusammenfassend ist festzuhalten, dass durch Schule und Unterricht eine ganze Reihe unterschiedlicher Zielsetzungen erreicht werden sollen. In inhaltlicher Hinsicht werden sowohl fachspezifische als auch fächerübergreifende Lernziele als relevant angesehen. Bei diesen Lernzielen werden jeweils, bei der Betrachtung einer Gruppe von Schüler/-innen, ein hoher durchschnittlicher Lernerfolg und geringe Unterschiede im Lernerfolg angestrebt. Zum letztgenannten Aspekt gehört auch das Ziel der Entkoppelung des Lernerfolgs von der sozialen Herkunft, welches von zentraler Bedeutung für faire Bildungschancen ist.

Welche Faktoren die Erreichung dieser Ziele beeinflussen, wird durch das Konzept der pädagogischen Wirksamkeit beschrieben. Pädagogische Wirksamkeit bezeichnet einen Ursache-Wirkungs-Zusammenhang, bei dem ein pädagogischer Faktor bei einem Individuum Lernaktivitäten anregt und dadurch Lernwirkungen erzeugt. Als pädagogische Faktoren können sowohl pädagogische Einheiten als auch pädagogische Merkmale betrachtet werden. Innerhalb der Kategorie der pädagogischen Merkmale kann zwischen Prozessmerkmalen und Strukturmerkmalen unterschieden werden. Weiterhin ist die Komplexität pädagogischer Wirksamkeit zu beachten, da sich Wirksamkeit auf verschiedene Aspekte wie Zielsetzungen, Zeiträume, Gruppen von Schüler/-innen und Kontexte beziehen kann.

Wissenschaftliche Untersuchungen zu pädagogischer Wirksamkeit werden unter dem Begriff pädagogische Wirksamkeitsforschung zusammengefasst. Das primäre Ziel liegt in der Bereitstellung eines empirisch fundierten Orientierungswissens, welches die Grundlage für eine evidenzbasierte pädagogische Praxis und Bildungspolitik bildet. Typischerweise werden für Wirksamkeitsstudien korrelative Designs verwendet. Experimentelle Designs weisen zwar eine höhere interne Validität auf, sind im Schulbereich aber nur begrenzt umsetzbar und daher auch seltener. Wichtige Standards für Wirksamkeitsstudien sind ein theoriegeleitetes Vorgehen, die Berücksichtigung mehrerer Dimensionen pädagogischer Wirksamkeit, geeignete Forschungsdesigns, Erhebungsmethoden und Auswertungsmethoden sowie die Analyse komplexer Wirkzusammenhänge.

3. Effekte der sozialen Herkunft

Im vorausgehenden Kapitel wurde mit der Betrachtung pädagogischer Wirksamkeit eine wichtige theoretische Grundlage für die weiteren Ausführungen gelegt. Im vorliegenden Kapitel wird die erste inhaltliche Fragestellung behandelt, nämlich die Frage, wie die soziale Herkunft von Schüler/-innen ihren Lernerfolg beeinflusst. Diese Thematik weist zwei inhaltliche Bezüge zum Konzept der pädagogischen Wirksamkeit auf. Erstens stellt die soziale Herkunft ein Strukturmerkmal dar, welches den Lernerfolg vermittelt über pädagogische Merkmale beeinflusst. Dieser Aspekt steht im Vordergrund des vorliegenden Kapitels. Zweitens stellen geringe Kompetenzunterschiede nach sozialer Herkunft ein Zielkriterium für die Wirksamkeit von Schule und Unterricht dar. Dieser Aspekt wird in den beiden nachfolgenden Kapiteln behandelt.

Kompetenzunterschiede nach sozialer Herkunft verweisen auf den Themenkomplex der *Bildungsungleichheiten* (Synonym: Bildungsdisparitäten). Bildungsungleichheiten beziehen sich auf systematische Unterschiede zwischen Gruppen von Lernenden (z.B. sozialer Herkunft, Migrationshintergrund, Geschlecht) bezüglich des Zugangs zu Bildungseinrichtungen und angeboten, der Nutzung von Bildungseinrichtungen und -angeboten, dem Erfolg von Lernprozessen und dem Erwerb von Bildungszertifikaten (Becker, 2011, S. 91; Gniewosz & Gräsel, 2015, S. 198). Die wissenschaftliche Beschäftigung mit und die gesellschaftliche Diskussion über Bildungsungleichheiten sind in Deutschland keine neuen Phänomene. Bereits im Rahmen der öffentlichen Bildungsdiskussionen der 1960er Jahre hatte beispielsweise Peisert (1967) mit der Sozialfigur der „Katholischen Arbeitertochter vom Land“ auf bestehende Bildungsungleichheiten aufmerksam gemacht. Im Anschluss an diese Debatte und die durch sie bedingten Bildungsreformen der 1960er und 1970er Jahre ebte das Interesse an der Thematik allerdings bedeutsam ab. In den nachfolgenden zwei Jahrzehnten erfuhren Bildungsungleichheiten kaum öffentliche und wissenschaftliche Aufmerksamkeit (Geißler & Weber-Menges, 2010, S. 155f.). Erst die PISA-Studie im Jahr 2000 (Baumert et al., 2001a) brachte Bildungsungleichheiten zurück auf die öffentliche Agenda und initiierte umfangreiche Forschungsaktivitäten, welche bis heute anhalten.

Für ein klares Verständnis ist auf bestimmte Begriffsverwendungen hinzuweisen. Der Begriff *Lernerfolg* wird im Weiteren gleichbedeutend mit Erwerb von Fachkompetenzen verwendet (in Abgrenzung zu anderen Lernzielen wie

sozialen Kompetenzen oder Wertorientierungen). Die Begriffe *Herkunft* und *herkunftsbedingt* beziehen sich im Weiteren ausschließlich auf die soziale Herkunft von Schüler/-innen (in Abgrenzung zur kulturellen Herkunft).

Nachfolgend wird zunächst eine Einführung in den Themenbereich der Bildungsungleichheiten nach sozialer Herkunft gegeben (Kapitel 3.1). Im Weiteren stehen dann herkunftsbedingte Unterschiede im Lernerfolg im Vordergrund. Hierfür wird ein theoretisches Modell entwickelt (Kapitel 3.2) und es werden Forschungsansätze zu vermittelnden Mechanismen erörtert (Kapitel 3.3). Der daran anschließende Hauptteil des Kapitels besteht in einem Überblick über Forschungsbefunde zu den Ursachen von Kompetenzunterschieden nach sozialer Herkunft. Behandelt werden systematische Unterschiede familiären Ressourcen (Kapitel 3.4), familiären Lernumwelten (Kapitel 3.5), frühpädagogischen Einrichtungen (Kapitel 3.6), Schulen (Kapitel 3.7), schulbezogenen Einrichtungen (Kapitel 3.8), Freizeitaktivitäten (3.9), Gleichaltrigengruppen (3.10) und Lernvoraussetzungen der Schüler/-innen (Kapitel 3.11).

Der Überblick über den Forschungsstand wird bewusst sehr ausführlich gestaltet. Zu der Frage nach vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg gibt es bisher kaum Übersichtsarbeiten. Vorhandene Arbeiten (z. B. Erikson & Jonsson, 1996, S. 10 ff.; Solga & Dombrowski, 2009, S. 21 ff.) begrenzen sich typischerweise auf ausgewählte Aspekte und verweisen nur beispielhaft auf empirische Untersuchungen. Der nachfolgende, auf einem theoretischen Modell basierende Forschungsüberblick ist durch ein höheres Maß an Systematik gekennzeichnet. Eine systematische Aufarbeitung der verfügbaren Evidenzen ist insbesondere aus zwei Gründen relevant. Erstens hängt die Ermittlung von wirksamen Maßnahmen zur Verringerung von herkunftsbedingten Kompetenzunterschieden wesentlich davon ab, auf welche Ursachen diese Kompetenzunterschiede zurückzuführen sind. Insofern bildet der Forschungsüberblick eine wichtige Grundlage für die im fünften Kapitel erfolgenden Überlegungen dazu, welche schulischen Faktoren einen Einfluss auf herkunftsbedingte Unterschiede im Lernerfolg haben (siehe Kapitel 5.2 „Mechanismen der Effekte von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg“). Zweitens ermöglicht der Forschungsüberblick eine differenzierte Einschätzung, inwieweit bei den einzelnen Mechanismen weiterer Forschungsbedarf besteht.

3.1 Bildungsungleichheiten nach sozialer Herkunft

Einführend soll ein grundlegendes Verständnis für Bildungsungleichheiten nach sozialer Herkunft erarbeitet werden. Zunächst wird dargestellt, wie die

soziale Herkunft von Schüler/-innen theoretisch konzeptualisiert und empirisch operationalisiert wird. Anschließend wird ein Überblick gegeben, welche Aspekte des Bildungserfolgs von der sozialen Herkunft beeinflusst werden.

3.1.1 Konzeptualisierung und Operationalisierung sozialer Herkunft

Die *soziale Herkunft* eines Schülers/einer Schülerin wird anhand des sozialen Status (Synonym: Sozialstatus, sozioökonomischer Status) seiner/ihrer Eltern ermittelt. Der *soziale Status* einer Person beschreibt ihre relative Position in der gesellschaftlichen Hierarchie (Ditton & Maaz, 2015, S. 229; Ehmke & Jude, 2010, S. 232). Gesellschaftliche Hierarchie bezieht sich auf eine Rangordnung vorhandener Positionen anhand von verfügbaren Ressourcen, sozialem Ansehen und Macht. Personen mit hohem sozialen Status verfügen demnach über mehr Ressourcen, erfahren mehr soziales Ansehen und haben mehr Macht als Personen mit niedrigem sozialen Status. Zentrale Dimensionen des sozialen Status sind Bildungsstand, Berufsposition und Einkommen (Sirin, 2005, S. 418). Demnach handelt es sich bei *sozial privilegierten Schüler/-innen* um Kinder von Eltern mit relativ hohem sozialen Status. Bei *sozial benachteiligten Schüler/-innen* handelt es sich um Kinder von Eltern mit relativ niedrigem sozialen Status.

Bezogen auf den Bildungserfolg erscheint insbesondere die *Verfügbarkeit von familiären Ressourcen* als ein wichtiger Aspekt des sozialen Status. Ressourcen bezeichnen materielle und immaterielle Güter, welche Handlungsmöglichkeiten erweitern und Handlungsbereitschaften erhöhen, wodurch sie als Mittel zur Erreichung von Zielen eingesetzt werden können (angelehnt an Baumert et al., 2003, S. 54). Unterschieden werden können ökonomische Ressourcen (z. B. Geld), persönliche Ressourcen (z. B. Wissen) und soziale Ressourcen (z. B. Beziehungen) (angelehnt an Bourdieu, 2012). Aus der Beschreibung des sozialen Status anhand von verfügbaren Ressourcen ergibt sich, dass der Zusammenhang zwischen Schülerkompetenzen und sozialer Herkunft inhaltlich als *Abhängigkeit des Lernerfolgs von familiären Ressourcen* interpretiert werden kann.

Der soziale Status wird in empirischen Studien mit unterschiedlichen *Erhebungsmethoden* erfasst (Ditton & Maaz, 2015). Das Einkommen stellt eine quantitative Variable dar und kann daher direkt gemessen werden. Für die Erfassung von Bildungsstand und Berufsposition wurden spezifische Verfahren entwickelt, welche verschiedenen Bildungsabschlüssen bzw. Berufen jeweils bestimmte Messwerte zuordnen. Relevante Klassifikationen für den Bildungsstand sind ISCED (UNESCO Institute for Statistics, 2012) und CASMIN (Lüt-

tinger & König, 1988). Bei der Berufsposition sind insbesondere ISEI (Ganzeboom, 2010) und EGP-Schema (Erikson & Goldthorpe, 1992) von Bedeutung.

3.1.2 Soziale Herkunft und Bildungserfolg

Die soziale Herkunft eines Schülers/einer Schülerin beeinflusst unterschiedliche Aspekte seiner/ihrer Bildungsprozesse und -ergebnisse. Zahlreiche Studien haben Bildungsungleichheiten untersucht und Effekte der sozialen Herkunft auf den Lernerfolg (z. B. Erwerb von Kompetenzen), die Bildungsbeteiligung (z. B. besuchte Schulform) und den erreichten Schulabschluss nachgewiesen. Im Weiteren werden überblicksartig zentrale Forschungsbefunde dargestellt.

Der Einfluss der sozialen Herkunft auf den *Lernerfolg* der Schüler/-innen äußert sich darin, dass sozial privilegierte Schüler/-innen über bessere Fachkompetenzen verfügen als sozial benachteiligte Schüler/-innen. Studien zum Zusammenhang zwischen sozialer Herkunft und Schülerkompetenzen reichen bis an den Anfang des 20. Jahrhunderts zurück (Kornhauser, 1918). Der internationale Forschungsstand ist so umfangreich, dass inzwischen mehrere Meta-Analysen vorliegen (Sirin, 2005; White, 1982). In der jüngeren Meta-Analyse von Sirin (2005) wird eine durchschnittliche Korrelation von $r = 0.28$ ermittelt (S. 432), was die hohe Bedeutung der sozialen Herkunft für den Lernerfolg verdeutlicht. Die Effekte der sozialen Herkunft auf unterschiedliche Fachkompetenzen variieren dabei nur geringfügig (Mathematik > Lesen > Naturwissenschaften, S. 435). Aufschlussreich sind in diesem Kontext die Schulleistungsstudien PISA, TIMSS und IGLU, da sie international vergleichende Analysen ermöglichen. So war ein zentraler Befund der ersten PISA-Studie, dass der Effekt der sozialen Herkunft auf den Lernerfolg in Deutschland deutlich stärker ausgeprägt ist als in anderen Ländern (Baumert & Schümer, 2001, S. 391). Es finden sich zwar Hinweise darauf, dass herkunftsbedingte Kompetenzunterschiede seitdem leicht abgenommen haben (Müller & Ehmke, 2016, S. 300; Stubbe et al., 2016, S. 311; abweichender Befund bei Hußmann et al., 2017, S. 204), aber der Effekt der sozialen Herkunft auf den Lernerfolg ist in Deutschland weiterhin überdurchschnittlich stark ausgeprägt.

Die soziale Herkunft eines Schülers/einer Schülerin hat weiterhin einen Einfluss auf seine/ihre *Bildungsbeteiligung*, also die Nutzung verfügbarer Bildungseinrichtungen. Von besonderer Bedeutung sind hierbei Übergänge zwischen Bildungseinrichtungen, welche die „Gelenkstellen“ des Bildungssystems darstellen (Maaz et al., 2010, S. 13). Herkunftseffekte zeigen sich im Schulbereich vor allem beim Übergang von der Grundschule auf die Sekundarschule. Sozial privilegierte Schüler/-innen besuchen deutlich häufiger ein Gymnasium

als sozial benachteiligte Schüler/-innen (Baumert & Schümer, 2001, S. 357 f.; Ditton et al., 2005, S. 293; Zielonka et al., 2014, S. 120). Um diese Unterschiede in der besuchten Schulform zu erklären, hat sich die Unterscheidung zwischen primären und sekundären Herkunftseffekten als nützlich erwiesen (Boudon, 1974). *Primäre Herkunftseffekte* beziehen sich auf die zuvor dargestellten Einflüsse der sozialen Herkunft auf den Kompetenzerwerb. Aus dieser Perspektive besuchen sozial privilegierte Schüler/-innen häufiger das Gymnasium als sozial benachteiligte Schüler/-innen, weil sie am Ende der Grundschulzeit über bessere Kompetenzen verfügen. *Sekundäre Herkunftseffekte* verweisen auf Einflüsse der sozialen Herkunft auf die besuchte Schulform, die über die Kompetenzen der Schüler/-innen hinausgehen (kompetenzunabhängige Effekte der sozialen Herkunft). Empirische Studien belegen, dass sozial privilegierte Schüler/-innen auch bei gleichem Kompetenzstand häufiger das Gymnasium besuchen als sozial benachteiligte Schüler/-innen (Baumert & Schümer, 2001, S. 357 f.; Ditton et al., 2005, S. 296; Zielonka et al., 2014, S. 121).

Unterschiede in der besuchten Schulform führen schließlich zu Unterschieden in den erreichten *Schulabschlüssen*. Heranwachsende aus Familien mit hohem sozialen Status erreichen im Durchschnitt deutlich günstigere Schulabschlüsse als Heranwachsende aus Familien mit niedrigem sozialen Status. Bei Eltern mit Abitur erreichen 67% der Kinder ebenfalls ein Abitur und weniger als 5% der Kinder einen Hauptschulabschluss. Bei Eltern mit Hauptschulabschluss erreichen demgegenüber nur 15% der Kinder ein Abitur und 37% der Kinder einen Hauptschulabschluss (Seifert, 2005, S. 3 f.). Angesichts dieses engen Zusammenhangs zwischen den Schulabschlüssen von Eltern und Kindern erscheint es zutreffend, von einer „Vererbung“ von Bildung zu sprechen (Ditton, 2010, S. 59).

Die angeführten Befunde verdeutlichen, dass Bildungsungleichheiten nach sozialer Herkunft ein umfassendes und anhaltendes gesellschaftliches Problem darstellen. Hieraus ergibt sich die Relevanz, durch empirische Forschung ein vertieftes Verständnis darüber zu entwickeln, aufgrund welcher Mechanismen Bildungsungleichheiten nach sozialer Herkunft entstehen und mit welchen Maßnahmen diese wirksam verringert werden können (Neumann et al., 2014, S. 193).

3.2 Theoretisches Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Nach der überblicksartigen Einführung in Bildungsungleichheiten im letzten Abschnitt werden im Weiteren Unterschiede im *Lernerfolg* tiefergehend betrachtet. Der Schwerpunkt liegt bei den *Ursachen* für die Entstehung von

Kompetenzunterschieden nach sozialer Herkunft. Nachfolgend werden einige theoretische Grundlagen erörtert, die die Basis für ein im Anschluss vorgestelltes theoretisches Modell bilden.

3.2.1 Theoretische Grundlagen des Modells zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

In der Literatur werden als Erklärungsansätze für Herkunftseffekte Unterschiede in genetischen Faktoren, Unterschiede in physiologischen Faktoren und Unterschiede in Sozialisationsfaktoren diskutiert (Erikson & Jonsson, 1996, S. 10 ff.; Schauenberg, 2007, S. 50). *Genetische Faktoren* beziehen sich darauf, dass statushohe Eltern über günstigere Persönlichkeitsmerkmale verfügen (z. B. höhere Intelligenz), welche sie durch Vererbung an ihre Kinder weitergeben. Demgegenüber verweisen *physiologische Faktoren* auf Unterschiede in Einflüssen, die auf der körperlichen Ebene wirken (z. B. Ernährung). *Sozialisationsfaktoren* beschreiben schließlich Unterschiede in den Lernumwelten von Heranwachsenden, welche zu unterschiedlichen Lernerfahrungen führen (z. B. Erziehung durch die Eltern). Sozialisationsfaktoren sind für die Ermittlung von schulischen Maßnahmen zur Verringerung von herkunftsbedingten Kompetenzunterschieden von besonderer Relevanz, da die Schule diese beeinflussen bzw. in Wechselwirkung mit ihnen treten kann. Die weiteren Ausführungen beschränken sich daher auf Sozialisationsfaktoren.

In der Literatur findet sich eine ganze Reihe von *theoretischen Ansätzen*, welche die Entstehung von Bildungsungleichheiten nach sozialer Herkunft durch Sozialisationsfaktoren erklären. Wichtige theoretische Ansätze sind schichtspezifische Sozialisationsforschung, Reproduktionstheorie, institutionelle Diskriminierung und rationale Bildungsentscheidungen. Im Rahmen der vorliegenden Arbeit ist es weder möglich noch sinnvoll, die genannten Ansätze im Detail darzustellen. Stattdessen basieren die weiteren Ausführungen auf einer integrativen Argumentation, bei der auf Komponenten verschiedener Theorien zurückgegriffen wird. Für eine tiefergehende Betrachtung der genannten Ansätze sei der Leser/die Leserin auf die einschlägige Literatur verwiesen (Becker & Lauterbach, 2010; Brake & Büchner, 2012; Maaz et al., 2010).

Nachfolgend wird ein theoretisches Modell vorgestellt, welches bestehende Erklärungsansätze für Kompetenzunterschiede nach sozialer Herkunft aufgreift und zusammenführt. Eine zentrale Annahme ist hierbei, dass eine angemessene theoretische Betrachtung von Herkunftseffekten einen *interdisziplinären Ansatz* erfordert. Aus der *bildungssoziologischen Theorie* werden Bourdieus (2012) Kapitalarten aufgegriffen, die insbesondere für die Beschreibung von

Unterschieden in familiären Ressourcen hilfreich sind. Weiterhin wird auf das Konzept der Bildungsentscheidungen (Boudon, 1974) Bezug genommen, mit dem sich Unterschiede in der Nutzung von Bildungseinrichtungen analysieren lassen. *Pädagogisch-psychologische Theorien* werden demgegenüber für die Betrachtung von Lehr-Lern-Prozessen, den sie einbettenden Lernumwelten und den Lernvoraussetzungen von Heranwachsenden verwendet. Wichtige Grundlagen stellen hierbei die Modelle zu Einflussfaktoren auf Schulleistungen von Helmke und Weinert (1997a, S. 86) und Helmke (2014, S. 28) dar. Ergänzt werden diese Modelle durch das Konzept der Lernumwelten (Bäumer et al., 2011). Weiterhin werden einige Überlegungen zu pädagogischer Wirksamkeit aus dem vorherigen Kapitel aufgegriffen.

3.2.2 Beschreibung des Modells zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Die Funktion des nachfolgend dargestellten Modells liegt darin, eine *theoretische Ordnung* der sozialisationsbezogenen Einflüsse der sozialen Herkunft auf den Lernerfolg zu leisten. Hierfür werden Kategorien, also Gruppen von inhaltlich verwandten Merkmalen, gebildet und Wirkzusammenhänge zwischen den Kategorien postuliert. Entsprechend den inhaltlichen Schwerpunkten der vorliegenden Arbeit bezieht sich das Modell nur auf Lernerfolg beim Erwerb von Fachkompetenzen. Das Modell bildet sowohl die Grundlage für den nachfolgenden Forschungsüberblick als auch für die empirischen Analysen.

Abbildung 4 stellt das Modell dar. Erkennbar stellt das Modell einen *mehrschrittigen Mediationsprozess* dar. Sozialstatus und familiäre Ressourcen stellen Strukturmerkmale dar. Diese haben einen Einfluss auf Lernumwelten und Lernaktivitäten, bei denen Prozessmerkmale im Vordergrund stehen. Die Lernaktivitäten der Kinder wirken sich direkt auf den Erwerb von Fachkompetenzen aus, aber auch indirekt vermittelt über die Förderung von Lernvoraussetzungen. Während zweifellos gewisse Unterschiede zwischen verschiedenen Fachkompetenzen und den sie beeinflussenden Faktoren bestehen (z. B. hat die Lesehäufigkeit einen größeren Einfluss auf die Lesekompetenz als auf die Mathematikkompetenz), liegt dem Modell die Annahme zugrunde, dass zumindest die gleichen Kategorien fächerübergreifend relevant sind (d. h., Freizeitaktivitäten beeinflussen alle Arten von Fachkompetenzen) (ähnlich wie bei Walberg, 1984, S. 20).

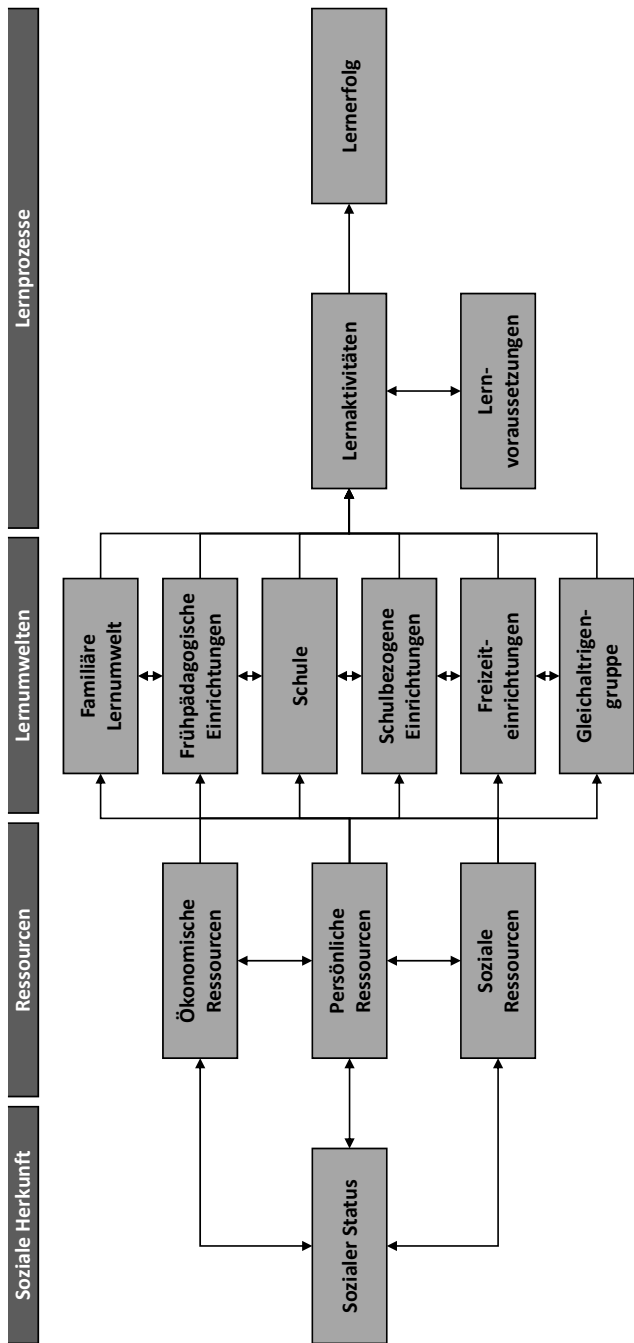


Abb. 4: Theoretisches Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Im Weiteren werden die einzelnen Komponenten des Modells und ihre Beziehungen zueinander schrittweise von der linken zur rechten Seite erläutert.² Der *soziale Status* der Eltern bzw. der Familie umfasst die klassischen Dimensionen Bildungsstand, Berufsposition und Einkommen. Der soziale Status steht in einem direkten Verhältnis zu den *familiären Ressourcen*. Es werden drei Arten von Ressourcen unterschieden, angelehnt an die Kapitalarten von Bourdieu (2012): ökonomische, persönliche und soziale Ressourcen. Familiäre Ressourcen werden hierbei inhaltlich enger gefasst als die Kapitalarten bei Bourdieu. Erstens sind nur lernrelevante Ressourcen gemeint, d. h. Ressourcen, die einen unmittelbaren Bezug zu Einflussfaktoren auf das kindliche Lernen haben (weshalb z. B. vorhandene Kunstwerke, die einen Aspekt des kulturellen Kapitals darstellen, nicht als Ressource angesehen werden). Zweitens stellen Ressourcen durchwegs Strukturmerkmale dar, während die Kapitalarten teilweise auch Prozessmerkmale umfassen (Jungbauer-Gans, 2004, S. 386 f.). *Persönliche Ressourcen* beziehen sich auf bestimmte Persönlichkeitsmerkmale der Eltern (z. B. Fachkompetenzen, Bildungsaspirationen). *Ökonomische Ressourcen* umfassen den Wohnort und die Bildungsausgaben der Familie. Zu den *sozialen Ressourcen* gehören die Familienstruktur sowie die näheren Bezugspersonen der Familie (Verwandte, Freund/-innen), welche ihrerseits über ökonomische, persönliche und soziale Ressourcen verfügen, auf die im Rahmen der sozialen Beziehung bis zu einem gewissen Grad zugegriffen werden kann. Angenommen wird ein wechselseitiger Einfluss zwischen sozialem Status und familiären Ressourcen: einerseits bedingen die verfügbaren Ressourcen den erreichten sozialen Status (z. B. beeinflussen die Fachkompetenzen der Eltern ihren Bildungsstand), andererseits bedingt der soziale Status die verfügbaren Ressourcen (z. B. führt ein höheres Einkommen zu einem günstigeren Wohnort).

Zentral ist die Annahme, dass Unterschiede in den familiären Ressourcen zu Unterschieden in den von den Heranwachsenden erfahrenen Lernumwelten führen (Maaz et al., 2010, S. 14 f.; Neumann et al., 2014, S. 170). Hinsichtlich des Einflusses der familiären Ressourcen ist hierbei von mehreren *Wirkmechanismen* auszugehen. Eine wesentliche Rolle spielen Unterschiede im Verhalten der Eltern. Demnach gestalten statushohe Eltern für ihre Kinder anregendere

2 Es ist anzumerken, dass im Modell die kulturelle Herkunft von Schüler/-innen, also ein Migrationshintergrund der Familie, nicht berücksichtigt wird. Es bestehen zwar Zusammenhänge zwischen kultureller Herkunft und sozialer Herkunft, da Migrant/-innen im Durchschnitt über niedrigere soziale Positionen verfügen. Als Wirkrichtung ist hierbei allerdings primär anzunehmen, dass der Migrationshintergrund den sozialen Status bedingt. Insofern stellt der soziale Status eher einen Mechanismus des Migrationshintergrundes dar als umgekehrt (Garcia, 2015, S. 4).

familiäre Lernumwelten als statusniedrige Eltern (z. B. Erziehungsstil). Außerdem gelingt es ihnen besser, anregende außerfamiliäre Lernumwelten auszuwählen (z. B. Bildungsentscheidungen). Ein weiterer Mechanismus ist in Unterschieden im Verhalten von pädagogischen Fachkräften zu sehen, wenn diese Entscheidungen in Abhängigkeit von familiären Ressourcen fällen (z. B. Übergangsempfehlungen für Sekundarschulen, siehe Stahl, 2009, S. 101 f.). Schließlich ergibt sich ein weiterer Mechanismus daraus, dass sich außerfamiliäre Lernumwelten gegenseitig beeinflussen. Zum Beispiel können herkunftsspezifische Unterschiede in den besuchten Bildungseinrichtungen zu unterschiedlichen Freundeskreisen der Heranwachsenden führen.

Herkunftsspezifische Unterschiede in *Lernumwelten* sind bei Familien, frühpädagogischen Einrichtungen, Grund- und Sekundarschulen, schulbegleitenden Einrichtungen (z. B. Nachhilfe), Freizeiteinrichtungen (z. B. Sportverein) und Gleichaltrigengruppen denkbar. Unterschiede in der familiären Lernumwelt beziehen sich primär auf das Erziehungshandeln der Eltern. Hierzu gehören der Erziehungsstil, die kognitive Anregung und die schulische Beteiligung. Bei Bildungseinrichtungen können sich Unterschiede darauf beziehen, ob und wie lange die Kinder bestimmte Bildungseinrichtungen besuchen (z. B. Kindergarten), welche Bildungseinrichtungen die Kinder besuchen (z. B. Sekundarschulform) und welche Lernerfahrungen die Kinder innerhalb der Bildungseinrichtungen machen (z. B. Teilnahme an Ganztagsangeboten) (Maaz et al., 2010, S. 12 f.). Bezüglich der Gleichaltrigengruppe können Unterschiede in der Anzahl der Freund/-innen und der Art der Freund/-innen bestehen. Allen Lernumwelten ist gemein, dass ihre Wirkweise durch Strukturmerkmale, welche ihrerseits Prozessmerkmale bedingen, beschrieben werden kann (Bäumer et al., 2011, S. 94 f.).

Aus einer *zeitlichen Perspektive* ist zwischen Einflüssen von Lernumwelten vor dem Schuleintritt und Einflüssen von Lernumwelten während dem Schulbesuch zu unterscheiden (Schauenberg, 2007, S. 50). Unterschiedliche Lernumwelten während der frühen Kindheit (z. B. frühe familiäre Lernumwelt, frühe Gleichaltrigengruppe, Kindergarten) führen dazu, dass sich Kinder verschiedener sozialer Herkunft bereits bei Schuleintritt in ihren Lernvoraussetzungen unterscheiden. Es ist anzunehmen, dass sozial privilegierte Schüler/-innen dadurch die schulischen Anforderungen besser bewältigen können als sozial benachteiligte Schüler/-innen. Mit dem Schuleintritt können Unterschiede in den sich neu eröffnenden Lernumwelten entstehen (z. B. Schule, schulbegleitende Einrichtungen), während Unterschiede in den bereits zuvor relevanten Lernumwelten fortbestehen (z. B. spätere familiäre Lernumwelt, spätere Gleichaltrigengruppe). Hierbei ist anzunehmen, dass herkunftsspezifische

Unterschiede in verschiedenen Lebensphasen nicht unabhängig voneinander wirken, sondern kumulativ. Demnach führen Unterschiede in den biografisch früh erfahrenen Lernumwelten zu Unterschieden in den biografisch später erfahrenen Lernumwelten (z. B. Unterschiede im Kindergartenbesuch führen zu Unterschieden im Sekundarschulbesuch, siehe Becker, 2010b, S. 150 ff.).

Die herkunftsspezifischen Unterschiede in den genannten Lernumwelten resultieren in unterschiedlichen *Lernaktivitäten* der Heranwachsenden. Zu den Lernaktivitäten gehören fachliche Lernaktivitäten, die einen direkten Einfluss auf den Erwerb von Fachkompetenzen haben, als auch Lernaktivitäten, die vermittelt über die Förderung von Lernvoraussetzungen wirken (z. B. Sport fördert Konzentrationsfähigkeit). Innerhalb der Schule erfolgen Lernaktivitäten sowohl im Unterricht als auch im Rahmen von Ganztagsangeboten. Außerhalb der Schule umfassen Lernaktivitäten beispielsweise die Bearbeitung von Hausaufgaben, die Vorbereitung von Prüfungen und lernrelevante Freizeitaktivitäten (z. B. Lesen, Musizieren). Aus den unterschiedlichen Lernaktivitäten ergeben sich schließlich unterschiedliche Kompetenzen.

3.3 Forschungsansätze zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Im weiteren Verlauf des Kapitels werden vorhandene Forschungsbefunde zu den einzelnen Bereichen des Modells dargestellt. Um diese einordnen zu können, wird zunächst knapp erörtert, wie vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg empirisch untersucht werden können.

Grundsätzlich gilt, dass eine Variable mindestens *zwei empirische Voraussetzungen* erfüllen muss, um als vermittelnder Mechanismus in Betracht zu kommen: die Variable muss sowohl mit der sozialen Herkunft als auch mit dem Lernerfolg eine empirisch nachweisbare Korrelation aufweisen (Steinmayr et al., 2012, S. 337). Eine bivariate Korrelation zwischen einer Variable und dem Lernerfolg liefert allerdings nur begrenzte Evidenz dafür, dass diese Variable wirksam ist (im Sinne eines Kausaleffekts auf den Lernerfolg), weil die Variable in diesem Fall auch nur ein Korrelat anderer, wirksamer Einflussfaktoren sein könnte. Aussagekräftiger sind daher multivariate Analysen, bei denen parallel die Effekte mehrerer Variablen überprüft werden. Neben empirischen Befunden ist weiterhin das Vorhandensein einer *theoretischen Erklärung* relevant für die Interpretation eines Kausaleffekts. Anzumerken ist, dass bei vielen Variablen des obigen Modells experimentelle Studien grundsätzlich nicht möglich sind (z. B. Erziehungsstil, Gleichaltrigengruppe). Insofern ist bei der Untersu-

chung von vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg häufig nur ein moderates Maß an kausaler Evidenz zu erreichen.

Die Evidenz für einen vermittelnden Mechanismus hängt weiterhin davon ab, ob die zwei relevanten Zusammenhänge im Rahmen *einer Studie* belegt werden oder in *zwei verschiedenen Studien*. Der letztgenannte Fall bezieht sich beispielsweise darauf, dass in einer Studie ein Zusammenhang zwischen sozialem Status und Erziehungsstil nachgewiesen wird und in einer anderen Studie ein Zusammenhang zwischen Erziehungsstil und Lernerfolg. Da sich die Stichproben verschiedener Studien typischerweise voneinander unterscheiden, ist in diesem Szenario nicht gewiss, ob die beiden Zusammenhänge auch zusammen auftreten würden. Aussagekräftiger ist es daher, wenn eine Studie beide Zusammenhänge in der gleichen Stichprobe belegt. Werden beide Zusammenhänge in der gleichen Stichprobe untersucht, besteht außerdem die Möglichkeit einer Analyse mit *Pfadmodellen*. Dieses statistische Verfahren ist für die Überprüfung von Mediationen geeignet, da mehrere Zusammenhänge in einem gemeinsamen Modell analysiert werden können (Geiser, 2011, S. 75 ff.). Eine Analyse mit Pfadmodellen liefert somit eine stärkere Evidenz für einen vermittelnden Mechanismus als zwei getrennt berechnete Korrelationen. Es liegt bereits eine Reihe von Studien vor, die Effekte der sozialen Herkunft mithilfe von Pfadmodellen untersuchen (Ehmke et al., 2006b; Linberg et al., 2020; Schauenberg, 2007; Steinmayr et al., 2012).

Weitere Herausforderungen bei der Analyse von vermittelnden Mechanismen ergeben sich aus der Komplexität möglicher Wirkzusammenhänge. Eine besondere Situation bezieht sich darauf, dass eine Variable *positiv mit der sozialen Herkunft*, aber *negativ mit dem vorausgehenden Kompetenzstand* des Schülers/der Schülerin korreliert. Beispielsweise ist anzunehmen, dass es statushöheren Familien in finanzieller Hinsicht leichter möglich ist, Nachhilfe in Anspruch zu nehmen (Guill & Wendt, 2016, S. 247). Andererseits wird Nachhilfe insbesondere bei schlechten Fachkompetenzen initiiert, was wiederum bei sozial privilegierten Schüler/-innen seltener der Fall ist. In diesem Fall gilt es zu analysieren, ob die soziale Herkunft unter Kontrolle der Kompetenzen einen Effekt auf die Nutzung von Nachhilfe hat, was einem sekundären Herkunftseffekt entspricht (Guill & Wendt, 2016, S. 253).

Ein weiterer Sonderfall liegt vor, wenn eine Variable bei Schüler/-innen mit verschiedener sozialer Herkunft unterschiedlich ausgeprägt ist, die Variable aber gleichzeitig auch *differenziell wirksam* ist, d. h., bei verschiedenen Schülergruppen unterschiedlich starke Effekte hat. Ein denkbare Beispiel wäre, dass sozial privilegierte Kinder häufiger einen Kindergarten besuchen. Wenn der Kindergartenbesuch bei sozial privilegierten Kindern stärkere Effekte hat,

werden Kompetenzunterschiede zusätzlich verstärkt. Wenn der Kindergartenbesuch dagegen bei sozial benachteiligten Kindern stärkere Effekte hat, werden die Unterschiede im Besuch des Kindergartens möglicherweise kompensiert, so dass der Kindergarten insgesamt keinen vermittelnden Mechanismus darstellt. Im zweiten Fall hängt der Gesamteffekt davon ab, wie stark die zwei Sachverhalte (Nutzungsunterschiede und differenzielle Wirksamkeit) jeweils ins Gewicht fallen. Vor diesem Hintergrund erscheint es bei bestimmten Variablen angezeigt, differenzielle Wirksamkeit zu überprüfen, indem Interaktionseffekte modelliert werden.

Diese Ausführungen verdeutlichen, dass die Analyse vermittelnder Mechanismen zwischen sozialer Herkunft und Lernerfolg eine methodisch anspruchsvolle Aufgabe darstellt. Bei der Literaturrecherche für den nachfolgend dargestellten Forschungsüberblick zeigte sich, dass nur ein kleinerer Teil der vorhandenen Studien die genannten methodischen Anforderungen erfüllt. Bei der großen Mehrheit der vorhandenen Studien handelt es sich um einfache bivariate Analysen. Dies gilt es bei der Interpretation der nachfolgend berichteten Befunde zu bedenken.

3.4 Unterschiede bei familiären Ressourcen

Im Weiteren wird ein Überblick über Forschungsbefunde zu den einzelnen Bereichen des theoretischen Modells gegeben. Begonnen wird mit Zusammenhängen zwischen Sozialstatus und familiären Ressourcen. Zahlreiche Studienergebnisse belegen, dass statushohe Eltern über größere *persönliche Ressourcen* verfügen als statusniedrige Eltern. Persönliche Ressourcen beziehen sich auf Persönlichkeitsmerkmale von Eltern, die die Bereitschaften und Möglichkeiten bei der Förderung der eigenen Kinder bedingen. Festzuhalten ist zunächst, dass statushohe Eltern bessere *Fachkompetenzen* haben als statusniedrige Eltern, beispielsweise bessere Lese- und Mathematikkompetenzen (Maehler et al., 2013, S. 96 ff.). Auf dieser Grundlage ist es für statushohe Eltern leichter, ihre Kinder kognitiv anzuregen und beim schulischen Lernen zu begleiten. In Bezug auf *pädagogische Kompetenzen* zeigt sich, dass statushohe Eltern über mehr Wissen über die kindliche Entwicklung und förderliches Erziehungshandeln verfügen als statusniedrige Eltern (Rowe, 2008, S. 197; September et al., 2016, S. 1066). Anzunehmen ist, dass die pädagogischen Kompetenzen der Eltern nicht nur die direkten Interaktionen mit den Kindern bedingen, sondern auch ihre Einflussnahme auf außerfamiliäre Lernumwelten. Neben den Kompetenzen der Eltern, welche primär ihre Handlungsmöglichkeiten bedingen, spielen motivationale Merkmale der Eltern eine Rolle, welche ihre Handlungs-

bereitschaften bedingen. Eine wichtige persönliche Ressource sind hierbei die *Bildungsaspirationen* der Eltern, also ihre Zielvorstellung bezüglich der angestrebten Bildungsergebnisse ihrer Kinder. Statushohe Eltern wünschen sich für ihre Kinder bessere Noten und höhere Schulabschlüsse als statusniedrige Eltern (Dumont et al., 2014, S. 8 f.; Goodman & Gregg, 2010, S. 27 f.). Im Einklang damit messen statushohe Eltern den Erziehungszielen vielseitige Bildung (durch das Elternhaus), Neugier und Lesefreude eine größere Bedeutung bei als statusniedrige Eltern (Süßlin, 2015, S. 8). Es ist anzunehmen, dass die höheren Zielvorstellungen dazu führen, dass statushohe Eltern eher dazu bereit sind, einen bedeutsamen Aufwand für die Förderung ihrer Kinder zu erbringen.

Neben den persönlichen Ressourcen sind die *ökonomischen Ressourcen* der Familie von Bedeutung. Diese sind dadurch gekennzeichnet, dass sie einen starken Bezug zur finanziellen Ausstattung der Familie haben. Bezüglich des *Wohnorts* ist festzuhalten, dass statushohe Haushalte häufiger in Bezirken mit einer hohen Konzentration statushoher Haushalte wohnen, während statusniedrige Haushalte häufiger in Bezirken mit einer hohen Konzentration statusniedriger Haushalte wohnen (soziale Segregation, siehe Goebel & Hoppe, 2015, S. 8). Der Wohnort bedingt sowohl das Angebot an Bildungseinrichtungen als auch Möglichkeiten zur Entwicklung von Freundschaften. Weiterhin tätigen statushohe Eltern höhere *Bildungsausgaben* im Sinne von finanziellen Investitionen in die Förderung ihrer Kinder (OECD, 2019, S. 110; Schröder et al., 2015, S. 160 f.). Bildungsausgaben sind relevant für die Nutzung kostenpflichtiger Bildungseinrichtungen (z. B. Nachhilfe), für die Nutzung kostenpflichtiger Elternbildung (z. B. Erziehungsratgeber) und für die Beschaffung von Lernmaterial. Im Einklang damit zeigt sich, dass statushohe Familien über mehr Lernmaterial verfügen als statusniedrige Familien (z. B. Spielzeug, Kinderbücher, Lernsoftware) (Bradley et al., 1988, S. 66 ff.; Ehmke et al., 2006b, S. 240).

Die dritte Art familiärer Ressourcen sind *soziale Ressourcen*. Damit gemeint ist das soziale Beziehungsgefüge der Familie in seiner Bedeutung für die kindliche Sozialisation. Dies betrifft zunächst einmal die *Familienstruktur*. Unter statusniedrigen Familien ist der Anteil alleinerziehender Eltern überdurchschnittlich hoch (Statistisches Bundesamt, 2018, S. 22, 39). Außerdem gibt es Hinweise darauf, dass statusniedrige Personen mehr Kinder bekommen als statushohe Personen (Balbo et al., 2013, S. 8 f.; Kim, 2016, S. 1). Beide Aspekte der Familienstruktur führen dazu, dass in statusniedrigen Familien pro Kind weniger Ressourcen zur Verfügung stehen, was sich sowohl auf das elterliche Erziehungs Handeln als auch auf die Nutzung von Bildungseinrichtungen auswirken kann. Hinsichtlich der *Verwandtschaft* ist festzuhalten, dass statushohe Eltern ihrerseits häufiger Eltern und Geschwister mit hohem sozialem Status

haben (Autorengruppe Bildungsberichterstattung, 2008, S. 211; Seifert, 2005, S. 3f.), sodass die Schüler/-innen von den Ressourcen des erweiterten familiären Umkreises profitieren könnten. Bezüglich des *Freundeskreises* der Eltern zeigt sich, dass statushohe Eltern sowohl einen größeren Freundeskreis haben (Splendid Research, 2017, S. 11) als auch einen höheren Anteil an Freund/-innen mit hohem sozialem Status (soziale Homophilie, siehe McPherson et al., 2001, S. 426 ff.). Somit stehen statushohen Familien zusätzliche persönliche und ökonomische Ressourcen zur Verfügung.

Insgesamt betrachtet sind sowohl bei persönlichen als auch bei ökonomischen und sozialen Ressourcen bedeutsame Unterschiede zwischen Familien mit verschiedenem sozialem Status festzustellen.

3.5 Unterschiede bei familiären Lernumwelten

Die zuvor aufgezeigten Unterschiede in familiären Ressourcen führen zu Unterschieden in den von den Heranwachsenden erfahrenen Lernumwelten. Diese werden im Weiteren genauer betrachtet und bezüglich der vorhandenen Evidenzen als vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg eingeordnet. Zunächst wird auf die familiäre Lernumwelt eingegangen. Das elterliche Erziehungshandeln wird hierbei bezüglich des Erziehungsstils, der kognitiven Anregung und der schulischen Beteiligung betrachtet.

3.5.1 Unterschiede beim Erziehungsstil der Eltern

Ein wichtiges Konzept zur Beschreibung von familiären Interaktionen ist der elterliche *Erziehungsstil*, also bereichsübergreifende Tendenzen in Eltern-Kind-Interaktionen. Die Betrachtung des elterlichen Erziehungsstils erfolgt in der Literatur sowohl mithilfe eines dimensionalen Ansatzes als auch mithilfe eines typologischen Ansatzes. Der *dimensionale Ansatz* erfasst das elterliche Erziehungshandeln in der Regel anhand der zwei Dimensionen emotionale Wärme (auch: Wertschätzung, Responsivität) und Kontrolle (auch: Lenkung, Struktur) (Pinquart, 2016, S. 476). Zu emotionaler Wärme gehören Verhaltensweisen wie ein liebevoller Umgang, emotionale Unterstützung und Eingehen auf Bedürfnisse. Kontrolle beschreibt dagegen, in welchem Ausmaß und auf welche Art die Eltern das Verhalten des Kindes beeinflussen. Hinsichtlich der Art der Kontrolle können proaktive Kontrolle (z. B. konsistente Regeln und Erwartungen, Monitoring des kindlichen Verhaltens), dominante Kontrolle (z. B. Machtausübung, Bestrafen, körperliche Züchtigung) und psychologische Kontrolle (z. B. Erzeugen von Schuldgefühlen und Scham, Liebesentzug)

unterschieden werden (Pinquart, 2016, S. 476). Der *typologische Ansatz* greift diese Dimensionen auf und benennt mehrere Typen von Erziehungsstilen (Baumrind, 1991). Unterschieden werden ein autoritativer Erziehungsstil (viel emotionale Wärme, viel proaktive Kontrolle), ein autoritärer Erziehungsstil (wenig emotionale Wärme, viel dominante Kontrolle), ein permissiver Erziehungsstil (viel emotionale Wärme, wenig Kontrolle) und ein vernachlässigender Erziehungsstil (wenig emotionale Wärme, wenig Kontrolle).

Zusammenhänge zwischen *sozialem Status* und Erziehungsstil wurden in zahlreichen Studien untersucht, unter anderem im Rahmen der sogenannten schichtspezifischen Sozialisationsforschung (Brake & Büchner, 2012, S. 90 ff.). Die Befundlage ist nicht gänzlich konsistent (Schauenberg, 2007, S. 238), aber gewisse Zusammenhänge wurden wiederholt berichtet. In Studien mit dimensionaler Erfassung des Erziehungsstils sprechen die Befunde insbesondere für Unterschiede bei der Kontrolldimension: statushohe Eltern weisen in ihrer Einflussnahme mehr Konsistenz auf und verwenden seltener dominante Kontrollformen (Gülseven et al., 2018, S. 569; Hosokawa & Katsura, 2017, S. 13; Roubinov & Boyce, 2017, S. 3). In typologischen Studien zeigt sich, dass statushohe Eltern häufiger einen autoritativen Erziehungsstil und seltener einen autoritären oder vernachlässigenden Erziehungsstil einsetzen als statusniedrige Eltern (Hoff et al., 2002, S. 237 f.). Hinsichtlich des permissiven Erziehungsstils sind die Befunde unklar, da er in manchen Studien häufiger in statushohen Familien (Glasgow et al., 1997, S. 516) und in manchen Studien häufiger in statusniedrigen Familien (September et al., 2016, S. 1068) verwendet wird.

Es ist anzunehmen, dass der elterliche Erziehungsstil den *Lernerfolg* von Heranwachsenden durch mehrere Mechanismen beeinflusst. Denkbar ist ein Einfluss auf die Quantität fachlicher Lernaktivitäten (z. B. Lenkung beeinflusst Schuleschwänzen) und auf die Lernvoraussetzungen der Heranwachsenden (z. B. emotionale Wärme beeinflusst Selbstkonzept). Empirische Befunde zu den Auswirkungen des elterlichen Erziehungsstils auf den Erwerb schulischer Kompetenzen wurden durch zahlreiche Studien erbracht. Pinquart (2016) hat eine Meta-Analyse auf der Grundlage von 308 Einzelstudien vorgelegt. Bei dimensionaler Betrachtung des Erziehungsstils zeigt sich, dass emotionale Wärme, das Zugestehen von Autonomie und proaktive Kontrolle positive Zusammenhänge zum Lernerfolg aufweisen, während dominante Kontrolle und psychologische Kontrolle negative Zusammenhänge aufweisen. In Bezug auf Erziehungsstil-Typen hat der autoritative Erziehungsstil positive Effekte und autoritärer, permissiver und vernachlässigender Erziehungsstil jeweils negative Effekte (S. 481 f.).

Zusammenfassend ist festzuhalten, dass sowohl statusspezifische Unterschiede bei bestimmten Aspekten des Erziehungsstils als auch Zusammenhänge zwischen Erziehungsstil und Lernerfolg durch empirische Befunde belegt werden.

3.5.2 Unterschiede bei der kognitiven Anregung durch die Eltern

Da sich der Erziehungsstil konzeptuell auf bereichsübergreifende Aspekte des Erziehungshandelns bezieht, erscheint es sinnvoll, das elterliche Erziehungshandeln zusätzlich aus einer bereichsspezifischen Perspektive zu betrachten (z. B. kognitiver Bereich, emotionaler Bereich, sozialer Bereich). Für den Erwerb von Fachkompetenzen ist primär die kognitive Anregung durch die Eltern von Interesse. *Kognitive Anregung* bezieht sich auf das Ausmaß, in dem Eltern durch Interaktionen und Aktivitäten lernförderliche kognitive Prozesse beim Kind initiieren. Betrachtet werden kann sowohl die allgemeine kognitive Anregung als auch die kognitive Anregung in bestimmten Themenbereichen, z. B. sprachliche Anregung und mathematische Anregung (Kluczniok et al., 2013).

Besonders viel Beachtung in der Forschung haben *statusspezifische Unterschiede* in der *sprachlichen Anregung* durch die Eltern erhalten. Die Befunde belegen mit großer Konsistenz bedeutende Unterschiede zwischen Familien mit verschiedenem sozialem Status. Statushohe Eltern sprechen insgesamt mehr mit ihren Kindern als statusniedrige Eltern, sie verwenden ein vielfältigeres Vokabular (insbesondere mehr bildungssprachliche Begriffe) und sie richten mehr Fragen an ihr Kind (Hoff et al., 2002, S. 234 ff.; Lehl et al., 2012, S. 123; Rowe, 2008, S. 196; Schwab & Lew-Williams, 2016, S. 3). Weitere Unterschiede zugunsten von statushohen Familien sind bei sprachlich anregenden Eltern-Kind-Aktivitäten wie Vorlesen, Wortspielen oder Singen festzustellen (Caro, 2018, S. 11; Kluczniok et al., 2013, S. 430 f.). Auch bezüglich der *mathematischen Anregung* weisen Forschungsbefunde auf systematische Unterschiede hin. So sprechen statushohe Eltern mit ihren Kindern häufiger über Zahlen und Formen und beschäftigen sich gemeinsam mit den Kindern häufiger mit mathematischem Spielmaterial (z. B. Bauklötze, Brettspiele) (Caro, 2018, S. 11; Kluczniok et al., 2013, S. 430 f.; Susperreguy et al., 2020, S. 10 f.).

Die in der Familie erfahrene kognitive Anregung wirkt sich ihrerseits auf kindliche *Lernprozesse* aus. Plausible Mechanismen sind darin zu sehen, dass die Quantität und/oder die Qualität fachlicher Lernaktivitäten gefördert wird. Auch bei bestimmten Lernvoraussetzungen von Heranwachsenden sind günstige Effekte denkbar (z. B. Intelligenz, Fachinteresse). Lernförderliche Effekte

der kognitiven Anregung können sowohl bezüglich sprachlicher Kompetenzen (Lehrl et al., 2012, S. 123; Rowe, 2008, S. 196; Schmerse et al., 2018, S. 346) als auch bezüglich mathematischer Kompetenzen (Anders et al., 2012, S. 237; Susperreguy et al., 2020, S. 10 f.) belegt werden, insbesondere in der frühen Kindheit.

Resümierend ist festzustellen, dass sowohl für statusspezifische Unterschiede in der kognitiven Anregung als auch für lernförderliche Effekte der kognitiven Anregung Evidenzen vorliegen. In einer jüngeren Studie wird die Annahme, dass die kognitive Anregung einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt, explizit mithilfe eines *Pfadmodells* überprüft und bestätigt (Linberg et al., 2020, S. 23 f.).

3.5.3 Unterschiede bei der schulischen Beteiligung der Eltern

Ab dem Schuleintritt wird das schulische Lernen der Kinder zu einem wichtigen Thema der familiären Interaktionen. Die *schulische Beteiligung* der Eltern umfasst als Teilaspekte die Begleitung schulischer Lernaktivitäten (z. B. Hausaufgaben), Gespräche über die schulischen Erfahrungen der Kinder (z. B. Zeigen von Interesse) und den Kontakt zur Schule (z. B. Elternabende) (Castro et al., 2015, S. 34; Tazuti & Jarlégan, 2019, S. 251 f.). Besonders viel Beachtung in der Forschung hat die elterliche *Hausaufgabenbegleitung* gefunden, da dies die typischste Form der Schulbeteiligung von Eltern darstellt (Tazuti & Jarlégan, 2019, S. 252). Konzeptuell ist als quantitatives Merkmal die Häufigkeit bzw. der Umfang der Hausaufgabenbegleitung und als qualitatives Merkmal die Art und Weise der Hausaufgabenbegleitung zu unterscheiden. Die Qualität der Hausaufgabenbegleitung umfasst ihrerseits mehrere Aspekte (Dumont, 2012, S. 84 f., 111 f.; Helmke et al., 1991, S. 10). Eine autonomieförderliche Unterstützung ist dadurch gekennzeichnet, dass die Initiierung von Hilfe dem Kind überlassen wird, und kann somit von ungefragter Einmischung abgegrenzt werden. Bei einer prozessbezogenen Begleitung unterstützen die Eltern die Verstehensprozesse der Kinder, während sie bei einer produktbezogenen Begleitung kontrollieren, ob die Hausaufgaben (richtig) erledigt wurden. Eine strukturierende Begleitung besteht in klaren Vorgaben, wann und wo das Kind die Hausaufgaben erledigen soll (z. B. nicht vor dem laufenden Fernseher). Als emotionaler Aspekt kann außerdem betrachtet werden, inwieweit es im Rahmen der Hausaufgabenerledigung zu Streit zwischen Eltern und Kind kommt.

Vielfach findet sich in der Literatur die Annahme, dass der *Sozialstatus* der Eltern ihre Hausaufgabenbegleitung bedingt (Dumont, 2012, S. 79; Trautwein et al., 2001, S. 705 f.). Die empirischen Befunde sind diesbezüglich allerdings

inkonsistent. Wird die Häufigkeit der elterlichen Hausaufgabenbegleitung betrachtet, zeigen sich in manchen Studien keine statusspezifischen Unterschiede (Autorengruppe Bildungsberichterstattung, 2014, S. 84; Dumont, 2012, S. 87; Süßlin, 2015, S. 11 f.), in anderen Studien zeigt sich, dass statusniedrige Eltern ihre Kinder häufiger bei ihren Hausaufgaben begleiten (Ehmke et al., 2006b, S. 240; Schauenberg, 2007, S. 244). Bei der Interpretation dieser Befunde ist zu berücksichtigen, dass die Häufigkeit der Hausaufgabenbegleitung auch vom Kompetenzstand des Kindes abhängt, d. h., leistungsschwächere Schüler/-innen werden häufiger bei den Hausaufgaben begleitet (Dumont, 2012, S. 106). Dies stellt eine plausible Erklärung für die in manchen Studien berichtete höhere Häufigkeit der Hausaufgabenbegleitung durch statusniedrige Eltern dar. Aufschlussreich wären insofern Analysen, in denen der Zusammenhang zwischen sozialem Status und Häufigkeit der Hausaufgabenhilfe unter Kontrolle des kindlichen Kompetenzstandes ermittelt wird. Derartige Analysen wurden im Rahmen der Literaturrecherche allerdings nicht gefunden. In Bezug auf die Qualität der Hausaufgabenbegleitung berichten mehrere Studien, dass statushohe Eltern ihre Kinder häufiger in einer autonomieförderlichen Art unterstützen (Cooper et al., 2000, S. 478 ff.; Dumont, 2012, S. 86 f., 179). Hinsichtlich der Kontrolle der Hausaufgabenerledigung sind die Befunde inkonsistent (Dumont, 2012, S. 120; Schauenberg, 2007, S. 245; Trautwein et al., 2001, S. 714 f.), ebenso wie in Bezug auf Streit während der Hausaufgabenerledigung (Dumont, 2012, S. 121, 148, 172).

Effekte der Hausaufgabenbegleitung der Eltern auf den *Lernerfolg* der Schüler/-innen sind theoretisch zu erwarten, weil die Hausaufgabenbegleitung einen Einfluss auf die Quantität und/oder der Qualität fachlicher Lernaktivitäten hat. Daneben ist ein förderlicher Effekt auf die Lernmotivation denkbar. Die Befundlage ist allerdings sehr heterogen. Die Hausaufgabenbegleitung hat in Meta-Analysen teilweise positive Effekte, teilweise negative Effekte und teilweise keine Effekte (Castro et al., 2015, S. 39). Diese inkonsistente Befundlage ist möglicherweise darauf zurückzuführen, dass in Meta-Analysen in der Regel nicht zwischen der Quantität und der Qualität der Hausaufgabenbegleitung unterschieden wird. Für die Quantität der Hausaufgabenbegleitung wurden mehrfach negative Zusammenhänge zu Schülerkompetenzen ermittelt (Klieme et al., 2006, S. 139; Roßbach, 1995, S. 109; Schauenberg, 2007, S. 246). Eine naheliegende Erklärung hierfür ist ein reziproker Effekt, d. h., die Eltern reagieren auf schlechte Kompetenzen durch eine intensivere Hausaufgabenbegleitung (Wilder, 2014, S. 392). Hinsichtlich der Qualität der Hausaufgabenbegleitung zeigen sich positive Effekte einer autonomieförderlichen, prozessbezogenen und strukturierenden Begleitung, während Einmischung, Kontrolle und Streit

negative Effekte haben (Dumont, 2012, S. 204; Helmke et al., 1991, S. 13; Trautwein et al., 2001, S. 717; Patall et al., 2008, S. 1085).

Zusammenfassend finden sich nur teilweise Evidenzen für die Annahme, dass herkunftsbedingte Kompetenzunterschiede auf Unterschiede in der schulischen Beteiligung der Eltern zurückzuführen sind, allerdings spielen hierbei vermutlich auch methodische Schwierigkeiten eine Rolle.

3.6 Unterschiede bei frühpädagogischen Einrichtungen

Die Bedeutung der Eltern erstreckt sich nicht auf die Gestaltung der familiären Lernumwelt. Daneben üben die Eltern einen wesentlichen Einfluss auf die kindliche Entwicklung aus, indem sie gezielt Bildungseinrichtungen für ihre Kinder auswählen. Nachfolgend werden empirische Befunde dargestellt, inwieweit Kinder verschiedener sozialer Herkunft unterschiedliche Bildungseinrichtungen nutzen und wie sich Bildungseinrichtungen auf den Lernerfolg auswirken. Die Nutzung von Bildungseinrichtungen hängt wesentlich davon ab, welche Einrichtungen am Wohnort der Familie vorhanden sind (Gelegenheitsstrukturen). Innerhalb des vorhandenen Angebots haben Familien gewisse Entscheidungsspielräume. Entscheidungsprozesse dieser Art werden durch das Konzept der *Bildungsentscheidungen* beschrieben (Boudon, 1974). Aus dieser Perspektive stellt die Auswahl von Bildungseinrichtungen eine rationale Entscheidung dar. Familien verschiedener sozialer Herkunft unterscheiden sich hierbei darin, wie sie den Nutzen, die Kosten und die Erfolgswahrscheinlichkeit des Besuchs einer Bildungseinrichtung einschätzen (Schauenberg, 2007, S. 147 ff.; Stocké, 2007, S. 512 ff.). Im Weiteren wird auf Unterschiede bei frühpädagogischen Einrichtungen, Schulen, schulbegleitenden Einrichtungen und Freizeiteinrichtungen eingegangen.

Die Frage, ob bei der Nutzung von frühpädagogischen Einrichtungen Unterschiede nach dem *sozialen Status* bestehen, wurde vielfach untersucht. Unterschiede können hierbei sowohl bei der Nutzungsquote und -dauer (quantitative Aspekte) als auch bei der Art der besuchten Einrichtung (qualitative Aspekte) bestehen. Hinsichtlich der Nutzungsquote und -dauer haben Studien in Deutschland wiederholt Nutzungsunterschiede aufgezeigt (Autorengruppe Bildungsberichterstattung, 2016, S. 60; Burghardt, 2018, S. 887; Geier & Riedel, 2008, S. 18 f.; Fuchs-Rechlin & Bergmann, 2014, S. 106 ff.). Demnach nutzen statushohe Familien häufiger und länger frühpädagogische Einrichtungen als statusniedrige Familien. Die Unterschiede sind bei Kindern bis zum Alter von drei Jahren deutlich ausgeprägt, während bei Kindern über drei Jahren nur geringe Unterschiede bestehen. Die Befundlage zu qualitativen Unterschieden

zwischen den von verschiedenen Statusgruppen genutzten Einrichtungen fällt weniger einheitlich aus. In deutschen Studien zeigen sich vorrangig Unterschiede bezüglich der Komposition der frühpädagogischen Einrichtungen. Kinder von statushohen Eltern besuchen häufiger Einrichtungen mit einem hohen Anteil an sozial privilegierten Kindern und einem niedrigen Anteil an Kindern mit Migrationshintergrund (Becker, 2010a, S. 37 f.; Becker & Schober, 2017, S. 1023 f.). Bei der Prozessqualität der Einrichtungen, also den Interaktionen zwischen Erzieher/-innen und Kindern, werden in deutschen Studien keine herkunftsspezifischen Unterschiede gefunden (Becker & Schober, 2017, S. 1023 f.; Lehl et al., 2014, S. 16), während internationale Studien Qualitätsunterschiede zugunsten von sozial privilegierten Kindern berichten (McCartney et al., 2007, S. 7; Pianta et al., 2005, S. 153; Torquati et al., 2011, S. 460).

Hinsichtlich Effekten von frühpädagogischen Einrichtungen auf den Erwerb von *Fachkompetenzen* kommen unterschiedliche Mechanismen in Frage. Plausibel erscheint die Erhöhung der Quantität fachlicher Lernaktivitäten, insbesondere in Bezug auf bereichsspezifische Vorläuferfähigkeiten (z. B. phonologische Bewusstheit, Mengenvergleich). Zudem können sich frühpädagogische Einrichtungen günstig auf Lernvoraussetzungen wie Intelligenz oder Selbstregulation auswirken (Protzko et al., 2013, S. 32; Sammons et al., 2014, S. 34 f.). In empirischen Studien zeigt sich weitestgehend konsistent, dass eine längere Besuchsdauer und eine höhere Qualität der Einrichtung mit besseren sprachlichen und mathematischen Vorläuferfähigkeiten einhergehen (Anders et al., 2012, S. 238; McCartney et al., 2007, S. 23; Roßbach et al., 2008, S. 142 ff.; Schmerse et al., 2018, S. 346; Tietze et al., 1998, S. 312 ff.). Die Effekte können dabei teilweise noch bis in die Schulzeit hinein festgestellt werden (Anders et al., 2013, S. 203 ff.; Sammons et al., 2008, S. 187).

Vielfach wird an frühpädagogische Einrichtungen die Erwartung gestellt, dass sie zur Kompensation herkunftsbedingter Kompetenzunterschiede beitragen können. Dies stützt sich insbesondere auf die Annahme *differenzieller Effekte*, wonach sozial benachteiligte Kinder stärker von einem Besuch frühpädagogischer Einrichtungen bzw. von einer hohen Qualität profitieren als sozial privilegierte Kinder (Roßbach et al., 2008, S. 146). Die Befundlage fällt für Besuch und Qualität unterschiedlich aus. Beim Besuch einer frühpädagogischen Einrichtung zeigt sich konsistent, dass sozial benachteiligte Kinder stärker vom Besuch bzw. einer hohen Besuchsdauer profitieren als sozial privilegierte Kinder (Anders, 2013, S. 259; Raudenbush & Eschmann, 2015, S. 455 ff.; Roßbach et al., 2008, S. 146 f.). Bezüglich differenzieller Effekte der Qualität von frühpädagogischen Einrichtungen fällt die Befundlage dagegen inkonsistent aus. Teilweise hat eine hohe Einrichtungsqualität stärkere Effekte bei sozial benach-

teiligten Kindern (McCartney et al., 2007, S. 8 f.; Peisner-Feinberg et al., 2001, S. 1544 f.), teilweise bestehen keine Unterschiede (ECCE Study Group, 1999, S. 220; Sammons et al., 2002, S. 62), teilweise hat eine hohe Einrichtungsqualität stärkere Effekte bei sozial privilegierten Kindern bzw. bei Kindern, die viel familiäre Anregung erfahren (Anders et al., 2012, S. 239 f.; Schmerse et al., 2018, S. 350).

Insgesamt betrachtet zeigen sich im frühpädagogischen Bereich Herkunftsunterschiede bei der Besuchsdauer und teilweise auch bei der Einrichtungsqualität. Besuchsdauer und Qualität weisen ihrerseits Zusammenhänge zum kindlichen Kompetenzerwerb auf. Weiterhin liegen Hinweise auf differenzielle Effekte vor, wobei die Befunde beim Besuch relativ konsistent ausfallen, bei der Qualität dagegen inkonsistent. Beim Besuch von frühpädagogischen Einrichtungen liegen somit Hinweise auf gegenläufige Mechanismen vor (Unterschiede beim Besuch vs. differenzielle Effekte des Besuchs). Der Gesamteffekt frühpädagogischer Einrichtungen, der sich daraus ergibt, wie stark die beiden Mechanismen jeweils ins Gewicht fallen, kann auf der Grundlage der vorhandenen Studien nicht abschließend eingeschätzt werden.

3.7 Unterschiede bei Schulen

Der Bildungseinrichtung Schule kommt eine zentrale Bedeutung für den Bildungserfolg von Heranwachsenden zu. Insofern stellt sich die Frage, inwieweit die Schule selbst zu herkunftsbedingten Bildungsungleichheiten beiträgt. Dies wird insbesondere von der Reproduktionstheorie angenommen (Bourdieu & Passeron, 1977). Konzeptuell sind *Unterschiede zwischen Schulen* (d. h., Schüler/-innen verschiedener sozialer Herkunft besuchen unterschiedliche Schulen) und *Unterschiede innerhalb von Schulen* (d. Schüler/-innen verschiedener sozialer Herkunft machen innerhalb der gleichen Schulen unterschiedliche Lernerfahrungen) voneinander abzugrenzen (Maaz et al., 2010, S. 13). Nachfolgend werden relevante Forschungsbefunde in den zwei Bereichen dargestellt.

3.7.1 Unterschiede in den besuchten Schulen

Zahlreiche Studien belegen, dass sich Schüler/-innen verschiedener sozialer Herkunft nicht gleichmäßig auf die vorhandenen Schulen verteilen. Hierfür spielt zunächst einmal die *soziale Segregation* in Städten eine wichtige Rolle (Goebel & Hoppe, 2015, S. 8). Besonders bei Grundschulen, bei denen die Zuordnung von Schüler/-innen in Deutschland anhand des Wohnorts erfolgt,

führt dies dazu, dass sich *Schulen in ihrer sozialen Schülerkomposition unterscheiden* (Stancel-Piatak, 2017, S. 188). Die soziale Schülerkomposition stellt ihrerseits einen wichtigen *Einflussfaktor auf den Lernerfolg* der Schüler/-innen dar (Holzberger et al., 2020, S. 11).

In der Sekundarschule sind in Deutschland vor allem die verschiedenen *Schulformen* von Bedeutung. Wie bereits erwähnt wurde, besuchen sozial privilegierte Schüler/-innen auch bei gleichem Kompetenzstand *häufiger ein Gymnasium* als sozial benachteiligte Schüler/-innen (Baumert & Schümer, 2001, S. 357 f.; Ditton et al., 2005, S. 296; Zielonka et al., 2014, S. 121). Weiterhin gibt es Hinweise darauf, dass *Schulformen unterschiedlich anregende Lernmilieus* darstellen. Mehrere Studien zeigen, dass Schüler/-innen in Gymnasien einen größeren Lernerfolg aufweisen als vergleichbare Schüler/-innen in anderen Schulformen (Neumann et al., 2014, S. 187 f.). Die Befundlage ist nicht gänzlich konsistent, aber insbesondere für die Fächer Mathematik, Naturwissenschaften und Englisch liegen klare Evidenzen vor. Als Erklärung für dieses Befundmuster kommen neben Kompositionseffekten (d. h., Schulformen unterscheiden sich in ihrer Schülerkomposition) auch institutionelle Effekte (d. h., Schulformen unterscheiden sich in Lehrplänen, Unterrichtsqualität, Qualifikationen von Lehrkräften) in Betracht (Maaz et al., 2010, S. 30; Neumann et al., 2014, S. 188).

Eine weitere Möglichkeit für Unterschiede zwischen Schulen besteht in Herkunftseffekten bei der *Nutzung von Privatschulen*. Mit großer Konsistenz zeigt sich, dass sozial privilegierte Schüler/-innen *häufiger Privatschulen* besuchen als sozial benachteiligte Schüler/-innen (Görlitz et al., 2018, S. 1107 f.; Jungbauer-Gans et al., 2012, S. 72; Köppe, 2012, S. 210 f.). Hinsichtlich der *Wirksamkeit von Privatschulen* fallen die Befunde nicht ganz eindeutig aus. Es finden sich zwar Hinweise darauf, dass Schüler/-innen an Privatschulen bessere Kompetenzen aufweisen als Schüler/-innen an öffentlichen Schulen, aber dies scheint primär an günstigeren Ausgangsbedingungen und Effekten der Schülerkomposition zu liegen (und nicht an der pädagogischen Qualität) (Dronkers & Avram, 2009, 904 f.; Hoffmann et al., 2019, S. 399 ff.; Jungbauer-Gans et al., 2012, S. 77 ff.).

Insgesamt betrachtet sprechen die Befunde dafür, dass Unterschiede in den besuchten Schulen Mechanismen zwischen sozialer Herkunft und Lernerfolg darstellen, wobei insbesondere Unterschiede in der Schülerkomposition und Unterschiede in der Schulform eine Rolle spielen.

3.7.2 Unterschiede in schulischen Lernaktivitäten

Auch wenn Schüler/-innen verschiedener sozialer Herkunft die gleichen Schulen besuchen, ist es möglich, dass sie systematisch unterschiedliche Lernerfahrungen sammeln. Entsprechend der Angebots-Nutzungs-Struktur von Schule können unterschiedliche Lernerfahrungen sowohl auf Angebotsunterschiede als auch auf Nutzungsunterschiede zurückzuführen sein (Maaz et al., 2010, S. 23). *Angebotsunterschiede* beziehen sich darauf, dass Lehrkräfte Schüler/-innen verschiedener sozialer Herkunft systematisch unterschiedlich behandeln. *Nutzungsunterschiede* beziehen sich darauf, dass Schüler/-innen verschiedener sozialer Herkunft aus den vorhandenen Möglichkeiten systematisch unterschiedliche Lernaktivitäten auswählen. Im theoretischen Modell verweist dies auf die mediierende Rolle der Kategorie „Lernaktivitäten des/der Heranwachsenden“. Unterschiede in den schulischen Lernaktivitäten sind für die vorliegende Arbeit von besonderem Interesse und werden daher tiefergehend behandelt. Als Aspekte schulischer Lernaktivitäten werden die Anwesenheit im Unterricht und die Nutzung von Ganztagsangeboten betrachtet.

Die *Anwesenheit im Unterricht* stellt ein gut untersuchtes Forschungsfeld dar. Die Regelmäßigkeit des Schulbesuchs bzw. die Häufigkeit von Fehlzeiten stellt einen grundlegenden Aspekt der Nutzung von schulischen Lernangeboten durch Schüler/-innen dar. Fehlzeiten können hierbei sowohl auf gesundheitliche Probleme als auch auf bewusstes Fernbleiben („Schwänzen“) zurückzuführen sein. Zusammenhänge zwischen Anwesenheit im Unterricht und *sozialer Herkunft* wurden mehrfach belegt. In einer jüngeren Meta-Analyse berichten Gubbels und Kolleg/-innen (2019), dass sozial benachteiligte Schüler/-innen der Schule signifikant häufiger fernbleiben als sozial privilegierte Schüler/-innen (S. 1657). Die Anwesenheit im Unterricht bedingt ihrerseits die Quantität fachlicher Lernaktivitäten, weshalb ein Effekt auf den *Lernerfolg* naheliegt. Mehrere Studien belegen erwartungsgemäß, dass höhere Fehlzeiten mit niedrigeren Schülerkompetenzen einhergehen (Helmke, 2007a, S. 106; Petrides et al., 2005, S. 248; Strand, 2014, S. 235). Die Anwesenheit im Unterricht erfüllt somit die empirischen Voraussetzungen für einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg.

Ein weiterer Bereich schulischer Lernaktivitäten ist die *Nutzung von Ganztagsangeboten*. Im Zuge des Ausbaus von Ganztagschulen erfährt dieses Thema zunehmend Aufmerksamkeit, auch unter dem Gesichtspunkt von *Nutzungsunterschieden nach sozialer Herkunft*. Die empirische Befundlage fällt heterogen aus. In einigen Studien wird ermittelt, dass sozial privilegierte Schüler/-innen häufiger Ganztagsangebote nutzen als sozial benachteiligte Schüler/-innen (Züchner & Fischer, 2014, S. 353). Demgegenüber berichten

andere Studien, dass sozial benachteiligte Schüler/-innen Ganztagsangebote häufiger nutzen als sozial privilegierte Schüler/-innen (Holtappels et al., 2010, S. 182). Interessant sind in diesem Kontext Analysen, die zwischen verschiedenen Ganztagsangeboten differenzieren. Demnach werden Hausaufgabenbetreuungen und Förderunterricht häufiger von leistungsschwachen und sozial benachteiligten Schüler/-innen besucht, während Freizeitangebote häufiger von leistungsstarken und sozial privilegierten Schüler/-innen besucht werden (Linberg et al., 2015, S. 33 ff.; Züchner & Fischer, 2014, S. 354).

Bezüglich der *Wirkweise von Ganztagsangeboten* ist von unterschiedlichen Mechanismen auszugehen. Je nach Ganztagsangebot ist sowohl eine Erhöhung der Quantität fachlicher Lernaktivitäten (z. B. Förderunterricht) als auch eine Förderung von Lernvoraussetzungen (z. B. Freizeitangebote) denkbar. Die vorhandenen empirischen Befunde fallen unterschiedlich aus. Wird die Nutzung von Ganztagsangeboten allgemein betrachtet (d. h. unabhängig von der Art der Ganztagsangebote), zeigt sich mehrfach, dass Schüler/-innen, die an Ganztagsangeboten teilnehmen, einen höheren Lernerfolg aufweisen als Schüler/-innen, die nicht an Ganztagsangeboten teilnehmen (Schüb bach, 2014, S. 107; Schüb bach et al., 2013, S. 164). In der letztgenannten Studie werden auch differenzielle Effekte der Nutzung von Ganztagsangeboten überprüft. Es zeigt sich, dass Schüler/-innen verschiedener sozialer Herkunft gleichermaßen von Ganztagsangeboten profitieren, aber dass Schüler/-innen aus Familien mit hoher Anregungsqualität stärker von Ganztagsangeboten profitieren als Schüler/-innen aus Familien mit niedriger Anregungsqualität (Schüb bach, 2014, S. 108 f.). In einer weiteren Studie wird ermittelt, dass sich eine hohe Nutzungsintensität (Anzahl der Tage pro Woche) günstig auf den Lernerfolg auswirkt (Kuhn & Fischer, 2011, S. 221).

Bei der Betrachtung von *Förderunterricht* und *Hausaufgabenbetreuungen* sind die Befunde inkonsistent. Teilweise werden positive Zusammenhänge zwischen einer Nutzung und Schülerkompetenzen ermittelt, teilweise keine Zusammenhänge und teilweise negative Zusammenhänge (Bellin & Wegner, 2010, S. 92; Linberg et al., 2018, S. 1219 ff.; Steinmann et al., 2019, S. 165 ff.). Die negativen Zusammenhänge sind wahrscheinlich auf Selektionseffekte zurückzuführen (d. h., leistungsschwächere Schüler/-innen besuchen häufiger Förderunterricht und Hausaufgabenbetreuungen), welche durch korrelative Studien nur bedingt kontrolliert werden können.

Insgesamt betrachtet sprechen empirische Befunde dafür, dass die Schulanwesenheit einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt. Bei der Nutzung von Ganztagsangeboten kann dies gegen-

wärtig nicht abschließend beurteilt werden, da sowohl die Befunde zu Nutzungsunterschieden als auch die Befunde zu Lerneffekten heterogen ausfallen.

3.8 Unterschiede bei schulbezogenen Einrichtungen

Eine weitere Kategorie von Bildungseinrichtungen, die von Eltern gezielt ausgewählt werden, sind schulbezogene Einrichtungen wie Kinderhorte, Hausaufgabenbetreuungen und Nachhilfe. Im Kontext von Bildungsungleichheiten hat insbesondere Nachhilfe Beachtung in der Forschung gefunden. Unter Nachhilfe werden außerschulische Unterstützungsmaßnahmen, welche auf die Verbesserung von Fachkompetenzen abzielen und in der Regel kostenpflichtig sind, verstanden (Klemm & Hollenbach-Biele, 2016, S. 10 f.).

Die Annahme, dass *Unterschiede in der Nutzung von Nachhilfe* einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen, wird in der Literatur mehrfach formuliert (Hille et al., 2016, S. 112 f.; Solga & Dombrowski, 2009, S. 22). Die empirischen Belege für diesen Mechanismus sind allerdings nicht eindeutig. Dies betrifft zunächst einmal herkunftsspezifische Unterschiede in der Nutzung von Nachhilfe. Die Befundlage für Deutschland ist ausgesprochen inkonsistent (Dohmen et al., 2008, S. 44 ff.; Klemm & Hollenbach-Biele, 2016, S. 14 f.). In einigen Studien nutzen sozial privilegierte Schüler/-innen häufiger Nachhilfe (Hille et al., 2016, S. 115; Schneider, 2005, S. 372 f.), in anderen Studien nutzen sozial privilegierte Schüler/-innen seltener Nachhilfe (Abele & Liebau, 1998, S. 42 f.; Guill & Wendt, 2016, S. 253; Luplow & Schneider, 2014, S. 40). Unter theoretischen Gesichtspunkten lassen sich gegenläufige Argumente für Nutzungsunterschiede anbringen. Einerseits verfügen statushohe Familien über höhere Bildungsaspirationen und mehr finanzielle Ressourcen, was für eine höhere Bereitschaft für Nachhilfe spricht. Andererseits haben statushohe Familien einen niedrigeren Bedarf an Nachhilfe, weil die Eltern ihre Kinder bei Lernproblemen häufiger selbst unterstützen können und weil die Kinder seltener schlechte Kompetenzen aufweisen. Bezüglich des letztgenannten Aspekts sind Studien interessant, in denen der Zusammenhang zwischen sozialem Status und Nachhilfe unter Kontrolle des Kompetenzstandes des Kindes analysiert wird. In zwei Studien wird bei gleichem Kompetenzstand kein Zusammenhang zwischen sozialem Status und Nachhilfe ermittelt (Guill & Wendt, 2016, S. 253; Luplow & Schneider, 2014, S. 42), während in einer Studie sozial privilegierte Schüler/-innen bei gleichen Kompetenzen häufiger Nachhilfe erhalten als sozial benachteiligte Schüler/-innen (Schneider, 2005, S. 376 f.). Klare Befunde, dass Schüler/-innen aus statushohen Familien häufiger Nachhilfe nutzen, liegen für Deutschland somit nicht

vor. Im Gegensatz dazu zeigen sich in internationalen Studien mit größerer Konsistenz Nutzungsunterschiede zugunsten von statushohen Familien (Park et al., 2016, S. 236 f.).

Für die *Wirksamkeit von Nachhilfe* spricht aus einer theoretischen Perspektive insbesondere, dass durch Nachhilfe die Quantität fachlicher Lernaktivitäten erhöht wird. Trotz dieses plausiblen theoretischen Mechanismus fallen empirische Ergebnisse zur Wirksamkeit von Nachhilfe nicht eindeutig aus. In deutschen Längsschnittstudien werden sowohl positive Effekte (Haag, 2001; Mischo & Haag, 2002, S. 269) als auch keine Effekte (Guill & Bos, 2014, S. 52; Hosenfeld, 2011; Luplow & Schneider, 2014, S. 43 ff.) und negative Effekte (Guill & Bos, 2014, S. 51; Guill et al., 2020, S. 289 ff.; Ömerogullari et al., 2020, S. 5 f.) berichtet. Allerdings handelt es sich bei allen drei Studien, die negative Effekte berichten, um korrelative Studien. Insofern erscheint es plausibel, dass der negative Zusammenhang auf einen Selektionseffekt zurückzuführen ist (leistungsschwache Schüler/-innen erhalten häufiger Nachhilfe), welcher durch die in den Studien berücksichtigten Variablen nicht vollständig kontrolliert wird (Ömerogullari et al., 2020, S. 6). Für diese Interpretation spricht auch, dass sich in zwei Studien mit Propensity-Score-Matching, einem statistischen Verfahren durch welches Selektionseffekte gut kontrolliert werden können, keine negativen Effekte zeigen (Guill & Bos, 2014, S. 52; Luplow & Schneider, 2014, S. 43 ff.). Der internationale Forschungsstand wird in Übersichtsarbeiten ebenfalls als heterogen eingeschätzt, wobei manche Studien positive Effekte berichten und andere Studien keine Effekte (Park et al., 2016, S. 241).

Insgesamt betrachtet ist sowohl der Zusammenhang zwischen sozialem Status und Nachhilfe als auch der Zusammenhang zwischen Nachhilfe und Lernerfolg nicht eindeutig belegt. Möglicherweise ist dies darauf zurückzuführen, dass bei beiden Zusammenhängen auch der Kompetenzstand eine Rolle spielt und folglich komplexe Analysestrategien benötigt werden.

3.9 Unterschiede bei Freizeiteinrichtungen und Freizeitaktivitäten

Der Alltag von Schüler/-innen wird neben dem schulischen Lernen wesentlich durch ihre Freizeit geprägt. Konzeptuell wird im theoretischen Modell zwischen dem Zugang zu Freizeiteinrichtungen als Aspekt der Lernumwelten (z.B. Mitgliedschaft in einem Sportverein) und den tatsächlich ausgeführten Freizeitaktivitäten als Aspekt der Lernaktivitäten (z.B. Häufigkeit von Sport) unterschieden. Die beiden Kategorien werden nachfolgend gemeinsam betrachtet.

Zu *herkunftsspezifischen Unterschieden* in den Freizeitaktivitäten von Heranwachsenden liegen zahlreiche Studien vor, die eine konsistente Befundlage erbracht haben. Kinder aus statushohen Familien lesen häufiger, machen häufiger Sport bzw. sind häufiger Mitglied in einem Sportverein, spielen häufiger ein Musikinstrument bzw. besuchen häufiger eine Musikschule, schauen seltener fern und spielen seltener Videospiele (Autorengruppe Bildungsberichterstattung, 2014, S. 49, 85; DeGarmo et al., 1999, S. 1237 ff.; Leven & Schneekloth, 2010, S. 108 ff.; Schauenberg, 2007, S. 226). Diese Unterschiede in den Freizeitaktivitäten werden maßgeblich durch das Erziehungshandeln der Eltern bedingt (Leven & Schneekloth, 2010, S. 120, 135; Schaffner et al., 2013, S. 136; Süßlin, 2015, S. 8).

Hinsichtlich der Effekte von Freizeitaktivitäten auf den *Lernerfolg* ist anzunehmen, dass diese primär über die Beeinflussung von Lernvoraussetzungen wirken (z. B. Konzentrationsfähigkeit, Sprachkompetenzen). Eine Ausnahme stellt das Lesen dar, welches für den Erwerb von Lesekompetenz eine fachliche Lernaktivität darstellt. In empirischen Studien können lernförderliche Effekte können sowohl für Lesen (Pfof et al., 2010, S. 171 f.; Stutz et al., 2016, S. 109) als auch für Musizieren bestätigt werden (Vaughn, 2000, S. 151 ff.). Bei Sport ist die Befundlage weniger eindeutig, da einige Studien schwache positive Zusammenhänge zwischen sportlicher Aktivität und Fachkompetenzen berichten, andere Studien dagegen keine Zusammenhänge (Singh et al., 2012; Taras, 2005). Für Fernsehen (Williams et al., 1982) und Videospiele (Adelantado-Renau et al., 2019) zeigen sich in Meta-Analysen jeweils negative Zusammenhänge zu Fachkompetenzen.

Insgesamt betrachtet decken sich die Befunde mit der Annahme, dass Freizeitaktivitäten einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen.

3.10 Unterschiede bei Gleichaltrigengruppen

Neben der Familie und Bildungseinrichtungen wird die kindliche Entwicklung maßgeblich von der Gleichaltrigengruppe geprägt. Zur Gleichaltrigengruppe eines Kindes gehören alle Heranwachsenden in einem ähnlichen Alter, mit dem das Kind mehr oder weniger regelmäßigen Kontakt hat. Die Kontakte zwischen Gleichaltrigen erfolgen sowohl im Kontext von Bildungseinrichtungen (z. B. Schule, Sportverein) als auch in familiären Kontexten oder ganz ohne Betreuung Erwachsener.

Herkunftseffekte auf die Gleichaltrigengruppe können sich sowohl auf die Anzahl der Freund/-innen (quantitatives Merkmal) als auch auf die Art der

Freund/-innen beziehen (qualitatives Merkmal). Hinsichtlich der Anzahl der Freund/-innen berichteten mehrere Studien, dass sozial privilegierte Schüler/-innen im Durchschnitt über etwas größere Freundeskreise verfügen als sozial benachteiligte Schüler/-innen (Delgado et al., 2016, S. 1117; Hjalmarson & Mood, 2015, S. 205; Süßlin, 2015, S. 8). In qualitativer Hinsicht bestehen ebenfalls Unterschiede. Demnach entwickeln sich häufiger Freundschaften zwischen Heranwachsenden mit ähnlicher sozialer Herkunft als zwischen Heranwachsenden mit verschiedener sozialer Herkunft (Burgess et al., 2011, S. 17; Kramer & Wagner, 2012, S. 221 ff.). Die Unterschiede in der sozialen Herkunft der Freund/-innen gehen mit Unterschieden in den Persönlichkeitsmerkmalen der Freund/-innen einher. Beispielsweise haben sozial privilegierte Schüler/-innen häufiger Freund/-innen mit positiven Einstellungen zur Schule und hohen Bildungsaspirationen (Behtui & Neergaard, 2016, S. 956; Carolan, 2016, S. 1046).

Hinsichtlich der Art und Weise, wie die Gleichaltrigengruppe den *Lernerfolg* beeinflusst, erscheinen mehrere Mechanismen plausibel. Ein direkter Bezug zum fachlichen Lernen ist gegeben, wenn Heranwachsende gemeinsam Schularbeiten erledigen. Daneben spielen die Gespräche der Heranwachsenden über die Schule eine Rolle, da sie einen Einfluss auf die Lernmotivation haben können. Schließlich werden auch die ausgeübten Freizeitaktivitäten von Freund/-innen beeinflusst. Im Einklang mit diesen Überlegungen werden Effekte der Gleichaltrigengruppe auf den Lernerfolg durch empirische Befunde bestätigt. Der Zusammenhang zwischen der Anzahl der Freunde und dem Lernerfolg ist gut untersucht. Wentzel und Kolleg/-innen (2018) haben die Befunde meta-analytisch ausgewertet und einen positiven Effekt ermittelt (S. 1255). Ebenfalls viel Aufmerksamkeit in der Forschung hat die Bedeutung der Einstellungen der Freund/-innen gefunden. Konsistent zeigt sich, dass Heranwachsende über bessere Kompetenzen verfügen, wenn ihre Freund/-innen positive Einstellungen zur Schule und hohe Bildungsaspirationen haben (Behtui & Neergaard, 2016, S. 961; Johnson, 2000, S. 6 f.; Schmitt, 2012, S. 128 f., 159 f.).

Zusammengenommen stützt die Befundlage die Annahme, dass die Gleichaltrigengruppe einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt.

3.11 Unterschiede bei Lernvoraussetzungen

Die in den vorausgehenden Abschnitten berichteten Unterschiede in Lernumwelten und Lernaktivitäten führen dazu, dass sich Heranwachsende verschiedener sozialer Herkunft in ihren Lernvoraussetzungen unterscheiden. In der

Literatur herrscht Einigkeit darüber, dass bereits bei Schuleintritt bedeutende herkunftsspezifische Unterschiede zwischen den Schüler/-innen bestehen (Ditton, 2011, S. 245; Maaz et al., 2010, S. 32; Neumann et al., 2014, S. 170). Im Weiteren wird erörtert, inwieweit empirische Befunde Zusammenhänge zwischen sozialer Herkunft und Lernvoraussetzungen einerseits und Zusammenhänge zwischen Lernvoraussetzungen und Lernerfolg andererseits belegen. Unterteilt werden die Lernvoraussetzungen in kognitive Merkmale, motivationale Merkmale und psychische Dispositionen.

3.11.1 Unterschiede bei kognitiven Merkmalen der Heranwachsenden

Nachfolgend werden Merkmale behandelt, welche verschiedene Aspekte des kognitiven Leistungspotenzials von Kindern beschreiben. Hierzu zählen mathematische Ausgangskompetenzen, Intelligenz, metakognitive Kompetenzen und sprachliche Kompetenzen. Bei allen Merkmalen aus dem kognitiven Bereich ist anzunehmen, dass sie insbesondere die Qualität von Lernprozessen beeinflussen, d. h., sie erleichtern die Verarbeitung und Einspeicherung von Lerninhalten.

Kinder verschiedener sozialer Herkunft verfügen bereits bei Schuleintritt über *unterschiedliche mathematische Ausgangskompetenzen*. Unterschiede in mathematischen Ausgangskompetenzen zeigen sich hierbei sowohl beim Eintritt in die Grundschule (Anders et al., 2012, S. 237; Garcia, 2015, S. 4) als auch beim Eintritt in die Sekundarschule (Richter et al., 2012, S. 199 ff.; Stubbe et al., 2016, S. 303 ff.). Dieser Kompetenzvorsprung ist wesentlich für Unterschiede im *Lernerfolg*, da der vorausgehende Kompetenzstand im Allgemeinen den besten Prädiktor für den später erreichten Kompetenzstand darstellt (Ditton & Krüskens, 2009, S. 37). Zahlreiche Studien belegen den starken Zusammenhang zwischen Ausgangskompetenz und nachfolgend erreichter Kompetenz. Einen Überblick über vorhandene Studien und eine meta-analytische Zusammenfassung gibt Marzano (2000, S. 70).

Neben fachspezifischen Ausgangskompetenzen spielt die *Intelligenz* eine zentrale Rolle für den Lernerfolg. Intelligenz bezieht sich nachfolgend primär auf fluide Intelligenz als bereichsübergreifende kognitive Leistungsfähigkeit. Zur fluiden Intelligenz gehören Aspekte wie schlussfolgerndes Denken, Verarbeitungsgeschwindigkeit und Kapazität des Arbeitsgedächtnisses (Weinert et al., 2011, S. 71). Verwandte Konzepte sind kognitive Grundfähigkeit und das inzwischen veraltete Konzept der Begabung (Englisch: aptitude). Die Frage nach *herkunftsbedingten Intelligenzunterschieden* ist hoch politisch und sollte

daher mit Vorsicht betrachtet werden. Empirische Befunde legen nahe, dass in der frühesten Kindheit kaum bedeutsame Unterschiede in kognitiven Vorläuferfähigkeiten bestehen (Weinert et al., 2017), während in der Grund- und Sekundarschulzeit konsistent Unterschiede in der Intelligenz festzustellen sind (Evans & Schamberg, 2009, S. 6546; Ramseier & Brühwiler, 2003, S. 34 ff.; Stahl, 2009, S. 137). Dieses Befundmuster legt nahe, dass nicht Unterschiede in genetischen Faktoren, sondern Unterschiede in Sozialisationsfaktoren maßgeblich sind. Die Intelligenz weist ihrerseits bedeutsame Zusammenhänge zum *Lernerfolg* auf, was durch mehrere Meta-Analysen belegt wird (Marzano, 2000, S. 74; Walberg, 1984, S. 23). Zur Rolle der Intelligenz liegen außerdem mehrere *pfadanalytische Studien* vor. Die Annahme, dass die Intelligenz einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt, wird weitestgehend konsistent bestätigt (Kriegbaum & Spinath, 2016, S. 52; Steinmayr et al., 2010, S. 542; Steinmayr et al., 2012, S. 343). Dieser Befund kann und sollte allerdings nicht zur Legitimation sozialer Ungleichheit genutzt werden (Brake & Büchner, 2012, S. 86 ff.), da die Intelligenz eine, bis zu einem gewissen Grad, veränderbare Größe darstellt (Protzko et al., 2013).

Eine weitere relevante kognitive Lernvoraussetzung stellen die *metakognitiven Kompetenzen* der Heranwachsenden dar. Metakognitionen sind Kognitionen höherer Ordnung wie beispielsweise Denken über Denken oder Wissen über Wissen. Im schulischen Kontext ist insbesondere das Nachdenken über die eigenen Lernaktivitäten relevant. Wichtige metakognitive Kompetenzen sind die Fähigkeit zur Planung eigener Lernaktivitäten, die Fähigkeit zur Überwachung eigener Lernaktivitäten und die Fähigkeit zur Einschätzung des eigenen Lernerfolgs (Muijs et al., 2014, S. 239). In mehreren Studien zeigen sich systematische Unterschiede, wonach sozial privilegierte Schüler/-innen über *bessere metakognitive Kompetenzen* verfügen als sozial benachteiligte Schüler/-innen (Karlen et al., 2014, S. 786 f.; Schneider et al., 2017, S. 295; Yerdelen-Damar & Pesman, 2013, S. 287). Forschungsbefunde zur Bedeutung metakognitiver Kompetenzen für den *Lernerfolg* wurden von Ohtani und Hisasaka (2018) in einer Meta-Analyse zusammengefasst. Es zeigt sich ein positiver Zusammenhang (S. 193), auch wenn die Intelligenz der Schüler/-innen kontrolliert wird (S. 200).

Der Lernerfolg von Kindern wird schließlich auch durch ihre *sprachlichen Kompetenzen* bedingt. Diese beziehen sich auf die Produktion und Rezeption von Sprache in mündlicher und schriftlicher Form (z. B. Wortschatz, Lesekompetenz, Schreibkompetenz). Sozial privilegierte Schüler/-innen verfügen in mehreren Studien über *bessere sprachliche Kompetenzen* als sozial benachteiligte Schüler/-innen (Hufmann et al., 2017, S. 201 ff.; Letourneau et al., 2013,

S. 218 f.; Stahl, 2009, S. 138 f.; Weinert & Ebert, 2013, S. 313 ff.). Da Unterricht in allen Fächern auf Sprache basiert, stellen sprachliche Kompetenzen eine fächerübergreifend bedeutsame Lernvoraussetzung dar. Bezüglich des *Lernerfolgs* in Mathematik kann festgestellt werden, dass sowohl der Wortschatz als auch die Lesekompetenz bedeutsame Zusammenhänge zur Mathematikkompetenz aufweisen (Alloway & Passolunghi, 2011, S. 135 f.; Hornung et al., 2014, S. 12 f.; Kunter et al., 2006, S. 185).

Zusammenfassend ist festzuhalten, dass sich bei allen betrachteten kognitiven Schülermerkmalen konsistente Zusammenhänge sowohl zur sozialen Herkunft als auch zum Lernerfolg zeigen.

3.11.2 Unterschiede bei motivationalen Merkmalen der Heranwachsenden

Neben kognitiven Merkmalen sind motivationale Merkmale der Kinder von wesentlicher Bedeutung für den Lernerfolg. Lernmotivation bezeichnet die Bereitschaft zur Ausführung von Lernaktivitäten. Ein wichtiger theoretischer Ansatz für die Betrachtung von Lernmotivation sind Erwartungs-Wert-Modelle (Eccles & Wigfield, 2002, S. 118 ff.). Gemäß diesem Ansatz weisen Schüler/-innen Lernmotivation auf, wenn das Lernen für sie einen Wert hat (Anreizkomponente) und wenn sie davon ausgehen, dass sie beim Lernen erfolgreich sein werden (Erwartungskomponente). Die Lernmotivation bedingt sowohl die Quantität fachlicher Lernaktivitäten (z. B. Dauer des schulischen Lernens) als auch die Qualität fachlicher Lernaktivitäten (z. B. Anstrengung beim schulischen Lernen).

Ein wichtiges Konzept zur Beschreibung von Lernanreizen ist *Interesse*. Interesse bezieht sich auf die Einstellung einer Person zu einem Gegenstandsbereich (z. B. Mathematik) in der Form von kognitiven und emotionalen Bewertungen (Spinath, 2015, S. 58). Eine positive kognitive Bewertung zeigt sich dadurch, dass Schüler/-innen den Gegenstandsbereich als wichtig oder nützlich ansehen. Eine positive emotionale Bewertung kommt dagegen darin zum Ausdruck, dass sich Schüler/-innen gerne mit dem Gegenstandsbereich beschäftigen. Ein verwandtes Konzept ist der subjektive Aufgabenwert (Eccles & Wigfield, 2002, S. 119 f.). Unter dem Gesichtspunkt von *Herkunftseffekten* ist zunächst relevant, inwieweit sich Schüler/-innen verschiedener sozialer Herkunft in ihrem Interesse bzw. ihrem Aufgabenwert unterscheiden. Mehrfach werden systematische Unterschiede zugunsten von sozial privilegierten Schüler/-innen berichtet (Dotterer et al., 2009, S. 9; OECD, 2007, S. 130 ff.), teilweise aber auch keine bzw. vernachlässigbare Unterschiede (Caro, 2018, S. 11). Der Zu-

sammenhang zwischen Interesse und *Lernerfolg* ist seinerseits gut belegt und meta-analytisch abgesichert (Schiefele et al., 1993, S. 131). Aufschlussreich sind in diesem Kontext *pfadanalytische Studien*, welche eine mediiierende Rolle von Interesse bzw. Aufgabenwert überprüfen und diese weitestgehend konsistent bestätigen (Kriegbaum & Spinath, 2016, S. 52; Steinmayr et al., 2012, S. 342 f.).

Die subjektiven Erfolgserwartungen beim Lernen werden wesentlich durch das *akademische Selbstbild* bedingt. Wichtige Konzepte sind das fähigkeitsbezogene Selbstkonzept, also die subjektive Einschätzung der eigenen Fähigkeiten (Steinmayr et al., 2012, S. 339), und die akademische Selbstwirksamkeitserwartung, also die Überzeugung, beim Lernen Herausforderungen bewältigen und Ziele erreichen zu können (Karlen et al., 2014, S. 784). In Bezug auf *Herkunftseffekte* wird mehrfach berichtet, dass Selbstkonzepte und Selbstwirksamkeitserwartungen bei sozial privilegierten Schüler/-innen günstiger ausgeprägt sind als bei sozial benachteiligten Schüler/-innen (Kalaycioglu, 2015, S. 1396; Karlen et al., 2014, S. 787; Muijs, 1997, S. 270). Vereinzelt finden sich aber auch Befunde, wonach beim Selbstbild keine bedeutsamen Unterschiede nach sozialer Herkunft bestehen (Yerdelen-Demar & Pesman, 2013, S. 287). Der positive Zusammenhang zwischen Selbstbild und *Schülerkompetenzen* wird wiederum durch mehrere Meta-Analysen belegt (Hansford & Hattie, 1982, S. 132; Valentine et al., 2004, S. 120). Zur Rolle des Selbstbildes liegen weiterhin einige *pfadanalytische Studien* vor, die die Mediationsannahme mehrheitlich bestätigen (Steinmayr et al., 2012, S. 342 f.; Kriegbaum & Spinath, 2016, S. 52).

Insgesamt betrachtet liegen für den Zusammenhang zwischen sozialer Herkunft und Lernmotivation weitestgehend konsistente Befunde vor, während der Zusammenhang zwischen Lernmotivation und Lernerfolg durch eine konsistente Befundlage belegt wird.

3.11.3 Unterschiede bei psychischen Dispositionen der Heranwachsenden

Die dritte Kategorie von Lernvoraussetzungen bezieht sich auf psychische Dispositionen. Damit gemeint sind situationsübergreifende Tendenzen im Erleben und Verhalten. Ein wichtiger theoretischer Ansatz ist das Fünf-Faktoren-Modell (auch „Big Five“ genannt), welches Extraversion, Neurotizismus, Gewissenhaftigkeit, Verträglichkeit und Offenheit umfasst (Asendorpf, 2007, S. 155).

Mit großer Konsistenz zeigen empirische Studien *herkunftsspezifische Unterschiede* bei der Disposition *Offenheit*. Sozial privilegierte Schüler/-innen weisen eine höhere Offenheit auf, d. h., sie sind im stärkeren Maße aufgeschlos-

sen gegenüber neuen Erfahrungen und neuen Ideen, neugierig, einfallsreich und künstlerisch-kulturell interessiert (Dahmann & Anger, 2014, S. 19; Peter & Storck, 2015, S. 6; Steinmayr et al., 2010, S. 543). Weiterhin für den schulischen Kontext relevant ist die Disposition *Gewissenhaftigkeit*. Zu dieser Disposition gehören Subfacetten wie Pflichtbewusstsein, Selbstregulation, Besonnenheit, Zuverlässigkeit und Ordentlichkeit. Wird diese Disposition global betrachtet, fallen die Zusammenhänge zur sozialen Herkunft eher inkonsistent aus (Dahmann & Anger, 2014, S. 19; Peter & Storck, 2015, S. 6; Steinmayr et al., 2010, S. 543). Systematische Unterschiede zeigen sich allerdings, wenn die Subfacette Selbstregulation (auch: Selbstdisziplin, Fokussiertheit) betrachtet wird, welche sich auf die eigenständige Steuerung von Kognitionen (z. B. Konzentrationsfähigkeit), Emotionen (z. B. Frustrationstoleranz) und Verhalten (z. B. Belohnungsaufschub) bezieht. Die Selbstregulation von Schüler/-innen steht konsistent in einem positiven Zusammenhang zu ihrer sozialen Herkunft (Kaiser & Diewald 2014, S. 101; Miech et al., 2001, S. 108; Sammons et al., 2014, S. 14).

Als Mechanismus, wie Offenheit und Gewissenhaftigkeit den *Lernerfolg* beeinflussen, ist insbesondere ein Einfluss auf den quantitativen Umfang von Lernaktivitäten plausibel (z. B. Neugier als Aspekt von Offenheit, Selbstdisziplin als Aspekt von Gewissenhaftigkeit). Die Bedeutung von psychischen Dispositionen für den Lernerfolg wurden von zahlreichen Studien untersucht. Hinsichtlich der Rolle von Gewissenhaftigkeit für den Lernerfolg kommen zwei Meta-Analysen zu dem Ergebnis, dass ein deutlicher Zusammenhang besteht (O'Connor & Paunonen, 2007, S. 976; Poropat, 2009, S. 328). Die Subfacette Selbstregulation weist hierbei eine besonders hohe Korrelation auf (O'Connor & Paunonen, 2007, S. 979). Für Offenheit wird in einer Meta-Analyse ein schwacher Zusammenhang zum Lernerfolg ermittelt (Poropat, 2009, S. 328), in der anderen Meta-Analyse dagegen ein vernachlässigbarer Zusammenhang (O'Connor & Paunonen, 2007, S. 977). Erwähnenswert ist außerdem eine *pfadanalytische Studie*, in der sowohl für Offenheit als auch für Gewissenhaftigkeit eine mediiierende Rolle bestätigt wird (Steinmayr et al., 2010, S. 544).

Zusammenfassend liegen klare Evidenzen dafür vor, dass die Gewissenhaftigkeits-Facette Selbstregulation einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt. Bei der Offenheit zeigen sich zwar Zusammenhänge zur sozialen Herkunft, die Bedeutung für den Lernerfolg ist allerdings nicht eindeutig.

3.12 Zusammenfassung des dritten Kapitels

Im vorliegenden Kapitel wurden Effekte der sozialen Herkunft auf den Lernerfolg unter theoretischen und empirischen Gesichtspunkten betrachtet. Zunächst wurde ein theoretisches Modell entwickelt, welches die bildungssoziologische und die pädagogisch-psychologische Perspektive verbindet. Auf dieser Grundlage wurden relevante Befunde zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg aufgearbeitet. Empirisch belegt werden können statusspezifische Unterschiede in den persönlichen, ökonomischen und sozialen Ressourcen von Familien. Bei Betrachtung der familiären Lernumwelt zeigen sich Unterschiede bei bestimmten Aspekten des Erziehungsstils und der kognitiven Anregung. Bei der schulischen Beteiligung der Eltern fallen die Befunde nicht eindeutig aus, auch wegen methodischen Schwierigkeiten. Die Rolle frühpädagogischer Einrichtungen kann nicht abschließend bewertet werden, da einerseits Hinweise auf Nutzungsunterschiede vorliegen, andererseits aber auch Hinweise auf differenzielle Effekte. Im schulischen Bereich liegen zahlreiche Befunde zu Unterschieden zwischen Schulen vor. Nutzungsunterschiede können bezüglich der Anwesenheit im Unterricht belegt werden, während die Befunde bei der Nutzung von Ganztagsangeboten unklar sind. Ebenfalls unklar ist bisher, ob Nachhilfe einen Mediator zwischen sozialer Herkunft und Lernerfolg darstellt, wobei auch hier methodische Schwierigkeiten von Bedeutung sind. Im Freizeitbereich legen die Befunde nahe, dass die Freizeitaktivitäten von Heranwachsenden vermittelnde Mechanismen darstellen. Bei Gleichaltrigengruppen zeigen sich konsistente Unterschiede sowohl bei der Größe des Freundeskreises als auch bei der Art der Freund/-innen. Schließlich zeigen sich bei den Lernvoraussetzungen der Schüler/-innen konsistente Unterschiede bei kognitiven Merkmalen und weitestgehend konsistente Unterschiede bei motivationalen Merkmalen und psychischen Dispositionen. Insgesamt betrachtet sprechen die Befunde dafür, dass die Entstehung von herkunftsbedingten Kompetenzunterschieden auf eine große Anzahl von vermittelnden Mechanismen mit jeweils relativ kleinen Effekten zurückzuführen ist (und nicht auf wenige Mechanismen mit relativ großen Effekten) (Becker & Lauterbach, 2010, S. 27).³

3 Dies gilt umso mehr, als dass weitere Mechanismen theoretisch plausibel und empirisch belegt sind, welche in diesem Literaturüberblick nicht berücksichtigt wurden, weil sie für die weiteren Ausführungen nicht von wesentlicher Bedeutung sind (z. B. familiärer Stress, siehe Evans & Schamberg, 2009; Ernährung, siehe Urke, Bull & Mittelmark, 2011; kindliche Gesundheit, siehe Becker & Kurz, 2011).

Der Forschungsüberblick zeigt, dass inzwischen eine ganze Reihe von empirischen Befunden vorhanden ist. Gleichwohl findet sich in der Literatur die Einschätzung, dass der empirische Forschungsstand zu den Ursachen von herkunftsbedingten Kompetenzunterschieden bisher ungenügend ist und folglich weiterer Forschungsbedarf besteht (Becker & Lauterbach, 2010, S. 17 f.; Maaz et al., 2010, S. 36; Solga & Dombrowski, 2009, S. 20). Bei einer differenzierten Betrachtung kommt man zu der Einschätzung, dass in quantitativer Hinsicht zwar viele Studien zur Verfügung stehen, dass die Studien aber in qualitativer Hinsicht zu einem großen Teil nur bedingt aussagekräftig sind. Zwei methodische Limitationen sind hervorzuheben. Erstens handelt es sich bei einem großen Teil der dargestellten Studien um bivariate korrelative Analysen. In vielen Fällen sind allerdings komplexe Wirkzusammenhänge denkbar (z. B. Drittvariablen als gemeinsame Ursachen, reziproke Effekte). Vor diesem Hintergrund ist davon auszugehen, dass bivariate Analysen den tatsächlichen Kausaleffekt in der Regel überschätzen. Zweitens basiert der größte Teil der hier berichteten Evidenzen darauf, dass Zusammenhänge zwischen sozialer Herkunft und Mediatoren einerseits und Zusammenhänge zwischen Mediatoren und Lernerfolg andererseits in verschiedenen Studien und somit in verschiedenen Stichproben nachgewiesen werden. Nur bei einem kleinen Teil der Studien werden beide Zusammenhänge in der gleichen Stichprobe untersucht und bei einem noch kleineren Teil der Studien erfolgt eine Analyse mithilfe von Pfadmodellen. Dies schränkt die Evidenz für vermittelnde Mechanismen weiter ein.

Vor diesem Hintergrund besteht primär Bedarf an hochwertigen Studien, welche die Komplexität von vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg in methodisch angemessener Weise untersuchen. Aussagekräftig wären insbesondere multivariate Längsschnittstudien, welche parallel mehrere Mechanismen analysieren, Mediationen mithilfe von Pfadmodellen analysieren und gegebenenfalls sekundäre Herkunftseffekte sowie differenzielle Wirksamkeit berücksichtigen. Unter inhaltlichen Gesichtspunkten bestehen Forschungsdefizite insbesondere in den Bereichen schulische Beteiligung der Eltern, Nutzung von frühpädagogischen Einrichtungen, Nutzung von Ganztagsangeboten und Nutzung von Nachhilfe. In diesen Bereichen erlauben es die vorhandenen Studien nicht, eine abschließende Einschätzung möglicher Mediationsprozesse vorzunehmen. Die genannten Punkte werden in der vorliegenden Arbeit aufgegriffen. Im empirischen Teil wird eine auf dem theoretischen Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg basierende Analyse vorgelegt.

4. Effekte von Klassen und Schulen

Im letzten Kapitel wurde der Effekt der sozialen Herkunft auf den Lernerfolg behandelt und ein breiter Überblick über mögliche Mediatoren gegeben. Im vorliegenden Kapitel wird nun genauer auf die Bildungseinrichtung Schule eingegangen. Im Vordergrund stehen die Fragen, welche Bedeutung die Schule für den Lernerfolg von Schüler/-innen hat und in welchem Maße Schule herkunftsbedingte Unterschiede im Lernerfolg beeinflusst. Dabei handelt es sich um zentrale Fragestellungen der schulbezogenen Wirksamkeitsforschung und der Bildungssoziologie, welche auch in der vorliegenden Arbeit untersucht werden. Die erstgenannte Frage nach globalen Effekten hat dabei in der Forschung bisher deutlich mehr Aufmerksamkeit erfahren als die Frage nach Effekten auf Unterschiede im Lernerfolg (Sammons et al., 2016, S. 74; Strand, 2010, S. 291).

Wissenschaftliche Untersuchungen zu den Effekten von Schule sind mit gewissen *methodischen Herausforderungen* verbunden. Eine zentrale Schwierigkeit liegt darin, dass in den meisten Ländern aufgrund der gesetzlichen Schulpflicht (fast) alle Heranwachsenden die Schule besuchen und daher keine Studien mit Kontrollgruppen ohne Schulbesuch möglich sind (Luyten et al., 2017, S. 376 f.). Vor diesem Hintergrund wurden alternative Forschungsansätze entwickelt, um die Größe der Effekte von Schule im Rahmen nicht experimenteller Studien abzuschätzen. Diese werden nachfolgend vorgestellt und diskutiert.

Die Perspektive im vorliegenden Kapitel fokussiert auf die Effekte von pädagogischen Einheiten, während die Effekte von pädagogischen Merkmalen erst im nächsten Kapitel thematisiert werden. Im Kontext pädagogischer Effekte ist auf die *Mehrebenenstruktur von Schule* hinzuweisen (Ditton, 2000, S. 76 f.; Scheerens & Bosker, 1997, S. 45). Damit ist gemeint, dass das Schulsystem durch eine hierarchische Struktur zusammenhängender Personengruppen gekennzeichnet ist: mehrere Schüler/-innen bilden zusammen eine Klasse, mehrere Klassen bilden zusammen eine Schule, mehrere Schulen bilden zusammen ein Schulsystem. Hieraus ergibt sich, dass Effekte von pädagogischen Einheiten auf verschiedenen Ebenen betrachtet werden können. *Klasseneffekte* beziehen sich darauf, welche Lernwirkungen bei den Schüler/-innen durch den Besuch einer bestimmten Klasse entstehen. Entsprechend beziehen sich *Schuleffekte* auf die Lernwirkungen durch den Besuch einer bestimmten Schule und *Systemeffekte*

te auf die Lernwirkungen durch den Besuch eines bestimmten Schulsystems. Nachfolgend wird der Begriff „schulische Effekte“ als Oberkategorie verwendet und umfasst somit alle drei genannten Arten von Effekten.

Im Einklang mit der Begriffsverwendung im letzten Kapitel bezieht sich der Begriff *Lernerfolg* auch weiterhin auf den Erwerb von Fachkompetenzen und der Begriff *Herkunft* bezieht sich auf die soziale Herkunft von Schüler/-innen. Somit verweist die Formulierung *herkunftsbedingte Unterschiede im Lernerfolg* auf Kompetenzunterschiede nach sozialer Herkunft.

Nachfolgend werden zunächst relevante theoretische Positionen zur Rolle der Schule dargestellt (Kapitel 4.1). Anschließend werden schulische Effekte auf den durchschnittlichen Lernerfolg behandelt (Kapitel 4.2). Danach werden schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg behandelt (Kapitel 4.3). Abschließend wird auf die Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg eingegangen (Kapitel 4.4). In den drei letztgenannten Abschnitten werden jeweils Überblicke über relevante Forschungsbefunde in den entsprechenden Themenbereichen gegeben. Dieser breite Forschungsüberblick ermöglicht eine Einschätzung, in welchen Bereichen das Potential der Bildungseinrichtung Schule bereits fundiert beurteilt werden kann und in welchen Bereichen noch Forschungsdefizite bestehen.

4.1 Theoretische Positionen zur Rolle der Schule

Über die Rolle, welche die Schule für die Entstehung von Bildungsungleichheiten spielt, wird seit langer Zeit eine wissenschaftliche Kontroverse geführt. In der Literatur lassen sich ausgesprochen unterschiedliche, teilweise sogar gegensätzliche Positionen finden (Downey et al., 2004, S. 613 ff.). Die Extrempole bilden einerseits die Position, dass *Schule zur Vergrößerung von Bildungsungleichheiten beitrage*, und andererseits die Position, dass die *Schule als „große Gleichmacherin“ Bildungsungleichheiten verringere*. Die Einschätzung, dass die Schule selbst eine Ursache für Kompetenzunterschiede darstellt, erhält ihre Grundlage insbesondere durch die im letzten Kapitel berichteten Befunde, wonach sich Schüler/-innen verschiedener sozialer Herkunft in ihren schulbezogenen Lernerfahrungen unterscheiden. Hierbei sind sowohl Unterschiede zwischen Schulen als auch Unterschiede innerhalb von Schulen von Bedeutung. In beiden Fällen wird die Frage aufgeworfen, wer die Verantwortung für die unterschiedlichen Lernerfahrungen trägt. Die Reproduktionstheorie (Bourdieu & Passeron, 1977) schreibt diese Verantwortung der Schule zu, weil Schule primär an den Interessen statushoher Gruppen ausgerichtet sei. Diese

Haltung kommt auch in Konzepten wie „Schule als Mittelschichtinstitution“ oder „institutionelle Diskriminierung“ zum Ausdruck (Brake & Büchner, 2012, S. 95 ff.). Alternativ kann die Verantwortung bei der Familie bzw. den Schüler/-innen gesucht werden. Demnach wären primär familiäre Bildungsentscheidungen (für Unterschiede zwischen Schulen) bzw. die Lernvoraussetzungen der Schüler/-innen (für Unterschiede innerhalb von Schulen) entscheidend.

Wird eine optimistische Position eingenommen, wonach die Schule als „Gleichmacherin“ zumindest versucht, herkunftsbedingte Kompetenzunterschiede gering zu halten, wird unmittelbar die Frage aufgeworfen, inwieweit die Schule Kompetenzunterschiede überhaupt beeinflussen *kann*. Dies verweist zunächst einmal auf die grundlegendere Frage, welche *Bedeutung der Schule überhaupt für den Lernerfolg von Schüler/-innen* zukommt (im Vergleich zu anderen Faktoren wie Familienmerkmalen oder Schülermerkmalen). Wissenschaftliche Einschätzungen zu dieser Frage unterliegen erkennbar einem historischen Wandel (Creemers et al., 2010, S. 5). Die frühen Studien von Coleman (1966) und Jencks (1972) resultierten in einer pessimistischen Einschätzung, wonach Schulen nur einen geringen Einfluss auf den Lernerfolg von Schüler/-innen haben. Als Reaktion versuchte die frühe Schulwirksamkeitsforschung nachzuweisen, dass Schulen einen erheblichen Unterschied machen können (Brookover et al., 1979; Rutter et al., 1979). Die zugrunde liegende optimistische Einschätzung konnte nur teilweise bestätigt werden. Gegenwärtig wird mehrheitlich eine moderate Position vertreten, wonach Schulen einen begrenzten, aber bedeutsamen Einfluss auf den Lernerfolg der Schüler/-innen haben (Kyriakides & Creemers, 2011, S. 248; Kyriakides, 2004, S. 142).

Im Einklang mit dieser Einschätzung wird in der Literatur davon ausgegangen, dass dem Einfluss von Schulen auf Kompetenzunterschiede nach sozialer Herkunft *deutliche Grenzen* gesetzt sind (Helmke & Weinert, 1997a, S. 90, 138; Lee & Burkam, 2002, S. 1). Argumentativ wird dies damit begründet, dass ein bedeutender Teil der Sozialisation schon vor Schuleintritt erfolgt, wodurch die Schüler/-innen deutlich unterschiedliche Lernvoraussetzungen aufweisen. Außerdem bestehen parallel zum Schulbesuch weiterhin Unterschiede in außerschulischen Lernumwelten (z. B. Familie, Freizeiteinrichtungen). Die hohe Bedeutung außerschulischer Einflüsse wird insbesondere daran deutlich, dass Heranwachsende nur 13% ihrer wachen Lebenszeit in der Schule verbringen (Walberg, 1984, S. 22). Angesichts dieser Rahmenbedingungen ist nicht zu erwarten, dass Kompetenzunterschiede nach sozialer Herkunft ohne Weiteres durch den regulären Schulbesuch abgebaut werden können. Realistischer erscheint die Einschätzung, dass bereits das *Verhindern einer weiteren Zunahme*

von Kompetenzunterschieden im Laufe der Schulzeit als eine pädagogische Leistung von Schule zu bewerten ist.

4.2 Schulische Effekte auf den durchschnittlichen Lernerfolg

Im Weiteren werden relevante Forschungsbefunde zu schulischen Effekten behandelt. Begonnen wird mit schulischen Effekten auf den durchschnittlichen Lernerfolg. Zunächst werden Befunde berichtet, welche für die Einschätzung von *absoluten Effekten* von Bedeutung sind. Anschließend wird auf Befunde zu *relativen Effekte* eingegangen, differenziert nach verschiedenen Ebenen, Zielsetzungen und Schülergruppen.

4.2.1 Absolute schulische Effekte auf den durchschnittlichen Lernerfolg

Absolute schulische Effekte auf den durchschnittlichen Lernerfolg beziehen sich darauf, wie stark sich der *Kompetenzerwerb der Schüler/-innen durch den Schulbesuch verbessert* (im Vergleich zu keinem Schulbesuch). Eine erste Annäherung an die absoluten Effekte von Schule liefern Studien, die untersuchen, wie stark die Schülerkompetenzen im Laufe eines Schuljahres zunehmen. Nur wenn sich zeigen lässt, dass Schüler/-innen am Ende des Schuljahres bessere Kompetenzen aufweisen als am Anfang des Schuljahres, kann überhaupt von schulischen Effekten ausgegangen werden. Empirische Befunde belegen konsistent bedeutsame Kompetenzzuwächse im Laufe eines Schuljahres. Typischerweise bewegt sich der durchschnittliche Kompetenzzuwachs pro Schuljahr in der Größenordnung zwischen $d = 0.25$ und $d = 0.40$ (Ehmke et al., 2006a, S. 72 ff.; Hattie, 2013, S. 20 f.; Lehmann et al., 1999, S. 68; Walter et al., 2006, S. 100 ff.). Das bedeutet, der Kompetenzmittelwert am Ende des Schuljahres ist zwischen 25% und 40% einer Standardabweichung höher als der Kompetenzmittelwert am Anfang des Schuljahres. Der Kompetenzzuwachs fällt hierbei in den unteren Jahrgangsstufen in der Regel etwas größer aus und in den oberen Klassenstufen etwas kleiner (Ditton & Krüsken, 2009, 43 ff.; Moser et al., 2017, S. 68). Die vorhandenen Befunde sind somit im Einklang mit der Annahme, dass der Schulbesuch insgesamt einen bedeutsamen Effekt auf den Lernerfolg hat. Insbesondere bei den Fächern, die primär durch die Schule vermittelt werden (z. B. Mathematik, Naturwissenschaften), erscheint es theoretisch plausibel, dass der festgestellte Kompetenzzuwachs ursächlich auf die Schule zurückzuführen ist. Einen klaren Beleg für einen Kausaleffekt können

diese Befunde allerdings nicht leisten, da der Kompetenzzuwachs theoretisch auch durch andere Lernumwelten verursacht werden könnte.

Eine diesbezüglich aussagekräftigere Forschungsstrategie besteht darin, den Lernfortschritt von Schüler/-innen *während der Schulzeit und während der Ferienzeit* miteinander zu vergleichen. Während der Schulzeit wirken alle Lernumwelten parallel (z. B. Schule, Familie, Freizeiteinrichtungen), während in der Ferienzeit die Schule keinen Einfluss ausübt. Insofern kann durch die Gegenüberstellung von Schulzeit und Ferienzeit der spezifische Beitrag der Schule ermittelt werden. Derartige Studien wurden insbesondere im US-amerikanischen Raum durchgeführt. Befunde zum sogenannten Ferieneffekt (Englisch: summer setback) zeigen konsistent, dass der Kompetenzzuwachs von Schüler/-innen während der Schulzeit deutlich größer ausfällt als während der Sommerferien (Alexander et al., 2001, S. 176 f.; Downey et al. 2004, S. 621). Der Rückgang des Lernens während der Ferienzeit ist hierbei in Mathematik stärker ausgeprägt als im Lesen (Alexander et al. 2001, S. 177; Cooper et al., 1996, S. 253 f.). Dies ist insofern plausibel, als dass die Beschäftigung mit Mathematik primär im Kontext der Schule erfolgt, während Lesen auch eine typische Freizeitbeschäftigung von Heranwachsenden darstellt. Insgesamt betrachtet liefern die Befunde zu Kompetenzzuwächsen und zum Ferieneffekt überzeugende Evidenzen, dass der Schulbesuch bedeutsame Effekte auf den Lernerfolg von Schüler/-innen hat.

4.2.2 Relative schulische Effekte auf den durchschnittlichen Lernerfolg

Eine Einschränkung der zuvor angeführten Befunde liegt darin, dass sie die Effekte von Schule insgesamt betrachten, d. h., Unterschiede zwischen Klassen/Schulen bleiben unberücksichtigt. *Unterschiede zwischen Klassen/Schulen* verweisen auf *relative schulische Effekte* und somit auf die Frage, welche Bedeutung es für den Lernerfolg eines Schülers/einer Schülerin hat, welche Klasse/Schule er/sie besucht. Die zugrunde liegende Idee ist, dass sich Klassen/Schulen in der Art und Weise, wie sie durch die pädagogischen Fachkräfte ausgestaltet werden, voneinander unterscheiden (z. B. Unterrichtsmethoden, Schulklima). Wenn diese pädagogische Ausgestaltung einen Einfluss auf den Lernerfolg der Schüler/-innen hat, sollten sich die Schüler/-innen verschiedener Klassen/Schulen systematisch im durchschnittlichen Lernerfolg unterscheiden, d. h., in manchen Klassen/Schulen sollte der Mittelwert der Schülerkompetenzen höher ausfallen als in anderen Klassen/Schulen.

Aufbauend auf dieser Idee werden relative schulische Effekte anhand von *Varianzzerlegungen* untersucht. Mithilfe der Intraclass Correlation (ICC) kann überprüft werden, wie stark die Unterschiede zwischen Klassen/Schulen in den von den Schüler/-innen erreichten Kompetenzen ausfallen (Ditton, 1998, S. 61f.; Hox, 2010, S. 15). Beispielsweise kann man die Gesamtvarianz der Mathematikkompetenzen von Schüler/-innen in die Varianz zwischen den Schüler/-innen (also die Abweichung der individuellen Mathematikkompetenz vom Schulmittelwert der Mathematikkompetenz) und die Varianz zwischen Schulen (also die Abweichung des Schulmittelwerts der Mathematikkompetenzen vom Gesamtmittelwert der Mathematikkompetenzen) zerlegen. Die Varianz zwischen Schüler/-innen ist ein Indikator für die Bedeutung aller individuellen Einflussfaktoren und die Varianz zwischen Schulen ein Indikator für die Bedeutung aller schulischen Einflussfaktoren. Die ICC stellt den Anteil der Varianz zwischen Schulen an der Gesamtvarianz dar und zeigt somit an, wie viel Varianz in den Mathematikkompetenzen durch die Schulzugehörigkeit erklärt werden kann.

Die frühe schulbezogene Wirksamkeitsforschung hatte ein starkes Interesse daran, die *Größe von Schuleffekten* zu ermitteln und zu zeigen, dass Schulen einen Unterschied machen können (Creemers et al., 2010, S. 7f.). In den frühen Studien wurde häufig nur die Schulebene modelliert, d. h., es wurden lediglich Unterschiede zwischen Schulen betrachtet (während Unterschiede zwischen Klassen unberücksichtigt blieben). Der vorhandene Forschungsstand zu Schuleffekten wurde von Scheerens und Bosker (1997) in einer Meta-Analyse zusammengefasst. Sie ermitteln eine durchschnittliche unadjustierte ICC von 19%, d. h., dass im Durchschnitt 19% der Varianz von Schülerkompetenzen zwischen Schulen liegen und 81% der Varianz zwischen Schüler/-innen (S. 77). Aussagekräftiger für Schuleffekte ist eine adjustierte ICC, bei der mithilfe eines Mehrwertmodells die Ausgangsbedingungen der Schüler/-innen berücksichtigt werden (da sonst nicht schulische Einflussfaktoren wie Intelligenz oder soziale Herkunft miteinfließen). Bei dieser Vorgehensweise verringert sich die durchschnittliche ICC auf 8%. In einer jüngeren Analyse ermitteln Hedges und Hedberg (2007) geringfügig höhere Werte: Die unadjustierte ICC beträgt durchschnittlich 22%, die adjustierte ICC beträgt durchschnittlich 11% (S. 68f.). Somit fallen die Unterschiede zwischen Schüler/-innen (innerhalb von Schulen) deutlich größer aus als die Unterschiede zwischen Schulen. Hieraus kann geschlossen werden, dass individuelle Einflussfaktoren (z. B. Intelligenz, soziale Herkunft) *insgesamt* einen größeren Einfluss auf den Lernerfolg haben als die besuchte Schule. Allerdings werden auch kleine Schuleffekte als bedeutsam angesehen, weil Schulen eine große Anzahl von Schüler/-innen

erreichen und weil sie kumulativ über mehrere Jahre hinweg wirken (Reynolds et al., 2014, S. 206; Scheerens & Bosker, 1997, S. 79).

Eine differenziertere Betrachtung von schulischen Effekten ermöglichen Analysen, welche *zusätzlich zur Schulebene auch die Klassenebene berücksichtigen*. Das heißt, die Varianz der Schülerkompetenzen wird in Unterschiede zwischen Schüler/-innen, Unterschiede zwischen Klassen und Unterschiede zwischen Schulen zerlegt. In Übersichtsarbeiten zeigt sich mehrheitlich, dass der Varianzanteil zwischen Klassen größer ist als der Varianzanteil zwischen Schulen (Luyten, 2003, S. 42 ff.; Marzano, 2000, S. 59 ff., Scheerens et al., 1989, S. 794). Marzano (2000) kommt in seiner Übersicht zu der Einschätzung, dass 13% der Varianz zwischen Klassen liegen und 7% der Varianz zwischen Schulen (S. 62). Aus diesen Befunden kann geschlussfolgert werden, dass relative Klasseneffekte größer ausfallen als relative Schuleffekte. Das heißt, für den Lernerfolg eines Schülers/einer Schülerin ist entscheidender, welche Klasse er/sie besucht und somit von welcher Lehrkraft er/sie unterrichtet wird, als welche Schule er/sie besucht (Hattie, 2013, S. 85). Dies ist theoretisch plausibel, da die Klasse denjenigen Ausschnitt der schulischen Lernumwelt darstellt, der am unmittelbarsten auf den Schüler/die Schülerin einwirkt. Weiterhin sprechen die Befunde dafür, dass Schulen in Bezug auf Wirksamkeit keine homogenen Einheiten darstellen, sondern dass innerhalb von Schulen bedeutende Unterschiede zwischen den Klassen bestehen (Ditton & Kreckler, 1995, S. 517; Walberg, 1984, S. 21).

Es ist anzunehmen, dass schulische Effekte *bei verschiedenen Lernzielen* unterschiedlich stark ausfallen. Da schulische Lernangebote zum überwiegenden Teil auf die Aneignung und Anwendung von Wissen ausgerichtet sind, sollten die Effekte im kognitiven Bereich besonders groß ausfallen. Tatsächlich zeigt sich weitestgehend konsistent, dass bei kognitiven Zielen (z. B. Fachkompetenzen) mehr Varianz zwischen Klassen/Schulen liegt als bei motivationalen Zielen (z. B. Interesse) oder emotionalen Zielen (z. B. Wohlbefinden) (Kunter & Voss, 2011, S. 103; Muijs et al., 2014, S. 243; Opdenakker & Van Damme, 2000, S. 175). Schulen haben demnach einen größeren Einfluss auf die kognitive Entwicklung als auf die motivationale und emotionale Entwicklung der Schüler/-innen. Werden innerhalb des kognitiven Bereichs verschiedene Fachkompetenzen verglichen, zeigt sich in mehreren Studien, dass schulische Effekte in den Bereichen Mathematik und Naturwissenschaften größer ausfallen als im Bereich Lesen (Brandsma & Knuver, 1989, S. 782; Scheerens & Bosker, 1997, S. 77; Strand, 2016, S. 135).

4.2.3 Relative schulische Effekte auf den durchschnittlichen Lernerfolg bei verschiedenen Schülergruppen

Von besonderem Interesse für die vorliegende Arbeit ist die Frage, ob relative schulische Effekte *bei verschiedenen Schülergruppen* unterschiedlich stark ausfallen. Dieser Sachverhalt wäre gegeben, wenn es bei einer Schülergruppe eine größere Rolle spielt, welche Schule sie besuchen, als bei einer anderen Schülergruppe. In Bezug auf Schüler/-innen verschiedener sozialer Herkunft können unterschiedliche Argumente angeführt werden. Einerseits kann argumentiert werden, dass sozial privilegierte Schüler/-innen anregende Lernangebote besser nutzen können als sozial benachteiligte Schüler/-innen, weil sie über günstigere Lernvoraussetzungen verfügen (Raudenbush & Eschmann, 2015, S. 453). Andererseits kann argumentiert werden, dass sozial benachteiligte Schüler/-innen, gerade weil sie ungünstigere Lernvoraussetzungen haben, stärker von hochwertigen schulischen Lernangeboten abhängig sind (Ditton, 2011, S. 250; Opendakker et al., 2002, S. 419).

Um diese Frage empirisch zu untersuchen, besteht eine naheliegende Forschungsstrategie darin, die untersuchte Stichprobe aufzuteilen und *für verschiedene Schülergruppen getrennte ICCs* zu berechnen. Die Aufteilung der Stichprobe kann dabei entweder anhand des individuellen Merkmals (z. B. Schüler/-innen von Eltern mit hohem Sozialstatus vs. Schüler/-innen von Eltern mit niedrigem Sozialstatus) oder anhand des Schulmittelwerts (z. B. Schulen mit hohem Sozialstatus-Mittelwert vs. Schulen mit niedrigem Sozialstatus-Mittelwert) erfolgen. Für beide Verfahren finden sich Beispiele in der Literatur (Sammons et al., 1993; Pfeifer, 2011). Aussagekräftiger für differenzielle Effekte bei verschiedenen Schülergruppen ist die Aufteilung anhand des individuellen Merkmals, weil die Schüler/-innen der zwei Gruppen dann (tendenziell) die gleichen Schulen besuchen und somit geklärt werden kann, ob die gleichen Schulen differenzielle Effekte haben. Bei der Aufteilung anhand des Schulmittelwerts sind Unterschiede in der ICC dagegen nicht eindeutig auf differenzielle Effekte zurückzuführen, sondern können auch durch Unterschiede in der Varianz der pädagogischen Qualität zustande kommen (Nye et al., 2004, S. 251).⁴

4 Wenn beispielsweise die ICC bei Schulen mit niedrigem Sozialstatus-Mittelwert größer ausfällt als bei Schulen mit hohem Sozialstatus-Mittelwert, könnte das darauf zurückzuführen sein, dass innerhalb der Gruppe der Schulen mit niedrigem Sozialstatus-Mittelwert die Unterschiede in der pädagogischen Qualität größer ausfallen als innerhalb der Gruppe der Schulen mit hohem Sozialstatus-Mittelwert.

Im Rahmen der Literaturrecherche wurden vier Studien ermittelt, in denen schulische Effekte bei Schüler/-innen verschiedener sozialer Herkunft verglichen werden. Bei allen vier Studien erfolgt die Aufteilung der Stichprobe anhand des Klassen- bzw. Schulmittelwerts, wodurch ihre Aussagekraft begrenzt ist. Die Ergebnisse sind zudem nicht ganz konsistent. In der Studie von Pfeifer (2011) fallen Klasseneffekte innerhalb der Gruppe der Klassen mit niedrigem Sozialstatus-Mittelwert stärker aus als innerhalb der Gruppe der Klassen mit hohem Sozialstatus-Mittelwert (S. 146). Ähnliche Befunde berichten Nye und Kolleg/-innen (2004) für Klasseneffekte, während die Befunde für Schuleffekte sehr heterogen ausfallen (teilweise kein Unterschied, teilweise größere Schuleffekte bei Schulen mit niedrigem Sozialstatus-Mittelwert, teilweise größere Schuleffekte bei Schulen mit hohem Sozialstatus-Mittelwert) (S. 250). In einer anderen Studie fallen Schuleffekte bei Schulen mit niedrigem Sozialstatus-Mittelwert fast doppelt so groß aus wie bei Schulen mit hohem Sozialstatus-Mittelwert (Palardy, 2008, S. 32). In der mehrere Stichproben umfassenden Analyse von Hedges und Hedberg (2007) sind dagegen mehrheitlich keine bedeutenden Unterschiede zwischen Schulen mit durchschnittlichem Sozialstatus-Mittelwert und Schulen mit niedrigem Sozialstatus-Mittelwert in Bezug auf die Stärke von Schuleffekten festzustellen (S. 68 ff.). Somit liefern die vorliegenden Befunde nur in der Tendenz Hinweise darauf, dass schulische Effekte bei sozial benachteiligten Schüler/-innen größer ausfallen als bei sozial privilegierten Schüler/-innen.

Abschließend ist hervorzuheben, dass die zuvor dargestellten Studien *relative Effekte* untersuchen. Das heißt, selbst wenn relative Effekte bei sozial benachteiligten Schüler/-innen stärker ausfallen, bedeutet das nicht, dass durch Schule bestehende Kompetenzunterschiede verringert werden (das wäre der Fall, wenn *absolute Effekte* bei sozial benachteiligten Schüler/-innen stärker ausfallen). Die Aussagekraft von Untersuchungen zu relativen Effekten ist eher darin zu sehen, das überprüft werden kann, ob bestimmte Schülergruppen besonders sensitiv auf die pädagogische Ausgestaltung von Schule reagieren. Unter praktischen Gesichtspunkten ist dies relevant für Entscheidungen über die Verteilung von pädagogischen Ressourcen, da Investitionen von Ressourcen in bestimmte Schülergruppen dann besonders effizient sein sollten.

4.3 Schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg

Nachfolgend werden schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg behandelt. Wiederum werden sowohl absolute Effekte als auch relative Effekte betrachtet. Die Darstellung erfolgt in dieser Reihenfolge.

4.3.1 Absolute schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg

Bei der Betrachtung von Unterschieden im Lernerfolg beziehen sich *absolute schulische Effekte* darauf, inwieweit der Schulbesuch die zwischen den Schüler/-innen *bestehenden Kompetenzunterschiede durch den Schulbesuch verändert*. Es wird also gefragt, ob Kompetenzunterschiede durch den Schulbesuch zunehmen, unverändert bleiben oder abnehmen (im Vergleich zu keinem Schulbesuch).

Die Untersuchung dieser Frage erfordert die längsschnittliche Erfassung von herkunftsbedingten Unterschieden im Lernfortschritt. Hierfür steht eine Reihe von Indikatoren zur Verfügung (Nachbauer & Kyriakides, 2020). Während verschiedene Indikatoren insgesamt relativ hoch miteinander korrelieren (S. 319 f.), unterscheiden sich die Indikatoren teilweise in ihrer Bedeutung: Einige Indikatoren zeigen Unterschiede im unbedingten Lernfortschritt an (unabhängig von der Ausgangskompetenz), während andere Indikatoren Unterschiede im bedingten Lernfortschritt anzeigen (im Verhältnis zur Ausgangskompetenz) (S. 315).⁵ *Unterschiede im unbedingten Lernfortschritt* werden beispielsweise durch den Regressionskoeffizienten der sozialen Herkunft auf den Differenzwert zwischen Endkompetenz und Ausgangskompetenz (Gainscore) erfasst. *Unterschiede im unbedingten Lernfortschritt* werden ebenfalls erfasst, wenn für beide Messzeitpunkte querschnittlich ein Regressionskoeffizient der sozialen Herkunft auf die Kompetenz berechnet wird und der Unterschied zwischen den beiden Regressionskoeffizienten betrachtet wird. Demgegenüber ist ein Indikator, der *Unterschiede im bedingten Lernfortschritt* anzeigt,

5 Im Artikel von Nachbauer und Kyriakides (2020) werden die Formulierungen „Unterschiede im absoluten Lernfortschritt“ und „Unterschiede im relativen Lernfortschritt“ verwendet (S. 315). An dieser Stelle werden stattdessen die Formulierungen „Unterschiede im unbedingten Lernfortschritt“ (gleichbedeutend mit „Unterschiede im absoluten Lernfortschritt“) und „Unterschiede im bedingten Lernfortschritt“ (gleichbedeutend mit „Unterschiede im relativen Lernfortschritt“) verwendet. Diese Formulierungen werden gewählt, um eine klare inhaltliche Abgrenzung von absoluten und relativen schulischen Effekten zu gewährleisten.

der Regressionskoeffizient der sozialen Herkunft auf die Endkompetenz unter Kontrolle der Ausgangskompetenz. Ein veranschaulichendes Beispiel findet sich bei Nachbauer und Kyriakides (2020) in Appendix 2. Aufgrund der unterschiedlichen inhaltlichen Bedeutungen können sich in bestimmten Fällen abweichende Befunde ergeben, je nachdem, welcher Indikator betrachtet wird (Ditton & Krüsken, 2009, S. 51; Quinn, 2015, S. 63 f.).

Aufschlussreich für schulische Effekte auf Kompetenzunterschiede sind zunächst einmal Studien zur *Veränderung von Kompetenzunterschieden nach sozialer Herkunft* im Laufe der Schulzeit. Die Studienlage zu dieser Frage ist recht umfangreich (Neumann et al., 2014), allerdings unterscheiden sich die Befunde in Abhängigkeit vom verwendeten Indikator. Werden *Unterschiede im bedingten Lernfortschritt* betrachtet, zeigen sich konsistent Vorteile für sozial privilegierte Schüler/-innen (Anders et al., 2013, S. 205 f.; Ditton & Krüsken, 2009, S. 51 f.; Kiemer et al., 2017, S. 139 ff.; Lehmann & Lenkeit, 2008, S. 42 ff.; Schneider & Pfost, 2013, S. 173 f.). Das bedeutet, wenn Schüler/-innen mit gleichen Ausgangskompetenzen betrachtet werden, erreichen sozial privilegierte Schüler/-innen höhere Endkompetenzen als sozial benachteiligte Schüler/-innen. Werden demgegenüber *Unterschiede im unbedingten Lernfortschritt betrachtet*, fallen die Befunde unterschiedlich aus: teilweise haben sozial privilegierte Schüler/-innen einen höheren Lernfortschritt (Ivanov & Nikolova, 2010, S. 96 ff.; Lehmann & Lenkeit, 2008, S. 43 f.; Schnabel et al., 2002, S. 188 f.), teilweise bestehen keine Unterschiede (Ditton & Krüsken, 2009, S. 49; Ehmke et al., 2006a, S. 80; Guill et al., 2011, S. 47 ff.) und vereinzelt haben sozial benachteiligte Schüler/-innen einen höheren Lernfortschritt (Ivanov, 2011, S. 101 ff.; Lehmann & Lenkeit, 2008, S. 41 f.).

Zusammenfassend ist festzuhalten, dass sich mehrfach Hinweise darauf finden, dass sozial privilegierte Schüler/-innen einen höheren Lernfortschritt erreichen als sozial benachteiligte Schüler/-innen. Die Befunde sind dabei konsistenter bei Unterschieden im bedingten Lernfortschritt als bei Unterschieden im unbedingten Lernfortschritt (Neumann et al., 2014, S. 178).⁶ Außerdem

6 Dieser Befund ist möglicherweise auf zwei gegenläufige Tendenzen zurückzuführen. Einerseits verfügen sozial privilegierte Schüler/-innen über günstigere Lernvoraussetzungen (z. B. höhere Intelligenz, höhere Lernmotivation) und anregendere familiäre Lernumwelten. Dies begünstigt einen höheren Lernfortschritt der sozial privilegierten Schüler/-innen. Andererseits haben sozial privilegierte Schüler/-innen eine höhere Ausgangskompetenz, was eher einen niedrigeren Lernfortschritt wahrscheinlich macht (Baumert et al., 2012, S. 1357; Ditton & Krüsken, 2009, S. 44 ff.). Möglicherweise neutralisieren sich diese zwei Tendenzen gegenseitig. Hierdurch ließe sich erklären, warum häufig keine Unterschiede im unbedingten Lernfortschritt

können systematische Unterschiede eher in der Grundschule als in der Sekundarschule beobachtet werden (Neumann et al., 2014, S. 189 f.). Das Befundmuster zeigt, dass die bei Schulbeginn vorhandenen Kompetenzunterschiede nach sozialer Herkunft durch die Schule nicht abgebaut werden. Somit kann allein der universelle Zugang zur Schule den Einfluss der sozialen Herkunft nicht neutralisieren, zumindest nicht bei der aktuell gegebenen Ausgestaltung von Schule.

Die zuvor berichteten Befunde werfen die Frage auf, inwieweit herkunftsbedingte Unterschiede im Lernerfolg auf die Schule oder auf außerschulische Lernumwelten (z. B. Familie, Freizeiteinrichtungen) zurückzuführen sind. Um dies zu untersuchen, bietet sich wiederum die *Gegenüberstellung von Schulzeit und Ferienzeit* an. Die bereits erwähnten Studien zum Ferieneffekt haben eine weitestgehend konsistente Befundlage erbracht (Alexander et al., 2001, S. 177 ff.; Cooper et al., 1996, S. 253 ff.; Downey et al., 2004, S. 621 ff.; Quinn et al., 2016, S. 448 ff.). Beim Vergleich von Schulzeit und Ferienzeit zeigt sich, dass Kompetenzunterschiede nach sozialer Herkunft während der Ferienzeit stärker zunehmen als während der Schulzeit. Während der Schulzeit bleiben die Kompetenzunterschiede je nach Studie entweder konstant oder nehmen geringfügig zu (aber weniger als während der Ferienzeit). Dies legt nahe, dass herkunftsbedingte Unterschiede im Lernerfolg stärker durch nicht schulische Faktoren verursacht werden als durch schulische Faktoren.

Für Deutschland liegen bisher nur zwei Studien zum Ferieneffekt in Abhängigkeit von der sozialen Herkunft vor. In einer Studie wird im Einklang mit US-amerikanischen Befunden ermittelt, dass sozial benachteiligte Schüler/-innen während der Sommerferien einen geringeren bedingten Lernfortschritt aufweisen als sozial privilegierte Schüler/-innen (Becker et al., 2008, S. 266 ff.). Demgegenüber zeigen sich in der anderen Studie keine Unterschiede nach sozialer Herkunft während der Ferien (Coelen & Siewert, 2008, S. 438 f.). Einschränkend ist darauf hinzuweisen, dass die Wahrscheinlichkeit für einen differenziellen Ferieneffekt in Deutschland wesentlich kleiner ist als in den USA, da die Sommerferien hier deutlich kürzer sind als in den USA, wo sie bis zu drei Monate dauern.

Die Befunde zum Ferieneffekt stehen im deutlichen Widerspruch zu der Annahme der Reproduktionstheorie, wonach die Schule Kompetenzunterschiede nach sozialer Herkunft vergrößert. Im Gegenteil scheint die Schule dafür zu sorgen, dass sich der Lernfortschritt von Schüler/-innen verschiede-

gefunden werden, aber durchaus Unterschiede im bedingten Lernfortschritt (bei dem die Ausgangskompetenz kontrolliert wird).

ner sozialer Herkunft angleicht. In Verbindung mit den zuvor berichteten Befunden, die tendenziell für eine Vergrößerung von herkunftsbedingten Kompetenzunterschieden im Laufe der Schulzeit sprechen, entsteht der Eindruck, dass die Schule zumindest das Auseinanderdriften von Kompetenzunterschieden abbremst (d. h., Kompetenzunterschiede nehmen während des Schulbesuchs zwar zu, aber weniger stark, als wenn die Heranwachsenden keine Schule besuchen würden) (Downey & Condrón, 2016, S. 212).

4.3.2 Relative schulische Effekte auf herkunftsbedingte Unterschiede im Lernerfolg

Die zuvor berichteten Befunde konzentrierten sich darauf, wie sich die Schule insgesamt auf Kompetenzunterschiede nach sozialer Herkunft auswirkt. Wiederum erscheint es allerdings relevant, ob hierbei *Unterschiede zwischen Klassen/Schulen* bestehen. Diese Perspektive erlaubt eine Einschätzung, inwieweit die Ausgestaltung von Schule und Unterricht einen Einfluss auf Kompetenzunterschiede hat. Damit angesprochen sind *relative schulische Effekte* auf herkunftsbedingte Unterschiede im Lernerfolg. Die zugrunde liegende Idee ist, dass sich Klassen/Schulen in ihrer pädagogischen Ausgestaltung unterscheiden und dass, wenn diese Ausgestaltung einen Einfluss auf Kompetenzunterschiede hat, das Ausmaß der Kompetenzunterschiede in verschiedenen Klassen/Schulen unterschiedlich stark ausgeprägt sein sollte (angelehnt an Baumert & Schümer, 2001, S. 360). *Relative Klasseneffekte* beziehen sich auf Unterschiede zwischen Klassen im Zusammenhang zwischen sozialer Herkunft und Lernerfolg. *Relative Schuleffekte* beziehen sich somit auf Unterschiede zwischen Schulen im Zusammenhang zwischen sozialer Herkunft und Lernerfolg. Inhaltlich kommt in diesen relativen Effekten zum Ausdruck, welche Bedeutung es für die Kompetenzunterschiede innerhalb einer Gruppe von Schüler/-innen hat, welche Klasse/Schule diese Gruppe von Schüler/-innen besucht.

Forschungsansätze für die Untersuchung von relativen Effekten auf Unterschiede im Lernerfolg werden nachfolgend am Beispiel von Schuleffekten veranschaulicht. Die grundsätzliche Vorgehensweise besteht darin, für jede Schule einen Kennwert zu berechnen, welcher das Ausmaß von Kompetenzunterschieden innerhalb dieser Schule anzeigt (z. B. Mittelwertsdifferenz, Korrelation) und anschließend die Kennwerte verschiedener Schulen miteinander zu vergleichen (Nachbauer & Kyriakides, 2020). Ein häufig verwendeter Indikator ist der sogenannte *soziale Gradient*. Dabei handelt es sich um den Regressionskoeffizient der sozialen Herkunft zur Vorhersage der Schülerkompetenz (Baumert & Schümer, 2001, S. 386 f.; Züchner & Fischer, 2014, S. 360). Eine

elegante und zunehmend verwendete statistische Methode stellen in diesem Kontext *Random-Slope-Modelle* dar. Mit diesem Verfahren kann im Rahmen eines integrativen Mehrebenenmodells für jede Schule ein eigener Regressionskoeffizient der sozialen Herkunft berechnet werden (Raudenbush & Bryk, 2002, S. 75 ff.). Die Varianz der schulspezifischen Regressionskoeffizienten kann dabei auch inferenzstatistisch getestet werden, d. h., es wird überprüft, ob die Unterschiede zwischen den Schulen signifikant ausfallen.

Im Rahmen der Literaturrecherche wurde eine Reihe von Studien ermittelt, die relative *Klasseneffekte* auf herkunftsbedingte Unterschiede im Lernerfolg untersuchen. In der Mehrheit der Studien zeigt sich, dass der Regressionskoeffizient der sozialen Herkunft signifikant zwischen Klassen variiert, sowohl in deutschen Studien (Pfeifer, 2011, S. 119; Schwippert, 2001, S. 127 ff.; Stahl, 2009, S. 176; Stancel-Piatak, 2017, S. 214) als auch in internationalen Studien (Kyriakides et al., 2018b, S. 14 f.; Nye et al., 2004, S. 252). In zwei deutschsprachigen Studien ist die Varianz des Regressionskoeffizienten dagegen nicht signifikant (Carigiet Reinhard, 2012, S. 319; Schulz-Heidorf, 2016, S. 169). Zu relativen *Schuleffekten* sind ebenfalls mehrere Studien verfügbar, wobei es sich hierbei zum größten Teil um internationale Studien handelt. Die Befundlage fällt heterogener aus. Es finden sich mehrere Studien, in denen der Regressionskoeffizient der sozialen Herkunft signifikant zwischen Schulen variiert (Kyriakides et al., 2018a, S. 115 ff.; Kyriakides et al., 2018b; S. 14 f. Strand, 2010, S. 307; Thomas et al., 1997, S. 461; Wenglinsky, 1998, S. 276), aber auch mehrere Studien, in denen die Varianz nicht signifikant ausfällt (Brandsma & Knuver, 1989, S. 783; Kyriakides, 2004, S. 155; Radisch et al., 2006, S. 45 f.; Strand, 2016, S. 130; Young & Fraser, 1993, S. 277). Besondere Beachtung verdient in diesem Kontext die Studie von Gustafsson und Kolleg/-innen (2018), in welcher auf der Grundlage von TIMSS-Daten Schuleffekte auf herkunftsbedingte Kompetenzunterschiede in verschiedenen Ländern verglichen werden. Es zeigt sich, dass in 12 von 50 Ländern signifikante Unterschiede zwischen Schulen bestehen (S. 28 f.).

Anzumerken ist, dass die Aussagekraft von vielen der genannten Studien insofern begrenzt ist, als dass lediglich Signifikanztests durchgeführt werden. Diese sind abhängig von der Stichprobengröße, was eine mögliche Erklärung für die heterogenen Befunde darstellt. Außerdem machen Signifikanztests keine Aussagen darüber, *wie stark* die Unterschiede zwischen Klassen/Schulen ausfallen (dies wird bei Effekten auf den durchschnittlichen Lernerfolg durch die Intra-class Correlation geleistet).⁷ In den wenigen Studien, in denen inhalt-

7 Hieran wird eine gewisse Inkonsistenz bei den üblichen Methoden zur Untersuchung von Effekten auf den durchschnittlichen Lernerfolg und auf Unterschiede im Lernerfolg deutlich. Während bei Effekten auf den durchschnittlichen Lernerfolg üb-

liche Überlegungen zur Größe der Effekte angestellt werden, sind die Unterschiede zwischen Schulen durchgängig als inhaltlich bedeutsam einzuschätzen (Strand, 2010, S. 307; Strand, 2016, S. 130; Young & Fraser, 1993, S. 284). Bemerkenswerterweise werden teilweise auch dann inhaltlich bedeutsame Unterschiede zwischen Schulen gefunden, wenn der Signifikanztest negativ ausfällt.

Das typische *Spektrum* der schulspezifischen Regressionskoeffizienten umfasst als niedrigste Werte schwach positive Regressionskoeffizienten oder Regressionskoeffizienten nahe null (unter Kontrolle der Ausgangskompetenz). Negative Regressionskoeffizienten sind nicht zu beobachten. Demnach kann der Kompetenzzuwachs von Schüler/-innen verschiedener sozialer Herkunft durch eine geeignete Gestaltung von Schule zwar angeglichen werden, aber bestehende Kompetenzunterschiede können nicht abgebaut werden (in dem Sinne, dass sozial benachteiligte Schüler/-innen einen größeren Kompetenzzuwachs aufweisen als sozial privilegierte Schüler/-innen).

Insgesamt betrachtet sprechen die Befunde mehrheitlich dafür, dass bedeutsame relative Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg bestehen. Bei relativen Schuleffekten sind die Befunde dagegen als heterogen einzuschätzen. Eine mögliche Interpretation dieser Befundlage ist, dass relative Klasseneffekte größer ausfallen als relative Schuleffekte, ähnlich wie beim durchschnittlichen Lernerfolg. Zwei Studien, die gleichzeitig Klassen- und Schuleffekte auf Unterschiede im Lernerfolg untersuchen, stützen diese Annahme (Kyriakides et al., 2018b, S. 14; Nye et al., 2004, S. 252).

4.4 Die Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg

Die bisher dargestellten Befunde sprechen dafür, dass sowohl der durchschnittliche Lernerfolg als auch herkunftsbedingte Unterschiede im Lernerfolg bedeutsam zwischen Klassen variieren und teilweise auch zwischen Schulen. Dies wirft die Frage auf, in welchem Zusammenhang diese zwei pädagogischen Effekte zueinanderstehen. Wenn ein hoher durchschnittlicher Lernerfolg in der

licherweise die ICC und somit eine Effektstärke verwendet wird, wird bei Effekten auf Unterschiede im Lernerfolg üblicherweise der Signifikanztest verwendet. Eine naheliegende Erklärung hierfür ist, dass für die Varianz von Regressionskoeffizienten gegenwärtig keine standardisierte Effektstärke zur Verfügung steht. Gleichwohl wäre eine solche Effektstärke von hohem inhaltlichem Interesse, da hierdurch ermittelt werden könnte, *wie stark* herkunftsbedingte Kompetenzunterschiede zwischen Klassen/Schulen variieren.

Regel mit geringen Unterschieden im Lernerfolg einhergeht, sind die beiden *Zielsetzungen vereinbar*. Wenn andererseits ein hoher durchschnittlicher Lernerfolg in der Regel mit großen Unterschieden im Lernerfolg einhergeht, besteht ein *Zielkonflikt*. Eine wesentliche Rolle spielt hierbei der Mechanismus, durch den Klassen/Schulen geringe Unterschiede im Lernerfolg erreichen (Baumert & Schümer, 2001, S. 389). Werden geringe Unterschiede im Lernerfolg durch die Steigerung des Lernerfolgs von sozial benachteiligten Schüler/-innen erreicht, resultiert ein hoher durchschnittlicher Lernerfolg (Zielvereinbarkeit). Werden geringe Unterschiede im Lernerfolg hingegen dadurch erreicht, dass sozial privilegierte Schüler/-innen in ihrem Lernerfolg beeinträchtigt werden, resultiert ein niedriger durchschnittlicher Lernerfolg (Zielkonflikt).

Im Rahmen der Literaturrecherche wurden mehrere Studien ermittelt, die das Verhältnis zwischen den beiden Zielkriterien untersuchen. Die *analytische Vorgehensweise* besteht darin, für jede Klasse/Schule einen Kennwert für den durchschnittlichen Lernerfolg (z. B. adjustierter Kompetenzmittelwert) und einen Kennwert für Unterschiede im Lernerfolg (z. B. Regressionskoeffizient des Effekts der sozialen Herkunft auf die Kompetenz) zu berechnen und diese Kennwerte zu korrelieren.

Die Befundlage fällt insgesamt sehr heterogen aus. Zunächst werden Studien auf der *Klassenebene* vorgestellt. Schwippert (2001) berichtet in seiner mehrere Schulformen umfassenden Analyse, dass höhere Kompetenzmittelwerte mehrheitlich mit geringeren Kompetenzunterschieden nach sozialer Herkunft einhergehen (S. 119 ff.). Zu einem ähnlichen Ergebnis kommen Kyriakides und Kolleg/-innen (2018b) für verschiedene Fächer und verschiedene Schulstufen. Die genannten Studien sprechen für die Vereinbarkeit der zwei Zielsetzungen. Demgegenüber ermittelt Carigiet Reinhard (2012), dass höhere Kompetenzmittelwerte mit größeren Kompetenzunterschieden nach sozialer Herkunft einhergehen (S. 321). Diese Studie spricht also für einen Zielkonflikt.

Erwähnenswert sind in diesem Kontext einige ältere Studien aus Deutschland, die das Verhältnis zwischen Kompetenzmittelwert und Kompetenzegalisierung untersuchen. Unter *Egalisierung* wird hierbei entweder eine Abnahme der Varianz der Kompetenzen oder ein niedriger Zusammenhang zwischen Ausgangskompetenz und Endkompetenz verstanden. Die auf die Klassenebene bezogenen Befunde fallen ebenfalls sehr heterogen aus: teilweise gehen höhere Kompetenzmittelwerte mit größeren Kompetenzunterschieden einher (Baumert et al., 1986, S. 646; Treiber & Schneider, 1980, S. 271 f.), teilweise besteht kein Zusammenhang (Helmke, 1988, S. 58 f.), teilweise gehen höhere Kompetenzmittelwerte mit geringeren Kompetenzunterschieden einher (Treinies & Einsiedler, 1996, S. 297).

In der internationalen Literatur finden sich weiterhin mehrere Studien, die die Vereinbarkeit von hohem durchschnittlichem Lernerfolg und geringen herkunftsbedingten Unterschieden im Lernerfolg auf der *Schulebene* betrachten. Auch hier zeigt sich eine heterogene Befundlage. In mehreren Studien gehen höhere Kompetenzmittelwerte von Schulen mit geringeren Kompetenzunterschieden nach sozialer Herkunft einher (Kyriakides et al., 2018a, S. 120 f.; Kyriakides et al., 2018b, S. 15; Strand, 2014, S. 239), aber ebenfalls in mehreren Studien gehen höhere Kompetenzmittelwerte mit größeren Kompetenzunterschieden nach sozialer Herkunft einher (Lee & Bryk, 1989, S. 181; Thomas et al., 1997, S. 465; Strand, 2010, S. 304).

Insgesamt betrachtet kann die Frage, ob die Zielsetzungen eines hohen durchschnittlichen Lernerfolgs und geringer Unterschiede im Lernerfolg vereinbar, unabhängig oder konfligierend sind, nicht abschließend beantwortet werden. Es steht nur eine überschaubare Anzahl an Studien zur Verfügung und die Befunde fallen sehr heterogen aus. Dass keine klaren Belege dafür vorliegen, dass ein hoher durchschnittlicher Lernerfolg automatisch mit geringen Unterschieden im Lernerfolg einhergeht, kann folgendermaßen interpretiert werden: Allgemeine Qualitätsverbesserungen, die auf eine Steigerung des durchschnittlichen Lernerfolgs abzielen, werden vermutlich nicht ausreichen, um (praktisch nebenbei) auch das Problem der Unterschiede im Lernerfolg zu beheben (Mortimore, 1998, S. 310 f.; Sammons, 2010, S. 56). Es ist anzunehmen, dass hierfür gezielte Maßnahmen notwendig sind, die spezifisch auf die Abschwächung von Unterschieden im Lernerfolg ausgerichtet sind.

4.5 Zusammenfassung des vierten Kapitels

Im vorliegenden Kapitel wurden relevante Forschungsbefunde zu den Effekten von Klassen/Schulen auf den durchschnittlichen Lernerfolg und auf herkunftsbedingte Unterschiede im Lernerfolg dargestellt. Hinweise auf *absolute Effekte auf den durchschnittlichen Lernerfolg* liefert der Kompetenzzuwachs von Schüler/-innen im Laufe eines Schuljahres und der Vergleich des Kompetenzzuwachses während der Schulzeit und während der Ferienzeit. *Relative Effekte auf den durchschnittlichen Lernerfolg* werden durch Unterschiede zwischen Klassen/Schulen im adjustierten Kompetenzmittelwert der Schüler/-innen belegt. Relative Klasseneffekte sind hierbei stärker als relative Schuleffekte. Weiterhin liegen Hinweise darauf vor, dass relative Effekte bei sozial benachteiligten Schüler/-innen stärker ausfallen als bei sozial privilegierten Schüler/-innen, aber die Befundlage ist nicht ganz eindeutig. *Absolute Effekte auf herkunftsbedingte Unterschiede im Lernerfolg* können durch die Veränderung

von Kompetenzunterschieden nach sozialer Herkunft nachgezeichnet werden, insbesondere durch die Gegenüberstellung von Schulzeit und Ferienzeit. Die Befunde sprechen dafür, dass die Schule das Auseinanderdriften von Kompetenzunterschieden abbremst. *Relative Effekte auf herkunftsbedingte Unterschiede im Lernerfolg* zeigen sich demgegenüber darin, dass der Zusammenhang zwischen sozialer Herkunft und Lernerfolg in verschiedenen Klassen/Schulen unterschiedlich stark ausfällt. Relative Klasseneffekte können hierbei konsistenter belegt werden als relative Schuleffekte. Zur Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs und geringer Unterschiede im Lernerfolg fallen die Befunde ausgesprochen heterogen aus, so dass eine abschließende Einschätzung nicht möglich ist.

Auf der Grundlage der bisher dargestellten Befunde sollen nun die eingangs dargestellten *theoretischen Positionen* zur Rolle der Schule erneut beleuchtet werden. Für die Frage, ob Schule unterschiedsverringernde oder unterschiedsvergrößernde Effekte hat, spielen verschiedene Argumente eine Rolle (Raudenbush & Eschenbach, 2015, S. 453 ff.). Ein wichtiges *Argument für einen unterschiedsverringernenden Effekt von Schule* ist der Befund, dass Kinder verschiedener sozialer Herkunft durch ihre Eltern unterschiedlich viel Anregung erfahren (siehe Kapitel 3.5 „Unterschiede bei familiären Lernumwelten“). Vor diesem Hintergrund kann angenommen werden, dass es durch den gemeinsamen Schulbesuch zu einer Angleichung der von sozial privilegierten Kindern und sozial benachteiligten Kindern insgesamt erfahrenen Anregung kommt. Demgegenüber ist ein *wichtiges Argument für einen unterschiedsvergrößernden Effekt*, dass sozial privilegierte Kinder innerhalb der Lernumwelt Schule mehr Anregung erfahren als sozial benachteiligte Kinder. Dies wird durch Befunde belegt, wonach sozial privilegierte Kinder bessere Schulen besuchen und schulische Lernangebote besser nutzen als sozial benachteiligte Kinder (siehe Kapitel 3.7 „Unterschiede bei Schulen“).

Für eine Einschätzung, wie stark die beiden Argumente jeweils ins Gewicht fallen, sind die Untersuchungen zum Ferieneffekt aufschlussreich. Dass Kompetenzunterschiede während der Ferienzeit stärker zunehmen als während der Schulzeit, spricht dafür, dass der *Schulbesuch insgesamt einen unterschiedsverringernenden Effekt* hat. Das Befundmuster zeigt, dass der lernförderliche Effekt des Schulbesuchs bei sozial benachteiligten Kindern stärker ausfällt als bei sozial privilegierten Kindern. Hieraus kann geschlossen werden, dass die Angleichung der insgesamt erfahrenen Anregung ein größeres Gewicht hat als die schulbezogenen Vorteile der sozial privilegierten Kinder. Somit gibt es zwar Unterschiede in den schulischen Lernerfahrungen zugunsten von sozial privilegierten Kindern, aber die Unterschiede in den familiären Lernerfahrungen

sind sogar noch größer (Downey et al. 2004, S. 614; Raudenbush & Eschmann, 2015, S. 453).

Aufschlussreich ist in diesem Kontext die Perspektive, dass den theoretischen Positionen, wonach die Schule entweder als „Gleichmacherin“ oder als „Verursacherin von Unterschieden“ angesehen wird, *verschiedene Vergleichsmaßstäbe* zugrunde liegen. Vergleicht man das bestehende Schulsystem mit einer fiktiven Gesellschaft ohne Schulsystem (analog zum Vergleich von Schulbesuch vs. kein Schulbesuch), hat die Schule einen unterschiedsverringenden Effekt. Vergleicht man das bestehende Schulsystem dagegen mit einer fiktiven Gesellschaft mit einem Schulsystem, in dem sich Schüler/-innen verschiedener sozialer Herkunft gleichmäßig auf die vorhandenen Schulen verteilen, hat die Schule unterschiedsvergrößernde Effekte. Hieran wird deutlich, dass sich die beiden Positionen nicht zwangsläufig widersprechen, wie es auf den ersten Blick erscheinen mag.

Abschließend ist anzumerken, dass die Frage nach den absoluten Effekten von Schule (Schule ja/nein) auf Kompetenzunterschiede nach sozialer Herkunft zwar theoretisch interessant, aber unter praktischen Gesichtspunkten weniger relevant ist. Dass die Gesellschaft weiterhin öffentliche Schulen betreiben wird, kann als gesichert angesehen werden. Eine größere Bedeutung für praktische Entscheidungen kommt daher den relativen Effekten von Klassen/Schulen zu. Empirische Befunde zu Unterschieden zwischen Klassen/Schulen lassen erkennen, inwieweit die Ausgestaltung von Schule und Unterricht einen Einfluss auf Kompetenzunterschiede nach sozialer Herkunft hat. Dies ermöglicht eine Einschätzung, ob die Ausgestaltung von Schule und Unterricht einen aussichtsreichen Ansatzpunkt für Maßnahmen zur Verringerung von Kompetenzunterschieden darstellt.

Der zuvor gegebene Forschungsüberblick zeigt, dass nur eine überschaubare Anzahl an Studien zu relativen Effekten auf herkunftsbedingte Kompetenzunterschiede vorliegt und dass die Befundlage teilweise sehr heterogen ausfällt. Forschungsdefizite bestehen insbesondere bei der Frage nach der Größe von relativen schulischen Effekten auf herkunftsbedingte Unterschiede im Lernerfolg und bei der Frage nach der Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg. Die vorliegende Arbeit leistet in diesen Bereichen einen empirischen Beitrag.

5. Effekte von pädagogischen Merkmalen

Die im letzten Kapitel berichteten Forschungsbefunde sprechen dafür, dass die Bildungseinrichtung Schule einen Einfluss auf Kompetenzunterschiede nach sozialer Herkunft hat. Der Schwerpunkt der Betrachtung lag bei den Effekten von pädagogischen Einheiten wie Klassen und Schulen. Die Befunde zu relativen Effekten legen nahe, dass sich die Ausgestaltung von Schule und Unterricht auf herkunftsbedingte Unterschiede im Lernerfolg auswirkt. Die Perspektive auf pädagogische Einheiten ist allerdings insofern begrenzt, als dass keine Rückschlüsse darüber möglich sind, *wie* Schule und Unterricht gestaltet werden sollten, um herkunftsbedingte Unterschiede im Lernerfolg gering zu halten. Um diese unter dem Gesichtspunkt evidenzbasierter Entscheidungen hoch relevante Frage zu bearbeiten, muss die Perspektive auf *pädagogische Merkmale* gerichtet werden. Durch die Analyse von pädagogischen Merkmalen kann erklärt werden, warum es bestimmten Klassen/Schulen besser gelingt, herkunftsbedingte Unterschiede im Lernerfolg gering zu halten, als anderen Klassen/Schulen. Hierfür wird untersucht, bei welchen Merkmalen von Klassen/Schulen unterschiedliche Ausprägungen zu unterschiedlich starken Kompetenzunterschieden nach sozialer Herkunft führen (z. B. bei häufigem Einsatz von Gruppenarbeiten fallen Kompetenzunterschiede geringer aus als bei seltenem Einsatz von Gruppenarbeiten).

In der Literatur finden sich diverse Überlegungen zu geeigneten pädagogischen bzw. bildungspolitischen *Maßnahmen zur Verringerung von Kompetenzunterschieden nach sozialer Herkunft* (Expertenrat Herkunft und Bildungserfolg, 2011, S. 10 ff.; Fend, 2013, S. 129, 137 ff.; Geißler & Weber-Menges, 2010, S. 161 f.; Neumann et al., 2014, S. 192 f.). Diese Überlegungen werden nachfolgend zusammengefasst. Im Bereich der *Familie* wird die flächendeckende Einführung von frühzeitiger Elternberatung mit aufsuchendem und dadurch niedrigschwelligem Ansatz als aussichtsreiche Maßnahme angesehen. Im *frühpädagogischen Bereich* wird in quantitativer Hinsicht auf den weiteren Ausbau frühpädagogischer Einrichtungen (insbesondere in Wohngebieten mit hohem Anteil an sozial benachteiligten Familien) und in qualitativer Hinsicht auf die Verbesserung der Qualität der Einrichtungen gesetzt. Als zentrale Aspekte frühpädagogischer Qualität werden hierbei eine systematische Entwicklungsdiagnostik und darauf aufbauende individuelle Fördermaßnahmen angesehen (insbesondere bezüglich kognitiven Lernvoraussetzungen). Notwendig

hierfür erscheinen verbindliche Qualitätsstandards sowie eine entsprechende Aus- und Weiterbildung frühpädagogischer Fachkräfte. Auch im *Schulbereich* wird der Verbesserung der Diagnostik und Förderung von Schüler/-innen mit Lerndefiziten eine hohe Bedeutung beigemessen. Als Intervention bei Lernproblemen wird insbesondere schulischer Förderunterricht empfohlen. Um die notwendigen Rahmenbedingungen für individuelle Förderung zu schaffen, werden die Aus- und Weiterbildung von Lehrkräften (zu Themen wie Diagnostik und Fördermaßnahmen), der vermehrte Einsatz von spezialisierten Fachkräften (z. B. Schulsozialarbeiter/-innen, Schulpsycholog/-innen) und der Ausbau von Ganztagschulen (insbesondere in Wohngebieten mit hohem Anteil an sozial benachteiligten Familien) als sinnvoll angesehen. Weiterhin in Betracht gezogen wird eine Verlängerung der Unterrichtszeit, entweder im Rahmen von Ganztagschulen oder in der Form von Sommerschulen. Schließlich werden auch Schulstrukturreformen als mögliche Maßnahmen diskutiert, d. h. eine Verlängerung der Grundschulzeit und eine Verringerung der Anzahl der Schulformen in der Sekundarschule.

Auf zwei Sachverhalte ist an dieser Stelle hinzuweisen. Erstens liegen nur bei einem Teil der in der Literatur diskutierten Maßnahmen *empirische Evidenzen* für deren Wirksamkeit zur Verringerung von Kompetenzunterschieden nach sozialer Herkunft vor (Expertenrat Herkunft und Bildungserfolg, 2011, S. 23), während bei einem anderen Teil die Wirksamkeit aufgrund theoretischer Argumente angenommen wird. Zweitens ist augenfällig, dass im Schulbereich die *Unterrichtsgestaltung* bisher kaum als möglicher Ansatzpunkt für unterschiedsverringende Maßnahmen diskutiert wird. Dies steht im Kontrast zu Befunden der schulbezogenen Wirksamkeitsforschung, wonach Unterricht der wichtigste schulische Einflussfaktor auf den Lernerfolg der Schüler/-innen ist (Hattie, 2013, S. 22, 27 f.; Muijs et al., 2014, S. 231).

Die Frage nach empirischen Evidenzen für Effekte von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg wird im vorliegenden Kapitel eingehender behandelt, auch unter Berücksichtigung der Rolle von Unterricht. Im Weiteren werden zunächst einige wichtige theoretische Grundlagen entwickelt. Dies umfasst einerseits ein theoretisches Modell der Wirksamkeit von Schule (Kapitel 5.1) und andererseits Überlegungen zu den Mechanismen, durch die pädagogische Merkmale herkunftsbedingte Unterschiede im Lernerfolg beeinflussen können (Kapitel 5.2). Anschließend wird ein Überblick über relevante Forschungsansätze gegeben (Kapitel 5.3). Im Weiteren werden vorhandene Befunde zu Effekten von pädagogischen Merkmalen in verschiedenen Bereichen von Schule berichtet. Behandelt werden Effekte

von Unterrichtsmerkmalen (Kapitel 5.4), Effekte von Ganztagsangeboten (Kapitel 5.5) und Effekte von Merkmalen der Lerngruppe (Kapitel 5.6).

Der Forschungsüberblick wird bewusst sehr ausführlich gehalten. Es sind zwar bereits zahlreiche Überblicksarbeiten zu den Effekten von pädagogischen Merkmalen vorhanden (z. B. Hattie, 2013; Lipowsky, 2006; Muijs & Reynolds, 2001), dies berücksichtigen allerdings nur Befunde zu Effekten auf den durchschnittlichen Lernerfolg. Der nachfolgende Forschungsüberblick stellt insofern eine Erweiterung dar, als dass auch Befunde zu Effekten auf herkunftsbedingte Unterschiede im Lernerfolg systematisch aufbereitet werden. Auf dieser Grundlage kann eingeschätzt werden, inwieweit schulische Maßnahmen zur Verringerung von herkunftsbedingten Kompetenzunterschieden durch empirische Evidenzen abgesichert sind bzw. inwieweit weiterer Forschungsbedarf besteht.

5.1 Theoretisches Modell der Wirksamkeit von Schule

Einleitend wird ein theoretisches Modell der Wirksamkeit von Schule entwickelt. Zunächst werden einige Grundlagen zur theoretischen Modellierung von Schule erläutert. Anschließend wird das Modell beschrieben.

5.1.1 Theoretische Grundlagen des Modells der Wirksamkeit von Schule

Es existieren mittlerweile zahlreiche *theoretische Modelle zu Schule und Unterricht* (Übersichten finden sich bei Helmke & Weinert, 1997a, S. 80 ff. und Scheerens & Bosker, 1997, S. 46 ff.). Die verschiedenen Modelle weisen zwar jeweils gewisse Ähnlichkeiten zueinander auf, setzen aber teilweise auch recht unterschiedliche Schwerpunkte. In der internationalen Literatur finden sich sowohl *integrative Modelle*, welche die Schule als Organisation im Ganzen betrachten (Scheerens & Bosker, 1997, S. 46 ff.) als auch *Unterrichtsmodelle*, welche auf Lehr-Lern-Prozesse fokussieren (Caroll, 1963; Harnischfeger & Wiley, 1978). In der deutschsprachigen Literatur dominiert gegenwärtig die zweitgenannte Modellart, insbesondere in der Form von *Angebots-Nutzungs-Modellen von Unterricht* (Helmke, 2014, S. 71; Klieme et al., 2006, S. 131; Seidel & Reiss, 2014, S. 256). Bei der Entwicklung des theoretischen Rahmens zeigte sich, dass keines der genannten Modelle eine befriedigende Passung zur vorliegenden Untersuchung aufwies. Vor diesem Hintergrund wurde ein Modell der Wirksamkeit von Schule entwickelt, welches auf die Spezifika der vorliegenden

Untersuchung abgestimmt ist. Als Vorlage diente insbesondere das Modell von Helmke (2014, S. 71), welches inhaltlich erweitert wurde.

Das Modell greift weiterhin einige der in Kapitel 2 erfolgten theoretischen Überlegungen zu pädagogischer Wirksamkeit auf. Eine wichtige konzeptionelle Grundlage ist die Unterscheidung verschiedener *Arten von pädagogischen Merkmalen*. Zentral ist die Gegenüberstellung von Prozess- und Strukturmerkmalen (in Anlehnung an Bäumler et al., 2011, S. 94 f.; Scheerens et al., 2007, S. 27 f.). Weiterhin von Bedeutung ist die Unterscheidung verschiedener *Ebenen von Schule*. Aus der Mehrebenenstruktur von Schule ergibt sich, dass pädagogische Merkmale auf unterschiedlichen Analyseebenen verortet sein können (Ditton, 2000, S. 76). Merkmale auf der *Schülerebene* können bei allen Schüler/-innen unterschiedliche Ausprägungen aufweisen (z. B. Häufigkeit eines Lobs durch die Lehrkraft). Merkmale auf der *Klassenebene* weisen bei allen Schüler/-innen einer Klasse die gleiche Ausprägung auf, können aber bei verschiedenen Klassen unterschiedlich ausgeprägt sein (z. B. pädagogische Überzeugungen der Lehrkraft). Merkmale auf der *Schulebene* weisen bei allen Schüler/-innen einer Schule die gleiche Ausprägung auf, können aber bei verschiedenen Schulen unterschiedlich ausgeprägt sein (z. B. pädagogische Überzeugungen des Schulleiters/der Schulleiterin). Im nachfolgenden Modell werden, aus Gründen einer übersichtlichen Darstellung, die Klassenebene und die Schulebene zusammengefasst, es werden aber differenzierende Erläuterungen im Text gegeben.

5.1.2 Beschreibung des Modells der Wirksamkeit von Schule

Die primäre Funktion des Modells ist eine theoretische Ordnung der Lernumwelt Schule durch die Unterscheidung von Kategorien pädagogischer Merkmale und Annahmen über Wirkzusammenhänge zwischen diesen Kategorien. Es bildet sowohl die Grundlage für den nachfolgenden Forschungsüberblick als auch für die empirischen Analysen. Abbildung 5 stellt das Modell der Wirksamkeit von Schule dar.

Aus dem Modell von Helmke (2014) übernommene Komponenten sind die Modellierung von Unterricht anhand einer Angebots-Nutzungs-Struktur, die Berücksichtigung von Einflussfaktoren auf Unterricht und die Berücksichtigung von Schüler- und Familienmerkmalen. Ein wesentlicher Unterschied besteht darin, dass sich das vorliegende Modell nicht nur auf Unterricht bezieht, sondern auch auf Ganztagsangebote. Weiterhin ergänzt wurden die Kategorie der Hintergrundprozesse sowie Interaktionseffekte zwischen schulischen Lernangeboten und Familienmerkmalen bzw. Lernvoraussetzungen von Schüle-

r/-innen. Anhand der oberen Leiste wird außerdem die Schul- und Klassenebene der Schülerebene gegenübergestellt. Auf die Darstellung der Systemebene wird verzichtet, weil diese für die vorliegende Arbeit nicht weiter von Bedeutung ist.

Nachfolgend wird das Modell erläutert. An dieser Stelle erfolgt nur eine überblicksartige Darstellung, während ausgewählte Kategorien des Modells in den späteren Abschnitten noch genauer betrachtet werden. Die *Komponenten auf der Schülerebene* finden sich in gleicher Form auch im theoretischen Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg und wurden dort bereits ausführlich erläutert. Außerschulische Einflussfaktoren werden im vorliegenden Modell nur vereinfacht dargestellt. Beispielsweise werden außerschulische Lernumwelten wie Gleichaltrigengruppe oder freizeitbezogene Einrichtungen nicht abgebildet, weil sie in diesem Kontext nur eine untergeordnete Rolle spielen (ähnlich wie im Modell von Helmke). Weiterhin ist die Darstellung der familiären Lernumwelt simplifizierend, weil eine differenzierte Betrachtung bereits durch das andere Modell geleistet wird. Hervorzuheben ist, dass zu den familiären Merkmalen auch der soziale Status zählt, weshalb der von der Familie ausgehende Pfeil den *Effekt der sozialen Herkunft auf den Lernerfolg* repräsentiert.

Der Lernerfolg der Schüler/-innen wird wesentlich von schulischen Lernangeboten beeinflusst. Bei schulischen Merkmalen dieser Kategorie handelt es sich um Prozessmerkmale mit direkten Effekten auf die Lernaktivitäten der Schüler/-innen. Die wichtigsten Lernangebote sind der reguläre, fachbezogene *Unterricht* sowie ergänzende *Ganztagsangebote* (z. B. Hausaufgabenbetreuung, Freizeitangebote). Unterricht und Ganztagsangebote können jeweils bezüglich ihres Umfangs, ihrer Inhalte und ihrer methodischen Gestaltung betrachtet werden. Bei beiden Arten von Lernangeboten gilt, dass sich die Prozessmerkmale sowohl auf das Verhalten der durchführenden pädagogischen Fachkraft (z. B. Lehrkraft-Schüler-Interaktionen) als auch auf das Verhalten der anwesenden Mitschüler/-innen (z. B. Schüler-Schüler-Interaktionen) beziehen können.

Hinsichtlich der *Einordnung anhand von Ebenen* verdeutlicht die Lage im Modell, dass schulische Lernangebote auf unterschiedlichen Ebenen betrachtet werden können. Unterricht wird typischerweise auf der Klassenebene betrachtet und dann als geteilte Lernumwelt verstanden (Lüdtke et al., 2006, S. 85 f.). Ganztagsangebote sind demgegenüber typischerweise auf der Schulebene verortet, da sie Schüler/-innen aus mehreren Klassen zugänglich sind. Weiterhin können bestimmte Merkmale von schulischen Lernangeboten auf der Schülerebene liegen, da sie bei verschiedenen Schüler/-innen unterschiedlich ausgeprägt sind. Ein Beispiel wäre, dass eine Lehrkraft verschiedene Schüler/-

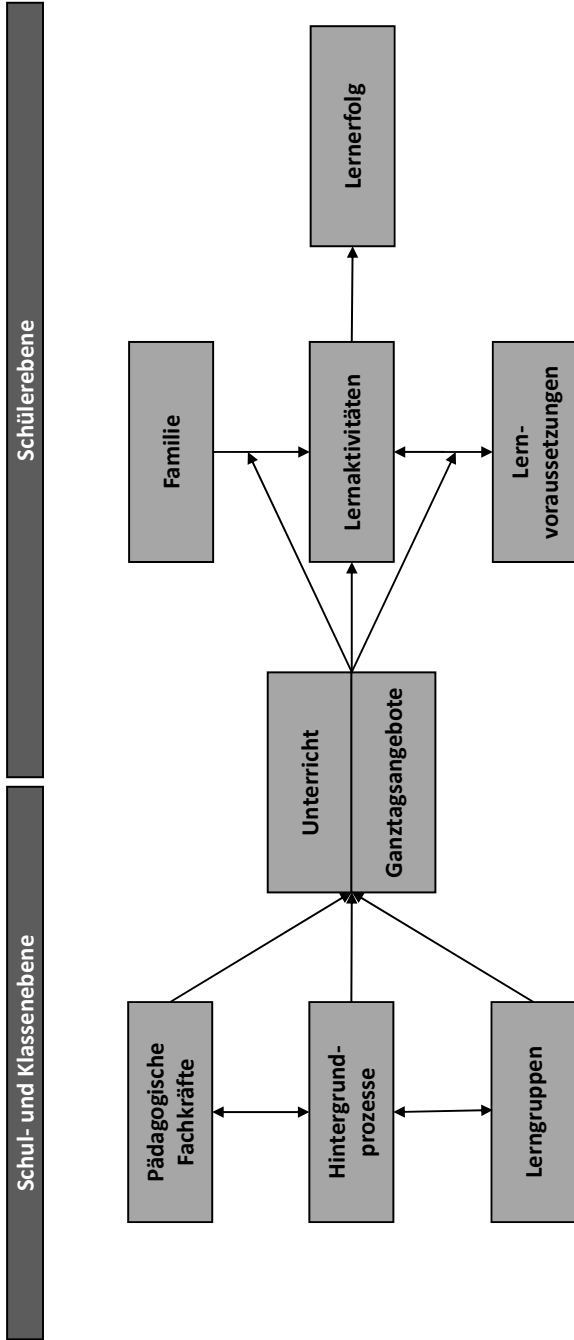


Abb. 5: Theoretisches Modell der Wirksamkeit von Schule

innen im Unterricht unterschiedlich behandelt (z. B. Ausmaß an individueller Unterstützung).

Hervorzuheben ist, dass Unterricht und Ganztagsangebote im Modell *unterschiedliche Effekte* haben. Der *mittlere Pfeil* auf die Lernaktivitäten steht für einen Haupteffekt der schulischen Lernangebote und bezieht sich somit auf den durchschnittlichen Lernerfolg. Der *obere und untere Pfeil* beziehen sich dagegen auf Interaktionseffekte mit den Merkmalen der Familie bzw. mit Merkmalen des Schülers/der Schülerin. Der *obere Pfeil* steht für Effekte von schulischen Lernangeboten auf herkunftsbedingte Unterschiede im Lernerfolg. Dies repräsentiert die Annahme, dass die Ausgestaltung von schulischen Lernangeboten einen Einfluss darauf hat, wie stark Familienmerkmale (wie die soziale Herkunft) den Lernerfolg von Schüler/-innen bedingen. In ähnlicher Weise steht der *untere Pfeil* für die Annahme, dass die Ausgestaltung von schulischen Lernangeboten einen Einfluss darauf hat, wie stark die Lernvoraussetzungen der Schüler/-innen ihren Lernerfolg bedingen (z. B. wie stark die Intelligenz den Lernerfolg bedingt, siehe Opdenakker & Van Damme, 2000, S. 181).

Die Ausgestaltung von Unterricht und Ganztagsangeboten hängt von einer Reihe von pädagogischen Merkmalen ab. Insofern beziehen sich die nachfolgend behandelten Kategorien auf Merkmale mit *indirekten Effekten*. Von zentraler Bedeutung sind Merkmale der durchführenden *pädagogischen Fachkräfte*. Bei Unterricht sind dies Lehrkräfte, bei Ganztagsangeboten können es neben Lehrkräften auch andere Fachkräfte sein (z. B. Erzieher/-innen, Sozialpädagog/-innen). Weiterhin zu den pädagogischen Fachkräften zählen Schulleiter/-innen. Relevante Merkmale von pädagogischen Fachkräften sind Fachkompetenzen, pädagogische Kompetenzen, pädagogische Überzeugungen und berufliche Motivation (angelehnt an Baumert & Kunter, 2006).

Weitere Einflussfaktoren auf schulische Lernangebote sind Merkmale der *Lerngruppen*. Diese Kategorie verweist auf die Mitschüler/-innen. Lerngruppen können auf unterschiedlichen Ebenen betrachtet werden. Bei Unterricht bezieht sich die Lerngruppe auf die Klasse, bei Ganztagsangeboten auf die Gruppe der teilnehmenden Schüler/-innen. Daneben können auf der Schulebene alle Schüler/-innen einer Schule betrachtet werden. Wichtige Merkmale zur Beschreibung von Lerngruppen sind die Größe (z. B. Klassengröße, Schulgröße) und die Komposition (z. B. Klassenkomposition, Schulkomposition).

Eine dritte Kategorie von Einflussfaktoren auf schulische Lernangebote sind *Hintergrundprozesse*. Dabei handelt es sich um Aktivitäten von pädagogischen Fachkräften, welche außerhalb der eigentlichen Lernangebote erfolgen und die darauf abzielen, Lernangebote zu ermöglichen, zu erleichtern oder zu verbessern. Relevante Hintergrundprozesse sind Vor- und Nachbereitung, Eva-

luation, Weiterbildung und Schulleitungshandeln. Hintergrundprozesse liegen teilweise auf der Klassenebene (z. B. Unterrichtsvorbereitung der Lehrkraft) und teilweise auf der Schulebene (z. B. Führungsstil der Schulleitung). Diese Kategorie stellt insofern einen Sonderfall dar, als dass es sich um Prozessmerkmale mit indirekten Effekten handelt (die restlichen Kategorien des Modells sind entweder Prozessmerkmale mit direkten Effekten oder Strukturmerkmale mit indirekten Effekten).

Im Weiteren wird der Begriff „pädagogisches Merkmal“ als Oberkategorie für alle Merkmale innerhalb von Schulen verwendet. Pädagogische Merkmale schließen also Unterrichtsmerkmale, Merkmale von Ganztagsangeboten, Merkmale von Fachkräften, Merkmale von Lerngruppen und Merkmale von Hintergrundprozessen ein.

5.2 Mechanismen der Effekte von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg

Neben der Modellierung der Lernumwelt Schule sollen durch den theoretischen Rahmen des vorliegenden Kapitels auch Erklärungen ermöglicht werden, aufgrund welcher Mechanismen pädagogische Merkmale wirksam sind. Im zweiten Kapitel wurden bereits *Mechanismen* eingeführt, durch die *Effekte auf den durchschnittlichen Lernerfolg* erklärt werden können (siehe Kapitel 2.2.4 „Lernaktivitäten und Lernvoraussetzungen“). Demnach beeinflussen schulische Lernangebote den durchschnittlichen Lernerfolg der Schüler/-innen, indem sie entweder die fachlichen Lernaktivitäten (Quantität oder Qualität) oder die Lernvoraussetzungen der Schüler/-innen (z. B. Lernmotivation, metakognitive Kompetenzen) fördern.

Im Weiteren werden einige theoretische Überlegungen dazu angestellt, durch welche *Mechanismen* schulische Lernangebote *herkunftsbedingte Unterschiede im Lernerfolg* beeinflussen können. Anhand dieser theoretischen Mechanismen können pädagogische Merkmale ermittelt werden, bei denen ein Effekt auf herkunftsbedingte Unterschiede im Lernerfolg zu erwarten ist (a priori). Weiterhin kann überprüft werden, ob bei einem vorfindbaren empirischen Zusammenhang ein Kausaleffekt auf herkunftsbedingte Unterschiede im Lernerfolg plausibel erscheint (a posteriori). Theoretische Erklärungen dieser Art können Kausalinterpretationen fundieren und sind somit relevant, um die vorliegende Evidenz einzuschätzen. Nachfolgend werden zwei Gruppen von Mechanismen betrachtet, nämlich einerseits die Förderung verschiedener Schülergruppen und andererseits die Gegenüberstellung von differenziell wirksamen Lerngelegenheiten und unterschiedlichen Lerngelegenheiten.

5.2.1 Förderung verschiedener Schülergruppen

Im letzten Kapitel wurde bereits ausgeführt, dass pädagogische Einheiten wie Klassen oder Schulen geringe Unterschiede im Lernerfolg sowohl durch die besondere Förderung von sozial benachteiligten Schüler/-innen als auch durch die Beeinträchtigung von sozial privilegierten Schüler/-innen erreichen können. Diese Überlegung kann auf pädagogische Merkmale übertragen werden. Demnach haben pädagogische Merkmale einen unterschiedsverringenden Effekt, wenn sie entweder den *Lernerfolg von sozial benachteiligten Schüler/-innen besonders steigern* (z. B. durch intensive Unterstützung dieser Schülergruppe) oder den *Lernerfolg von sozial privilegierten Schüler/-innen beeinträchtigen* (z. B. durch ein für diese Schülergruppe zu langsames Lerntempo). Unzweifelhaft stellt nur die erste Variante eine pädagogisch vertretbare Vorgehensweise dar (Baumert & Schümer, 2001, S. 371; Strand, 2010, S. 292).

Grundsätzlich sind zahlreiche Möglichkeiten denkbar, in welchem Verhältnis die Lernfortschritte der beiden Schülergruppen zueinanderstehen können. Einige anschauliche Szenarien sollen herausgegriffen und genauer betrachtet werden. Zunächst werden Szenarien betrachtet, bei denen pädagogische Merkmale *unterschiedsverringende Effekte* haben. Der günstigste Fall liegt vor, wenn ein pädagogisches Merkmal den Lernerfolg beider Schülergruppen fördert, aber der Effekt bei sozial benachteiligten Schüler/-innen stärker ausfällt als bei sozial privilegierten Schüler/-innen. Die für das Lernen günstigen Ausprägungen des pädagogischen Merkmals sind somit bei beiden Schülergruppen identisch. Pädagogische Merkmale dieser Art stellen das Mittel der Wahl dar, denn sie sind sowohl für einen hohen durchschnittlichen Lernerfolg als auch für geringe Unterschiede im Lernerfolg wirksam. Eine weitere Möglichkeit besteht darin, dass ein pädagogisches Merkmal den Lernerfolg von sozial benachteiligten Schüler/-innen fördert, während es bei sozial privilegierten Schüler/-innen keinen Effekt hat. Das pädagogische Merkmal ist somit wirksam für geringe Unterschiede im Lernerfolg und in begrenztem Maß auch für einen hohen durchschnittlichen Lernerfolg. Es besteht schließlich auch die Möglichkeit, dass ein pädagogisches Merkmal einen positiven Effekt auf den Lernerfolg von sozial benachteiligten Schüler/-innen hat, aber einen negativen Effekt auf den Lernerfolg von sozial privilegierten Schüler/-innen. Für das Lernen der beiden Schülergruppen sind somit verschiedene Ausprägungen des pädagogischen Merkmals günstig, woraus sich ein Konflikt ergibt (z. B. für sozial benachteiligte Schüler/-innen sind viele Wiederholungen günstig, für sozial privilegierte Schüler/-innen sind wenige Wiederholungen günstig). In diesem Fall kann sich sogar ein negativer Effekt auf den durchschnittlichen Lernerfolg ergeben,

nämlich wenn der negative Effekt bei sozial privilegierten Schüler/-innen stärker ausfällt als der positive Effekt bei sozial benachteiligten Schüler/-innen.

Ebenso sind Szenarien denkbar, bei denen pädagogische Merkmale *unterschiedsvergrößernde Effekte* haben. Erwähnenswert ist das Szenario, bei dem ein pädagogisches Merkmal bei beiden Schülergruppen positive Effekte hat, aber der Effekt bei sozial privilegierten Schüler/-innen stärker ausfällt. Weiterhin interessant ist das Szenario, bei dem ein pädagogisches Merkmal den Lernerfolg von sozial privilegierten Schüler/-innen fördert, aber keinen Effekt bei sozial benachteiligten Schüler/-innen hat. In beiden Fällen würde der durchschnittliche Lernerfolg gesteigert werden, während gleichzeitig Unterschiede im Lernerfolg zunehmen würden. Derartige pädagogische Merkmale bei der Gestaltung von Schule und Unterricht grundsätzlich nicht zu nutzen, um Unterschiede im Lernerfolg gering zu halten, erscheint wiederum pädagogisch nicht vertretbar, da in beiden Fällen sozial benachteiligte Schüler/-innen nicht beeinträchtigt werden.

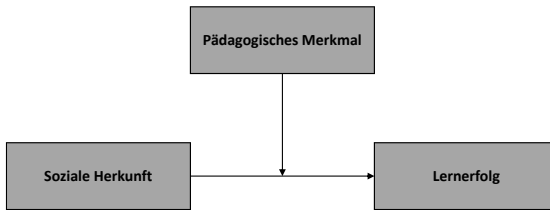
Diese Ausführungen verdeutlichen, dass das Ziel der Entkoppelung des Lernerfolgs von der sozialen Herkunft nicht absolut gesetzt werden sollte. Demnach erscheinen *nur bestimmte Strategien* der Zielerreichung als pädagogisch vertretbar. Für die Wirksamkeitsforschung ergibt sich hieraus, dass die Mechanismen von pädagogischen Merkmalen in Bezug auf verschiedene Schülergruppen empirisch überprüft werden sollten, da nur so ein fundiertes evidenzbasiertes Orientierungswissen gewonnen werden kann.

5.2.2 Differenziell wirksame Lerngelegenheiten und unterschiedliche Lerngelegenheiten

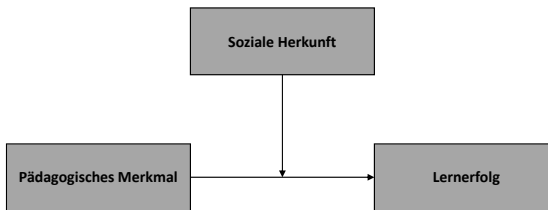
Unter dem Gesichtspunkt theoretischer Mechanismen erscheint es relevant zu klären, welche Wirkzusammenhänge zwischen pädagogischen Merkmalen innerhalb von Schulen, der sozialen Herkunft und dem Lernerfolg bestehen können. Grundsätzlich sind zwei Wirkzusammenhänge zu unterscheiden: *Moderation* (Synonym: Interaktion) und *Mediation* (Baron & Kenny, 1986). Beide Wirkzusammenhänge wurden in den vorausgehenden Ausführungen schon erwähnt: im Modell der Wirksamkeit von Schule (Abbildung 5) haben schulische Lernangebote eine moderierende Funktion, während im Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg (Abbildung 4) die Schule einen Mediator darstellt.

Zunächst sollen einige in der Literatur verwendete Konzepte inhaltlich geklärt werden. Zwei häufig verwendete Begrifflichkeiten sind „Effekt auf Unterschiede“ (im Lernerfolg) (z. B. Hanushek & Woessmann, 2006) und „dif-

Effekt auf Unterschiede im Lernerfolg (pädagogisches Merkmal als Moderator)



Differenziell wirksame Lerngelegenheiten (soziale Herkunft als Moderator)



Unterschiedliche Lerngelegenheiten (pädagogisches Merkmal als Mediator)

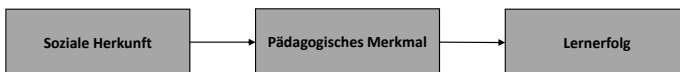


Abb. 6: Mögliche Wirkzusammenhänge zwischen pädagogischen Merkmalen, sozialer Herkunft und Lernerfolg

ferenzieller Effekt“ bzw. „differenzielle Wirksamkeit“ (z. B. Kyriakides, 2004; Vanlaar et al., 2015). Werden pädagogische Einheiten wie Klassen oder Schulen betrachtet, sind die beiden Begrifflichkeiten gleichbedeutend. Wenn beispielsweise eine Schule *differenziell wirksam* ist, wird der Lernerfolg einer Schülergruppe (z. B. sozial privilegierte Schüler/-innen) stärker gefördert als der Lernerfolg einer anderen Schülergruppe (z. B. sozial benachteiligte Schüler/-innen)

und somit hat die Schule einen *Effekt auf Unterschiede im Lernerfolg* zwischen diesen Schülergruppen.

Bei der Betrachtung von pädagogischen Merkmalen ergeben sich demgegenüber unterschiedliche Bedeutungen. Ein *Effekt auf Unterschiede im Lernerfolg* liegt vor, wenn verschiedene Ausprägungen eines pädagogischen Merkmals zu verschiedenen starken Unterschieden im Lernerfolg führen. Ein Beispiel wäre, dass an Schulen, an denen Förderunterricht angeboten wird, herkunftsbedingte Unterschiede im Lernerfolg niedriger ausfallen als an Schulen ohne Förderunterricht. Statistisch formuliert liegt eine Moderation vor, und zwar spezifisch eine Cross-Level-Interaktion (Aguinis et al., 2013). Das bedeutet, ein pädagogisches Merkmal auf einer höheren Ebene (Förderunterricht) wirkt als Moderator des Zusammenhangs zwischen sozialer Herkunft und Lernerfolg auf der Schülerebene. Eine grafische Veranschaulichung findet sich im oberen Teil von Abbildung 6.

Entsprechend der Mehrebenenstruktur von Schule können sich Cross-Level-Interaktionen auf *verschiedene Ebenen* beziehen. Im genannten Beispiel des Förderunterrichts liegt das moderierende Merkmal auf der Schulebene. Weitere Möglichkeiten wären, dass ein Merkmal auf der Klassenebene die Unterschiede im Lernerfolg innerhalb von Klassen beeinflusst (z. B. in Klassen mit stark strukturiertem Unterricht fallen Unterschiede im Lernerfolg geringer aus als in Klassen mit wenig strukturiertem Unterricht) oder dass ein Merkmal auf der Systemebene die Unterschiede im Lernerfolg innerhalb von Ländern beeinflusst (z. B. in Ländern mit Gesamtschulsystem fallen Unterschiede im Lernerfolg geringer aus als in Ländern mit gegliedertem Schulsystem).

Ein Effekt auf Unterschiede im Lernerfolg kann auf zwei Mechanismen zurückzuführen sein: differenziell wirksame Lerngelegenheiten und unterschiedliche Lerngelegenheiten. Eine *differenziell wirksame Lerngelegenheit* liegt vor, wenn die gleiche Lerngelegenheit bei Schüler/-innen mit verschiedener sozialer Herkunft unterschiedliche Lernwirkungen erzeugt. Ein verwandtes Konzept ist Aptitude-Treatment-Interaktion (Snow, 1989).⁸ Im genannten

8 Das Konzept der differenziellen Wirksamkeit ist inhaltlich weiter gefasst als das Konzept der Aptitude-Treatment-Interaktion. Differenzielle Wirksamkeit kann sich auf verschiedene Arten von pädagogischen Merkmalen (z. B. Unterrichtsmerkmale, Schulmerkmale) und auf verschiedene Arten von Schülermerkmalen beziehen (z. B. Lernvoraussetzungen, äußere Merkmale). Aptitude-Treatment-Interaktion bezieht sich dagegen spezifisch auf Interaktionen zwischen Unterrichtsmerkmalen („treatment“) und Lernvoraussetzungen der Schüler/-innen („aptitude“). In diesem Sinne stellt Aptitude-Treatment-Interaktion einen Sonderfall von differenzieller Wirksamkeit dar.

Beispiel könnte der unterschiedsverringende Effekt von Förderunterricht darauf zurückzuführen sein, dass Förderunterricht bei sozial benachteiligten Schüler/-innen stärkere Effekte auf den Lernerfolg hat als bei sozial privilegierten Schüler/-innen. Im statistischen Sinne liegt ebenfalls eine Moderation vor, allerdings wirkt hierbei die soziale Herkunft als Moderator des Zusammenhangs zwischen pädagogischem Merkmal und Lernerfolg (siehe mittlerer Teil von Abbildung 6). Spezifisch handelt es sich um einen Interaktionseffekt auf der Schülerebene, d. h., sowohl die soziale Herkunft als auch das pädagogische Merkmal werden auf der Schülerebene betrachtet.

Der zweite Mechanismus, der einem Effekt auf Unterschiede im Lernerfolg zugrunde liegen kann, sind *unterschiedliche Lerngelegenheiten*. Dieser Mechanismus ist dadurch gekennzeichnet, dass das pädagogische Merkmal bei Schüler/-innen verschiedener sozialer Herkunft unterschiedliche Ausprägungen aufweist. Der unterschiedsverringende Effekt von Förderunterricht könnte beispielsweise darauf zurückzuführen sein, dass sozial benachteiligte Schüler/-innen häufiger den Förderunterricht besuchen als sozial privilegierte Schüler/-innen. Statistisch formuliert liegt eine Mediation vor, weil einerseits die soziale Herkunft das pädagogische Merkmal beeinflusst und andererseits das pädagogische Merkmal den Lernerfolg (siehe unterer Teil von Abbildung 6).

Zusammenfassend ist festzuhalten, dass ein pädagogisches Merkmal einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg hat, wenn entweder der Mechanismus differenziell wirksamer Lerngelegenheiten oder der Mechanismus unterschiedlicher Lerngelegenheiten gegeben ist. Weiterhin ist es möglich, dass ein pädagogisches Merkmal gleichzeitig über beide Mechanismen wirkt. Der resultierende Effekt auf Unterschiede im Lernerfolg hängt dann davon ab, ob die beiden Mechanismen gleichförmig oder gegensätzlich wirken.⁹ Nachfolgend werden die beiden Mechanismen tiefergehend betrachtet.

9 Die Mechanismen wären beispielsweise gleichförmig, wenn Förderunterricht häufiger von sozial benachteiligten Schüler/-innen besucht wird und bei dieser Schülergruppe auch stärkere Effekte hat. Bei Gleichförmigkeit fällt der unterschiedsverringende Effekt folglich besonders stark aus. Dagegen wären die Mechanismen gegenläufig, wenn sozial benachteiligte Schüler/-innen den Förderunterricht häufiger besuchen, aber die Effekte bei sozial privilegierten Schüler/-innen stärker ausfallen (oder anders herum). Ob und in welche Richtung in diesem Fall ein Effekt auf Unterschiede im Lernerfolg resultiert, hängt davon ab, wie stark die beiden Mechanismen jeweils ins Gewicht fallen.

5.2.3 Vertiefte Betrachtung differenziell wirksamer Lerngelegenheiten

Im Weiteren wird theoretisch erörtert, warum schulische Lerngelegenheiten bei Schüler/-innen verschiedener sozialer Herkunft unterschiedliche Effekte haben können. Es erscheint wenig plausibel, dass pädagogische Merkmale in direkter Weise mit der sozialen Herkunft interagieren, da es sich beim sozialen Status um ein Strukturmerkmal handelt, welches selbst nur indirekte Effekte auf den Lernerfolg hat. Überzeugender ist die Annahme, dass pädagogische Merkmale bei der Beeinflussung des Lernerfolgs mit vermittelnden Merkmalen interagieren, welche bei Schüler/-innen verschiedener sozialer Herkunft unterschiedlich ausgeprägt sind und ihren Lernerfolg bedingen. Unter Rückgriff auf die Ausführungen im dritten Kapitel kommen hierfür insbesondere die Lernvoraussetzungen der Schüler/-innen, das Erziehungshandeln der Eltern und die Lernaktivitäten der Schüler/-innen außerhalb der Schule in Frage.

Bezüglich der Lernvoraussetzungen der Schüler/-innen erscheint es plausibel, dass schulische Lernangebote unterschiedlich hohe *Anforderungen an die Lernvoraussetzungen der Schüler/-innen* stellen. Hohe Anforderungen sind gegeben, wenn die Schüler/-innen für erfolgreiches Lernen sehr günstige Lernvoraussetzungen aufweisen müssen (z. B. hohe Intelligenz, gute Selbstregulation), und niedrige Anforderungen, wenn auch Schüler/-innen mit ungünstigen Lernvoraussetzungen erfolgreich lernen können (z. B. niedrige Intelligenz, schlechte Selbstregulation) (Klauer & Leutner, 2012, S. 181). Wenn die Lehrkraft beispielsweise im Unterricht auf eine effiziente Klassenführung achtet (z. B. klare Regeln, angemessene Interventionen bei Störverhalten), können auch Schüler/-innen mit schlechter Selbstregulation einen hohen Lernerfolg erreichen. Wenn die Lehrkraft andererseits die Klassenführung vernachlässigt (z. B. keine Regeln, keine Interventionen bei Störverhalten), stellt der Unterricht hohe Anforderungen, weil erfolgreiches Lernen eine gute Selbstregulation voraussetzt. Da die Selbstregulation mit der sozialen Herkunft zusammenhängt (Sammons et al., 2014, S. 14), kann die Annahme abgeleitet werden, dass Klassenführung bei Schüler/-innen verschiedener sozialer Herkunft differenziell wirksam ist.

In ähnlicher Weise können schulische Lernangebote unterschiedlich hohe *Anforderungen an das Erziehungshandeln der Eltern* stellen. Diese Anforderungen hängen davon ab, ob die Schule die Verantwortung für die Sicherstellung günstiger Lernbedingungen an die Eltern überträgt (hohe Anforderungen) oder diese Verantwortung selbst übernimmt (niedrige Anforderungen). Der zweitgenannte Fall entspricht einer ausgeprägten „Förderverantwortung der

Schule“ (Fend, 2013, S. 129). Als Beispiel können Hausaufgabenbetreuungen in der Schule genannt werden. Wenn die Schule keine Hausaufgabenbetreuung anbietet, werden hohe Anforderungen an das Erziehungs Handeln der Eltern gestellt, weil das erfolgreiche Bearbeiten der Hausaufgaben dann auch davon abhängt, ob die Eltern ihre Kinder bei den Hausaufgaben angemessen begleiten. Wenn die Schule demgegenüber für alle Schüler/-innen eine Hausaufgabenbetreuung anbietet, werden niedrige Anforderungen an die Eltern gestellt, weil dann auch diejenigen Schüler/-innen ihre Hausaufgaben erfolgreich bearbeiten können, bei denen sich die Eltern nur in geringem Maße an den Hausaufgaben beteiligen. Da die elterliche Hausaufgabenbegleitung teilweise mit dem Sozialstatus zusammenhängt (Cooper et al., 2000, S. 478 ff.; Ehmke et al., 2006b, S. 240), liefert dieser Mechanismus ein Argument, warum bei Hausaufgabenbetreuungen differenzielle Effekte nach sozialer Herkunft anzunehmen sind.

Eine weitere Möglichkeit ist darin zu sehen, dass schulische Lernangebote zu einer *Angleichung der Lernaktivitäten von Schüler/-innen* verschiedener sozialer Herkunft führen. Die Grundlage hierfür ist, dass sozial privilegierte Schüler/-innen außerhalb der Schule häufiger anregende Lernaktivitäten ausüben als sozial benachteiligte Schüler/-innen. Insofern ist anzunehmen, dass eine Erhöhung des Zeitumfangs schulbezogener Aktivitäten, welche in der Regel durch einen hohen Anregungsgehalt gekennzeichnet sind, zu einer Verringerung von herkunftsbedingten Unterschieden im Lernerfolg führt (Raudenbush & Eschmann, 2015, S. 454). Differenzielle Effekte von schulischen Lernangeboten sind insbesondere dann zu erwarten, wenn der Einfluss der jeweiligen Lernaktivität auf den Lernerfolg durch einen *Sättigungseffekt* gekennzeichnet ist. Dies soll am Beispiel der Lesehäufigkeit veranschaulicht werden. Angenommen sozial benachteiligte Schüler/-innen lesen drei Stunden pro Woche und sozial privilegierte Schüler/-innen lesen vier Stunden pro Woche. Weiterhin wird ein Sättigungseffekt angenommen, d. h., der lernförderliche Effekt der Lesehäufigkeit nimmt mit zunehmender Lesehäufigkeit ab (Beispiel: der Effekt durch eine Steigerung von drei auf vier Stunden Lesen ist größer als der Effekt durch eine Steigerung von vier auf fünf Stunden Lesen). Unter diesen Bedingungen hätte eine Erhöhung des Umfangs des schulischen Leseunterrichts um eine Stunde pro Woche einen differenziellen Effekt, d. h., der lernförderliche Effekt des zusätzlichen Leseunterrichts wäre bei sozial benachteiligten Schüler/-innen größer als bei sozial privilegierten Schüler/-innen.

Aus den zuvor angestellten Überlegungen kann geschlussfolgert werden, dass geringe Unterschiede im Lernerfolg am ehesten erreicht werden können, wenn schulische Lernangebote eine ausreichende *Passung* zu den inneren und

äußeren Lernbedingungen von sozial benachteiligten Schüler/-innen aufweisen. Insbesondere gilt es, die Anforderung an die Lernvoraussetzungen der Schüler/-innen und an das Erziehungs Handeln der Eltern gering zu halten, da sonst sozial benachteiligte Schüler/-innen einen niedrigeren Lernerfolg erreichen als sozial privilegierte Schüler/-innen. Eine Überforderung von sozial benachteiligten Schüler/-innen sollte also vermieden werden. Eine Unterforderung von sozial benachteiligten Schüler/-innen sollte allerdings ebenfalls vermieden werden, da sonst das Lernpotential der Schüler/-innen nicht ausgeschöpft wird (Kyriakides, 2007, S. 43). Insofern erscheinen schulische Lernangebote günstig, bei denen sozial benachteiligte Schüler/-innen so weit gefordert werden, wie es ihre Lernbedingungen zulassen.

5.2.4 Vertiefte Betrachtung unterschiedlicher Lerngelegenheiten

Auch zum Mechanismus unterschiedlicher Lerngelegenheiten sollen einige tiefere Überlegungen angestellt werden. Es wird erörtert, in welchen Varianten unterschiedliche Lerngelegenheiten auftreten können und worauf sie zurückzuführen sind. Grundlegend ist festzuhalten, dass Unterschiede sowohl bei der *Quantität von Lerngelegenheiten* (d. h., eine Schülergruppe erfährt mehr Lerngelegenheiten als die andere Schülergruppe) als auch bei der *Qualität von Lerngelegenheiten* (d. h., eine Schülergruppe erfährt andere Lerngelegenheiten als die andere Schülergruppe) bestehen können.¹⁰

Weiterhin können Unterschiede auf verschiedenen *Ebenen von Schule* verortet sein (Downey et al., 2004, S. 615; Maaz et al., 2010, S. 28). Dieser Sachverhalt wurde bereits im dritten Kapitel thematisiert (Unterschiede zwischen Schulen vs. Unterschiede innerhalb von Schulen). In diesem Kontext ist hervorzuheben, dass unterschiedliche Lerngelegenheiten auf einer bestimmten Ebene häufig durch ein pädagogisches Merkmal auf der nächsthöheren Ebene bedingt werden. Dies soll durch einige Beispiele veranschaulicht werden. Das Systemmerkmal „gegliedertes Schulsystem“ führt dazu, dass Schüler/-innen verschiedener sozialer Herkunft unterschiedliche Sekundarschulen besuchen. Das Schulmerkmal „Bildung von leistungshomogenen Klassen“ führt dazu,

10 Genau genommen müssen weitere Bedingungen erfüllt sein, damit sich unterschiedliche Lerngelegenheiten auf Unterschiede im Lernerfolg auswirken. *Unterschiede in der Quantität einer Lerngelegenheit* haben nur dann einen Effekt auf Unterschiede im Lernerfolg, wenn die jeweilige Lerngelegenheit auch wirksam ist. *Unterschiede in der Qualität von Lerngelegenheiten* haben nur dann einen Effekt auf Unterschiede im Lernerfolg, wenn sich die verschiedenen Arten von Lerngelegenheiten in ihrer Wirksamkeit unterscheiden.

dass Schüler/-innen verschiedener sozialer Herkunft Klassen mit unterschiedlicher Schülerkomposition besuchen. Das Unterrichtsmerkmal „Lob nach sozialer Bezugsnorm“ führt dazu, dass Schüler/-innen verschiedener sozialer Herkunft unterschiedlich viel Lob von der Lehrkraft erhalten.

Aus der *Angebots-Nutzungs-Struktur* von Schule ergibt sich, dass unterschiedliche Lerngelegenheiten sowohl auf Angebotsunterschiede als auch auf Nutzungsunterschiede zurückzuführen sein können (Maaz et al., 2010, S. 23). Von besonderer Relevanz ist hierbei, dass die Angebots- und Nutzungsunterschiede durch pädagogische Merkmale bedingt sein können. Pädagogische Merkmale können mit Angebotsunterschieden einhergehen, d.h., Lehrkräfte behandeln Schüler/-innen verschiedener sozialer Herkunft systematisch unterschiedlich (z.B. das Unterrichtsmerkmal Individualisierung führt dazu, dass die Lehrkraft den sozial privilegierten Schüler/-innen anspruchsvollere Aufgaben gibt als den sozial benachteiligten Schüler/-innen). Empirische Befunde sprechen dafür, dass Lehrkräfte Schüler/-innen unterschiedlich behandeln in Abhängigkeit davon, welches Leistungsvermögen sie bei den einzelnen Schüler/-innen erwarten (Babad, 1993, S. 350 ff.). Neben Angebotsunterschieden können pädagogische Merkmale auch mit Nutzungsunterschieden einhergehen. Dies ist der Fall, wenn Schüler/-innen verschiedener sozialer Herkunft systematisch unterschiedliche Lernaktivitäten auswählen (z.B. das Unterrichtsmerkmal offener Unterricht führt dazu, dass die sozial privilegierten Schüler/-innen anspruchsvollere Aufgaben auswählen als die sozial benachteiligten Schüler/-innen).

Die verschiedenen Varianten und Ursachen von unterschiedlichen Lerngelegenheiten verdeutlichen, dass sich *mehrere Ansatzpunkte* für pädagogische und bildungspolitische Maßnahmen zur Verringerung von herkunftsbedingten Unterschieden im Lernerfolg bieten. Ein besonderer Fokus ist darauf zu richten, bestehende Unterschiede bei schulischen Lerngelegenheiten zuungunsten von sozial benachteiligten Schüler/-innen abzubauen (d.h., wenn sozial benachteiligte Schüler/-innen weniger oder schlechtere schulische Lerngelegenheiten erfahren als sozial privilegierte Schüler/-innen). In der Literatur findet sich allerdings mehrfach die Annahme, dass auch bei gleichen schulischen Lerngelegenheiten sozial privilegierte Schüler/-innen einen höheren Lernerfolg erreichen als sozial benachteiligte Schüler/-innen (Aktionsrat Bildung, 2007, S. 79 f.; Raudenbush & Eschmann, 2015, S. 453). Für diese Annahme spricht, dass sozial privilegierte Schüler/-innen über günstigere Lernvoraussetzungen verfügen, wodurch sie schulische Lerngelegenheiten besser nutzen können, und dass sie in außerschulischen Lernumwelten mehr Anregung erfahren. Folglich ist anzunehmen, dass eine bedeutsame Angleichung

von Schülerkompetenzen nur erreicht werden kann, wenn sozial benachteiligte Schüler/-innen mehr oder bessere schulische Lerngelegenheiten erhalten als sozial privilegierte Schüler/-innen (Expertenrat Herkunft und Bildungserfolg, 2011, S. 22; Fend, 2013, S. 138 f.).

5.3 Forschungsansätze zu Effekten von pädagogischen Merkmalen

In den vorausgehenden Abschnitten wurde eine theoretische Grundlage erarbeitet, welche es ermöglicht, empirische Zusammenhänge zwischen pädagogischen Merkmalen und herkunftsbedingten Unterschieden im Lernerfolg inhaltlich zu interpretieren. Im Weiteren wird zunächst ein Überblick über den bestehenden Forschungsstand und üblicherweise verwendete Forschungsansätze der schulbezogenen Wirksamkeitsforschung gegeben. Anschließend werden empirische Befunde zu den einzelnen Kategorien pädagogischer Merkmale vorgestellt.

In den vorhandenen Wirksamkeitsstudien werden die verschiedenen *Zielsetzungen von Schule und Unterricht* in unterschiedlichem Maße als Kriterien für Wirksamkeit berücksichtigt. Erkennbar ist eine starke Ausrichtung auf kognitive Ziele, während motivationale, soziale und emotionale Ziele eher eine untergeordnete Rolle spielen (Reynolds et al., 2014, S. 205; Sammons, 1999, S. 79). Weiterhin ist festzustellen, dass in den meisten Studien nur der durchschnittliche Lernerfolg als Zielkriterium berücksichtigt wird. Relativ wenige Studien untersuchen Effekte auf Unterschiede im Lernerfolg (Gustafsson et al., 2018, S. 16; Kyriakides, 2007, S. 41; Sammons et al., 2016, S. 74; Strand, 2010, S. 291). Obwohl die Förderung sozial benachteiligter Schüler/-innen ursprünglich ein zentrales Anliegen der schulbezogenen Wirksamkeitsforschung war, wurde diesem Thema in den nachfolgenden Jahrzehnten relativ wenig Beachtung geschenkt (Kyriakides & Creemers, 2018, S. 1; Sammons, 1999, S. 75). Erst in jüngerer Zeit werden wieder vermehrt Studien zum Zielkriterium geringer Unterschiede im Lernerfolg gefordert (Ditton, 2002, S. 206; Kyriakides & Creemers, 2011, S. 248; Pfeifer, 2011, S. 188; Strand, 2010, S. 310).

Untersuchungen zu den Effekten von pädagogischen Merkmalen auf den *durchschnittlichen Lernerfolg* verwenden typischerweise korrelative Studiendesigns. Die übliche Analysestrategie besteht darin, für jede Klasse/Schule einen adjustierten Kompetenzmittelwert zu berechnen (Mehrwertmodell) und dieses Kriterium mit pädagogischen Merkmalen zu korrelieren. In gegenwärtigen Studien werden hierfür überwiegend Mehrebenenanalysen benutzt (z. B. Kunter & Voss, 2011), während in älteren Studien mehrheitlich Ein-Ebenen-

Analysen auf der Klassenebene verwendet werden (z. B. Helmke & Weinert, 1997b). Teilweise werden anhand der adjustierten Kompetenzmittelwerte auch Extremgruppen gebildet (hoch wirksame Klassen/Schulen vs. niedrig wirksame Klassen/Schulen) und anschließend in Bezug auf pädagogische Merkmale miteinander verglichen (z. B. Stringfield, 1994). Neben korrelativen Studien werden, in deutlich geringerem Umfang, auch experimentelle Studien durchgeführt (z. B. Möller et al., 2002).

Auch bei Untersuchungen zu den Effekten von pädagogischen Merkmalen auf *Unterschiede im Lernerfolg* haben die meisten Studien ein korrelatives Design. Die methodischen Vorgehensweisen sind allerdings durch eine große Heterogenität gekennzeichnet. Eine *erste Analysestrategie* weist Parallelen zu Studien zum durchschnittlichen Lernerfolg auf. Hierbei wird für jede Klasse/Schule ein Indikator für Unterschiede im Lernerfolg berechnet und dieser Indikator wird mit pädagogischen Merkmalen korreliert. Als Indikatoren für Unterschiede im Lernerfolg werden beispielsweise der Zusammenhang zwischen sozialer Herkunft und Kompetenzen (z. B. Schwippert, 2001), der Zusammenhang zwischen Migrationshintergrund und Kompetenzen (z. B. Seiz et al., 2016), der Zusammenhang zwischen Ausgangskompetenzen und Endkompetenzen (z. B. Treiber & Schneider, 1980) oder die Veränderung der Varianz der Kompetenzen (z. B. Baumert et al., 1986) verwendet. Einen Sonderfall dieser Analysestrategie stellen Studien zu Optimalklassen dar, bei denen Klassen untersucht werden, die sowohl einen hohen durchschnittlichen Lernerfolg als auch geringe Unterschiede im Lernerfolg erreichen (z. B. Helmke, 1988). Eine *zweite Analysestrategie* besteht darin, die Stichprobe anhand von Schülermerkmalen aufzuteilen und anschließend die Zusammenhänge zwischen pädagogischen Merkmalen und Lernerfolg für verschiedene Schülergruppen getrennt zu berechnen. Die Gegenüberstellung von Schülergruppen bezieht sich beispielsweise auf Schüler/-innen verschiedener Leistungsstärke (z. B. Treiber & Weinert, 1985) oder Schüler/-innen verschiedener sozialer Herkunft (z. B. Palardy, 2008). Neben korrelativen Studien finden sich auch einige wenige experimentelle Studien zu Effekten auf Unterschiede im Lernerfolg (z. B. Charalambous et al., 2018). Unabhängig von Design und Analysestrategie ist wiederum der Trend zur zunehmenden Verwendung von Mehrebenenanalysen festzustellen (z. B. Charalambous et al., 2018; Palardy, 2008; Schwippert, 2001).

Nachfolgend werden empirische Befunde zu Effekten von pädagogischen Merkmalen dargestellt. Es werden sowohl Befunde zu Effekten auf den durchschnittlichen Lernerfolg als auch zu Effekten auf Unterschiede im Lernerfolg berücksichtigt, allerdings jeweils nur in Bezug auf den Erwerb von Fachkom-

petenzen (d.h., motivationale, soziale und emotionale Lernziele bleiben unberücksichtigt). Da die Studienlage zum durchschnittlichen Lernerfolg sehr umfangreich ist, kann für die Aufbereitung des Forschungsstandes zu einem großen Teil auf Meta-Analysen (z. B. Kyriakides et al., 2010; Kyriakides et al., 2013; Marzano et al., 2000; Scheerens et al., 2007) und Reviews (z. B. Brophy & Good, 1986; Lipowsky, 2006; Muijs & Reynolds, 2001) zurückgegriffen werden. Demgegenüber ist die Studienlage zu Unterschieden im Lernerfolg deutlich weniger umfangreich, weshalb nur Einzelstudien berichtet werden können. Das primäre Interesse des Forschungsüberblicks liegt bei Studien zu Kompetenzunterschieden nach sozialer Herkunft. Da die Anzahl der verfügbaren Studien hierzu relativ gering ausfällt, werden teilweise auch Befunde zu anderen Schülergruppen berichtet (z. B. Schüler/-innen verschiedener Leistungsstärke, Schüler/-innen mit und ohne Migrationshintergrund).

Die inhaltliche Gliederung basiert auf dem theoretischen Modell der Wirksamkeit von Schule. Der Forschungsüberblick begrenzt sich auf die Kategorien des theoretischen Modells, die im Rahmen der empirischen Analysen untersucht werden.¹¹ Dabei handelt es sich um Unterrichtsmerkmale, Ganztagsangebote und Merkmale der Lerngruppe. Bei jedem Merkmal erfolgen einleitend theoretische Überlegungen zu Konzeptualisierung und Wirkmechanismen. Bei Effekten auf den durchschnittlichen Lernerfolg wird auf die Wirkmechanismen Förderung von fachlichen Lernaktivitäten und Förderung von Lernvoraussetzungen Bezug genommen (siehe Kapitel 2.2.4), bei Effekten auf herkunftsbedingte Unterschiede im Lernerfolg auf die Wirkmechanismen differenziell wirksame Lerngelegenheiten und unterschiedliche Lerngelegenheiten (siehe Kapitel 5.2.2). Auf die Effekte von Hintergrundprozessen und Merkmalen von pädagogischen Fachkräften wird aus Platzgründen nicht weiter eingegangen. Der interessierte Leser/die interessierte Leserin sei an dieser Stelle auf die einschlägige Literatur verwiesen (Baumert & Kunter, 2006; Kellermann et al., 2022; Kraft et al., 2018; Lipowsky, 2006; Robinson et al., 2008).

11 Neben den nachfolgend behandelten pädagogischen Merkmalen können anhand der vorgestellten Mechanismen weitere pädagogische Merkmale ermittelt werden, bei denen ein Effekt auf herkunftsbedingte Unterschiede im Lernerfolg plausibel erscheint. Hierzu zählen beispielsweise die Verteilung von Interaktionen im Unterricht (z. B. Aufrufverhalten der Lehrkraft), Hausaufgaben, spezifische Fördermaßnahmen (z. B. Motivationsförderung, Sprachförderung), Unterrichts- und Schulklima, Schüler-Lehrkraft-Beziehungen, Schulleitungshandeln und Elternarbeit. Auf diese Merkmale kann im Rahmen der vorliegenden Arbeit aus Platzgründen nicht eingegangen werden.

5.4 Effekte von Unterrichtsmerkmalen

Nachfolgend wird ein Überblick über den empirischen Forschungsstand zu den Effekten von Unterrichtsmerkmalen gegeben. Unterrichtsmerkmale können in drei Bereiche unterteilt werden: der Unterrichtsumfang (d.h. das zeitliche Ausmaß des Unterrichts), die Unterrichtsgestaltung (d.h. die Unterrichtsmethoden und Unterrichtsprinzipien, die verwendet werden) und die Unterrichtsinhalte (d.h. die Themen, die besprochen werden). Die vorliegende Arbeit fokussiert auf die beiden erstgenannten Bereiche, die im Weiteren genauer betrachtet werden.

5.4.1 Effekte des Unterrichtsumfangs

Nachfolgend werden Effekte des Unterrichtsumfangs behandelt. Konzeptuell kann der zeitliche Umfang von unterrichtlichen Lehr-Lern-Prozessen in mehrere Komponenten zerlegt werden (Helmke, 2014, S. 78; Scheerens et al., 2013, S. 5). Die *vorgesehene Unterrichtszeit* in einem Fach ergibt sich aus der Anzahl der Schultage pro Jahr, der Anzahl der Unterrichtsstunden pro Woche und der Dauer der Unterrichtsstunden. Die *tatsächliche Unterrichtszeit* bezieht sich auf die Zeit, in der tatsächlich Unterricht stattfindet, also vorgesehene Unterrichtszeit abzüglich Stundenausfällen, Zeitaufwand für Organisatorisches und Zeitverlust durch Unterrichtsstörungen. Neben diesen beiden Zeitkomponenten auf der Klassenebene kann außerdem auf der Schülerebene die *aktive Lernzeit* (Englisch: time on task) betrachtet werden. Dabei handelt es sich um den Zeitanteil, in dem ein bestimmter Schüler/eine bestimmte Schülerin sich tatsächlich mit den Lerninhalten beschäftigt (und nicht mit anderen Dingen, z. B. gedankliches Abschweifen, Schwätzen mit Mitschüler/-innen).

Die Bedeutung des zeitlichen Umfangs von Lehr-Lern-Prozessen für den *Lernerfolg* wurde bereits im Modell von Carroll (1963) betont. Die Annahme, dass ein höherer Unterrichtsumfang zu einem höheren Kompetenzmittelwert führt, kann naheliegenderweise durch den Mechanismus der Erhöhung der Quantität fachlicher Lernaktivitäten begründet werden. Gleichwohl ist anzunehmen, dass zusätzliche Unterrichtszeit nicht automatisch zu besseren Lernergebnissen führt, sondern dass die Zeit auch in der Form von anregenden Lernaktivitäten genutzt werden muss (Scheerens et al., 2013, S. 1). Der empirische Forschungsstand zum Zusammenhang zwischen Unterrichts- bzw. Lernzeit und Lernerfolg ist so umfangreich, dass inzwischen mehrere Meta-Analysen vorliegen (Scheerens et al., 2013, S. 6 ff.). Meta-analytisch belegt werden können lernförderliche Effekte der tatsächlichen Unterrichtszeit (Kyriakides et al., 2010, S. 817; Scheerens et al., 2013, S. 58) und der aktiven Lernzeit (Fraser et

al., 1987, S. 195). Es liegen außerdem Hinweise darauf vor, dass der Zusammenhang zwischen Unterrichtszeit und Lernerfolg durch einen Sättigungseffekt gekennzeichnet ist, d. h., ab einem gewissen Unterrichtsumfang führt zusätzliche Unterrichtszeit nur noch zu minimalen Verbesserungen des Lernerfolgs (Helmke, 2007a, S. 105).

Eine Erhöhung des Unterrichtsumfangs ist häufig mit der Erwartung verbunden, dass *Unterschiede im Lernerfolg* verringert werden, weil sozial benachteiligte Schüler/-innen im besonderen Maße von der zusätzlichen Unterrichtszeit profitieren sollen (Scheerens et al., 2013, S. 18). Für diese globale Annahme spricht, dass ein höherer Unterrichtsumfang zu einer Angleichung der Lernaktivitäten von Schüler/-innen verschiedener sozialer Herkunft führt. Es finden sich mehrfach empirische Hinweise darauf, dass der Unterrichtsumfang einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg hat. Pfeifer (2011) berichtet, dass eine höhere Anzahl an Unterrichtsstunden mit geringeren herkunftsbedingten Kompetenzunterschieden innerhalb von Klassen einhergeht (S. 127). In ihrer ländervergleichenden Analyse zeigen Gustafsson und Kolleg/-innen (2018), dass eine höhere Unterrichtszeit auf der Schulebene in vier Ländern herkunftsbedingte Kompetenzunterschiede innerhalb von Schulen verringert (S. 28 f.). In der großen Mehrheit der Länder besteht allerdings kein systematischer Zusammenhang. Zwei Studien belegen außerdem Effekte auf der Systemebene, d. h., in Ländern mit längerer Unterrichtszeit fallen Kompetenzunterschiede nach sozialer Herkunft geringer aus als in Ländern mit kürzerer Unterrichtszeit (Ammermüller, 2005, S. 22 f.; Burger, 2016, S. 62). Im Einklang damit berichtet Kawaguchi (2013), dass nach der Verkürzung der Schulzeit in Japan (Einführung von schulfreien Samstagen) Kompetenzunterschiede nach sozialer Herkunft zugenommen haben (S. 23 ff.).

Insgesamt betrachtet sprechen die Befunde dafür, dass die Unterrichtszeit sowohl zum durchschnittlichen Lernerfolg als auch zu herkunftsbedingten Unterschieden im Lernerfolg Zusammenhänge aufweist.

5.4.2 Effekte von Sozialformen

Für den Lernerfolg der Schüler/-innen spielt es nicht nur eine Rolle, wie viel Unterricht stattfindet, sondern auch, wie der Unterricht gestaltet wird. In den weiteren Ausführungen werden Effekte der *Unterrichtsgestaltung* betrachtet. Einen wichtigen Aspekt der Unterrichtsgestaltung stellen die verwendeten Unterrichtsmethoden dar. Unter *Unterrichtsmethode* wird im Weiteren eine Lehr-Lern-Interaktion verstanden, die durch spezifische Aktivitäten der Lehrkraft und der Schüler/-innen gekennzeichnet ist. Unterrichtsmethoden beschreiben

somit die Form, in der sich Lehrkräfte und Schüler/-innen mit den Lerninhalten beschäftigen. Nachfolgend werden Unterrichtsmethoden zunächst anhand von *Sozialformen* und anschließend anhand der Gegenüberstellung von *instruktionalen* und *konstruktivistischen Ansätzen* gegliedert (eine ähnliche Klassifikation findet sich bei Klauer & Leutner, 2012).

Sozialformen beziehen sich auf verschiedene Muster der sozialen Interaktion zwischen Lehrkraft und Schüler/-innen. Beim *Klassenunterricht* (Synonym: Frontalunterricht; Englisch: whole-class teaching) interagiert die Lehrkraft mit der gesamten Klasse. Bei *Einzelarbeit* gehen Schüler/-innen ihren Lernaktivitäten weitestgehend ohne soziale Interaktion nach, bei *Partner- und Gruppenarbeit* interagiert jeweils ein Paar bzw. eine Gruppe von Schüler/-innen miteinander.

Theoretische Überlegungen zu Effekten von Sozialformen auf den *Lernerfolg* werden dadurch erschwert, dass es sich bei Sozialformen um inhaltlich weit gefasste Kategorien handelt (beispielsweise umfasst Klassenunterricht u. a. Lehrervortrag, Fragen, Unterrichtsgespräch und kognitives Modellieren). Das heißt, auch bei gleicher Sozialform kann der Unterricht sehr unterschiedlich ausgestaltet werden (Kunter & Voss, 2011, S. 87). Dessen ungeachtet können theoretische Argumente abgewogen werden, die für oder gegen bestimmte Sozialformen sprechen. Bei der Gegenüberstellung von Klassenunterricht und Einzel-/Partner-/Gruppenarbeiten spricht für den Klassenunterricht vor allem, dass die Lehrkraft bei dieser Sozialform die meisten Einflussmöglichkeiten hat und somit ihre pädagogische und fachliche Expertise am direktesten in den Lernprozess einbringen kann, was zu einer hohen Quantität und Qualität fachlicher Lernaktivitäten führen sollte (Muijs & Reynolds, 2001, S. 5). Demgegenüber kann eingewendet werden, dass Schüler/-innen im Klassenunterricht zu einem größeren Zeitanteil passiv sind, was sowohl die Qualität der fachlichen Lernaktivitäten (z. B. Verarbeitungstiefe) als auch die Lernmotivation der Schüler/-innen (z. B. Lernfreude) einschränken kann. Empirische Wirksamkeitsstudien liegen insbesondere für den Vergleich von *Klassenunterricht* und *Einzelarbeiten* vor. Mehrfach zeigt sich, dass in Klassen mit hohem durchschnittlichen Lernerfolg ein höherer Zeitanteil für Klassenunterricht und ein niedrigerer Zeitanteil für Einzelarbeit verwendet wird (Brophy & Good, 1986, S. 361; Driessen & Sleegers, 2000, S. 74 f.; Muijs & Reynolds, 2001, S. 3 ff.; Schwerdt & Wuppermann, 2011, S. 372; Teodorovic, 2011, S. 230). Hieraus kann geschlussfolgert werden, dass direkte Interaktionen zwischen Lehrkraft und Schüler/-innen ein hohes Anregungspotential für das fachliche Lernen haben (Muijs & Reynolds, 2001, S. 5).

Eine weitere relevante Gegenüberstellung bezieht sich auf *Einzelarbeit* versus *Partner-/Gruppenarbeit*. Theoretisch spricht für Partner-/Gruppenarbeit, dass sich die Schüler/-innen gegenseitig unterstützen können, was die Qualität des fachlichen Lernens erhöhen sollte. Außerdem ist anzunehmen, dass soziale Interaktion einen motivationsförderlichen Effekt hat. Demgegenüber besteht bei Partner-/Gruppenarbeit die Gefahr, dass sich Schüler/-innen gegenseitig ablenken und dass sich einzelne Schüler/-innen nicht aktiv beteiligen (Trittbrettfahrer-Effekt), wodurch die Quantität fachlicher Lernaktivitäten verringert werden kann. Johnson und Kolleg/-innen (2000) vergleichen in einer Meta-Analyse die Effekte von kooperativem Lernen und individuellem Lernen. Unter kooperativem Lernen versteht man strukturierte Partner-/Gruppenarbeiten, bei der Schüler/-innen zusammenarbeiten müssen, um gemeinsame Ziele zu erreichen (z. B. Gruppenpuzzle, Gruppenrallye). Es zeigt sich, dass Schüler/-innen bei kooperativem Lernen einen höheren durchschnittlichen Lernerfolg aufweisen als bei individuellem Lernen (S. 11). Somit überwiegen, bei angemessener Durchführung, die Vorteile von Partner-/Gruppenarbeiten.

Hinsichtlich Effekten von Sozialformen auf *herkunftsbedingte Unterschiede im Lernerfolg* erscheint es denkbar, dass verschiedene Sozialformen unterschiedlich hohe Anforderungen an die Lernvoraussetzungen der Schüler/-innen stellen. Bei *Einzelarbeit* ist anzunehmen, dass durch das selbstständige Lernen tendenziell eher hohe Anforderungen an die Lernvoraussetzungen der Schüler/-innen gestellt werden, insbesondere an metakognitive Kompetenzen, Lernmotivation und Selbstregulation. Im Rahmen der Literaturrecherche wurden allerdings keine Studie gefunden, die diese Annahme empirisch überprüft. Beim *Klassenunterricht* gilt, dass es wesentlich von der konkreten Ausgestaltung abhängt, ob hohe oder niedrige Anforderungen an die Lernvoraussetzungen der Schüler/-innen gestellt werden (z. B. Klarheit und Strukturiertheit). Klassenunterricht an sich scheint für sozial privilegierte und sozial benachteiligte Schüler/-innen ähnlich wirksam zu sein (Brophy & Good, 1986, S. 341).

Bei *Partner-/Gruppenarbeiten* kann argumentiert werden, dass die Interaktionen mit den Mitschüler/-innen bis zu einem gewissen Grad defizitäre Lernvoraussetzungen kompensieren können (z. B. niedrige Ausgangskompetenz, niedrige Lernmotivation). Allerdings hängt das kompensatorische Potential wesentlich von der Art der Gruppenbildung ab. Aussichtsreich erscheinen insbesondere leistungsheterogene Gruppen, da Schüler/-innen mit ungünstigen Lernvoraussetzungen dann mit Schüler/-innen mit günstigen Lernvoraussetzungen interagieren (Hattie, 2002, S. 451). Bei leistungshomogenen Gruppen ist dies nicht der Fall, sondern es ergeben sich im Gegenteil sogar unterschiedliche Lerngelegenheiten zugunsten von leistungsstarken Schüler/-innen, da

diese mit anregenderen Interaktionspartner/-innen zusammenarbeiten. Die theoretischen Annahmen zu den Effekten von Partner-/Gruppenarbeiten auf Kompetenzunterschiede werden durch eine Reihe von Befunden bestätigt. In einer Meta-Analyse werden die Effekte von „Peer-Assisted Learning“ untersucht, einer Oberkategorie für verschiedene Arten von Partner- und Gruppenarbeiten (Rohrbeck et al., 2003). In der großen Mehrheit der Interventionen werden leistungsheterogene Paare bzw. Gruppen gebildet (S. 245). Es zeigt sich, dass die lernförderlichen Effekte von Partner-/Gruppenarbeiten bei Schüler/-innen mit Migrationshintergrund stärker ausfallen als bei Schüler/-innen ohne Migrationshintergrund ($p < 0.05$) und bei sozial benachteiligten Schüler/-innen stärker als bei sozial privilegierten Schüler/-innen ($p < 0.10$) (S. 246 f.). In einer anderen Meta-Analyse werden die Effekte von leistungshomogener und leistungsheterogener Gruppenbildung direkt miteinander verglichen (Lou et al., 1996, S. 445 f.). Insgesamt fallen die lernförderlichen Effekte bei leistungshomogener Gruppenbildung geringfügig stärker aus als bei leistungsheterogener Gruppenbildung, allerdings nur beim Lesen. Weiterhin zeigen sich Unterschiede zwischen Schülergruppen: Leistungsschwache Schüler/-innen profitieren besonders von leistungsheterogenen Gruppen, leistungsmoderate Schüler/-innen profitieren besonders von leistungshomogenen Gruppen und bei leistungsstarken Schüler/-innen besteht kein Unterschied.

Zusammenfassend sprechen die Befunde dafür, dass ein hoher durchschnittlicher Lernerfolg durch Klassenunterricht und kooperatives Lernen gefördert wird, während geringe Unterschiede im Lernerfolg durch kooperatives Lernen mit heterogenen Paaren/Gruppen begünstigt werden. Hieraus kann allerdings nicht geschlossen werden, dass Lehrkräfte überhaupt keine Einzelarbeit einsetzen sollten. Muijs und Reynolds (2001) argumentieren, dass Einzelarbeit zwar für die Einführung neuer Inhalte weniger geeignet ist, aber für Übungsphasen durchaus sinnvoll (S. 27).

5.4.3 Effekte von instruktionalen und konstruktivistischen Unterrichtsmethoden

Eine andere Möglichkeit zur Klassifikation von Unterrichtsmethoden bezieht sich auf die Gegenüberstellung von *instruktionalen* und *konstruktivistischen Ansätzen*. In der Literatur finden sich zahlreiche weitere, inhaltlich ähnliche Begriffspaare wie lehrkraftzentriert vs. schülerzentriert, darbietend vs. entdecken-lassend oder direktiv vs. nicht direktiv (Caro et al., 2016, S. 31; Klauer & Leutner, 2012, S. 96). Der zentrale Unterschied zwischen instruktionalem und konstruktivistischem Unterricht liegt darin, ob die *Steuerung des Lernpro-*

zesses bei der Lehrkraft oder bei den Schüler/-innen liegt. Die Lernsteuerung betrifft den *Vermittlungsstil* (Lehrkraft erklärt neue Inhalte vs. Schüler/-innen erarbeiten sich neue Inhalte) und die *Entscheidungsstrukturen* (lehrergelenkter Unterricht, d. h., die Lehrkraft entscheidet über Inhalte/Methoden/Lernzeit vs. offener Unterricht, d. h., die Schüler/-innen entscheiden über Inhalte/Methoden/Lernzeit) (angelehnt an Wiechmann & Wildhirt, 2015, S. 18 f.). Die zahlreichen Varianten bei der Ausgestaltung der Lernsteuerung bilden ein Spektrum mit fließenden Übergängen. Bei einer vereinfachten Kategorisierung kann zwischen *rein instruktionalen Unterrichtsmethoden* (z. B. direkte Instruktion, zielerreichendes Lernen), *ausgewogenen Unterrichtsmethoden* (z. B. angeleitetes entdeckendes Lernen, Lernen mit Lösungsbeispielen, Lehrlingslernen) und *rein konstruktivistischen Unterrichtsmethoden* (z. B. reines entdeckendes Lernen, problembasiertes Lernen, Wochenplanunterricht) unterschieden werden (angelehnt an Klauer & Leutner, 2012, S. 96 ff.).¹²

Für Effekte der verschiedenen Unterrichtsansätze auf den *durchschnittlichen Lernerfolg* können unterschiedliche Argumente angebracht werden. Für instruktionale Unterrichtsmethoden spricht, dass die Lehrkraft die Lerninhalte vorstrukturieren kann, wodurch die kognitiven Kapazitäten der Schüler/-innen entlastet werden (Sweller, 1988). Außerdem kann die Lehrkraft aufgrund ihrer pädagogischen Kompetenzen fundierter einschätzen, welche Vorgehensweisen beim Lernen am sinnvollsten sind. Beide Aspekte sollten der Qualität der fachlichen Lernaktivitäten zugutekommen. Bezüglich dieses Mechanismus kann allerdings auch argumentiert werden, dass es bei konstruktivistischen Unterrichtsmethoden zu einer besonders tiefen Verarbeitung der Lerninhalte kommt (Craik & Lokhart, 1972). Außerdem sind günstige Effekte von konstruktivistischen Methoden auf Metakognition, Lernmotivation und Selbstregulation denkbar. Da beide Ansätze Vor- und Nachteile aufweisen, wird mehrfach argumentiert, dass eine angemessene Kombination von instruktionalen und konstruktivistischen Unterrichtsmethoden besonders lernförderlich ist (Helmke, 2014, S. 68 f.; Lipowsky, 2007, S. 28 f.).

Empirische Belege für lernförderliche Effekte von *instruktionalem Unterricht* liefern die zahlreichen Untersuchungen zur direkten Instruktion.

12 Instruktionaler und konstruktivistischer Unterricht weisen gewisse Bezüge zu den zuvor behandelten Sozialformen auf. Instruktionaler Unterricht ist typischerweise durch einen höheren Anteil an Klassenunterricht gekennzeichnet, konstruktivistischer Unterricht durch einen höheren Anteil an Einzel-, Partner- und Gruppenarbeit. Ungeachtet dieser Tendenzen können Einzel-, Partner- und Gruppenarbeit aber jeweils sowohl instruktional (d. h. Steuerung durch die Lehrkraft) als auch konstruktivistisch (d. h. Steuerung durch die Schüler/-innen) ausgestaltet werden.

Unterricht nach dieser Methode hat folgenden Ablauf: Wiederholung der letzten Stunde, Mitteilung der Lernziele der aktuellen Stunde, kleinschrittige Einführung neuer Inhalte mit Verständnisfragen sowie angeleitete Übung mit Feedback (Helmke, 2014, S. 264). Die höhere Wirksamkeit von direkter Instruktion im Vergleich zu regulärem Unterricht (d. h. Unterricht ohne die genannten Charakteristika) wird durch mehrere Meta-Analysen bestätigt (zuletzt Stockard et al., 2018).

Für den *direkten Vergleich* von instruktionalen, konstruktivistischen und ausgewogenen Unterrichtsmethoden ist die Meta-Analyse von Alfieri und Kolleg/-innen (2011) zum entdeckenden Lernen aufschlussreich. *Entdeckendes Lernen* ist dadurch gekennzeichnet, dass die Schüler/-innen eigenständig zu Erkenntnissen kommen, indem sie offene Aufgaben bearbeiten. Beim *reinen entdeckenden Lernen* hält sich die Lehrkraft weitestgehend aus dem Lernprozess heraus, während sie die Schüler/-innen beim *angeleiteten entdeckenden Lernen* aktiv begleitet (z. B. Vorgaben zum Ablauf, unterstützende Hinweise, Feedback). In der Meta-Analyse zeigt sich, dass die Wirksamkeit des reinen entdeckenden Lernens niedriger ausfällt als die Wirksamkeit von instruktionalem Unterricht, während die Wirksamkeit von angeleitetem entdeckendem Lernen höher ausfällt (S. 7). Dieser Befund spricht für die Wirksamkeit eines ausgewogenen Unterrichts. Zahlreiche weitere Befunde sprechen für die Wirksamkeit eines ausgewogenen Unterrichts, der instruktionale und konstruktivistische Elemente verbindet (Kirschner et al., 2006; Lipowsky, 2006; Möller et al., 2002; Mayer, 2004).

Hinsichtlich der Effekte verschiedener Unterrichtsansätze auf *herkunftsbedingte Unterschiede im Lernerfolg* sind insbesondere die Anforderungen an die Lernvoraussetzungen der Schüler/-innen von Bedeutung. Mehrfach findet sich in der Literatur die Einschätzung, dass konstruktivistische Unterrichtsmethoden hohe Anforderungen an Ausgangskompetenz, Intelligenz, metakognitive Kompetenzen und Lernmotivation der Schüler/-innen stellen (Caro et al., 2016, S. 32; Lipowsky & Lotz, 2015, S. 168; Pfeifer, 2011, S. 44). Insofern wird angenommen, dass konstruktivistische Methoden eher zur Vergrößerung von Kompetenzunterschieden führen (Helmke, 2014, S. 68; Sammons & Bakkum, 2011, S. 16).

Empirische Befunde stützen diese Annahme (Kyriakides, 2007, S. 42f.). Schon im Rahmen der frühen Forschung zu Aptitude-Treatment-Interaktionen hat sich gezeigt, dass Schüler/-innen mit ungünstigen kognitiven Lernvoraussetzungen bei instruktionalen Methoden bessere Lernergebnisse erzielen als bei konstruktivistischen Methoden, während es bei Schüler/-innen mit günstigen kognitiven Lernvoraussetzungen umgekehrt ist (Snow, 1989, S. 22).

In einer jüngeren Studie werden vergleichbare Ergebnisse berichtet (Connor et al., 2004). Interessant ist in diesem Kontext die Studie von Möller und Kolleg/-innen (2002), in der ausgewogener Unterricht und konstruktivistischer Unterricht miteinander verglichen werden. Es wird ermittelt, dass Schüler/-innen mit ungünstigen Lernvoraussetzungen bei ausgewogenem Unterricht einen größeren Lernerfolg haben als bei konstruktivistischem Unterricht, während es für Schüler/-innen mit günstigen Lernvoraussetzungen keinen Unterschied macht (S. 183). Das gleiche Ergebnismuster berichten Tuovinen und Sweller (1999). Es wurden weiterhin zwei Studien ermittelt, die die Effekte von konstruktivistischem Unterricht bei Schüler/-innen verschiedener sozialer Herkunft untersuchen. In einer dänischen Studie zeigt sich, dass konstruktivistischer Unterricht bei sozial benachteiligten Schüler/-innen negative Effekte auf den Lernerfolg hat, während er bei sozial privilegierten Schüler/-innen keinen Effekt auf den Lernerfolg hat (d. h., weder lernförderlich noch lernbeeinträchtigend ist) (Andersen & Andersen, 2017, S. 543 ff.). In einer belgischen Studie hat konstruktivistischer Unterricht bei Hoch-Risiko-Schüler/-innen (u. a. sozial benachteiligt, Migrationshintergrund) schwach negative Effekt auf den Lernerfolg, während bei Niedrig-Risiko-Schüler/-innen ein positiver Effekt auf den Lernerfolg festzustellen ist (Vanlaar et al., 2014, S. 417 f.).

Insgesamt betrachtet legen die Befunde nahe, dass der durchschnittliche Lernerfolg am besten durch ausgewogenen Unterricht gefördert werden kann. Unterschiede im Lernerfolg scheinen bei instruktionalem und ausgewogenem Unterricht kleiner auszufallen als bei konstruktivistischem Unterricht. Für den direkten Vergleich von instruktionalem und ausgewogenem Unterricht bezüglich Unterschieden im Lernerfolg liegen bisher keine Befunde vor.

Abschließend ist anzumerken, dass aus den berichteten Befunden zu Sozialformen und instruktionalen vs. konstruktivistischen Unterrichtsmethoden nicht geschlussfolgert werden sollte, dass Lehrkräfte primär mit einer bestimmten (wirksamen) Methode unterrichten sollten oder dass bestimmte Methoden überhaupt nicht eingesetzt werden sollten. In der Literatur herrscht Einigkeit darüber, dass guter Unterricht durch die didaktisch begründete Kombination verschiedener Methoden gekennzeichnet ist (Ditton, 2000, S. 83; Helmke, 2014, S. 263 ff.; Kunter & Trautwein, 2013, S. 138). Lernförderliche Effekte von Methodenvielfalt werden auch durch empirische Studien bestätigt (Brophy & Good, 1986, S. 330; Helmke, 2014, S. 270). Die berichteten Befunde zu den Effekten verschiedener Unterrichtsmethoden liefern Anhaltspunkte dafür, welche Unterrichtsmethoden im Rahmen eines methodenvariierenden Unterrichts priorisiert werden sollten. Zu berücksichtigen ist dabei auch die Passung zur unterrichteten Schülergruppe.

5.4.4 Effekte von Klassenführung

Neben den verwendeten Unterrichtsmethoden ist ein weiterer wichtiger Aspekt der Unterrichtsgestaltung die Umsetzung allgemeiner Unterrichtsprinzipien. *Unterrichtsprinzipien* sind situationsübergreifende Grundsätze einer lernförderlichen Unterrichtsgestaltung (Helmke, 2014, S. 168). Situationsübergreifend meint hierbei, dass Unterrichtsprinzipien jeweils Geltung für mehrere (wenn auch nicht unbedingt alle) Unterrichtssituationen und -methoden haben. Beispielsweise kann sich das Unterrichtsprinzip Klarheit auf Erklärungen, Fragen, Aufgabenstellungen, Rückmeldungen oder Verhaltensanweisungen beziehen. Inhaltlich verwandt ist das Konzept der Tiefenstrukturen von Unterricht (Kunter & Voss, 2011, S. 87). Im Weiteren werden die Unterrichtsprinzipien Klassenführung, Klarheit und Strukturiertheit, kognitive Aktivierung, Übung und Wiederholung, Unterstützung und Feedback sowie Individualisierung behandelt.

Unter *Klassenführung* (Synonym: Klassenmanagement) werden Maßnahmen verstanden, die auf eine optimale Nutzung der zur Verfügung stehenden Unterrichtszeit für fachliche Lernaktivitäten abzielen (Helmke & Helmke, 2014, S. 9). Für die Optimierung der Lernzeit müssen sowohl Zeitverluste durch die Lehrkraft als auch Zeitverluste durch die Schüler/-innen verhindert werden. Auf Seiten der Lehrkraft erfordert dies ein gutes Zeitmanagement (z. B. pünktlicher Unterrichtsbeginn), das Kurzhalten organisatorischer Tätigkeiten (z. B. Geld einsammeln) und die Vermeidung von externen Unterbrechungen (z. B. nicht funktionierende Technik). Auf Seiten der Schüler/-innen gilt es, problematische Verhaltensweisen geringzuhalten (z. B. Zwischenrufe, Schwätzen, gedankliches Abschweifen). Dies setzt angemessene Prävention und Intervention durch die Lehrkraft voraus (Helmke, 2014, S. 173 ff.). Präventive Strategien umfassen die Etablierung von Regeln und Routinen, nahtlose Übergänge zwischen verschiedenen Unterrichtsphasen und das kontinuierliche Monitoring des Schülerverhaltens. Angemessene Interventionen hängen von der Schwere des Problemverhaltens ab: kleine Störungen können beiläufig behoben werden, während schwerere Störungen verbale Ermahnung und gegebenenfalls auch Sanktionen erfordern (Gärtner et al., 2021, S. 536).

Der primäre Wirkmechanismus von Klassenführung bei der Förderung des *Lernerfolgs* liegt naheliegenderweise darin, dass die Quantität fachlicher Lernaktivitäten erhöht wird (Helmke & Helmke, 2014, S. 10). Daneben ist aber auch denkbar, dass Klassenführung einen förderlichen Effekt auf die Lernmotivation der Schüler/-innen hat. Bei einer effizienten Klassenführung bringt die Lehrkraft durch ihr Verhalten implizit zum Ausdruck, dass die Lerninhalte relevant sind, was wiederum das Interesse der Schüler/-innen fördern sollte

(Helmke & Helmke, 2014, S. 10; Klieme et al., 2009, S. 151). Lernförderliche Effekte von Klassenführung wurden in zahlreichen Studien untersucht und werden durch zwei Meta-Analysen bestätigt (Kyriakides et al., 2013, S. 147 f.; Marzano et al., 2003, S. 10).

Für die Annahme, dass Klassenführung einen Effekt auf *herkunftsbedingte Unterschiede im Lernerfolg* hat, spricht zunächst das bereits beim Unterrichtsumfang angebrachte Argument, dass eine Erhöhung der unterrichtlichen Lernzeit zu einer Angleichung der Lernaktivitäten von Schüler/-innen verschiedener sozialer Herkunft führt. Bei spezifischen Aspekten von Klassenführung erscheint es außerdem plausibel, dass sie die Anforderungen an die Lernvoraussetzungen der Schüler/-innen gering halten. Die klare Vorgabe von Regeln, Monitoring und angemessene Interventionen sollten bis zu einem gewissen Grad niedrige Lernmotivation und niedrige Selbstregulation auf Seiten der Schüler/-innen kompensieren können (Atlay et al., 2019, S. 3). Empirische Befunde stützen diese Überlegungen. Kontinuierliches Monitoring durch die Lehrkraft hat in Klassen mit vielen sozial benachteiligten Schüler/-innen einen größeren Effekt auf den Lernerfolg als in Klassen mit vielen sozial privilegierten Schüler/-innen (Brophy & Good, 1986, S. 349). Indirekte Evidenzen liefern Studien, die Effekte eines disziplinerorientierten Schulklimas untersuchen. Zwei Studien berichten, dass angemessenen und faire Disziplinierungsmaßnahmen bei sozial benachteiligten Schüler/-innen stärkere Effekte auf den Lernerfolg haben als bei sozial privilegierten Schüler/-innen (Lee & Bryk, 1989, S. 185; Palardy, 2008, S. 35). In einer jüngeren Studie zeigt sich außerdem, dass die Einführung eines Handyverbots in der Schule einen lernförderlichen Effekt hat und dass dieser Effekt bei sozial benachteiligten Schüler/-innen besonders stark ausgeprägt ist (Beland & Murphy, 2015, S. 12 f.). Werden globale Kategorien von Klassenführung untersucht, werden demgegenüber keine signifikanten Effekte auf Kompetenzunterschiede nach sozialer Herkunft festgestellt (Atlay et al., 2019, S. 7; Vanlaar et al., 2014, S. 420 f.).

Zusammenfassend sprechen die Befunde dafür, dass Klassenführung einen förderlichen Effekt auf den durchschnittlichen Lernerfolg hat. Bei bestimmten Aspekten von Klassenführung, welche einen Bezug zur Verhaltensregulation aufweisen, liegen außerdem Hinweise auf Effekte auf herkunftsbedingte Unterschiede im Lernerfolg vor.

5.4.5 Effekte von Klarheit und Strukturiertheit

Im Weiteren wird das Unterrichtsprinzip Klarheit und Strukturiertheit betrachtet. *Klarheit* bezieht sich darauf, inwieweit die Lehrkraft die Kommunikation

im Unterricht so gestaltet, dass die Schüler/-innen verstehen, was die Lehrkraft ihnen mitteilen möchte. Klarheit begünstigt somit erfolgreiche Kommunikation in dem Sinne, dass die empfangene Nachricht der gesendeten Nachricht entspricht. Teilaspekte von Klarheit sind eine angemessene Artikulation (z. B. Sprechlautstärke, Sprechgeschwindigkeit), Einfachheit (z. B. bekannte Begriffe, kurze Sätze), Anschaulichkeit (z. B. Beispiele, Analogien), Fokussierung (z. B. keine Füllwörter, keine Abschweifungen) und sprachliche Korrektheit (z. B. vollständige Sätze, korrekte Grammatik) (angelehnt an Helmke, 2014, S. 193 ff.). *Strukturiertheit* ist gegeben, wenn der Unterricht einen sinnvollen Aufbau aufweist. Dies umfasst die Sequenzierung von Unterrichtsstunden (z. B. Unterteilung in Einführung, Hauptteil und Abschluss), das Bieten von Orientierungshilfen für die Schüler/-innen (z. B. Lernziele explizit mitteilen, Ausblick über die Stunde geben, Übergänge anzeigen) und die Ordnung der Lerninhalte (z. B. Gliederung nach Themenbereichen, Hervorhebung von wichtigen Aspekten, Zusammenhänge zwischen Teilaspekten erklären) (angelehnt an Helmke, 2014, S. 197 f.).¹³

Als Mechanismus der Effekte von Klarheit und Strukturiertheit auf den *Lernerfolg* ist insbesondere anzunehmen, dass diese Unterrichtsmerkmale die Informationsverarbeitung der Schüler/-innen bedingen und somit die Qualität fachlicher Lernaktivitäten (Helmke, 2014, S. 169). Außerdem ist ein Effekt auf die Lernmotivation denkbar, vor allem weil klarer und strukturierter Unterricht die Erfolgserwartung der Schüler/-innen fördern sollte. Empirische Befunde aus zahlreichen Studien stehen im Einklang mit diesen Überlegungen. Die angenommen positiven Effekte auf den fachlichen Lernerfolg und auf motivational-emotionale Schülermerkmale werden durch Meta-Analysen bestätigt (Kyriakides et al., 2013, 147 f.; Titsworth et al., 2015, S. 15).

In Bezug auf *herkunftsbedingte Unterschiede im Lernerfolg* kann angenommen werden, dass klarer und strukturierter Unterricht niedrige Anforderun-

13 Es ist darauf hinzuweisen, dass das Konzept der Strukturiertheit in der Literatur in unterschiedlichen Bedeutungen verwendet wird. Teilweise wird der Sachverhalt, der hier als Steuerung des Lernens bezeichnet wird (Vermittlung von Inhalten vs. Erarbeiten von Inhalten, lehrkraftgelenkter Unterricht vs. offener Unterricht), als Strukturiertheit bezeichnet (z. B. Möller et al., 2002). Bei dieser Begriffsverwendung sind instruktionaler Unterricht und strukturierter Unterricht gleichzusetzen. Beim hier verwendeten Begriff von Strukturiertheit (Sequenzierung, Orientierung, Ordnung der Lerninhalte) ist instruktionaler Unterricht demgegenüber nicht zwangsläufig auch strukturiert (Beispiel: die Lehrkraft vermittelt Wissen und trifft lernrelevante Entscheidungen, aber sie gestaltet den Unterricht ohne erkennbare Phasen, ohne Mitteilug von Lernzielen und ohne Hervorhebung wichtiger Aspekte).

gen an die kognitiven Lernvoraussetzungen von Schüler/-innen stellt (Helmke, 2014, S. 200). Insbesondere Defizite bei Ausgangskompetenzen, Intelligenz und Sprachkompetenzen sollten bis zu einem gewissen Grad kompensiert werden können. Im Rahmen der Literaturrecherche wurden allerdings keine Studien gefunden, die die Effekte von Klarheit und Strukturiertheit bei Schüler/-innen verschiedener sozialer Herkunft vergleichen. Ermittelt werden konnten nur einige Studien, die die Effekte bei Schüler/-innen verschiedener Leistungsstärke vergleichen. Die Effekte von Klarheit werden in einigen älteren Studien zum Zielkriterium Egalisierung von Kompetenzunterschieden untersucht. Die Befunde fallen inkonsistent aus: teilweise besteht ein positiver Zusammenhang zur Egalisierung (Baumert et al., 1986, S. 652), teilweise kein Zusammenhang (Helmke, 1988, S. 61 ff.), teilweise besteht ein negativer Zusammenhang (Treiber & Weinert, 1985, S. 110). Hinsichtlich Strukturiertheit berichtet eine international angelegte Studie einen Interaktionseffekt mit dem Leistungsniveau der Klasse. Demnach hat Strukturiertheit bei leistungsschwachen Klassen stärkere Effekte als bei leistungsstarken Klassen (Vanlaar et al., 2015, S. 12 f.).

Insgesamt betrachtet können förderliche Effekte von Klarheit und Strukturiertheit auf den durchschnittlichen Lernerfolg nachgewiesen werden, während Effekte auf Unterschiede im Lernerfolg bislang nicht konsistent belegt werden können.

5.4.6 Effekte von kognitiver Aktivierung

Klarheit und Strukturiertheit beziehen sich primär auf die Art und Weise, wie Lehrkräfte Inhalte vermitteln. Die Art und Weise, wie Schüler/-innen die vermittelten Inhalte verarbeiten, steht im Vordergrund des Unterrichtsprinzips der kognitiven Aktivierung. *Kognitive Aktivierung* bezeichnet das Ausmaß, in dem die Lehrkraft die Schüler/-innen zum aktiven und tiefergehenden Nachdenken über die Lerninhalte anregt (Lipowsky, 2006, S. 60). Als kognitiv aktivierend werden insbesondere offene und herausfordernde Frage- und Aufgabenstellungen angesehen. Typische Komponenten sind beispielsweise, dass Vorwissen aktiviert und mit neuen Inhalten verknüpft wird, dass Erklärungen und Begründungen verlangt werden, dass Schlussfolgerungen gezogen werden, dass mehrere Lösungen und Lösungswege möglich sind, dass Vergleiche vorgenommen werden oder dass ein Transfer auf neue Situationen erfolgt (Lipowsky, 2006, S. 61; Praetorius et al., 2018, S. 414 f.).¹⁴

14 Obwohl das Konzept der kognitiven Aktivierung seinen Ursprung in einem konstruktivistischen Verständnis von Lernen hat (Praetorius et al., 2018, S. 410), ist das Konzept inhaltlich unabhängig von der Gegenüberstellung instruktionaler und kon-

Für Effekte von kognitiver Aktivierung auf den *durchschnittlichen Lernerfolg* spricht insbesondere, dass dieses Unterrichtsprinzip eine hohe Verarbeitungstiefe bei den Schüler/-innen begünstigt (Klieme et al., 2006, S. 131). Somit ist als ein Mechanismus die Förderung der Qualität fachlicher Lernaktivitäten anzunehmen. Empirisch überprüft wurden die Effekte von kognitiver Aktivierung in mehreren deutschsprachigen Studien (Praetorius et al., 2018). Für den Mathematikunterricht wurde wiederholt gezeigt, dass ein hohes Maß an kognitiver Aktivierung mit einem hohen durchschnittlichen Lernerfolg einhergeht (Fauth et al., 2014, S. 133; Kunter et al., 2006, S. 184f.; Lipowsky et al., 2009, S. 533; Spreitzer et al., 2022, S. 721). Weitere Belege liefert außerdem eine international vergleichende Studie (Caro et al., 2016). Bei bivariaten Analysen bestehen in etwa der Hälfte der Länder signifikante Zusammenhänge zwischen kognitiver Aktivierung und Mathematikkompetenzen, unter Kontrolle weiterer Unterrichtsmerkmale sogar in allen Ländern außer einem (S. 34 ff.). Interessanterweise zeigt sich in der großen Mehrheit der Länder neben einem linearen Effekt auch ein kurvilinearereffekt, d. h., die Mathematikkompetenzen fallen bei einem mittleren bis hohen Maß an kognitiver Aktivierung am günstigsten aus (S. 35).

In Bezug auf *Unterschiede im Lernerfolg* stellt sich die Frage, inwieweit das lernförderliche Potenzial von kognitiver Aktivierung zwischen Schüler/-innen verschiedener sozialer Herkunft variiert. Grundsätzlich ist anzunehmen, dass kognitiv aktivierende Fragen/Aufgaben eher hohe Anforderungen an die Lernvoraussetzungen von Schüler/-innen stellen. Die inhaltliche Offenheit der Frage-/Aufgabenstellungen beansprucht insbesondere Ausgangskompetenzen, Intelligenz und metakognitive Kompetenzen (Atlay et al., 2019, S. 2). Da diese Lernvoraussetzungen bei sozial benachteiligten Schüler/-innen weniger ausgeprägt sind, ist anzunehmen, dass ein hohes Maß an kognitiver Aktivierung mit großen Unterschieden im Lernerfolg einhergeht. Eine Reihe von empirischen Befunden steht im Einklang mit dieser Annahme. Bereits in der frühen Unterrichtsforschung wurde berichtet, dass in Klassen mit vielen sozial privilegierten Schüler/-innen kognitiv herausfordernde Aufgaben stärkere Effekte auf den Lernerfolg haben als in Klassen mit vielen sozial benachteiligten Schüler/-innen (Brophy & Good, 1986, S. 341). In einer jüngeren deutschen Studie zeigt sich, dass ein hohes Maß an kognitiver Aktivierung mit größeren

struktivistischer Unterrichtsmethoden (definiert über die Steuerung des Lernens). Demnach kann auch in instruktionalem Unterricht ein hohes Maß an kognitiver Aktivierung vorliegen, während im Gegenzug in konstruktivistischem Unterricht ein niedriges Maß an kognitiver Aktivierung gegeben sein kann (für ein Beispiel siehe Lipowsky & Lotz, 2015, S. 176f.).

Kompetenzunterschieden nach sozialer Herkunft einhergeht (Atlay et al., 2019, S. 7). Für sozial benachteiligte Schüler/-innen wird hierbei ein negativer Effekt von kognitiver Aktivierung auf den Lernerfolg ermittelt. Mehrere international angelegte Analysen auf der Grundlage von PISA 2012 sprechen ebenfalls für unterschiedsvergrößernde Effekte von kognitiver Aktivierung (Caro et al., 2016, S. 37; OECD, 2016b, S. 9 ff.; OECD, 2016c, S. 21 ff.). In diesen Studien liegt allerdings ein anderer Mechanismus zugrunde: kognitive Aktivierung hat bei allen Schüler/-innen positive Effekte, aber der Effekt fällt bei sozial privilegierten Schüler/-innen stärker aus als bei sozial benachteiligten Schüler/-innen (OECD, 2016b, S. 11; OECD, 2016c, S. 23). Im Kontrast zu den zuvor berichteten Studien wird in einer englischen Studie ermittelt, dass die Effekte von kognitiver Aktivierung bei sozial benachteiligten Schüler/-innen stärker ausfallen als bei sozial privilegierten Schüler/-innen (Burge et al., 2015, S. 11).

Zusammenfassend ist festzuhalten, dass empirische Befunde bestätigen, dass kognitive Aktivierung den durchschnittlichen Lernerfolg fördert. Die Befunde zu Unterschieden im Lernerfolg fallen nicht eindeutig aus, sprechen aber mehrheitlich für unterschiedsvergrößernde Effekte. Unklar bleibt bisher der zugrunde liegende Mechanismus (besondere Förderung sozial privilegierter Schüler/-innen vs. Beeinträchtigung sozial benachteiligter Schüler/-innen).

5.4.7 Effekte von Übung und Wiederholung

Im Anschluss an die Einführung von neuen Inhalten steht unter didaktischen Gesichtspunkten eine Unterrichtsphase, in der die behandelten Inhalte gefestigt werden. Damit angesprochen ist *Übung*, also das wiederholte Anwenden von Wissen und Kompetenzen (Helmke, 2014, S. 201). Übung zielt insbesondere auf die Verbesserung, die Automatisierung und das langfristige Behalten von Kompetenzen ab. Unterschieden werden können die Häufigkeit des Übens (Quantität der Übung) und die Art des Übens (Qualität der Übung). Die Häufigkeit des Übens stellt einen wichtigen Aspekt des Unterrichtstempos dar. Das Unterrichtstempo bezieht sich auf die Menge neuer Lerninhalte, die in einem festgelegten Zeitraum behandelt werden. Das Unterrichtstempo fällt umso niedriger aus, je mehr Unterrichtszeit für Übung und Wiederholung aufgewendet wird.

Als Wirkmechanismus des Effekts von Übung auf den durchschnittlichen Lernerfolg ist insbesondere das Bereitstellen zusätzlicher Lerngelegenheiten und somit die Erhöhung der Quantität fachlicher Lernaktivitäten anzunehmen. Hierbei erhöht Übung im Unterricht zwar nicht die Quantität der insgesamt ausgeführten Lernaktivitäten, aber Übung erhöht die spezifische Quantität

der Lernaktivitäten zu denjenigen Inhalten, die Gegenstand der Übung sind. Hinsichtlich des Zusammenhangs zwischen Anzahl von Wiederholungen und Lernergebnis ist unter theoretischen Gesichtspunkten von einem Sättigungseffekt auszugehen, d. h., der lernbezogene Nutzen pro Wiederholung nimmt mit jeder weiteren Wiederholung ab (Marzano et al., 2000, S. 66). Insofern erscheint ein mittleres Unterrichtstempo günstig: neue Inhalte werden ausreichend geübt (nicht zu wenig Übung), aber wenn die Inhalte zu einem Mindestmaß gefestigt sind, wird im Unterricht vorangeschritten (nicht zu viel Übung).

Die empirischen Befunde zu den Effekten von Übung fallen inkonsistent aus: in Meta-Analysen werden sowohl stark positive Effekte (Marzano et al., 2000, S. 4) als auch schwach positive Effekte (Kyriakides et al., 2013, S. 147 f.) und schwach negative Effekte (Scheerens et al., 2007, S. 119 ff.) auf den Lernerfolg berichtet. Für die heterogene Befundlage kommen zwei Erklärungen in Betracht. Erstens ist es denkbar, dass in verschiedenen Studien unterschiedliche Übungsintensitäten miteinander verglichen werden (z. B. wenn keine Übung mit moderater Übung verglichen wird, resultiert ein positiver Effekt; wenn moderate Übung mit viel Übung verglichen wird, resultiert kein Effekt). Diese Überlegung verweist auf den angenommenen Sättigungseffekt, der in einzelnen Studien auch berichtet wird (Brophy & Good, 1986, S. 337). Zweitens ist es möglich, dass in verschiedenen Studien unterschiedliche Arten von Übung betrachtet werden. Mehrfach finden sich Hinweise, dass repetitives Üben, also sich wiederholende Übungsaufgaben des immer gleichen Schemas, negative Effekte auf den Lernerfolg hat (Clausen, 2002, S. 172; Gruehn, 1998, S. 158). Als lernförderlich werden daher Übungsaufgaben angesehen, die vielfältige Anwendungsbeispiele aufzeigen und dadurch Transfer ermöglichen. Auf Verständnis abzielendes Üben ist der Vorzug gegenüber reinem Auswendiglernen zu geben, auch wenn letzteres manchmal unumgänglich ist (z. B. beim Vokabellernen) (Helmke, 2014, S. 202).

Weiterhin erscheint es plausibel, dass *verschiedene Schülergruppen* in unterschiedlichem Maß auf Übung und Wiederholung angewiesen sind. Demnach spielt Übung eine Rolle für die Anforderungen an die Lernvoraussetzungen der Schüler/-innen und an das Erziehungshandeln der Eltern. Werden im Unterricht nur wenige Übungsgelegenheiten geboten, müssen die Schüler/-innen entweder über eine hohe Ausgangskompetenz bzw. Intelligenz verfügen (so dass sie nur wenig Übung benötigen) oder über viel Motivation verfügen (so dass sie freiwillig außerhalb des Unterrichts üben) oder über engagierte Eltern verfügen (so dass die Eltern für ausreichend Übung sorgen). Da diese Bedingungen bei sozial benachteiligten Schüler/-innen häufiger nicht erfüllt sind,

sollte Übung im Unterricht einen Einfluss auf Kompetenzunterschiede nach sozialer Herkunft haben.

Empirische Studien zu Effekten von Übung auf *herkunftsbedingte Unterschiede im Lernerfolg* konnten im Rahmen der Literaturrecherche nicht ermittelt werden. Einige Studien untersuchen aber Effekte bei Schüler/-innen verschiedener Leistungsstärke. Die deutschen Studien zur Egalisierung berichten konsistent, dass eine hohe Übungsintensität mit der Egalisierung von Kompetenzunterschieden einhergeht (Baumert et al., 1986, S. 652; Helmke, 1988, S. 62; Treiber & Weinert, 1985, S. 110 ff.). In zwei der genannten Studien ist der egalisierende Effekt allerdings darauf zurückzuführen, dass eine hohe Übungsintensität einen lernbeeinträchtigenden Effekt bei leistungsstarken Schüler/-innen hat. Weitere Evidenzen für die Effekte von Übung liefert eine internationale Studie. Vanlaar und Kolleg/-innen (2015) berichten, dass Übung bei leistungsschwachen Klassen einen stärkeren Effekt hat als bei leistungsstarken Klassen (S. 12 f.).

Insgesamt betrachtet sprechen die Befunde dafür, dass Übung sowohl für einen hohen durchschnittlichen Lernerfolg als auch für geringe Unterschiede im Lernerfolg eine Rolle spielt, aber dass auch ungünstige Effekte wie das Beeinträchtigen des Lernerfolgs (insbesondere bei leistungsstarken Schüler/-innen) eintreten können.

5.4.8 Effekte von Unterstützung und Feedback

Die bisher thematisierten Unterrichtsprinzipien beziehen sich primär auf unterrichtliche Handlungen, bei denen die Lehrkraft die gesamte Klasse adressiert. Ein weiterer relevanter Bereich unterrichtlicher Handlungen betrifft die individuelle Lernbegleitung einzelner Schüler/-innen. Lernförderliche Interaktionen mit einzelnen Schüler/-innen können durch die Unterrichtsprinzipien Unterstützung und Feedback beschrieben werden. *Unterstützung* bezeichnet Maßnahmen zur Beseitigung von Lernproblemen (z. B. Verständnisschwierigkeiten, Fehlvorstellungen). Von zentraler Bedeutung ist hierbei ein konstruktiver Umgang mit Fehlern. Hierzu gehört es, Schüler/-innen Nachfragen zu ermöglichen und angemessen zu beantworten, Hinweise auf die richtige Lösung bzw. auf die richtige Lösungsstrategie zu geben und geeignete Lösungsstrategien zu modellieren (Muijs & Reynolds, 2001, S. 29). *Feedback* meint demgegenüber die Rückmeldung von Informationen an Schüler/-innen zu ihren Leistungen (Müller & Ditton, 2014, S. 15). Feedback folgt auf mündliche Schülerbeiträge, Übungsaufgaben, Hausaufgaben und Prüfungen. Sowohl bei Unterstützung als auch bei Feedback ist konzeptuell zwischen einer kognitiven

Komponente (z. B. Rückmeldung der korrekten Lösung) und einer emotionalen Komponente (z. B. Lob) zu unterscheiden (Kleickmann et al., 2020).

Für die Wirkweise von Unterstützung und Feedback bei der Förderung des *Lernerfolgs* kommen zwei Mechanismen in Betracht. Unterstützung und Feedback sollten die Qualität fachlicher Lernaktivitäten erhöhen, da Wissenslücken und Fehlvorstellungen gezielt behoben werden können. Theoretische Bezüge bestehen zur Zone der proximalen Entwicklung, wonach Schüler/-innen mit der Unterstützung einer kompetenten Person anspruchsvollere Aufgaben lösen können als ohne Unterstützung (Vygotsky, 1978). Als zweiter Mechanismus ist die Förderung von Motivation plausibel, wobei sich Unterstützung und Feedback insbesondere auf die Erfolgserwartung der Schüler/-innen günstig auswirken sollten. Positive Effekte auf die Lernmotivation sind vorrangig dann zu erwarten, wenn die Lernbegleitung in einer autonomieförderlichen Art erfolgt, so dass die Schüler/-innen die Lernbegleitung nicht als Kontrolle erleben (Ryan & Deci, 2000; Weinert & Helmke, 1987).

Empirische Befunde sprechen für lernförderliche Effekte von Unterstützung und Feedback. Scheerens und Kolleg/-innen (2007) berichten in ihrer Meta-Analyse bei Verwendung einer globalen Kategorie, die kognitive und emotionale Aspekte umfasst, einen schwachen positiven Effekt von Unterstützung (S. 120). In Studien, in denen nur die kognitive Komponente von Unterstützung erfasst wird (Helmke & Weinert, 1997b; Kleickmann et al., 2020), fallen die Effekte größer aus als in Studien, in denen Unterstützung global erfasst wird (Klieme et al., 2001b; Lipowsky et al., 2009). Für Feedback wird in Meta-Analysen ebenfalls ein positiver Effekt berichtet, aber die Effekte variieren deutlich zwischen verschiedenen Studien (Kluger & DeNisi, 1996, S. 258). Demnach ist die Ausgestaltung des Feedbacks entscheidend. Wirksames Feedback ist durch eine hohe Spezifität (statt Globalbewertungen), das Thematisieren von Ergebnissen und Vorgehensweisen sowie das Aufzeigen von Verbesserungsvorschlägen gekennzeichnet (Kellermann et al., 2023, S. 27 f.; Shute, 2008, S. 177 ff.).

In Bezug auf *herkunftsbedingte Unterschiede im Lernerfolg* sind mehrere Wirkmechanismen von Unterstützung und Feedback denkbar. Bezüglich der Anforderungen an die Lernvoraussetzungen der Schüler/-innen ist anzunehmen, dass Schüler/-innen mit niedrigen Ausgangskompetenzen, niedrigen metakognitiven Kompetenzen und geringen Erfolgserwartungen ein höheres Maß an Unterstützung und Feedback benötigen (Atlay et al., 2019, S. 3). Es kann auch argumentiert werden, dass Unterstützung und Feedback die Anforderungen an das Erziehungshandeln der Eltern gering halten, weil die Eltern dann nicht die Verantwortung dafür tragen, mögliche Lernprobleme ihrer

Kinder zu erkennen und zu beheben. Außerdem handelt es sich bei Unterstützung und Feedback um Unterrichtsmerkmale, bei denen der Mechanismus unterschiedlicher Lerngelegenheiten plausibel erscheint. Unterschiedliche Lerngelegenheiten liegen vor, wenn Lehrkräfte einzelne Schüler/-innen besonders intensiv beim Lernen begleiten. Unterschiedsverringernde Effekte sind zu erwarten, wenn Lehrkräfte bei der Verteilung von Interaktionen sozial benachteiligte Schüler/-innen priorisieren, was typischerweise der Fall zu sein scheint (Downey & Condrón, 2016, S. 213).

Es finden sich mehrere Studien zu Effekten von Unterstützung auf Unterschiede im Lernerfolg, aber die Befunde fallen nicht ganz einheitlich aus. In älteren US-amerikanischen Studien hatte sich gezeigt, dass in Klassen mit vielen sozial benachteiligten Schüler/-innen proaktive Unterstützung durch die Lehrkraft wichtig für den Lernerfolg ist, während in Klassen mit vielen sozial privilegierten Schüler/-innen die Initiierung von Unterstützung den Schüler/-innen überlassen werden kann (Brophy & Good, 1986, S. 349). Demgegenüber zeigt sich in einer jüngeren deutschen Studie, dass sozial privilegierte Schüler/-innen stärker von einem unterstützenden Unterrichtsklima profitieren als sozial benachteiligte Schüler/-innen (Atlay et al., 2019, S. 7). In einer anderen deutschen Studie wird berichtet, dass die Effekte von Unterstützung bei Schüler/-innen mit Migrationshintergrund stärker ausfallen als bei Schüler/-innen ohne Migrationshintergrund (Seiz et al., 2016, S. 244 f.).

Für die Effekte von Feedback finden sich ebenfalls empirische Belege. In der Grundschule ist festzustellen, dass das regelmäßige Nachbereiten von Texten mit niedrigeren herkunftsbedingten Unterschieden in der Lesekompetenz einhergeht (Pfeifer, 2011, S. 127). In ähnlicher Weise profitieren leistungsschwache Schüler/-innen stärker von der Besprechung der Hausaufgaben als leistungsstarke Schüler/-innen (Treiber & Weinert, 1985, S. 112). Neben der Häufigkeit des Feedbacks scheint auch die Art des Feedbacks eine Rolle zu spielen. In den frühen US-amerikanischen Studien wurde ermittelt, dass in Klassen mit vielen sozial privilegierten Schüler/-innen kritisches Feedback lernförderlich ist, während in Klassen mit vielen sozial benachteiligten Schüler/-innen Lob lernförderlich ist (Brophy, 1981, S. 16; Brophy & Good, 1986, S. 342, 349).

Zusammenfassend legen die Befunde nahe, dass Unterstützung und Feedback in angemessener Form den durchschnittlichen Lernerfolg fördern und tendenziell auch Unterschiede im Lernerfolg gering halten.

5.4.9 Effekte von Individualisierung

Im vorausgehenden Abschnitt wurde bereits kurz auf die Situation eingegangen, dass Lehrkräfte verschiedene Schüler/-innen im Unterricht unterschiedlich behandeln. Dieser Sachverhalt verweist auf das Unterrichtsprinzip der Individualisierung (Synonyme: Differenzierung, adaptiver Unterricht). Unter *Individualisierung* wird die Anpassung von unterrichtlichen Lernangeboten an die spezifischen Lernvoraussetzungen einzelner Schüler/-innen verstanden (Bohl et al., 2012, S. 44). Relevante Aspekte der Unterrichtsgestaltung sind individualisierte Erklärungen (z. B. zusätzliche Erläuterungen), individualisierte Fragen/Aufgaben (z. B. unterschiedlich schwierige Aufgaben) und individualisierte Lernzeiten (z. B. unterschiedliche Bearbeitungszeiten für Aufgaben). Auch die bereits thematisierten Praktiken der Bildung leistungshomogener Gruppen und der individuellen Lernbegleitung stellen Teilaspekte von Individualisierung dar.¹⁵

Die Relevanz von Individualisierung für den *Lernerfolg* wird dadurch begründet, dass Schüler/-innen verschiedene Lernvoraussetzungen aufweisen. Insofern können Unterschiede dahingehend bestehen, welche Lernangebote für die Schüler/-innen günstig sind. Erfahren alle Schüler/-innen das gleiche Lernangebot, können differenzielle Effekte resultieren, d. h., manche Schüler/-innen haben einen höheren Lernerfolg als andere Schüler/-innen (Aktionsrat Bildung, 2007, S. 79; Slavin, 1994, S. 145). Werden beispielsweise im Unterricht nur leichte Aufgaben gestellt, werden leistungsstarke Schüler/-innen unterfordert; werden dagegen nur schwierige Aufgaben gestellt, werden leistungsschwache Schüler/-innen überfordert. Durch Individualisierung kann dieser Konflikt gelöst werden, so dass im Idealfall alle Schüler/-innen für sie günstige Lernangebote erhalten. Für lernförderliche Effekte erscheinen hierbei zwei Mechanismen plausibel. Individualisierung sollte die Qualität fachlicher Lernaktivitäten erhöhen, weil die Auseinandersetzung mit den Lerninhalten bei allen Schüler/-innen auf einem angemessenen Schwierigkeitsniveau erfolgt (Ditton, 2000, S. 81f.). Im Einklang damit sollte Individualisierung

15 Es bestehen weiterhin inhaltliche Bezüge zwischen dem Unterrichtsprinzip Individualisierung und der Unterrichtsmethode offener Unterricht. Bei offenem Unterricht werden den Schüler/-innen Entscheidungsmöglichkeiten in Bezug auf die Ausgestaltung ihrer Lernaktivitäten zugestanden (z. B. Lernmethode, Lerntempo). Entscheidungsmöglichkeiten führen insofern zu individualisiertem Lernen, als dass die erfahrenen Lerngelegenheiten an die Präferenzen und Bedürfnisse der einzelnen Schüler/-innen angepasst sind. Anzumerken ist, dass dies nicht zwangsläufig auch bedeutet, dass die Schüler/-innen Lernaktivitäten wählen, die für ihren Lernerfolg günstig sind (Klauer & Leutner, 2012, S. 181 f.).

auch die Lernmotivation der Schüler/-innen fördern, weil alle Schüler/-innen sowohl Herausforderung als auch Erfolgserlebnisse erfahren (Helmke, 2014, S. 223).

Der Plausibilität dieser Argumentation stehen eher ernüchternde empirische Befunde gegenüber. In den verfügbaren Meta-Analysen werden teilweise schwach positive Effekte, teilweise keine Effekte und teilweise sogar schwach negative Effekte auf den Lernerfolg ermittelt (Lipowsky & Lotz, 2015, S. 162 f.). Negative Effekte von Individualisierung zeigen sich auch in mehreren deutschen Studien (Clausen, 2002, S. 172 f.; Gruehn, 1998, S. 158; Schulz-Heidorf, 2016, S. 175). Für den in empirische Studien festgestellten geringen Nutzen von Individualisierung kommen verschiedene Erklärungen in Betracht. Möglicherweise ist die Umsetzung individualisierender Maßnahmen durch die Lehrkräfte problematisch (Helmke, 2014, S. 263; Lipowsky & Lotz, 2015, S. 163). Beispielsweise ist denkbar, dass Lehrkräfte entweder das Anspruchsniveau von Aufgaben oder den Kompetenzstand von Schüler/-innen nicht richtig einschätzen, wodurch die Passung von Aufgabe zu Schüler/-in beeinträchtigt wird (d. h., die Qualität fachlicher Lernaktivitäten wird nicht erhöht). Allerdings bestehen auch bei gut umgesetzter Individualisierung einige Herausforderungen. Insbesondere gestaltet sich die Organisation des Unterrichtsablaufs deutlich schwieriger, wenn die Schüler/-innen verschiedene Lernaktivitäten ausführen (Muijs & Reynolds, 2001, S. 5 f.). Wenn Schüler/-innen beispielsweise verschiedene Aufgaben bearbeiten, ist die Besprechung einer bestimmten Aufgabe jeweils nur für diejenigen Schüler/-innen relevant, die diese Aufgabe bearbeitet haben, während für die restlichen Schüler/-innen ein Leerlauf entsteht. Dadurch kann die Quantität fachlicher Lernaktivitäten verringert werden. Alternativ kann auf die gemeinsame Besprechung von Aufgaben verzichtet werden, was aber die Qualität fachlicher Lernaktivitäten verringern kann (siehe Befunde zu den Effekten von Feedback).

An individualisierende Maßnahmen wird häufig auch die Erwartung gestellt, dass sie einen Beitrag zur Verringerung von *Unterschieden im Lernerfolg* leisten (Bohl et al., 2012, S. 42; Lipowsky & Lotz, 2015, S. 167; Schulz-Heidorf, 2016, S. 2). Konzeptuell hat Individualisierung einen starken Bezug zum Mechanismus unterschiedlicher Lerngelegenheiten, insofern erscheint ein Effekt durchaus plausibel. Allerdings ist anzunehmen, dass die Richtung des erwarteten Effekts maßgeblich vom Vergleichsmaßstab eines nicht individualisierten Unterrichts abhängt. Wenn Lehrkräfte den Unterricht nicht individualisieren, müssen sie sich entscheiden, ob sie das Anspruchsniveau eher an den leistungsschwachen Schüler/-innen ausrichten (wodurch leistungsstarke Schüler/-innen unterfordert werden) oder eher an den leistungsstarken Schüler/-in-

nen (wodurch leistungsschwache Schüler/-innen überfordert werden) (Treinies & Einsiedler, 1996, S. 302). In Abhängigkeit von den zwei möglichen Ausgangslagen eines nicht individualisierten Unterrichts ergeben sich unterschiedliche Effekte durch die Einführung individualisierender Maßnahmen. Im ersten Fall kommt die Individualisierung insbesondere den leistungsstarken Schüler/-innen zugute, wodurch sich Unterschiede im Lernerfolg vergrößern sollten. Demgegenüber profitieren im zweiten Fall vor allem die leistungsschwachen Schüler/-innen und somit sollten Unterschiede im Lernerfolg verringert werden.

Im Rahmen der Literaturrecherche wurden primär Studien ermittelt, die Effekte von Individualisierung auf Schüler/-innen verschiedener Leistungsstärke untersuchen. Helmke (1988) berichtet, dass Individualisierungsmaßnahmen, die insbesondere auf die Förderung leistungsschwacher Schüler/-innen fokussieren, mit der Egalisierung von Kompetenzunterschieden einhergehen (S. 61ff.). Im Einklang damit stehen Befunde zum zielerreichenden Lernen (Englisch: mastery learning). Diese Unterrichtsmethode ist durch regelmäßige Leistungstests und Fördermaßnahmen bei Nicht-Erfüllen der Mindestanforderungen gekennzeichnet. Meta-analytische Befunde sprechen dafür, dass leistungsschwache Schüler/-innen stärker von zielerreichendem Lernen profitieren als leistungsstarke Schüler/-innen (Hattie, 2013, S. 203). Es liegen allerdings auch gegensätzliche Befunde vor. In der Studie von Treiber und Weinert (1985) gehen individualisierende Maßnahmen mit größeren Kompetenzunterschieden einher (S. 151). In einer belgischen Studie zeigt sich, dass Niedrig-Risiko-Schüler/-innen stärker von Individualisierung profitieren als Hoch-Risiko-Schüler/-innen (u. a. sozial benachteiligt, Migrationshintergrund) (Vanlaar et al., 2014, S. 421).

Insgesamt betrachtet können förderliche Effekte von Individualisierung auf den durchschnittlichen Lernerfolg bisher nicht klar belegt werden. Weiterhin scheint Individualisierung teilweise unterschiedsverringernde und teilweise unterschiedsvergrößernde Effekte zu haben, was im Einklang mit den theoretischen Überlegungen steht.

5.5 Effekte von Ganztagsangeboten

Nachdem im letzten Abschnitt ein ausführlicher Überblick über empirische Befunde zum Unterricht gegeben wurde, wird nachfolgend auf Ganztagsangebote als weiterem relevantem schulischem Lernangebot eingegangen. Ganztagsangebote und Ganztagschulen erhalten seit der Jahrtausendwende vermehrte Aufmerksamkeit in der deutschen Bildungsforschung. Der Hintergrund liegt

darin, dass die Kultusministerkonferenz als Reaktion auf die Ergebnisse der ersten PISA-Studie unter anderem den flächendeckenden Ausbau von Ganztagschulen beschlossen hatte (Kultusministerkonferenz, 2002, S. 7).

Konzeptuell ist zwischen verschiedenen Merkmalsbereichen zu unterscheiden. Die *Organisationsform* einer Schule stellt ein Strukturmerkmal dar und bezieht sich auf das Ausmaß und die Verbindlichkeit des Ganztagsbetriebs (Aktionsrat Bildung, 2013, S. 23 ff.; Linberg et al., 2015, S. 7 f.). Unterschieden werden Halbtagschulen, offene Ganztagschulen (freiwilliger Besuch), teilgebundene Ganztagschulen (teilweise freiwilliger, teilweise verpflichtender Besuch) und vollgebundene Ganztagschulen (verpflichtender Besuch). Ein weiterer Bereich bezieht sich auf die eigentlichen *Ganztagsangebote* und somit auf Prozessmerkmale im Sinne von Handlungen pädagogischer Fachkräfte. Ganztagsangebote können bezüglich des Umfangs (z. B. Häufigkeit), der Inhalte (z. B. Hausaufgabenbetreuung, Förderunterricht, Freizeitangebote) und der Gestaltung (z. B. Lernmethoden) betrachtet werden. Diesen auf der Schulebene verorteten Merkmalen stehen auf der Schülerebene die *Nutzung von Ganztagsangeboten* (z. B. Häufigkeit des Besuchs) und die daraus resultierenden Lernwirkungen gegenüber.

Der Ausbau von Ganztagschulen ist unter anderem mit der Hoffnung verbunden, dass hierdurch eine Verbesserung des *Lernerfolgs* der Schüler/-innen erreicht werden kann (Aktionsrat Bildung, 2013, S. 20; Kultusministerkonferenz, 2015, S. 4; Linberg et al., 2015, S. 5). Effekte auf den durchschnittlichen Lernerfolg sind hierbei aufgrund von unterschiedlichen Mechanismen zu erwarten. Bestimmte Ganztagsangebote wirken eher über die Erhöhung der Quantität fachlicher Lernaktivitäten (z. B. Förderunterricht), andere über die Förderung von Lernvoraussetzungen (z. B. Freizeitangebote).

Neben der allgemeinen Förderung des Lernerfolgs wird an den Ausbau von Ganztagschulen auch die Erwartung gestellt, dass *herkunftsbedingte Unterschiede im Lernerfolg* verringert werden sollen (Aktionsrat Bildung, 2013, S. 20; Kultusministerkonferenz, 2015, S. 4). Unter theoretischen Gesichtspunkten sind unterschiedsverringende Effekte allerdings nur unter *bestimmten Bedingungen* zu erwarten. Aussichtsreich erscheinen *verpflichtende Ganztagsangebote*, weil sie über den Mechanismus der Angleichung von Lernaktivitäten wirken. Da sozial benachteiligte Schüler/-innen außerhalb der Schule weniger Anregung erfahren, sollte diese Schülergruppe stärker von verpflichtenden Ganztagsangeboten profitieren als sozial privilegierte Schüler/-innen (Raudenbush & Eschmann, 2015, S. 454). Bei *freiwilligen Ganztagsangeboten* spielt dagegen der Mechanismus unterschiedlicher Lerngelegenheiten eine Rolle. Unterschiedsverringende Effekte sind zu erwarten, wenn sozial benachtei-

ligte Schüler/-innen Ganztagsangebote häufiger nutzen als sozial privilegierte Schüler/-innen. Demgegenüber ist bei der häufigeren Nutzung durch sozial privilegierte Schüler/-innen mit unterschiedsvergrößernden Effekten zu rechnen. Die empirische Befundlage zu Nutzungsunterschieden fällt, wie im dritten Kapitel bereits dargestellt, bisher unklar aus. Schließlich ist anzunehmen, dass die *Inhalte der Ganztagsangebote* von Bedeutung sind. Da Effekte auf Fachkompetenzen am ehesten bei Ganztagsangeboten mit Bezug zum fachlichen Lernen zu erwarten sind, haben diese auch das größte Potential für die Geringhaltung von Kompetenzunterschieden. Spezifisch kann argumentiert werden, dass durch Hausaufgabenbetreuungen und Förderunterricht für leistungsschwache Schüler/-innen die Anforderungen sowohl an die Lernvoraussetzungen der Schüler/-innen (z. B. Ausgangskompetenz, Intelligenz) als auch an das Erziehungshandeln der Eltern (z. B. Hilfe bei Hausaufgaben, Initiierung von Nachhilfe) gering gehalten werden.

Empirische Befunde zu den Effekten der *Nutzung von Ganztagsangeboten* durch Schüler/-innen wurden bereits im dritten Kapitel behandelt. An dieser Stelle werden Untersuchungen zu den Effekten von schulischen Struktur- und Prozessmerkmalen genauer betrachtet. Zunächst wird auf Effekte der schulischen Organisationsform eingegangen und anschließend auf die Effekte von spezifischen Programmen.

5.5.1 Effekte der schulischen Organisationsform

In der deutschsprachigen Forschung zu Ganztagschulen dominieren Studien, in denen Schulen mit unterschiedlicher Organisationsform miteinander verglichen werden (Halbtagschulen vs. Ganztagschulen). Bei der schulischen Organisationsform handelt es sich um eine strukturelle Bedingung, die einen gewissen Rahmen für am Nachmittag erfolgende Lehr-Lern-Prozesse eröffnet. Allerdings kann dieser Rahmen sehr unterschiedlich ausgestaltet werden, z. B. bezüglich der inhaltlichen, methodischen und personellen Ausgestaltung der Ganztagsangebote. Diese Aspekte bleiben bei der Betrachtung der schulischen Organisationsform unberücksichtigt. Insofern hat die Organisationsform einer Schule vermutlich nur eine geringe Aussagekraft für die bei den Schüler/-innen erfolgenden Lernprozesse.

Empirische Befunde zu den Effekten der schulischen Organisationsform auf den *durchschnittlichen Lernerfolg* bestätigen diese Annahme. Bei der Gegenüberstellung von Halbtagschulen und Ganztagschulen zeigen sich in der Regel keine bedeutsamen Unterschiede bei den von den Schüler/-innen erreichten Kompetenzen (Aktionsrat Bildung, 2013, S. 60; Linberg et al., 2018,

S. 1217 ff.; Radisch et al., 2006, S. 46; Steinmann & Strietholt, 2018, S. 184; Strietholt et al., 2015, S. 754; abweichender Befund für vollgebundene Ganztagschulen bei Linberg et al., 2015, S. 39 ff.). Dieses Befundmuster legt nahe, dass die bloße Umstellung der Organisationsform von Halbtagschule auf Ganztagschule nicht automatisch zu besseren Schülerkompetenzen führt.

Mehrere deutschsprachige Studien untersuchen weiterhin Effekte der schulischen Organisationsform auf *herkunftsbedingte Unterschiede im Lernerfolg*. Mehrheitlich wird festgestellt, dass Kompetenzunterschiede nach sozialer Herkunft an Ganztagschulen genauso groß ausfallen wie an Halbtagschulen (Aktionsrat Bildung, 2013, S. 60; Linberg et al., 2015, S. 49 ff.; Strietholt et al., 2015, S. 754; Steinmann & Strietholt, 2018, S. 185). Lediglich in der Studie von Züchner und Fischer (2014) zeigt sich, dass Kompetenzunterschiede nach sozialer Herkunft an gebundenen Ganztagschulen signifikant niedriger ausfallen als an teilgebundenen und offenen Ganztagschulen (S. 360). Die Aussagekraft dieses Befunds ist insofern eingeschränkt, als dass in der untersuchten Stichprobe keine Halbtagschulen enthalten sind, sodass ein Vergleich zwischen gebundenen Ganztagschulen und Halbtagschulen nicht möglich ist.

Zusammenfassend ist festzuhalten, dass bisher kaum lernförderliche oder unterschiedsverringende Effekte von Ganztagschulen belegt werden können, wenn die schulische Organisationsform betrachtet wird.

5.5.2 Effekte von spezifischen Programmen

Die zuvor berichteten Befunde stützen die Annahme, dass lernförderliche Effekte wesentlich von der konkreten *Ausgestaltung der Ganztagsangebote* abhängen. Im Rahmen der Literaturrecherche wurden allerdings kaum deutschsprachige Studien gefunden, die die Effekte von Ganztagsangeboten (im Sinne von Prozessmerkmalen) untersuchen. Demgegenüber befasst sich eine Reihe von US-amerikanischen Studien mit *spezifischen Programmen* wie After-School Programs oder Tutoring Programs. Die meisten Studien untersuchen hierbei Effekte auf den *durchschnittlichen Lernerfolg*. Relevante Befunde werden nachfolgend zusammengefasst.

Bei sogenannten *After-School Programs* handelt es sich um eine inhaltlich weit gefasste Oberkategorie für verschiedene Arten von Nachmittagsangeboten wie Hausaufgabenhilfe, künstlerisch-kulturelle Aktivitäten oder Präventionsprogramme (Scott-Little et al., 2002, S. 394). Typischerweise werden After-School Programs in Schulen durchgeführt (vereinzelt aber auch in anderen Einrichtungen, siehe Zief et al., 2006, S. 9), daher haben sie auch Aussagekraft für die hier interessierenden Ganztagschulen. In Meta-Analysen und Re-

views werden teilweise schwache positive Effekte (Lauer et al., 2006, S. 292 ff.; Scott-Little et al., 2002, S. 405 ff.) und teilweise keine Effekte (Zief et al., 2006, 20 ff.) von After-School Programs auf den fachlichen Lernerfolg ermittelt.

Einen stärkeren Bezug zum schulischen Lernen haben sogenannte *Tutoring Programs*. Der englische Begriff Tutoring bezeichnet fachbezogene Förderangebote insbesondere für leistungsschwächere Schüler/-innen. Tutoring kann in unterschiedlichen Varianten durchgeführt werden (z. B. Einzelunterricht vs. Gruppenunterricht, Lehrkräfte als Tutoren vs. Gleichaltrige als Tutoren) (Pellegrini et al., 2020, S. 17 f.). Im Kontext von Ganztagschulen ist primär Tutoring von Interesse, bei dem eine Lehrkraft eine Gruppe von Schüler/-innen unterrichtet. Inhaltlich entspricht dies dem, was im Deutschen als *Förderunterricht* bezeichnet wird. Die Effekte eines solchen Tutorings wurden in einer Reihe von hochwertigen Studien mit experimentellem Design untersucht. Zwei Meta-Analysen synthetisieren den Forschungsstand und belegen positive Effekte auf den fachlichen Lernerfolg (Pellegrini et al., 2020, S. 18; Slavin et al., 2011, S. 12).

Erwähnenswert ist in diesem Kontext auch eine österreichische Studie zu den Effekten verschiedener Arten von Förderunterricht (Schmid et al., 2020, S. 162 ff.). Die Befunde legen nahe, dass Förderunterricht für leistungsschwache Schüler/-innen und Förderunterricht für Schüler/-innen mit Sprachproblemen positive Effekte auf den durchschnittlichen Lernerfolg haben. Bei Förderunterricht für leistungsstarke Schüler/-innen sind demgegenüber keine Effekte und teilweise sogar negative Effekte festzustellen.

Unter dem Gesichtspunkt *herkunftsbedingter Unterschiede im Lernerfolg* ist von besonderem Interesse, ob die Effekte von Ganztagsangeboten bei Schüler/-innen verschiedener sozialer Herkunft unterschiedlich stark ausfallen. In der internationalen Literatur finden sich hierzu einige wenige Studien. Diese Studien betrachten entweder die Dauer des Schultags (Halbtage vs. Ganztage) oder spezifische After-School Programs. Die Befunde fallen unterschiedlich aus. In einigen Studien profitieren sozial benachteiligte Schüler/-innen stärker von Ganztagsangeboten als sozial privilegierte Schüler/-innen, in anderen Studien bestehen keine Unterschiede zwischen den beiden Schülergruppen (Patall et al., 2010, S. 417; Raudenbush & Eschmann, 2015, S. 460 ff.).

Insgesamt betrachtet legen die Befunde nahe, dass die Effekte von verschiedenen Arten von Ganztagsangeboten sehr unterschiedlich ausfallen. Für die Wirksamkeit von Tutoring liegen überzeugende Evidenzen vor, während die Wirksamkeit von After-School Programs nur teilweise belegt werden kann. Ebenso können unterschiedsverringemde Effekte von Ganztagsangeboten bisher nur vereinzelt bestätigt werden. Hierbei ist allerdings zu beachten, dass zu

theoretisch aussichtsreichen Ganztagsangeboten wie Hausaufgabenbetreuungen oder Förderunterricht für leistungsschwache Schüler/-innen bisher kaum Studien vorliegen.

5.6 Effekte von Merkmalen der Lerngruppe

In den bisherigen Ausführungen zu Unterricht und Ganztagsangeboten standen zielgerichtete Maßnahmen von pädagogischen Fachkräften im Vordergrund. Einen weiteren wichtigen Aspekt der schulischen Lernumwelt stellen die Mitschüler/-innen und ihr Verhalten dar. Im Weiteren wird betrachtet, wie sich *Merkmale der Lerngruppe* auf den Lernerfolg von Schüler/-innen auswirken. Spezifisch wird auf die Rolle der Klassengröße und der Schülerkomposition eingegangen.

5.6.1 Effekte der Klassengröße

Ein Einfluss der Klassengröße auf unterrichtliche Lehr-Lern-Prozesse ist aufgrund mehrerer Mechanismen denkbar (Hattie, 2013, S. 102). Ein direkter Mechanismus wäre darin zu sehen, dass bei einer geringeren Anzahl von Schüler/-innen im Klassenzimmer weniger Unterrichtsstörungen auftreten. Insofern sollte die Klassengröße einen unmittelbaren Einfluss auf die Quantität fachlicher Lernaktivitäten haben. Als weiterer Mechanismus ist plausibel, dass der Zeitanteil für direkte Interaktionen zwischen Lehrkraft und einzelnen Schüler/-innen in kleineren Klassen größer ausfällt (weil die verfügbare Zeit auf weniger Schüler/-innen aufgeteilt werden muss). Dies sollte ein hohes Maß an Unterstützung und Feedback für die Schüler/-innen begünstigen. Schließlich sollte es der Lehrkraft in kleineren Klassen leichter möglich sein, anregende Unterrichtsmethoden wie kooperatives Lernen oder angeleitetes entdeckendes Lernen einzusetzen.

Empirische Befunde zu den Effekten der Klassengröße auf den *Lernerfolg* der Schüler/-innen fallen eher ernüchternd aus (Hanushek, 1986, S. 1161; Hattie, 2002, S. 467 ff.; Robinson, 1990, S. 82 ff.). In den meisten Studien werden entweder schwache positive Effekte oder keine Effekte durch kleinere Klassen ermittelt. Hierbei scheint der Zusammenhang zwischen Klassengröße und Lernerfolg durch einen Sättigungseffekt gekennzeichnet zu sein. Demnach entstehen bedeutsame Effekte auf den Lernerfolg erst, wenn die Klassengröße auf 15 Schüler/-innen oder weniger reduziert wird, während der Unterschied zwischen 20 und 40 Schüler/-innen vernachlässigbar ist (Hattie, 2002, S. 467). Eine naheliegende Erklärung für die eher schwachen Effekte der Klassengröße

bietet der zuvor berichtete Befund, dass Lehrkräfte ihre Unterrichtsgestaltung kaum an kleinere Klassen anpassen. Somit entsteht der Eindruck, dass den Kompetenzen der Lehrkraft eine deutlich größere Bedeutung zukommt als der Klassengröße: „A poor teacher with 30 students may remain a poor teacher with 15.“ (Hattie, 2002, S. 470).

Bei der Einschätzung der Rolle der Klassengröße ist allerdings auch zu beachten, inwieweit sie *herkunftsbedingte Unterschiede im Lernerfolg* beeinflusst. Mehrere Argumente sprechen dafür, dass sozial benachteiligte Schüler/-innen im besonderen Maße von kleineren Klassen profitieren. Einerseits ist anzunehmen, dass kleinere Klassen aufgrund des geringen Ausmaßes an Unterrichtsstörungen (und somit weniger ablenkenden Reizen) niedrigere Anforderungen an die Lernmotivation und Selbstregulation der Schüler/-innen stellen. Andererseits sollten sozial benachteiligte Schüler/-innen einen überproportionalen Anteil der in kleineren Klassen zusätzlich verfügbaren Zeit für direkte Lehrkraft-Schüler-Interaktionen erhalten (Blatchford et al., 2011, S. 718). Empirisch zeigt sich in mehreren Studien, dass die Effekte der Klassengröße bei sozial benachteiligten Schüler/-innen größer ausfallen als bei sozial privilegierten Schüler/-innen (Blatchford et al., 2011, S. 718; Gustafsson et al., 2018, S. 17; Robinson, 1990, S. 85 f.), in manchen Studien werden aber auch keine differenziellen Effekte gefunden (Rivkin et al., 2005, S. 446; Schwippert, 2001, S. 139 ff.). Erwähnenswert sind in diesem Kontext auch Befunde, wonach Schüler/-innen mit Migrationshintergrund stärker von kleineren Klassen profitieren als Schüler/-innen ohne Migrationshintergrund (Hattie, 2002, S. 469; Robinson, 1990, S. 85 f.).

Insgesamt betrachtet sprechen die Befunde dafür, dass die Klassengröße nur schwache Effekte auf den Lernerfolg hat, aber dass die Effekte bei verschiedenen Schülergruppen unterschiedlich stark ausfallen. Da es sich bei der Verringerung der Klassengröße um eine ausgesprochen teure Maßnahme handelt, wird die Frage nach der Effizienz, also der Verhältnismäßigkeit von Aufwand und Nutzen, aufgeworfen (Hattie, 2002, S. 472).

5.6.2 Effekte der Schülerkomposition

Während sich die Klassengröße darauf bezieht, wie viele Schüler/-innen in der Klasse sind, bezieht sich die *Klassenkomposition* darauf, aus welchen Schüler/-innen sich die Klasse zusammensetzt. Zentrale und vielfach untersuchte Merkmale sind die leistungsbezogene Komposition (durchschnittliche Intelligenz und/oder Ausgangskompetenz), die soziale Komposition (durchschnittlicher sozialer Status) und die ethnische Komposition (Anteil an Schüler/-innen

mit Migrationshintergrund) (Reynolds et al., 2014, S. 208). Die Zusammensetzung der Schülerschaft kann auch auf der Schulebene betrachtet werden und wird dann als *Schulkomposition* bezeichnet. Konzeptuell beziehen sich Kompositionseffekte darauf, dass unter Kontrolle des individuellen Schülermerkmals (z. B. Intelligenz) das aggregierte Schülermerkmal (z. B. Klassenmittelwert der Intelligenz) einen zusätzlichen Einfluss auf den Lernerfolg der Schüler/-innen hat (Ditton, 1998, S. 91). Im genannten Beispiel kann dies so interpretiert werden, dass eine höhere Intelligenz nicht nur zu besseren Fachkompetenzen beim einzelnen Schüler/bei der einzelnen Schülerin führt, sondern auch zu besseren Fachkompetenzen bei den Mitschüler/-innen.

Ein möglicher vermittelnder Mechanismus zwischen Schülerkomposition und unterrichtlichen Lehr-Lern-Prozessen ist das Lernverhalten der Mitschüler/-innen im Unterricht (Van Ewijk & Sleegers, 2010a, S. 241). Das Lernverhalten der Schüler/-innen umfasst einerseits das Ausmaß an Unterrichtsstörungen (beeinflusst die Quantität fachlicher Lernaktivitäten) und andererseits Schüler-Schüler-Interaktionen wie gegenseitige Unterstützung (beeinflusst die Qualität fachlicher Lernaktivitäten) und gegenseitige Motivierung (beeinflusst die Lernmotivation). Ein weiterer Mechanismus ist in der Anpassung der Unterrichtsgestaltung durch die Lehrkraft zu sehen. Demnach sollte die Lehrkraft die verwendeten Unterrichtsmethoden und prinzipien in Abhängigkeit von den Bedingungen in der jeweiligen Klasse auswählen (Van Ewijk & Sleegers, 2010a, S. 241). Somit würde die Schülerkomposition die Handlungsbereitschaften der Lehrkraft beeinflussen (z. B. in leistungsstarken Klassen sind die Lehrkräfte eher dazu bereit, konstruktivistische Unterrichtsmethoden einzusetzen).

Zahlreiche Studien befassen sich mit den Effekten der Schülerkomposition auf den *durchschnittlichen Lernerfolg*. Die konsistentesten Befunde zu Kompositionseffekten liegen für die leistungsbezogene Komposition von Klassen/Schulen vor. Demnach profitieren Schüler/-innen in erheblichem Maße von leistungsstarken Mitschüler/-innen (Baumert et al., 2006, S. 138; Kyriakides et al., 2018b, S. 14; Reynolds et al., 2014, S. 208). Auch für die soziale Komposition können bedeutsame Effekte nachgewiesen werden. Demnach begünstigen sozial privilegierte Mitschüler/-innen einen hohen Lernerfolg. Die Effekte werden aber deutlich reduziert und verschwinden teilweise sogar vollständig, wenn leistungsbezogene Merkmale der Schüler/-innen berücksichtigt werden (Baumert et al., 2006, S. 138; Holzberger et al., 2020, S. 11; Van Ewijk & Sleegers, 2010b, S. 143 ff.). Hinsichtlich der ethnischen Komposition liegen Befunde vor, dass sich ein sehr hoher Anteil von Mitschüler/-innen mit Migrationshintergrund ungünstig auf den Lernerfolg auswirkt. Die Effekte fallen allerdings

relativ schwach aus und verschwinden teilweise unter Kontrolle leistungsbezogener und sozialer Kompositionsmerkmale (Baumert et al., 2006, S. 141 f.; Reynolds et al., 2014, S. 209; Van Ewijk & Sleegers, 2010a, S. 249 ff.). Bei der Gegenüberstellung von Klassenkomposition und Schulkomposition sprechen die Befunde dafür, dass die Klassenkomposition größere Effekte auf den Lernerfolg hat als die Schulkomposition (Van Ewijk & Sleegers, 2010b, S. 145).

In Bezug auf mögliche Effekte der Schülerkomposition auf *herkunftsbedingte Unterschiede im Lernerfolg* kann mit unterschiedlichen Mechanismen argumentiert werden. Einerseits sollte eine günstige Komposition durch das förderliche Lernverhalten der Mitschüler/-innen (z. B. weniger Unterrichtsstörungen, mehr Unterstützung) die Anforderungen an die Lernvoraussetzungen des einzelnen Schülers/der einzelnen Schülerin gering halten. Andererseits kann eine günstige Komposition zu einer Unterrichtsgestaltung führen, die hohe Anforderungen an die Lernvoraussetzungen der Schüler/-innen stellt (z. B. hohes Maß an kognitiver Aktivierung, hohes Unterrichtstempo). Es liegen zahlreiche Studien zu den Effekten der Schülerkomposition auf Unterschiede im Lernerfolg vor, aber die Befunde fallen ausgesprochen heterogen aus. Mehrere Studien berichten, dass der Effekt einer günstigen Komposition bei sozial benachteiligten bzw. leistungsstarken Schüler/-innen größer ausfällt (Coleman et al., 1966, S. 22; Veldman & Sanford 1984, S. 634; Zimmer & Toma 2000, S. 84 ff.), aber ebenso berichten mehrere Studien, dass der Effekt einer günstigen Komposition bei sozial privilegierten bzw. leistungsstarken Schüler/-innen größer ausfällt (Baumert et al., 1987a, S. 45; Ma et al., 2016, S. 521; Opdenakker et al., 2002, S. 412). Dieses Bild zeigt sich auch in der international vergleichenden Studie von Gustafsson und Kolleg/-innen (2018): In 10 von 50 Ländern geht eine günstige Schülerkomposition mit geringeren Kompetenzunterschieden nach sozialer Herkunft einher, in 10 anderen Ländern geht eine günstige Schülerkomposition dagegen mit größeren Kompetenzunterschieden nach sozialer Herkunft einher (S. 22). Die Autor/-innen vermuten, dass diese Unterschiede in der Wirkrichtung auf die verschiedenen Schulsysteme der Länder zurückzuführen sind (S. 26).

Zusammenfassend ist festzuhalten, dass bedeutsame Kompositionseffekte auf den durchschnittlichen Lernerfolg belegt werden können, aber dass sich keine klare Tendenz abzeichnet, welche Schülergruppe am stärksten von einer günstigen Komposition profitiert.

5.7 Zusammenfassung des fünften Kapitels

Im vorliegenden Kapitel wurden die Effekte von pädagogischen Merkmalen innerhalb von Schulen behandelt. Hierfür wurde zunächst ein theoretisches Modell der Wirksamkeit von Schule entwickelt, in welchem mehrere Kategorien von pädagogischen Merkmalen unterschieden werden. Weiterhin wurden Mechanismen diskutiert, durch die pädagogische Merkmale herkunftsbedingte Unterschiede im Lernerfolg beeinflussen. Zentrale Mechanismen sind einerseits die Förderung verschiedener Schülergruppen und andererseits differenziell wirksame Lerngelegenheiten vs. unterschiedliche Lerngelegenheiten. Auf dieser Grundlage wurde ein Überblick über den empirischen Forschungsstand zur Wirksamkeit von pädagogischen Merkmalen gegeben.

Die Effekte von pädagogischen Merkmalen auf den *durchschnittlichen Lernerfolg* wurden in zahlreichen Studien untersucht. Belastbare Evidenzen für lernförderliche Effekte liegen beim Unterrichtsumfang, Klassenunterricht, kooperativem Lernen, instruktionalen Unterrichtsmethoden, ausgewogenen Unterrichtsmethoden (Verbindung von Instruktion und Konstruktion), Klassenführung, Klarheit und Strukturiertheit, kognitiver Aktivierung, Unterstützung und Feedback, Förderunterricht für leistungsschwache Schüler/-innen (Tutoring) und der Schülerkomposition vor.

Demgegenüber ist der Forschungsstand zu den Effekten von pädagogischen Merkmalen auf *herkunftsbedingte Unterschiede im Lernerfolg* deutlich begrenzter. Innerhalb der Kategorie der Unterrichtsmerkmale ist nur beim Unterrichtsumfang sowie bei kooperativen und konstruktivistischen Unterrichtsmethoden eine einigermaßen umfangreiche und konsistente Befundlage festzustellen. Deutliche Forschungsdefizite bestehen demgegenüber bezüglich der Effekte von übergreifenden Unterrichtsprinzipien (Klassenführung, Klarheit und Strukturiertheit, kognitive Aktivierung, Übung und Wiederholung, Unterstützung und Feedback, Individualisierung), da bei diesen Unterrichtsmerkmalen jeweils keine oder nur wenige Studien vorhanden sind. Bei Ganztagsangeboten liegen mehrere Studien zu den Effekten von schulischen Organisationsformen vor (und können diese nicht bestätigen), aber kaum Studien zu Prozessmerkmalen von Ganztagsangeboten. Dieses Forschungsdefizit ist insofern beachtenswert, als dass Effekte von Prozessmerkmalen theoretisch eher zu erwarten sind als Effekte von schulischen Organisationsformen. Zu den Effekten von Lerngruppen ist eine Reihe von Studien vorhanden. Ein Forschungsdefizit besteht allerdings bezüglich der Effekte der Schülerkomposition, da die Befundlage in diesem Bereich sehr inkonsistent ausfällt.

Insgesamt betrachtet ist der Forschungsstand zu den Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg als

unzureichend einzuschätzen. Die Anzahl der verfügbaren Studien ist überschaubar (selbst bei den besser untersuchten Merkmalen) und damit einhergehend sind kaum Übersichtsarbeiten vorhanden. Angesichts des gegebenen Forschungsstandes werden weitere, hochwertige Untersuchungen benötigt. Die vorliegende Arbeit leistet hierzu einen Beitrag.

6. Fragestellungen

In den bisherigen Kapiteln wurde eine theoretische Grundlage erarbeitet und ein Überblick über den bestehenden Forschungsstand gegeben. Damit ist die Aufarbeitung der vorhandenen Literatur abgeschlossen. Die nachfolgenden Kapitel stellen den empirischen Teil der Untersuchung dar. Hierfür gilt es zunächst, spezifische Fragestellungen festzulegen. Dies erfolgt im vorliegenden Kapitel. Zunächst werden die Fragestellungen hergeleitet (Kapitel 6.1). Nachfolgend werden die Fragestellungen formuliert (Kapitel 6.2). Abschließend wird dargelegt und begründet, in welchem pädagogischen Kontext die Untersuchung umgesetzt wird (Kapitel 6.3).

6.1 Herleitung der Fragestellungen

Für die Herleitung und Begründung der Fragestellungen werden im Weiteren zentrale Argumente aus den bisherigen Ausführungen noch einmal zusammenfassend dargestellt. Den Ausgangspunkt für die vorliegende Untersuchung stellen die bei Schüler/-innen wiederholt festgestellten *Kompetenzunterschiede nach sozialer Herkunft* dar. Problematisch sind diese systematischen Kompetenzunterschiede insbesondere aufgrund der weitreichenden Konsequenzen erfolgreicher Bildungsprozesse für die spätere Lebensführung. In diesem Sinne führen ungleiche Bildungschancen zu ungleichen Lebenschancen. Daher kann die Verringerung von herkunftsbedingten Kompetenzunterschieden als ein bildungspolitisches Ziel von hoher gesellschaftlicher Relevanz angesehen werden.

Für eine *evidenzbasierte pädagogische Praxis und Bildungspolitik* werden insbesondere fundierte Erkenntnisse darüber benötigt, auf welche Ursachen Kompetenzunterschiede nach sozialer Herkunft zurückzuführen sind und mit welchen Maßnahmen diese Kompetenzunterschiede wirksam verringert werden können (Neumann et al., 2014, S. 193). Bei beiden Themenbereichen ist der empirische Forschungsstand als ungenügend einzuschätzen. Bezüglich der *Ursachen von herkunftsbedingten Kompetenzunterschieden* kann zwar auf eine große Anzahl an bivariaten Analysen zurückgegriffen werden, die Aussagekraft dieser Studien ist aber aufgrund ihres Designs begrenzt. Gegenwärtig kann nur bei wenigen Sozialisationsfaktoren fundiert eingeschätzt werden, inwieweit sie eine Rolle als vermittelnder Mechanismus zwischen sozialer Herkunft und Lernerfolg spielen (siehe Kapitel 3). Forschungsdefizite bestehen insbesondere

in den Bereichen schulische Beteiligung der Eltern, Nutzung von frühpädagogischen Einrichtungen, Nutzung von Ganztagsangeboten und Nutzung von Nachhilfe.

In Bezug auf *Maßnahmen zur Verringerung von Kompetenzunterschieden nach sozialer Herkunft* stellt die Schule einen naheliegenden Ansatzpunkt dar. Einerseits kommt der Schule aufgrund der gesetzlichen Schulpflicht und der hohen Besuchsdauer eine zentrale Bedeutung für die Bildungsprozesse von Heranwachsenden zu, andererseits kann die Schule als öffentliche Bildungseinrichtung relativ gut politisch gesteuert werden (z. B. im Vergleich zur familiären Lernumwelt). Empirisch fundierte Einschätzungen zu der Frage, ob die Ausgestaltung von Schule und Unterricht einen geeigneten Ansatzpunkt darstellt, ermöglichen *Analysen mit dem Fokus auf pädagogische Einheiten*. Aufschlussreich sind insbesondere Studien zu relativen Effekten von Klassen und Schulen auf herkunftsbedingte Unterschiede im Lernerfolg und zur Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg. Die große Mehrheit der vorhandenen Studien zu pädagogischen Einheiten untersucht allerdings nur den durchschnittlichen Lernerfolg. Die wenigen Studien, die Unterschiede im Lernerfolg untersuchen, ergeben eine unklare Befundlage (siehe Kapitel 4). Eine fundierte Einschätzung des Potentials der Bildungseinrichtung Schule zur Verringerung von herkunftsbedingten Kompetenzunterschieden ist daher gegenwärtig nicht möglich.

Um konkrete Maßnahmen zu ermitteln, mit denen herkunftsbedingte Kompetenzunterschiede verringert werden können, bieten sich *Analysen mit dem Fokus auf pädagogische Merkmale* an. Auch in diesem Forschungsfeld ist festzustellen, dass die meisten Studien nur den durchschnittlichen Lernerfolg als Zielkriterium berücksichtigen. Die Anzahl der verfügbaren Studien zu den Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg ist überschaubar (siehe Kapitel 5). Vor diesem Hintergrund liegt nur bei wenigen pädagogischen Merkmalen ein Mindestmaß an Evidenz für Effekte auf herkunftsbedingte Unterschiede im Lernerfolg vor (und selbst bei diesen Merkmalen ist weitere Evidenz wünschenswert).

Innerhalb der Kategorie der pädagogischen Merkmale kommt schulischen Lernangeboten wie *Unterricht* und *Ganztagsangeboten* eine zentrale Bedeutung zu. Die auf der Klassen- und Schulebene verorteten Prozessmerkmale stellen die unmittelbarsten schulischen Einflussfaktoren auf die Lernaktivitäten der Schüler/-innen dar. Empirisch zeigt sich, dass die Ausgestaltung von schulischen Lernangeboten einen größeren Effekt auf den Lernerfolg hat als äußere Rahmenbedingungen (Wang et al., 1993, S. 276 ff.). Vor diesem Hintergrund ist

anzunehmen, dass bei schulischen Einflussfaktoren auf herkunftsbedingte Unterschiede im Lernerfolg ebenfalls Lernangebote auf der Klassen- und Schulebene eine zentrale Rolle spielen. Ihnen sollte daher in empirischen Wirksamkeitsstudien eine hohe Priorität beigemessen werden. Im Kontrast zu dieser Argumentation bestehen erhebliche Forschungsdefizite bezüglich der Effekte von übergreifenden Unterrichtsprinzipien (z. B. Klassenführung, Klarheit und Strukturiertheit, kognitive Aktivierung) und der Effekte von spezifischen Ganztagsangeboten (z. B. Förderunterricht, Hausaufgabenbetreuungen).

Für die schulischen Lernerfahrungen von Schüler/-innen sind neben den Interaktionen mit pädagogischen Fachkräften auch die Interaktionen mit Mitschüler/-innen von erheblicher Bedeutung. Dies verweist auf den Einfluss der *Lerngruppe*. Während Effekte der Lerngruppe auf den durchschnittlichen Lernerfolg gut untersucht sind (Dumont et al., 2013; Van Ewijk & Slegers, 2010b), ist bisher nicht ausreichend geklärt, inwieweit Effekte der Lerngruppe bei Schüler/-innen verschiedener sozialer Herkunft unterschiedlich ausfallen. Da Merkmale von Lerngruppen bis zu einem gewissen Grad pädagogisch gestaltet werden können, ist diese Frage auch unter dem Gesichtspunkt evidenzbasierter Maßnahmen zur Verringerung von herkunftsbedingten Unterschieden im Lernerfolg relevant. Hierbei besteht insbesondere bezüglich der Rolle der Schülerkomposition weiterer Forschungsbedarf, da die vorhandene Befundlage sehr inkonsistent ausfällt (siehe Kapitel 5).

Dieser Argumentation folgend wird in der vorliegenden Arbeit untersucht, auf welche Ursachen Kompetenzunterschiede nach sozialer Herkunft zurückzuführen sind und welchen Einfluss Unterricht, Ganztagsangebote und Lerngruppen auf Kompetenzunterschiede nach sozialer Herkunft haben. Schule und Unterricht sind allerdings nicht nur auf geringe Unterschiede im Lernerfolg, sondern auch auf einen hohen durchschnittlichen Lernerfolg ausgerichtet. Insofern erscheint es unerlässlich, auch das zweitgenannte Zielkriterium zu berücksichtigen. Insbesondere kann hierdurch überprüft werden, ob sich Strategien identifizieren lassen, mit denen beide Zielsetzungen gleichermaßen erreicht werden können (Helmke, 1988; Schwippert, 2001).

6.2 Formulierung der Fragestellungen

Die Formulierung der Fragestellungen erfolgt unter Berücksichtigung der verwendeten Datengrundlage. Für die vorliegende Untersuchung werden Daten des *Nationalen Bildungspanels* verwendet (Blossfeld & Roßbach, 2019). Aufgrund der spezifischen Eigenschaften dieses Datensatzes waren bei der Formulierung der Fragestellungen zwei Aspekte zu beachten. Erstens können von den

dargestellten theoretischen Kategorien jeweils nur *bestimmte Merkmale* untersucht werden (z. B. wird innerhalb der Kategorie der familiären Lernumwelt nur die elterliche Lernbegleitung untersucht). In den Fragestellungen wird jeweils die theoretische Kategorie benannt und zusätzlich werden in Klammern die spezifisch untersuchten Merkmale konkretisiert. Zweitens kann aufgrund der Stichprobenstruktur nur die *Klassenebene* modelliert werden (und nicht die Schulebene). Daher beziehen sich die Fragestellungen zu den Effekten von pädagogischen Einheiten ausschließlich auf Klasseneffekte.

Die Fragestellungen der vorliegenden Arbeit unterteilen sich in drei Bereiche: vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg, Effekte von Klassen und Effekte von pädagogischen Merkmalen.

Erster Bereich

Der erste Bereich von Fragestellungen bezieht sich auf vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg. Die Leitfrage in diesem Bereich lautet:

Welche Faktoren stellen vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg dar?

Der erste Bereich umfasst sechs Fragestellungen:

- F1.1 Inwieweit stellt die familiäre Lernumwelt (spezifisch die elterliche Lernbegleitung) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg dar?
- F1.2 Inwieweit stellen schulische Lernaktivitäten (spezifisch die Nutzung von Ganztagsangeboten) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg dar?
- F1.3 Inwieweit stellen schulbezogene Einrichtungen (spezifisch die Nutzung von Nachhilfe) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg dar?
- F1.4 Inwieweit stellen Freizeitaktivitäten (spezifisch Lesen, Musizieren und Sport) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg dar?
- F1.5 Inwieweit stellt die Gleichaltrigengruppe (spezifisch die Bildungsaspirationen der Freund/-innen) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg dar?
- F1.6 Inwieweit stellen die Lernvoraussetzungen von Schüler/-innen (spezifisch kognitive Merkmale, motivationale Merkmale und psychi-

sche Dispositionen) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg dar?

Zweiter Bereich

Der zweite Bereich von Fragestellungen bezieht sich auf die Effekte von Klassen. Somit wird die Wirksamkeit von pädagogischen Einheiten betrachtet. Die Leitfrage in diesem Bereich lautet:

Wie groß fallen die Effekte von Klassen auf den durchschnittlichen Lernerfolg bzw. auf herkunftsbedingte Unterschiede im Lernerfolg aus?

Der zweite Bereich umfasst drei Fragestellungen:

- F2.1 Wie groß fallen relative Klasseneffekte auf den durchschnittlichen Lernerfolg aus (insgesamt und getrennt für sozial privilegierte und sozial benachteiligte Schüler/-innen)?
- F2.2 Wie groß fallen relative Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg aus?
- F2.3 In welchem Verhältnis stehen relative Klasseneffekte auf den durchschnittlichen Lernerfolg und relative Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg zueinander?

Dritter Bereich

Der dritte Bereich von Fragestellungen bezieht sich auf die Effekte von pädagogischen Merkmalen. Die Leitfrage in diesem Bereich lautet:

Welche pädagogischen Merkmale (innerhalb von Schulen) haben Effekte auf den durchschnittlichen Lernerfolg bzw. auf herkunftsbedingte Unterschiede im Lernerfolg?

Der dritte Bereich umfasst sechs Fragestellungen:

- F3.1 Welche Unterrichtsmerkmale (spezifisch Unterrichtsumfang, Unterrichtsmethoden und Unterrichtsprinzipien) haben Effekte auf den durchschnittlichen Lernerfolg?
- F3.2 Welche Unterrichtsmerkmale (spezifisch Unterrichtsumfang, Unterrichtsmethoden und Unterrichtsprinzipien) haben Effekte auf herkunftsbedingte Unterschiede im Lernerfolg?

- F3.3 Welche Ganztagsangebote (spezifisch Förderunterricht, Hausaufgabenbetreuungen und fachbezogene Zusatzangebote) haben Effekte auf den durchschnittlichen Lernerfolg?
- F3.4 Welche Ganztagsangebote (spezifisch Förderunterricht, Hausaufgabenbetreuungen und fachbezogene Zusatzangebote) haben Effekte auf herkunftsbedingte Unterschiede im Lernerfolg?
- F3.5 Welche Merkmale der Lerngruppe (spezifisch Klassengröße und Klassenkomposition) haben Effekte auf den durchschnittlichen Lernerfolg?
- F3.6 Welche Merkmale der Lerngruppe (spezifisch Klassengröße und Klassenkomposition) haben Effekte auf herkunftsbedingte Unterschiede im Lernerfolg?

6.3 Auswahl des pädagogischen Kontexts

Im Weiteren wird dargelegt, in welchem pädagogischen Kontext die genannten Fragestellungen untersucht werden. Der untersuchte pädagogische Kontext bezieht sich hierbei auf das Unterrichtsfach, die Schulform und die Jahrgangsstufen. Teilweise verweisen diese Entscheidungen auf das Nationale Bildungspanel als verwendete Datengrundlage, weshalb ausgewählte Aspekte des Datensatzes bereits an dieser Stelle erläutert werden.

Eine zentrale Entscheidung betrifft die Frage, welches *Unterrichtsfach* untersucht werden soll. Für die vorliegende Untersuchung wurde das mathematische Lernen als Inhaltsbereich ausgewählt. Hierfür war ausschlaggebend, dass Mathematikkompetenzen primär im schulischen Kontext erworben werden (anders als Sprachkompetenzen, die maßgeblich durch außerschulische Kontexte beeinflusst werden). Dieser Umstand ist mit zwei bedeutenden Vorteilen verbunden. Erstens besteht eine relativ hohe Deckung zwischen den Inhalten des schulischen Mathematikunterrichts und den Inhalten der in empirischen Studien verwendeten Mathematiktests. Dies stellt eine wichtige Voraussetzung dafür dar, dass bei empirischen Zusammenhängen zwischen Unterrichtsmerkmalen und Mathematikkompetenzen überhaupt Kausaleffekte in Betracht gezogen werden können. Zweitens finden sich mehrfach empirische Hinweise darauf, dass schulische Effekte im mathematischen Bereich am stärksten ausfallen (Brandsma & Knuver, 1989, S. 782; Scheerens & Bosker, 1997, S. 77; Strand, 2016, S. 135). Folglich ist anzunehmen, dass kompensatorische Effekte von Schule und Unterricht am ehesten beim Erwerb von Mathematikkompetenzen festgestellt werden können.

Bezüglich der Frage nach der zu untersuchenden *Schulform* können unter inhaltlichen Gesichtspunkten verschiedene Argumente angebracht werden. Für die Grundschule spricht, dass die während der Grundschulzeit erworbenen Kompetenzen von erheblicher Bedeutung für den Übergang auf die Sekundarschule sind. Demgegenüber bedingen die während der Sekundarschulzeit erworbenen Kompetenzen wesentlich die Aufnahme einer Ausbildung bzw. eines Studiums. Hieran wird deutlich, dass die Zielsetzung der Entkoppelung des Lernerfolg von der sozialen Herkunft in allen Schulstufen relevant ist. Bei der Auswahl der Schulform waren letztlich auch pragmatische Gesichtspunkte zu berücksichtigen. Das Nationale Bildungspanel umfasst zwar Startkohorten sowohl in der Grund- als auch in der Sekundarschule, aber die benötigten Längsschnittdaten waren zum Zeitpunkt der Durchführung der Untersuchung nur für die Sekundarschule verfügbar. Vor diesem Hintergrund befasst sich die vorliegende Untersuchung mit Sekundarschulen.

Um einen realisierbaren Arbeitsumfang zu gewährleisten, beschränkt sich die vorliegende Untersuchung auf eine Schulform der Sekundarschule. Bei der Auswahl der Sekundarschulform war ausschlaggebend, dass die Stichprobe bestimmte Bedingungen erfüllen sollte, die eine methodisch angemessene Untersuchung der Fragestellungen ermöglichen. Relevant erschien insbesondere, dass die Stichprobe eine ausreichende Größe hat, dass die Schülerschaft hinsichtlich der sozialen Herkunft heterogen ist und dass bedeutende Kompetenzunterschiede nach sozialer Herkunft bestehen. Im Rahmen explorativer Analysen wurde überprüft, inwieweit diese Bedingungen in der Stichprobe des Nationalen Bildungspanels bei den einzelnen Sekundarschulformen gegeben sind. Schulen mit mehreren Bildungsgängen und Gesamtschulen wurden aufgrund zu geringer Stichprobenumfänge ausgeschlossen.¹⁶ Bei der Gegenüberstellung von Hauptschulen, Realschulen und Gymnasien zeigte sich, dass sowohl die soziale Heterogenität der Schülerschaft als auch Kompetenzunterschiede nach sozialer Herkunft im Gymnasium am stärksten ausfallen.¹⁷ Die

16 Die Stichprobe umfasst bei Hauptschulen N=91 Klassen, bei Realschulen N=87 Klassen, bei Gymnasien N=159 Klassen, bei Schulen mit mehreren Bildungsgängen N=53 Klassen und bei Gesamtschulen N=35 Klassen. Durch die Anwendung von Einschlusskriterien (siehe unten) wird der Stichprobenumfang noch verringert, so dass der für die Berechnung von Cross-Level-Interaktionen benötigte Stichprobenumfang von $N \geq 50$ (Nezlek et al., 2006, S. 222) bei Schulen mit mehreren Bildungsgängen und Gesamtschulen nicht gegeben ist.

17 Die Standardabweichung des Bildungsstandes der Eltern beträgt in der Hauptschule $\sigma=1.977$, in der Realschule $\sigma=2.097$ und im Gymnasium $\sigma=2.225$. Der standardisierte Regressionskoeffizient des Bildungsstandes der Eltern zur Vorhersage der

ausgeprägte soziale Heterogenität der Schülerschaft von Gymnasien wurde schon in früheren Studien festgestellt (Baumert & Schümer, 2001, S. 371) und verdeutlicht, dass das Gymnasium heutzutage keine reine Eliteschule mehr darstellt. Das wird auch daran deutlich, dass das Gymnasium gegenwärtig die am häufigsten besuchte Sekundarschulform ist. In der repräsentativen Stichprobe des Nationalen Bildungspanels hat die Teilstichprobe der Gymnasien dementsprechend auch den größten Umfang. Eingewendet werden kann, dass im Gymnasium der Mittelwert des sozialen Status am höchsten ausfällt und somit weniger Schüler/-innen mit sehr niedriger sozialer Herkunft berücksichtigt werden können als in der Haupt- und Realschule. Werden die Argumente einander vergleichend gegenübergestellt, überwiegen die Argumente für das Gymnasium (Stichprobengröße, soziale Heterogenität, Kompetenzunterschiede nach sozialer Herkunft vs. Mittelwert des sozialen Status). Aus diesem Grund wurde für die vorliegende Untersuchung das Gymnasium als Schulform ausgewählt.

Als *Untersuchungszeitraum* wurde der Zeitraum zwischen der fünften und der siebten Jahrgangsstufe gewählt. Unter inhaltlichen Gesichtspunkten kann argumentiert werden, dass den ersten Jahren nach dem Wechsel auf das Gymnasium eine besonders hohe Bedeutung für den Schulerfolg im Gymnasium zukommt. Beispielsweise zeigen Daten aus Bayern, dass die Anzahl der Schüler/-innen, die das Gymnasium abbrechen und auf die Realschule wechseln, in der sechsten und siebten Jahrgangsstufe besonders hoch ist und dann kontinuierlich abnimmt (Bayerisches Landesamt für Schule, 2018, S. 60). Unter methodischen Gesichtspunkten ermöglicht der gewählte Untersuchungszeitraum die längsschnittliche Betrachtung von Kompetenzentwicklungen über zwei Messzeitpunkte hinweg. Besonders vorteilhaft ist hierbei, dass die erste Erfassung der Schülerkompetenzen bereits am Anfang der fünften Jahrgangsstufe erfolgt und somit kurz nach dem Eintritt der Schüler/-innen in das Gymnasium. Hierdurch können die Ausgangsbedingungen der Schüler/-innen und die pädagogischen Effekte durch den Schulbesuch relativ gut voneinander getrennt werden (Sammons, 1999, S. 94). Vor diesem Hintergrund erscheint der gewählte Untersuchungszeitraum sowohl inhaltlich relevant als auch methodisch geeignet.

Abschließend ist anzumerken, dass die Auswahl des genannten pädagogischen Kontexts und damit verbunden die Verwendung des Nationalen Bildungspanels als Datengrundlage unweigerlich mit gewissen *Limitationen*

Mathematikkompetenz in der siebten Jahrgangsstufe beträgt in der Hauptschule $\beta=0.114$ ($p < 0.10$), in der Realschule $\beta=0.157$ ($p < 0.01$) und im Gymnasium $\beta=0.178$ ($p < 0.01$).

verbunden ist. Diese Limitationen sollen an dieser Stelle bereits überblicksartig benannt werden, damit eine realistische Vorstellung davon entsteht, was durch die vorliegende Untersuchung geleistet werden kann. Eine zentrale Limitation des gewählten korrelativen Forschungsdesigns besteht darin, dass bei der Interpretation von Kausaleffekten nur ein moderates Maß an interner Validität erreicht werden kann. Demnach kann die vorliegende Untersuchung Kausaleffekte nicht eindeutig beweisen, sondern lediglich mehr oder weniger fundierte Hinweise auf Kausaleffekte liefern. Eine weitere Limitation ist darin zu sehen, dass im Nationalen Bildungspanel die Erfassung von Unterricht und Ganztagsangeboten ausschließlich durch Befragungen von Lehrkräften und Schulleiter/-innen erfolgt (d. h., Befragungen von Schüler/-innen und Beobachtungen fehlen). Somit wird in der vorliegenden Arbeit nur eine einseitige und enggeführte Erfassung von Unterricht und Ganztagsangeboten geleistet. Die Auswahl der Schulform Gymnasium ist schließlich mit der Limitation verbunden, dass die Befunde der vorliegenden Untersuchung nicht ohne Weiteres auf andere Kontexte übertragen werden können. Die genannten Limitationen werden in der abschließenden Diskussion (Kapitel 9) noch einmal aufgegriffen und ausführlich erörtert.

7. Forschungsdesign und Forschungsmethoden

Nachdem im letzten Kapitel die Fragestellungen der empirischen Untersuchung formuliert wurden, wird im vorliegenden Kapitel aufgezeigt, mit welchem Forschungsdesign und mit welchen Forschungsmethoden diese Fragestellungen untersucht werden. Von zentraler Bedeutung ist zunächst einmal die Entscheidung für eine *korrelative Studie* mit multivariatem Längsschnittdesign. Wichtige Argumente für eine korrelative Studie sind insbesondere, dass eine sehr große Anzahl an pädagogischen Merkmalen untersucht werden kann und dass hierbei komplexe Wirkzusammenhänge zwischen pädagogischen Merkmalen analysiert werden können (wie im theoretischen Modell der Wirksamkeit von Schule angenommen). Eine wesentliche Einschränkung korrelativer Studien besteht darin, dass Kausaleffekte nicht eindeutig belegt werden können. Dem kann entgegengehalten werden, dass es angesichts des bisher sehr überschaubaren Forschungsstandes zu schulischen Effekten auf Kompetenzunterschiede zunächst von zentraler Bedeutung ist, mit korrelativen Studien eine breite Wissensgrundlage zu schaffen, d. h., viele pädagogische Merkmale zu überprüfen und dabei gewisse Limitationen in Kauf zu nehmen. Darauf aufbauend kann im Weiteren mit experimentellen Studien der Wissensstand zielgerichtet vertieft werden, d. h., ausgewählte pädagogische Merkmale werden einer fundierten Überprüfung unterzogen (Brophy & Good, 1986, S. 331). Nachfolgend wird die methodische Vorgehensweise bezüglich der Datengrundlage (Kapitel 7.1), der Variablen (Kapitel 7.2) und der Analysen (Kapitel 7.3) erläutert.

7.1 Datengrundlage

Eine angemessene Untersuchung der formulierten Fragestellungen stellt hohe *Anforderungen* an die verwendeten Daten. Benötigt werden Daten zur längsschnittlichen Entwicklung von Schülerkompetenzen, zum sozialen Status von Familien, zu nicht schulischen Einflussfaktoren auf den Lernerfolg und zur Ausgestaltung von Schule und Unterricht. Weiterhin sollte es sich um eine Zufallsstichprobe handeln, damit Repräsentativität gewährleistet ist und Schlussfolgerungen auf die Grundgesamtheit zulässig sind. Schließlich sollte die Stichprobe einen hohen Umfang aufweisen, sodass Interaktionseffekte mit ausreichender Teststärke überprüft werden können. Eine Datengrundlage, die

die genannten Anforderungen erfüllt, stellt das Nationale Bildungspanel dar, welches nachfolgend vorgestellt wird.

7.1.1 Das Nationale Bildungspanel

Beim *Nationalen Bildungspanel* (NEPS, von Englisch „National Educational Panel Study“) handelt es sich um ein großangelegtes Forschungsprojekt zur Untersuchung von Bildungsprozessen in Deutschland (Blossfeld & Roßbach, 2019). Im NEPS erfolgt die längsschnittliche Erfassung von Kompetenzentwicklungen, Lernumwelten, Bildungsentscheidungen und -übergängen, Bildungsungleichheiten sowie Bildungsrenditen über die gesamte Lebensspanne hinweg. Um dies zu leisten, verwendet das NEPS ein sogenanntes Multi-Kohorten-Sequenz-Design (Blossfeld et al., 2011, S. 13 f.). Das bedeutet, das NEPS umfasst sechs Startkohorten in unterschiedlichen Altersstufen (Neugeborene, Kindergarten, fünfte Jahrgangsstufe, neunte Jahrgangsstufe, Studierende, Erwachsene). Es handelt sich um Zufallsstichproben, welche repräsentativ für Deutschland sind. Die längsschnittlichen Erhebungen erfolgen bei allen Stichproben über mehrere Jahre, wodurch Bildungsverläufe über längere Zeiträume hinweg nachgezeichnet werden können.

Das NEPS wird von einem deutschlandweiten *Netzwerk* von Wissenschaftler/-innen und Forschungseinrichtungen durchgeführt (Blossfeld et al., 2011, S. 8 f.). Das Netzwerk ist interdisziplinär ausgerichtet und umfasst vorrangig die Disziplinen Erziehungswissenschaft, pädagogische Psychologie, Entwicklungspsychologie, Bildungssoziologie, Bildungsökonomie und Demographie. Koordiniert wird das NEPS vom Leibniz-Institut für Bildungsverläufe in Bamberg. Weitere beteiligte Forschungseinrichtungen sind unter anderem das Deutsche Institut für Internationale Pädagogische Forschung, das Max-Planck-Institut für Bildungsforschung, das Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik, das Deutsche Jugendinstitut und das Institut für Schulentwicklungsforschung.

Die vom NEPS bereitgestellten Daten sollen die Bearbeitung einer Reihe von *Fragestellungen* ermöglichen (Blossfeld et al., 2011, S. 12 f.). Zentrale Themenbereiche sind Einflüsse von Lernumwelten auf Kompetenzentwicklung und Bildungsbeteiligung, systematische Unterschiede bei Kompetenzentwicklung und Bildungsbeteiligung zwischen verschiedenen Personengruppen (z. B. nach sozialer Herkunft, Migrationshintergrund und Geschlecht) sowie die Bedeutung erworbener Kompetenzen für den Bildungs- und Berufserfolg.

7.1.2 Beschreibung der Stichprobe

Wie im vorausgehenden Kapitel bereits dargelegt wurde, befasst sich die vorliegende Untersuchung mit Sekundarschulen (spezifisch mit Gymnasien) im Zeitraum von der fünften bis zu siebten Jahrgangsstufe. Als Datengrundlage wird die Startkohorte 3 des NEPS verwendet (NEPS-Netzwerk, 2021). Bei der Startkohorte 3 des NEPS handelt es sich um eine *geschichtete Klumpenstichprobe* (Aßmann et al., 2011, S. 60f.). Als statistische Schichten wurden die verschiedenen Schulformen verwendet. Innerhalb jeder Schulform wurde eine mehrstufige Klumpenstichprobe gezogen. In einem ersten Schritt wurden zufällig Schulen gezogen (mit einer Ziehungswahrscheinlichkeit proportional zur Schulgröße). In einem zweiten Schritt wurden innerhalb der gezogenen Schulen zufällig zwei fünfte Klassen gezogen. Alle Schüler/-innen der gezogenen fünften Klassen wurden bezüglich der Teilnahme am NEPS angefragt.

Die Gymnasial-Stichprobe der Startkohorte 3 umfasst insgesamt 2415 Schüler/-innen in 159 Klassen in 78 Schulen. In der Gymnasial-Stichprobe sind Schulen aus 14 Bundesländern enthalten. Nicht enthalten sind Schulen aus Berlin und Mecklenburg-Vorpommern. Im Falle von Berlin ist dies vermutlich darauf zurückzuführen, dass der überwiegende Teil der Berliner Gymnasien erst mit der siebten Jahrgangsstufe beginnt. Im Falle von Mecklenburg-Vorpommern ist vermutlich die relativ geringe Einwohnerzahl und damit verbunden die relativ geringe Ziehungswahrscheinlichkeit ausschlaggebend.

Für die empirischen Analysen wurde eine Teilstichprobe ausgewählt. Diese Analysestichprobe ergibt sich aus der Anwendung bestimmter *Einschlusskriterien*. Diese Einschlusskriterien zielen darauf ab, die Plausibilität von Kausalinterpretationen von Zusammenhängen zwischen pädagogischen Merkmalen und Lernwirkungen zu erhöhen und den Einfluss von Störfaktoren zu verringern. Auf der Schülerebene werden nur Schüler/-innen eingeschlossen, die in der fünften und sechsten Jahrgangsstufe die gleiche Klasse besuchen.¹⁸ Aus methodischen Gründen können außerdem nur Schüler/-innen eingeschlossen werden, bei denen für die Variablen Bildungsstand und Berufsposition der Eltern gültige Werte vorliegen (siehe Kapitel 7.3.1 „Überblick über die verwendeten Analysemethoden“). Auf der Klassenebene werden nur Klassen eingeschlossen, in denen die gleiche Mathematiklehrkraft in der fünften und

18 Das heißt, es wurden folgende Schüler/-innen ausgeschlossen: Schüler/-innen, die die fünfte Jahrgangsstufe wiederholt haben; Schüler/-innen, die die Klasse oder Schule gewechselt haben; Schüler/-innen, die aufgrund einer fehlenden Klassen-ID in der sechsten Jahrgangsstufe keiner Klasse zugeordnet werden konnten.

sechsten Jahrgangsstufe unterrichtet, sowie Klassen, in denen mindestens fünf Schüler/-innen am NEPS teilnehmen.

Die *Analysestichprobe* umfasst insgesamt 1523 Schüler/-innen in 120 Klassen in 71 Schulen. Im Durchschnitt nehmen innerhalb von Klassen 16.7 Schüler/-innen am NEPS teil. Der durchschnittliche Anteil der am NEPS teilnehmenden Schüler/-innen pro Klasse liegt bei 65%. Das Minimum liegt bei 5 Schüler/-innen pro Klassen (bzw. Anteil von 25%), das Maximum bei 24 Schüler/-innen pro Klasse (bzw. Anteil von 100%).

Durch die Einschlusskriterien wird die Stichprobengröße deutlich verringert, erscheint für die geplanten Analysen aber immer noch als ausreichend (Nezlek et al., 2006, S. 222). Es stellt sich allerdings die Frage, ob sich durch den Ausschluss von Schüler/-innen und Klassen *systematische Unterschiede* zwischen der Analysestichprobe und der Gesamtstichprobe der Gymnasias/-innen ergeben. Hierfür wurden die deskriptiven Statistiken zentraler Schülervariablen in den beiden Stichproben miteinander verglichen. Tabelle 1 zeigt die Mittelwerte und Standardabweichungen der Variablen Mathematikkompetenz in der siebten Jahrgangsstufe, Mathematikkompetenz in der fünften Jahrgangsstufe, Geschlecht, Bildungsstand, Berufsposition und Migrationshintergrund sowie die standardisierten Mittelwertsdifferenzen (d bzw. h, je nach Skalenniveau der Variable).

Tab. 1: Vergleich der Analysestichprobe und der Gesamtstichprobe

Variable	Analysestichprobe (N=1523 Schüler/-innen)	Gesamtstichprobe (N=2415 Schüler/-innen)	Mittelwertsdifferenz
Mathematikkompetenz_7 \bar{x}	1.52	1.44	d=0.08
Mathematikkompetenz_7 σ	1.01	1.04	
Mathematikkompetenz_5 \bar{x}	0.74	0.64	d=0.10
Mathematikkompetenz_5 σ	0.96	0.99	
Geschlecht \bar{x}	0.50	0.49	h=0.02
Bildungsstand \bar{x}	15.48	15.37	d=0.05
Bildungsstand σ	2.18	2.23	
Berufsposition \bar{x}	63.96	63.10	d=0.05
Berufsposition σ	17.85	18.26	
Migrationshintergrund \bar{x}	0.15	0.18	h=-0.07

Anmerkungen: \bar{x} = Mittelwert. σ = Standardabweichung.

Beim Vergleich der Analysestichprobe und der Gesamtstichprobe sind keine bedeutsamen Unterschiede festzustellen (alle d bzw. h \leq 0.1). Die Verteilung der Schüler/-innen auf die einzelnen Bundesländer fällt in beiden Stichproben

ebenfalls sehr ähnlich aus. Es finden sich also keine Hinweise darauf, dass sich die Analysestichprobe systematisch von der Gesamtstichprobe unterscheidet. Vor diesem Hintergrund kann angenommen werden, dass die Analysestichprobe (wie die Gesamtstichprobe) repräsentativ für Deutschland ist.

7.2 Variablen

Im Weiteren werden die Variablen vorgestellt, welche in der vorliegenden Untersuchung verwendet werden. Es werden Variablen verwendet, die an drei verschiedenen Messzeitpunkten erhoben wurden (Leibniz Institut für Bildungsverläufe, 2019, S. 3 ff.). Der erste Messzeitpunkt erfolgte Ende 2010/Anfang 2011 (fünfte Jahrgangsstufe), der zweite Messzeitpunkt Ende 2011/Anfang 2012 (sechste Jahrgangsstufe) und der dritte Messzeitpunkt Ende 2012/Anfang 2013 (siebte Jahrgangsstufe). An allen Messzeitpunkten wurden als Erhebungsmethoden Kompetenztests und Fragebögen für Schüler/-innen, Fragebögen für Lehrkräfte, Fragebögen für Schulleiter/-innen und Telefoninterviews mit Eltern eingesetzt. Nachfolgend wird zunächst dargestellt, wie die Auswahl der Variablen erfolgte. Anschließend werden die Variablen beschrieben.

7.2.1 Auswahl der Untersuchungsvariablen

Die Auswahl von Variablen für die vorliegende Untersuchung erfolgt anhand der zwei *theoretischen Modelle* (Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg sowie Modell der Wirksamkeit von Schule). Auf dieser Grundlage werden auf der Schülerebene die Kategorien Schüler/-in, Familie, schulische Lernaktivitäten, schulbezogene Einrichtungen, Freizeitaktivitäten und Gleichaltrigengruppe unterschieden und auf der Klassen- und Schulebene die Kategorien Unterricht, Ganztagsangebote und Lerngruppe. Anhand dieser Kategorien wurde der Datensatz der Startkohorte 3 durchgesehen und inhaltlich passende Variablen wurden identifiziert.

Die ausgewählten Variablen sollen mehreren *Anforderungen* genügen. Einerseits sollen ausreichend Variablen berücksichtigt werden, um die theoretischen Modelle empirisch abbilden zu können. Andererseits soll die Anzahl der Variablen ein gewisses Maß nicht überschreiten, so dass der Arbeitsaufwand für die statistischen Analysen bewältigbar bleibt. Die *Variablen auf der Schülerebene* sollen zwei Funktionen erfüllen: Sie sollen Analysen zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg ermöglichen und sie sollen geeignete Kontrollvariablen für die Ermittlung von schulischen Effekten darstellen (Mehrwertmodell). Bei den Variablen auf der Klassen- und Schul-

ebene wurden Variablen zu schulischen Lernangeboten priorisiert (Unterricht, Ganztagsangebote). Daneben wurden auch ausgewählte Merkmale von Lerngruppen berücksichtigt (Klassengröße, Klassenkomposition). Insgesamt werden in der Untersuchung 27 Variablen auf der Schüler Ebene und 22 Variablen auf der Klassen- und Schulebene verwendet.

Die *Operationalisierung der sozialen Herkunft* der Schüler/-innen verdient angesichts des hohen Stellenwerts für die vorliegende Arbeit eine tiefergehende Betrachtung. Für die Erfassung der sozialen Herkunft stehen in der Startkohorte 3 die drei klassischen Indikatoren Bildungsstand, Berufsposition und Einkommen der Eltern zur Verfügung. Für die Operationalisierung der sozialen Herkunft können die Indikatoren entweder einzeln betrachtet werden oder es kann ein Index aus den drei Indikatoren gebildet werden (Ehmke & Siegle, 2005). Um die verschiedenen Möglichkeiten empirisch zu evaluieren, werden explorative Analysen zu den Effekten auf die Mathematikkompetenz durchgeführt. Es zeigt sich, dass der Effekt des Bildungsstandes ($r=0.165$) und der Effekt der Berufsposition ($r=0.178$) deutlich größer ausfallen als der Effekt des Einkommens ($r=0.085$). Ein aus den drei Indikatoren gebildeter Index weist eine vergleichbare Korrelation zur Mathematikkompetenz auf ($r=0.169$) wie Bildungsstand und Berufsposition. Der Index bietet also keine Vorteile für die Erfassung von Herkunftseffekten auf den Lernerfolg. Vor diesem Hintergrund wird kein Index gebildet. Stattdessen wird die soziale Herkunft anhand des Bildungsstandes und der Berufsposition operationalisiert und die beiden Indikatoren werden getrennt betrachtet. Diese Vorgehensweise hat zwei Vorteile: es werden eindeutige Interpretationen ermöglicht und die Ergebnisse für die zwei Variablen können miteinander verglichen werden. Diese Art der Operationalisierung der sozialen Herkunft wird auch in mehreren anderen Studien verwendet (z. B. Becker et al., 2008; Ditton et al., 2005; Kunter & Voss, 2011).

Die Korrelation zwischen Bildungsstand und Berufsposition beträgt $r=0.671$ und zeigt, dass die beiden Indikatoren ähnliche, aber nicht identische Aspekte der sozialen Herkunft reflektieren. Alle Analysen zur sozialen Herkunft werden einmal mit dem Bildungsstand und einmal mit der Berufsposition durchgeführt. Für eine übersichtliche Darstellung werden nachfolgend nur die Ergebnisse für den Bildungsstand berichtet. Für die Berufsposition werden nur ausgewählte Ergebnisse berichtet, um im Sinne von Robustheitsanalysen zu überprüfen, inwieweit die Ergebnisse invariant gegenüber der Operationalisierung der sozialen Herkunft sind. Dem Bildungsstand wird hierbei der Vorzug gegeben, weil diese Variable für die empirische Bildungsforschung von besonderem Interesse ist.

7.2.2 Variablen auf der Schülerebene

Im Weiteren werden die verwendeten Variablen inhaltlich beschrieben. Zunächst werden die Variablen auf der Schülerebene erläutert (siehe Tabelle 2). Diese Variablen werden durch Kompetenztests, Schülerfragebögen und Telefoninterviews mit Eltern erhoben. Eine ausführliche Darstellung der Erhebungsinstrumente findet sich in Anhang A.

Schüler/-in. Die erste Kategorie bezieht sich auf Merkmale der Schüler/-innen. Diese Kategorie umfasst die Mathematikkompetenzen, die Lernvoraussetzungen und das Geschlecht. Die vorliegende Untersuchung betrachtet den Lernerfolg von Schüler/-innen bezüglich des Erwerbs von Mathematikkompetenzen. Die Erfassung von Mathematikkompetenzen erfolgt im NEPS mithilfe eines Kompetenztests, der verschiedene Inhaltsbereiche (z. B. Raum und Form, Daten und Wahrscheinlichkeiten) und verschiedene Prozesse (z. B. Lösen von mathematischen Problemen, Verstehen von mathematischen Argumentationen) umfasst (Weinert et al., 2011, S. 77 f.). Für die empirischen Analysen werden Testwerte verwendet, die vom NEPS zur Verfügung gestellt werden. Dabei handelt es sich um Weighted Likelihood Estimates (WLEs), welche mithilfe von Item-Response-Theorie-Modellen berechnet wurden (Warm, 1989). Bezogen auf den Untersuchungszeitraum stellt die *Mathematikkompetenz am Anfang der siebten Jahrgangsstufe* (Mathematikkompetenz_7) die Endkompetenz und somit die abhängige Variable dar. Die *Mathematikkompetenz am Anfang der fünften Jahrgangsstufe* (Mathematikkompetenz_5) stellt die Ausgangskompetenz dar und ist somit ein Indikator für das mathematische Vorwissen. Weiterhin berücksichtigt wird die *Mathematiknote* am Ende der vierten Jahrgangsstufe, in der Funktion als Kontrollvariable. Für eine bessere Interpretierbarkeit wird die Mathematiknote invertiert (d. h., hohe Werte stehen für eine gute Note).

Drei weitere Variablen beziehen sich ebenfalls auf den kognitiven Bereich. Die *kognitive Grundfähigkeit* wird mithilfe eines Tests zum schlussfolgernden Denken gemessen. Die zwölf Testitems weisen in der Analysestichprobe erwartungswidrig niedrige Korrelationen auf. Die Bildung eines reliablen Skalenwerts ist daher nicht möglich. Um die kognitive Grundfähigkeit trotzdem zumindest als Kontrollvariable berücksichtigen zu können, wird ein einzelnes Testitem verwendet. Dabei handelt es sich um das Testitem, welches die höchste Korrelation mit der Mathematikkompetenz am Anfang der siebten Jahrgangsstufe aufweist. Es handelt sich um ein dichotomes Item (0 = Aufgabe nicht korrekt gelöst, 1 = Aufgabe korrekt gelöst). Die *metakognitive Kompetenz* bezieht sich auf die Fähigkeit zur Einschätzung des eigenen Lernerfolgs. Für die Erhebung haben die Schüler/-innen am Ende des Mathematiktests der fünften Jahr-

Tab. 2: Deskriptive Statistiken der Variablen auf der Schülerenebene

Variable	Items	Min	Max	\bar{x}	σ	Gültig (%)
<i>Schüler/-in</i>						
Mathematikkompetenz_7	23	-3.12	4.81	1.52	1.01	90
Mathematikkompetenz_5	25	-2.52	4.00	0.74	0.96	97
Mathematiknote	1	1	6	1.83	0.67	93
Kognitive Grundfähigkeit	1	0	1	0.52		96
Metakognitive Kompetenz	1	0	1	0.68		96
Lesekompetenz	33	-3.62	4.07	0.66	1.10	97
Interesse Mathematik	4	1	4	2.19	0.76	95
Selbstkonzept Mathematik	3	1	4	3.04	0.79	95
Gewissenhaftigkeit	2	1	5	3.27	0.86	91
Geschlecht	1	0	1	0.50		100
<i>Familie</i>						
Lernbegleitung Häufigkeit	1	1	5	3.62	0.96	94
Lernbegleitung Unterstützung	4	1	4	3.37	0.60	96
Lernbegleitung Erwartungen	3	1	4	3.47	0.54	96
Lernbegleitung Kontrolle	2	1	4	3.30	0.71	96
Bildungsstand	2	9	18	15.48	2.18	100
Berufsposition	2	14.21	88.96	63.96	17.85	100
Migrationshintergrund	2	0	1	0.15		100
<i>Schulische Lernaktivitäten</i>						
Nutzung Förderunterricht	1	0	1	0.10		91
Nutzung Hausaufgabenbetreuung	1	0	1	0.08		91
Nutzung fachbezogene Angebote	1	0	1	0.05		91
<i>Schulbezogene Einrichtungen</i>						
Nachhilfe Mathematik	2	0	1	0.07		100
<i>Freizeitaktivitäten</i>						
Häufigkeit Lesen	1	1	5	3.21	1.15	93
Häufigkeit Musizieren	1	0	30	11.11	11.01	87
Häufigkeit Sport	1	1	5	4.03	0.95	92
<i>Gleichaltrigengruppe</i>						
Freund/-innen Hauptschulaspiration	1	1	7	1.68	1.22	88
Freund/-innen Realschulaspiration	1	1	7	2.54	1.37	87
Freund/-innen Gymnasialaspiration	1	1	7	5.44	1.32	88

Anmerkungen: Items = Anzahl an Items, auf denen die Variable basiert. Min = Minimum. Max = Maximum. \bar{x} = Mittelwert. σ = Standardabweichung. Gültig (%) = Anteil an gültigen Werten. Alle Angaben beziehen sich auf N = 1523 Schüler/-innen.

gangsstufe eingeschätzt, wie viele Aufgaben des Tests sie richtig gelöst haben. In explorativen Analysen zeigt sich, dass insbesondere eine Überschätzung der eigenen Leistung negativ mit der Mathematikkompetenz in der siebten Jahrgangsstufe korreliert. Daher wird ein dichotomes Item mit den Ausprägungen 0 (niedrige metakognitive Kompetenz, d.h. Überschätzung um mehr als 20%) und 1 (hohe metakognitive Kompetenz, d.h. realistische Einschätzung) gebildet. Als Indikator für die sprachlichen Kompetenzen wird die *Lesekompetenz* benutzt. Sie wird durch einen Test zum Textverständnis erfasst. Als Testwerte werden die vom NEPS zur Verfügung gestellten WLEs verwendet.

Weitere Variablen beziehen sich auf motivationale Merkmale und psychische Dispositionen der Schüler/-innen. Anhand von zwei Skalen schätzen die Schüler/-innen ein, inwieweit sie *Interesse an Mathematik* haben und inwieweit sie über ein positives *mathematikbezogenes Selbstkonzept* verfügen (1 = trifft gar nicht zu, 4 = trifft völlig zu). Ebenfalls mit einer Skala geben die Schüler/-innen an, inwieweit sie sich als *gewissenhaft* einschätzen (1 = trifft gar nicht zu, 5 = trifft völlig zu). Zur Offenheit stehen im NEPS ebenfalls Items zur Verfügung, allerdings war eine Skalenbildung aufgrund einer zu niedrigen Reliabilität nicht möglich. Als letztes Schülermerkmal wird das *Geschlecht* berücksichtigt (0 = Junge, 1 = Mädchen).

Familie. Die zweite Kategorie stellen Merkmale der Familie dar. Die familiäre Lernumwelt wird in der Form der Lernbegleitung durch die Eltern untersucht. Die Messung erfolgt mithilfe des Schülerfragebogens. Ein Einzelitem erfasst als quantitativen Aspekt die *Häufigkeit der elterlichen Lernbegleitung* (1 = nie, 5 = immer). Drei Skalen beziehen sich auf qualitative Aspekte. Konkret beurteilen die Schüler/-innen, inwieweit ihre Eltern sie fachlich und emotional *unterstützen*, inwieweit ihre Eltern klare *Erwartungen* formulieren und inwieweit ihre Eltern die Erledigung ihrer schulischen Arbeiten *kontrollieren* (1 = stimme gar nicht zu, 4 = stimme völlig zu). Weiterhin werden familiäre Strukturmerkmale im Rahmen der Telefoninterviews mit den Eltern erhoben. Der *Bildungsstand* bezieht sich auf den höchsten Bildungsabschluss beider Eltern, erfasst mithilfe der CASMIN-Klassifikation (Lüttinger & König, 1988) und umgerechnet in Bildungsjahr-Äquivalente (9 = Hauptschulabschluss ohne Berufsausbildung, 18 = universitärer Hochschulabschluss). Diese Variable wird vom NEPS zur Verfügung gestellt. Die *Berufsposition* wird mit dem ISEI gemessen (Ganzeboom, 2010) und bezieht sich auf die höchste Berufsposition beider Eltern (14.21 = Reinigungskraft, 88.96 = Richter/in). Diese Variable ist ebenfalls im NEPS-Datensatz enthalten. Ein *Migrationshintergrund* ist gegeben, wenn mindestens ein Elternteil im Ausland geboren ist (0 = kein Migrationshintergrund, 1 = Migrationshintergrund).

Schulische Lernaktivitäten. Als Aspekt der Lernaktivitäten der Schüler/-innen innerhalb der Lernumwelt Schule wird die Nutzung von Ganztagsangeboten betrachtet. Diese wird mithilfe des Schülerfragebogens erhoben. Konkret wird erfasst, ob die Schüler/-innen in der sechsten Jahrgangsstufe an *Förderunterricht*, an der *Hausaufgabenbetreuung* oder an *fachbezogenen Angeboten* (z. B. fachspezifische Erweiterungskurse) teilnehmen. Es handelt sich jeweils um dichotome Items (0 = keine Nutzung, 1 = Nutzung).

Schulbezogene Einrichtungen. Innerhalb der Kategorie der schulbezogenen Einrichtungen wird auf die Nutzung von *Nachhilfe in Mathematik* fokussiert. Im Rahmen von Telefoninterviews werden die Eltern gefragt, ob der Schüler/die Schülerin Nachhilfe in Mathematik erhält (entweder in der fünften oder in der sechsten Jahrgangsstufe). Es handelt sich um ein dichotomes Item (0 = keine Nachhilfe, 1 = Nachhilfe).

Freizeitaktivitäten. Als weiterer Aspekt der Lernaktivitäten der Schüler/-innen werden außerschulische Freizeitaktivitäten betrachtet. Diese werden mithilfe des Schülerfragebogens erfasst. Konkret werden mit drei Einzelitems die Häufigkeiten des *Lesens* (1 = gar nicht außerhalb der Schule, 5 = mehr als 2 Stunden pro Tag), des *Musizierens* (Anzahl Tage pro Monat) und *sportlicher Aktivitäten* (1 = nie, 5 = [fast] täglich) erhoben.

Gleichaltrigengruppe. Die letzte Kategorie auf der Schülerebene sind Merkmale der Gleichaltrigengruppe. Mithilfe des Schülerfragebogens werden die Bildungsaspirationen der Freund/-innen erfasst, d. h., welche Schulabschlüsse die Freund/-innen anstreben. Die Schüler/-innen geben jeweils an, wie viele ihrer Freund/-innen eine *Hauptschulaspiration* (Hauptschulabschluss als Ziel), eine *Realschulaspiration* (Realschulabschluss als Ziel) oder eine *Gymnasialaspiration* (Abitur als Ziel) haben (1 = keine, 7 = alle).

Psychometrische Eigenschaften. Mehrere Variablen in den Kategorien Schüler/-in und Familie werden mithilfe von Skalen erfasst, d. h., für die Messung der Variablen werden jeweils mehrere Items verwendet. Skalen werden insbesondere für die Erhebung von latenten Merkmalen verwendet, also für Merkmale, die nicht direkt gemessen werden können (z. B. Interesse an Mathematik, unterstützende Lernbegleitung). Ob die Berechnung von Skalenwerten zulässig ist, wird mithilfe einer *konfirmatorischen Faktorenanalyse* (Bühner, 2011) überprüft. Modelliert werden die sechs Faktoren Interesse Mathematik, Selbstkonzept Mathematik, Gewissenhaftigkeit, Lernbegleitung Unterstützung, Lernbegleitung Erwartungen und Lernbegleitung Kontrolle. Die Modellpassung kann als gut angesehen werden (CFI = 0.964; RMSEA = 0.035) (Hox, 2010, S. 307). Alle standardisierten Faktorladungen weisen eine akzeptable Größe auf ($0.490 \leq \lambda \leq 0.915$) (Urban & Mayerl, 2014, S. 101). Die Faktorkorrelationen

bewegen sich in einer vertretbaren Größenordnung ($-0.037 \leq r \leq 0.678$) (Kline, 2011, S. 72). Die vollständigen Faktorladungen und Faktorkorrelationen werden in Anhang B berichtet. Insgesamt betrachtet bestätigen die Ergebnisse der konfirmatorischen Faktorenanalyse die angenommenen Messmodelle.

Weiterhin wurden die Reliabilitäten der sechs Skalen auf der Schülerebene überprüft. Hierfür wird die *interne Konsistenz* mithilfe des Reliabilitätskoeffizienten Ω (McDonald, 1999) analysiert. Den für Reliabilitätskoeffizienten typischerweise angesetzten Schwellenwert von $\Omega \geq 0.7$ (Bühner, 2011, S. 80) erreichen die Skalen Interesse Mathematik ($\Omega = 0.780$), Selbstkonzept Mathematik ($\Omega = 0.879$) und Lernbegleitung Unterstützung ($\Omega = 0.714$), während die interne Konsistenz der Skalen Gewissenhaftigkeit ($\Omega = 0.608$), Lernbegleitung Erwartungen ($\Omega = 0.628$) und Lernbegleitung Kontrolle ($\Omega = 0.642$) jeweils knapp darunter liegen. Da die Abweichungen nur gering ausfallen, erscheint die Verwendung der drei letztgenannten Skalen für die Untersuchung trotzdem zulässig.

7.2.3 Variablen auf der Klassen- und Schulebene

Nachfolgend werden die untersuchten Variablen auf der Klassen- und Schulebene beschrieben (siehe Tabelle 3). Die genannten Variablen werden mit Fragebögen für Lehrkräfte und Schulleiter/-innen erhoben. Die verwendeten Erhebungsinstrumente finden sich in Anhang A.

Unterricht. Alle Angaben zum Unterricht entstammen dem Fragebogen für Mathematiklehrkräfte. Der *Unterrichtsumfang* wird anhand der Anzahl der Mathematikstunden pro Woche erfasst. Fünf Variablen beziehen sich auf die im Unterricht verwendeten Sozialformen. Die Häufigkeiten von *Klassenunterricht* und *Einzelarbeit* werden mit Einzelitems gemessen, während die Häufigkeit von *Partner-/Gruppenarbeit* mit einer Skala bestehend aus drei Items zu Partnerarbeit, Gruppenarbeit und Gruppendiskussion erfasst wird (1 = nie, 6 = [fast] in jeder Stunde). Weiterhin werden die Lehrkräfte mit Einzelitems gefragt, inwieweit sie *leistungsheterogene Gruppen* oder *leistungshomogene Gruppen* bilden (1 = trifft gar nicht zu, 5 = trifft völlig zu).

Sieben Variablen beziehen sich auf Unterrichtsprinzipien. Die *Klassenführung* wird durch ein Einzelitem erfasst, welches sich auf den Zeitanteil bezieht, der im Unterricht für den Umgang mit Störungen und Unterbrechungen aufgewendet wird (Prozentangabe). Das Item wurde invertiert, so dass ein hoher Wert eine effiziente Klassenführung anzeigt. Das Ausmaß an *kognitiver Aktivierung* wird mithilfe einer Skala erhoben, in der die Lehrkräfte angeben, inwieweit sie den Schüler/-innen offene und komplexe Aufgaben stellen (1 = trifft

gar nicht zu, 5 = trifft völlig zu). *Übung und Wiederholung* wird mit einem Einzelitem zum entsprechenden Zeitanteil im Unterricht erfasst (Prozentangabe).

Tab. 3: Deskriptive Statistiken der Variablen auf der Klassen- und Schulebene

Variable	Items	Min	Max	\bar{x}	σ	Gültig (%)
<i>Unterricht</i>						
Unterrichtsumfang	1	4	6	4.29	0.56	63
Klassenunterricht	1	2	6	5.63	0.73	63
Einzelarbeit	1	1	6	4.71	1.08	63
Partner-/Gruppenarbeit	3	2	6	4.17	0.95	63
Leistungsheterogene Gruppen	1	1	5	3.28	1.07	63
Leistungshomogene Gruppen	1	1	3	2.11	0.79	63
Klassenführung	1	85	100	94.61	3.35	58
Kognitive Aktivierung	4	1.50	5	3.46	0.66	63
Übung und Wiederholung	1	0	35	14.36	6.99	58
Aufgaben mit Hilfe	1	5	30	15.54	6.71	59
Aufgaben ohne Hilfe	1	5	55	24.58	10.74	60
Besprechung Hausaufgaben	1	0	33	14.81	6.29	60
Individualisierung	3	1	4.67	3.11	0.84	63
<i>Ganztagsangebote</i>						
Förderunterricht Mathematik	1	0	2	0.47	0.68	63
Förderunterricht schwach	1	1	7	5.00	1.64	78
Förderunterricht stark	1	1	7	3.79	1.95	77
Hausaufgabenbetreuung	1	1	7	6.22	1.50	80
Mathematikangebote	1	1	7	3.75	1.87	73
<i>Lerngruppe</i>						
Klassengröße	1	17	33	26.97	3.74	58
Mittelwert Mathematik_5	1	-0.88	1.88	0.72	0.42	100
Mittelwert Bildungsstand	1	12.89	17.59	15.44	0.95	100
Anteil Migrationshintergrund	1	0	0.67	0.15	0.14	100

Anmerkungen: Items = Anzahl an Items, auf denen die Variable basiert. Min = Minimum. Max = Maximum. \bar{x} = Mittelwert. σ = Standardabweichung. Gültig (%) = Anteil an gültigen Werten. Alle Angaben beziehen sich auf N = 120 Klassen.

Drei Items zu Zeitanteilen für *Aufgaben mit Hilfe* der Lehrkraft, *Aufgaben ohne Hilfe* der Lehrkraft und die *Besprechung der Hausaufgaben* (jeweils Prozentangaben) werden als Indikatoren für Unterstützung und Feedback verwendet. Die *Individualisierung* des Unterrichts wird mit einer Skala erfasst. Die Lehrkräfte geben an, inwieweit sie Schüler/-innen verschiedener Leistungsstärke

unterschiedlich anspruchsvolle Aufgaben stellen (1 = trifft gar nicht zu, 5 = trifft völlig zu).

Ganztagsangebote. Die zweite Kategorie stellen Ganztagsangebote dar. Von den Mathematiklehrkräften wird erfragt, wie viele Stunden *Förderunterricht in Mathematik* pro Woche zur Verfügung stehen. Die restlichen Angaben stammen aus dem Fragebogen für Schulleiter/-innen. Die Schulleiter/-innen geben an, ob und wie oft *Förderunterricht für leistungsschwache Schüler/-innen*, *Förderunterricht für leistungsstarke Schüler/-innen*, *Hausaufgabenbetreuungen* und *Mathematikangebote* (z. B. Projekte, AGs) angeboten werden (1 = nein, 7 = 4–5-mal pro Woche).

Lerngruppe. Die letzte Kategorie bezieht sich auf Merkmale der Lerngruppe. Die *Klassengröße* wird durch Befragung der Klassenlehrkraft erhoben. Außerdem werden drei Variablen zur Klassenkomposition untersucht, nämlich das *mittlere mathematische Ausgangsniveau* (Mittelwert *Mathematik_5*), der *mittlere Bildungsstand* und der *Anteil an Schüler/-innen mit Migrationshintergrund*. Die drei Kompositionsmerkmale werden durch formative Aggregation der Variablen auf der Schülerebene berechnet (Lüdtke et al., 2008, S. 205).

Psychometrische Eigenschaften. Drei Variablen auf der Klassenebene werden mit Skalen gemessen. Es wird eine *konfirmatorische Faktorenanalyse* auf der Klassenebene berechnet mit den Faktoren Partner-/Gruppenarbeit, kognitive Aktivierung und Individualisierung. Es ergibt sich eine akzeptable Modellpassung (CFI = 0.919; RMSEA = 0.086). Eine Faktorladung von Partner-/Gruppenarbeit ($\lambda = 0.350$) und eine Faktorladung von Individualisierung ($\lambda = 0.424$) liegen unterhalb des Schwellenwerts von $\lambda \geq 0.5$. Die restlichen Faktorladungen sind angemessen ($0.530 \leq \lambda \leq 0.915$). Die Faktorkorrelationen bewegen sich in einer vertretbaren Größenordnung ($0.542 \leq r \leq 0.720$). Die vollständigen Faktorladungen und Faktorkorrelationen werden in Anhang B berichtet.

Die *interne Konsistenz* fällt bei der Skala zur kognitiven Aktivierung zufriedenstellend aus ($\Omega = 0.829$), während sie bei den anderen beiden Skalen knapp unterhalb des Schwellenwerts von $\Omega \geq 0.7$ liegt (Partner-/Gruppenarbeit $\Omega = 0.616$; Individualisierung $\Omega = 0.682$). Insgesamt betrachtet sind bei den Skalen auf der Klassenebene nicht alle psychometrischen Standards optimal erfüllt. Die Berechnung von Skalenwerten erscheint aber vertretbar.

7.3 Analysen

Im Weiteren wird erläutert, mit welchen Verfahren die beschriebenen Variablen untersucht werden. Zunächst wird ein Überblick über die verwendeten Analysemethoden gegeben. Dann wird auf die spezifischen Analysestrategien für die

einzelnen Fragestellungen eingegangen. Abschließend wird auf die praktische Umsetzung der Analysen und auf Robustheitsanalysen eingegangen.

7.3.1 Überblick über die verwendeten Analysemethoden

Nachfolgend werden die Analysemethoden Mehrebenenmodelle, Pfadmodelle und Strukturgleichungsmodelle überblicksartig dargestellt.

Mehrebenenmodelle. Bei Mehrebenenmodellen handelt es sich um ein Verfahren für die Analyse von hierarchischen Daten. *Hierarchische Daten* liegen vor, wenn die Versuchspersonen inhaltlich zusammenhängende Gruppen bilden (z. B. Schüler/-innen bilden Klassen, Klassen bilden Schulen). Die hier untersuchten Daten weisen eine hierarchische Struktur auf, aus der sich zwei Besonderheiten ergeben. Zum einen sind sich die Schüler/-innen einer Klasse/Schule ähnlicher als Schüler/-innen verschiedener Klassen/Schulen. Somit liegen *keine unabhängigen Beobachtungen* vor, weshalb es bei Signifikanztests im Rahmen regulärer Statistikverfahren zu einer Unterschätzung der Standardfehler kommt (Ditton, 1998, S. 13 ff.; Hox, 2010, S. 4 f.). Zum anderen sind die Variablen auf *verschiedenen Analyseebenen* verortet (z. B. Mathematikkompetenzen liegen auf der Schülerebene, Unterrichtsmerkmale liegen auf der Klassenebene). Um Variablen auf verschiedenen Analyseebenen mit regulären Statistikverfahren auswerten zu können, müssen sie auf eine gemeinsame Ebene gebracht werden. Sowohl die Aggregation von Variablen auf niedrigeren Analyseebenen als auch die Disaggregation von Variablen auf höheren Analyseebenen ist mit bedeutenden Nachteilen verbunden (Ditton, 1998, S. 31 f.; Hox, 2010, S. 2 f.). Aufgrund beider Besonderheiten sind reguläre Statistikverfahren für die Analyse hierarchischer Daten ungeeignet.

Die genannten Probleme haben zur Entwicklung sogenannter Mehrebenenmodelle geführt (Synonym: Mehrebenenanalyse) (Ditton, 1998; Hox, 2010; Raudenbush & Bryk, 2002). Mehrebenenmodelle gehören zur Familie der Regressionsmodelle. Zentrale Unterschiede zu regulären Regressionsmodellen bestehen darin, dass Regressionskoeffizienten zwischen Klassen/Schulen variieren können und dass bei der Schätzung der Standardfehler die Abhängigkeit der Beobachtungen berücksichtigt wird. Die *Vorteile von Mehrebenenmodellen* sind insbesondere die korrekte Schätzung von Standardfehlern, die korrekte Schätzung von Regressionskoeffizienten (durch Zerlegung in Individualeffekt und Gruppeneffekt) sowie die Ermöglichung von Analysen von Zusammenhängen zwischen Variablen auf verschiedenen Ebenen.

In der vorliegenden Untersuchung werden Mehrebenenmodelle für die Bearbeitung aller drei Bereiche von Fragestellungen genutzt. Spezifisch wer-

den *Zwei-Ebenen-Modelle* verwendet (Schülerebene und Klassenebene). Die Klassenebene und die Schulebene gleichzeitig zu modellieren ist aufgrund der Struktur der Stichprobe nicht möglich, da die Anzahl an Klassen pro Schule zu niedrig ausfällt (in den meisten Schulen zwei Klassen, in manchen Schulen nur eine Klasse). Bei der Gegenüberstellung von Klassen- und Schulebene erscheint die Klassenebene aussichtsreicher, da die meisten pädagogischen Merkmale auf der Klassenebene verortet ist. Aus diesem Grund wird der Klassenebene der Vorzug gegeben.

Pfadmodelle. Pfadmodelle stellen eine Analysemethode dar, mit der *komplexe Wirkzusammenhänge* zwischen Variablen modelliert werden können (Bortz & Schuster, 2010, S. 435 ff.; Geiser, 2011, S. 75 ff.). Bei regulären Regressionsmodellen erfolgt eine klare Einteilung der Variablen, d. h., eine abhängige Variable wird durch eine oder mehrere unabhängige Variablen erklärt. Bei Pfadmodellen können dagegen mehrere abhängige Variablen modelliert werden und Variablen können gleichzeitig abhängige und unabhängige Variablen sein. Die Kombination eines Pfadmodells mit einem Mehrebenenmodell wird als *Mehrebenen-Pfadmodell* bezeichnet (Christ & Schlüter, 2012, S. 139 ff.). Für die vorliegende Untersuchung werden Mehrebenen-Pfadmodelle für Analysen zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg verwendet.

Strukturgleichungsmodelle. Weiterhin zum Einsatz kommen Strukturgleichungsmodelle. Diese Verfahren werden für die Analyse *latenter Variablen* verwendet (Kline, 2011; Urban & Mayerl, 2014). Bei latenten Variablen handelt es sich um Variablen, welche nicht direkt gemessen werden können, weshalb für ihre Erhebung typischerweise Skalen eingesetzt werden. Bei einem manifesten Messmodell wird als Skalenwert der Mittelwert aller Items einer Skala verwendet, wodurch Messfehler in den Items vollständig in die Variable einfließen. Ein latentes Messmodell basiert demgegenüber auf einer Faktorenanalyse, wodurch Messfehler explizit modelliert werden und die Variable messfehlerkorrigiert geschätzt wird. Strukturgleichungsmodelle sind eine Kombination aus einem latenten Messmodell und einem Strukturmodell mit Zusammenhängen zwischen der latenten Variable und anderen Variablen (Christ & Schlüter, 2012, S. 48). Aus der Verbindung von Strukturgleichungsmodellen mit Mehrebenenmodellen ergeben sich *Mehrebenen-Strukturgleichungsmodelle* (Christ & Schlüter, 2012, S. 152 ff.). Diese erlauben es, Mess- und Strukturmodelle auf verschiedenen Analyseebenen zu spezifizieren. In der vorliegenden Untersuchung werden Mehrebenen-Strukturgleichungsmodelle für Analysen zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg (latente

Mediatoren) und für Analysen zu Effekten von pädagogischen Merkmalen (latente pädagogische Merkmale) genutzt.

Umgang mit fehlenden Werten. Wie in den rechten Spalten von Tabelle 2 und Tabelle 3 zu erkennen ist, weisen die meisten Variablen bei einem Teil der Versuchspersonen fehlende Werte auf. Bei den Variablen auf der Schülerebene beträgt der Anteil an fehlenden Werten zwischen 0% und 13%, bei den Variablen auf der Klassen- und Schulebene zwischen 20% und 42% (aggregierte Kompositionsmerkmale ausgeschlossen). Die relativ hohen Anteile an fehlenden Werten auf der Klassen- und Schulebene sind darauf zurückzuführen, dass bei der Rekrutierung der NEPS-Stichprobe Lehrkräfte und Schulleiter/-innen eigenständig entscheiden können, ob sie an der Studie teilnehmen wollen oder nicht. Das heißt, es werden auch Erhebungen in Klassen und Schulen durchgeführt, wenn die Lehrkraft oder der/die Schulleiter/-in die Befragung verweigert.

Fehlende Werte werden mithilfe der *Full-Information-Maximum-Likelihood-Methode* (FIML) behandelt (Enders, 2001). Dabei handelt es sich um eine der Methoden für den Umgang mit fehlenden Werten, die in der forschungsmethodischen Literatur empfohlen werden (Lüdtke et al., 2007, S. 116). Bei den Variablen Bildungsstand und Berufsposition müssen fehlende Werte allerdings anders behandelt werden. Diese Variablen werden als Prädiktoren mit Random Slope modelliert (siehe Kapitel 7.3.4 „Analysestrategie für Effekte von pädagogischen Merkmalen“). Bei Variablen mit Random Slope kann die FIML-Methode nicht genutzt werden (in der verwendeten Software Mplus 8.2). Daher müssen Fälle mit fehlenden Werten auf den Variablen Bildungsstand und Berufsposition ausgeschlossen werden.

7.3.2 Analysestrategie für vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg

Im Weiteren wird erläutert, mit welcher Analysestrategie vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg untersucht werden (Fragestellungen F1.1 bis F1.6). Zunächst wird ein analytisches Modell vorgestellt, anschließend wird auf die Analysemethoden eingegangen.

Analytisches Modell. Bei dem nachfolgend dargestellten analytischen Modell handelt es sich um eine vereinfachte Variante des in Kapitel 3 dargestellten theoretischen Modells, welche nur diejenigen Variablengruppen und Wirkungszusammenhänge umfasst, die einer empirischen Überprüfung unterzogen werden. Abbildung 7 zeigt das analytische Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg. Es werden sechs Variablengruppen

als potenzielle Mediatoren analysiert. Dabei handelt es sich um die familiäre Lernumwelt, schulbezogene Lernaktivitäten, die Nutzung von schulbezogenen Einrichtungen, Freizeitaktivitäten, die Gleichaltrigengruppe und die Lernvoraussetzungen der Schüler/-innen. Zusätzlich wird der verbleibende direkte Effekt der sozialen Herkunft auf den Lernerfolg überprüft, welcher nicht durch die genannten Mediatoren erklärt werden kann.

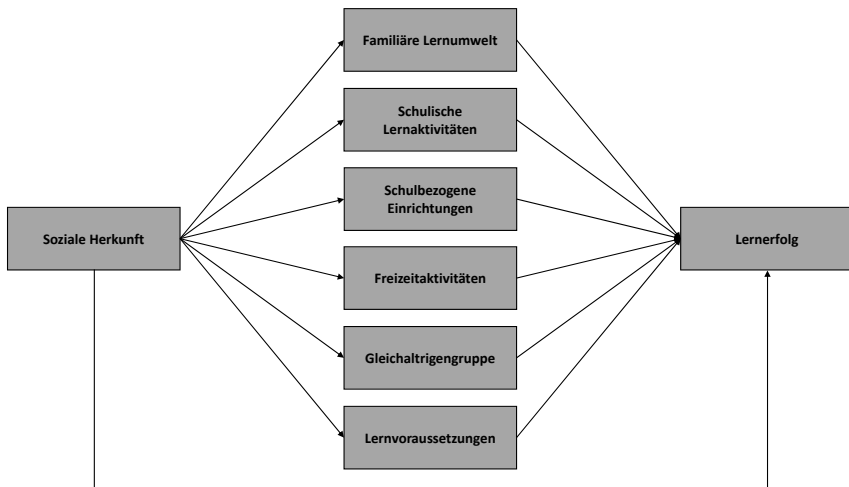


Abb. 7: Analytisches Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Analysemethoden. Die Vorgehensweise besteht in einer *schrittweisen Modellbildung*. Das bedeutet, zunächst werden bivariate Regressionsmodelle berechnet, danach werden einfache Pfadmodelle berechnet (mit jeweils einem Mediator) und abschließend wird ein multivariates Pfadmodell berechnet (mit mehreren Mediatoren). Eine ähnliche Vorgehensweise verwendet Stahl (2009, S. 132 ff.). Der Vorteil einer schrittweisen Modellbildung besteht darin, dass zunehmend strengere Kriterien bei der Überprüfung von Mediationseffekten angewendet werden. Dies ermöglicht eine *differenzierte Einschätzung*, wieviel Evidenz für einen Mediationseffekt vorliegt: Je mehr Modelle die Mediationsannahme bestätigen, desto fundierter ist die vorliegende Evidenz.

Im Rahmen der *bivariaten Regressionsmodelle* wird jeweils der Effekt der sozialen Herkunft auf den potenziellen Mediator (linke Seite des analytischen Modells) und der Effekt des potenziellen Mediators auf die Mathematikkompetenz in der siebten Jahrgangsstufe ermittelt (rechte Seite des analytischen Modells). Für diese Analysen werden lineare und logistische Zwei-Ebenen-Regressionsmodelle verwendet (je nach Skalenniveau des potenziellen Me-

diators). Wenn es aus theoretischen Gründen angezeigt ist, werden primäre (kompetenzabhängige) und sekundäre (kompetenzunabhängige) Effekte der sozialen Herkunft auf potenzielle Mediatoren isoliert, indem die Mathematikkompetenz in der fünften Jahrgangsstufe kontrolliert wird (Schneider, 2005, S. 376 f.). Bei den Effekten von potenziellen Mediatoren auf die Mathematikkompetenz in der siebten Jahrgangsstufe werden auch Interaktionseffekte zwischen sozialer Herkunft und potenziellen Mediatoren überprüft (differenzielle Wirksamkeit), indem Produktterme gebildet werden (Bortz & Schuster, 2010, S. 357).

Auf dieser Grundlage werden im nächsten Schritt *Pfadmodelle* berechnet. Alle Variablen, die sowohl einen Zusammenhang zur sozialen Herkunft als auch einen Zusammenhang zur Mathematikkompetenz in der siebten Jahrgangsstufe aufweisen, werden mithilfe eines Zwei-Ebenen-Pfadmodells analysiert. In diesem Modell ist die soziale Herkunft die unabhängige Variable, der potenzielle Mediator ist die vermittelnde Variable und die Mathematikkompetenz in der siebten Jahrgangsstufe ist die abhängige Variable. Alle Variablen, bei denen das einfache Pfadmodell die Mediationsannahme bestätigt, werden schließlich gemeinsam in einem multivariaten Zwei-Ebenen-Pfadmodell analysiert (d. h., es werden mehrere Mediatoren gleichzeitig modelliert).

Bei allen genannten Modellen handelt es sich um *Zwei-Ebenen-Modelle* (Schülerebene und Klassenebene) mit Random Intercept, damit die Standardfehler korrekt geschätzt werden. Alle Prädiktoren werden *um ihren Klassenmittelwert zentriert* (d. h., bei allen Variablenwerten wird der Klassenmittelwert der Variable abgezogen), so dass die Zusammenhänge den reinen Effekt auf der Schülerebene widerspiegeln (Raudenbush & Bryk, 2002, S. 139). Skalenbasierte Variablen werden mithilfe von Zwei-Ebenen-Strukturgleichungsmodellen *latent modelliert*.

7.3.3 Analysestrategie für Effekte von Klassen

Nachfolgend wird erläutert, mit welchen Analysestrategien der zweite Bereich von Fragestellungen untersucht wird (Fragestellungen F2.1 bis F2.3). Diese Fragestellungen beziehen sich auf *relative Effekte von Klassen*. Es wird also betrachtet, wie stark sich Klassen in ihrer Wirksamkeit unterscheiden. Für die Analysen werden sogenannte Effektivitätswerte verwendet (Ditton, 1998, S. 95 ff.; Kyriakides et al., 2018b, S. 11 f.). Bei *Effektivitätswerten* handelt es sich um statistische Indikatoren für die Wirksamkeit einzelner Klassen (also wie wirksam eine bestimmte Klasse im Vergleich zu anderen Klassen ist). Entsprechend den Fragestellungen wird jeweils ein Effektivitätswert für einen hohen

durchschnittlichen Lernerfolg und ein Effektivitätswert für geringe herkunftsbedingte Unterschiede im Lernerfolg berechnet. Nachfolgend werden zunächst die beiden Effektivitätswerte dargestellt und im Anschluss werden die Analysemethoden erläutert.

Effektivitätswerte für einen hohen durchschnittlichen Lernerfolg. Die Effektivitätswerte für einen hohen durchschnittlichen Lernerfolg werden mithilfe eines *Random-Intercept-Modells* berechnet (Raudenbush & Bryk, 2002, S. 75 ff., 152 ff.). Der *Intercept* (Deutsch: Achsenabschnitt) stellt in einer Regressionsgleichung die Konstante dar. Der Intercept gibt den erwarteten Wert der abhängigen Variable wider, wenn alle unabhängigen Variablen den Wert 0 annehmen. Bei einem Random-Intercept-Modell wird für jede Klasse ein eigener Intercept berechnet.

Aussagekräftige Effektivitätswerte basieren auf einem *Mehrwertmodell* (Sammons, 1999, S. 26 f.; Scheerens & Bosker, 1997, S. 53 ff.). Um den spezifischen Beitrag von Klassen zum Lernerfolg der Schüler/-innen zu erfassen, müssen die *Ausgangsbedingungen* der Schüler/-innen kontrolliert werden. Relevante nicht schulische Einflussfaktoren auf den Lernerfolg sind (entsprechend dem theoretischen Modell zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg) die Lernvoraussetzungen der Schüler/-innen, die Familie, schulbegleitende Einrichtungen, Freizeitaktivitäten und die Gleichaltrigengruppe. Dementsprechend wird zunächst ermittelt, welche Variablen in diesen Kategorien einen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe haben.¹⁹ Da die Variablen auf der Schülerebene in diesem Kontext nur Kontrollvariablen darstellen, werden alle Variablen manifest modelliert, um das Modell einfach zu halten. Zunächst werden bivariate Zusammenhänge analysiert (sowohl lineare als auch kurvilineare Effekte). Alle signifikanten Variablen werden anschließend in einem multivariaten Modell überprüft. Variablen, die sich im multivariaten Modell nicht mehr als signifikant erweisen, werden ausgeschlossen. Als relevante *Prädiktoren auf der Schülerebene* erweisen sich die Mathematikkompetenz in der fünften Jahrgangsstufe, die Mathematiknote, die kognitive Grundfähigkeit, die Lesekompetenz, das mathematikbezogene Selbstkonzept, das Geschlecht, eine kontrollierende Lernbegleitung der Eltern, der Bildungsstand der Eltern, Nachhilfe in Mathematik, die Häufigkeit des Lesens (linearer und kurvilinearereffekt) und der Anteil von Freund/-innen mit Hauptschulaspiration.

19 Das Interesse an Mathematik wird hierbei nicht berücksichtigt, da diese Variable erst in der sechsten Jahrgangsstufe erhoben wird und daher vom Unterricht beeinflusst sein könnte.

Auf der Grundlage der ermittelten nicht schulischen Einflussfaktoren werden mithilfe eines Random-Intercept-Modells klassenspezifische Effektivitätswerte berechnet. Die Gleichungen 1.1 und 1.2 beschreiben das verwendete Zwei-Ebenen-Regressionsmodell (L1 = Schülerebene, L2 = Klassenebene). Die Nomenklatur orientiert sich an Raudenbush und Bryk (2002).

$$\begin{aligned}
 \text{L1: } \quad & \text{MATHEMATIKKOMPETENZ}_{7ij} = \beta_0j + \beta_1 \text{ MATHEMATIKKOMPETENZ5}_{ij} + \\
 & \beta_2 \text{ MATHEMATIKNOTE}_{Eij} + \beta_3 \text{ KOGNITIVEGRUNDFÄHIGKEIT}_{ij} + \beta_4 \text{ LESEKOMPETENZ}_{ij} + \\
 & \beta_5 \text{ SELBSTKONZEPTMATHEMATIK}_{ij} + \beta_6 \text{ GESCHLECHT}_{ij} + \beta_7 \text{ LERNBEGLEITUNGSKONTROLLE}_{ij} + \\
 & \beta_8 \text{ BILDUNGSSTAND}_{ij} + \beta_9 \text{ NACHHILFEMATHEMATIK}_{ij} + \beta_{10} \text{ HÄUFIGKEITLESEN}_{ij} + \\
 & + \beta_{11} \text{ HÄUFIGKEITLESEN}_{ij}^2 + \beta_{12} \text{ FREUNDINNENHAUPTSCHULASPIRATION}_{ij} + r_{ij} \quad (1.1)
 \end{aligned}$$

$$\text{L2: } \quad \beta_{0j} = \gamma_{00} + u_{0j} \quad (1.2)$$

Die Koeffizienten β_1 bis β_{12} spiegeln die Effekte der Variablen auf der Schülerebene auf die Mathematikkompetenz in der siebten Jahrgangsstufe wider. Der Koeffizient β_{0j} bezeichnet die zwischen den Klassen variierenden Intercepts. Das heißt, für jede Klasse j wird ein *eigener Intercept* geschätzt. Um die Interpretation der klassenspezifischen Intercepts zu erleichtern, werden alle Variablen auf der Schülerebene *um ihren Gesamtmittelwert zentriert* (d. h., bei allen Variablenwerten wird der Gesamtmittelwert der Variable abgezogen). Die zentrierten Variablen haben jeweils den Mittelwert 0. Dadurch nimmt der klassenspezifische Intercept den Wert der Mathematikkompetenz an, den man erwarten würde, wenn alle Variablen auf der Schülerebene entsprechend ihrem Mittelwert ausgeprägt sind. Somit kann der klassenspezifische Intercept als *adjustierter Klassenmittelwert der Mathematikkompetenzen in der siebten Jahrgangsstufe* interpretiert werden (Enders & Tofighi, 2007, S. 125; Raudenbush & Bryk, 2002, S. 33).

Wie Gleichung 1.2 auf der Klassenebene zeigt, können die klassenspezifischen Intercepts β_{0j} als Abweichungen vom Stichprobenmittelwert ausgedrückt werden. γ_{00} repräsentiert den Stichprobenmittelwert der adjustierten Klassenmittelwerte (Enders & Tofighi, 2007, S. 125). u_{0j} sind klassenspezifische Intercept-Residuen, d. h., sie stellen die Abweichungen der adjustierten Klassenmittelwerte vom Stichprobenmittelwert dar. Die klassenspezifischen Intercept-Residuen u_{0j} werden als Effektivitätswerte für einen hohen durchschnittlichen Lernerfolg verwendet (Ditton, 1998, S. 97; Kyriakides et al., 2018b, S. 11; Raudenbush & Bryk, 2002, S. 152 ff.). Ein *positiver Wert* von u_{0j} bedeutet, dass der Mittelwert der Mathematikkompetenzen in Klasse j höher ausfällt, als anhand der Ausgangsbedingungen der Schüler/-innen zu erwarten wäre. Diese

Klassen sind *relativ wirksam* bezüglich der Förderung des durchschnittlichen Lernerfolgs. Ein *negativer Wert* von u_{0j} bedeutet, dass der Mittelwert der Mathematikkompetenzen in Klasse j niedriger ausfällt, als anhand der Ausgangsbedingungen der Schüler/-innen zu erwarten wäre. Diese Klassen sind *relativ unwirksam* bezüglich der Förderung des durchschnittlichen Lernerfolgs.

Effektivitätswerte für geringe herkunftsbedingte Unterschiede im Lernerfolg. Für die Berechnung der Effektivitätswerte für geringe herkunftsbedingte Unterschiede im Lernerfolg wird ein *Random-Intercept-Random-Slope-Modell* verwendet (Raudenbush & Bryk, 2002, S. 75 ff., 117 ff.). Ein *Slope* (Deutsch: Steigungskoeffizient) stellt in einer Regressionsgleichung den Koeffizienten einer unabhängigen Variable dar. Der Slope gibt die Richtung und die Stärke des Zusammenhangs zwischen der unabhängigen Variable und der abhängigen Variable wider (Beispiel: Wird der Zusammenhang zwischen elterlichem Bildungsstand und Mathematikkompetenz betrachtet, bedeutet ein Koeffizient von $\beta = 0.2$, dass eine Zunahme des elterlichen Bildungsstandes um 1 Einheit zu einer Zunahme der Mathematikkompetenz um 0.2 Einheiten führt). Bei einem Random-Intercept-Random-Slope-Modell wird für jede Klasse ein eigener Intercept und ein eigener Slope geschätzt.

Grundsätzlich wäre es möglich, beide Effektivitätswerte in einem gemeinsamen Modell zu berechnen. Diese Vorgehensweise erscheint allerdings unter theoretischen Gesichtspunkten als ungünstig, weil für die Ermittlung von Klasseneffekten auf den durchschnittlichen Lernerfolg und auf herkunftsbedingte Unterschiede im Lernerfolg *verschiedene Kontrollvariablen* auf der Schülerebene relevant sind. Für die Ermittlung von Klasseneffekten auf den durchschnittlichen Lernerfolg ist es sinnvoll, möglichst viele Ausgangsbedingungen von Schüler/-innen zu kontrollieren (Mehrwertmodell). Demgegenüber sind bei der Ermittlung von Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg nur bestimmte Kontrollvariablen angezeigt. Insbesondere sollten Variablen, die mögliche vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg darstellen (z. B. familiäre Lernumwelt, Freizeitaktivitäten, Lernvoraussetzungen), nicht kontrolliert werden, da diese Variablen den Koeffizienten der sozialen Herkunft verringern. Dadurch würden Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg unterschätzt werden.

Dieser Argumentation folgend werden Effektivitätswerte für Unterschiede im Lernerfolg mit einem eigenen Mehrebenenmodell berechnet. Es handelt sich um ein Random-Intercept-Random-Slope-Modell. Der Intercept wird als zwischen Klassen variierend modelliert, um das Modell korrekt zu spezifizieren, ist ansonsten aber nicht weiter von Bedeutung. Als *Prädiktoren auf der Schülerebene* werden der Bildungsstand der Eltern (Random Slope), die Mathe-

matikkompetenz in der fünften Klasse, das Geschlecht und der Migrationshintergrund modelliert, wobei alle Variablen *um ihren Gesamtmittelwert zentriert werden*. Die Mathematikkompetenz in der fünften Klasse wird kontrolliert, weil sich die Klassen bereits bei Eintritt der Schüler/-innen in das Gymnasium im Ausmaß an herkunftsbedingten Kompetenzunterschieden unterscheiden (Ausgangsunterschiede). Das bedeutet, in einigen Klassen haben sozial privilegierte Schüler/-innen deutlich bessere Ausgangskompetenzen als sozial benachteiligte Schüler/-innen, in anderen Klassen haben sozial privilegierte Schüler/-innen dagegen nur geringfügig bessere Ausgangskompetenzen als sozial benachteiligte Schüler/-innen.²⁰ Durch die Modellierung der Ausgangskompetenzen werden *Ausgangsunterschiede und Klasseneffekte voneinander getrennt* (Nachbauer & Kyriakides, 2020, S. 321f.). Geschlecht und Migrationshintergrund werden kontrolliert, um Unterschiede im Lernerfolg zu ermitteln, die *genuin auf die soziale Herkunft* zurückzuführen sind (und nicht auf andere äußere Merkmale).²¹ Weiterhin wird der klassenspezifische Mittelwert des Bildungsstandes kontrolliert, damit der Regressionskoeffizient des individuellen Bildungsstandes ausschließlich den *Zusammenhang innerhalb von Klassen* widerspiegelt (Aguinis et al., 2013, S. 1511f.). Gleichungen 2.1 bis 2.3 beschreiben das Random-Intercept-Random-Slope-Modell, welches für Effektivitätswerte für herkunftsbedingte Unterschiede im Lernerfolg verwendet wird.

$$\begin{aligned}
 \text{L1:} \quad & \text{MATHEMATIKKOMPETENZ}_{ij} = \beta_{0j} + \beta_{1j} \text{ BILDUNGSSTAND}_{ij} + \\
 & \beta_2 \text{ MATHEMATIKKOMPETENZ}_{5j} + \beta_3 \text{ GESCHLECHT}_{ij} + \\
 & \beta_4 \text{ MIGRATIONSHINTERGRUND}_{ij} + r_{ij}
 \end{aligned} \tag{2.1}$$

$$\text{L2:} \quad \beta_{0j} = \gamma_{00} + \gamma_{01} \text{ MITTELWERTBILDUNGSSTAND} + u_{0j} \tag{2.2}$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \tag{2.3}$$

20 Eine Überprüfung mit einem Random-Slope-Modell ergibt, dass der Regressionskoeffizient des Bildungsstandes auf die Mathematikkompetenz in der fünften Klasse signifikant zwischen den Klassen variiert ($p < 0.05$).

21 Dies soll an einem Beispiel veranschaulicht werden. Aufgrund der überschaubaren Anzahl an Schüler/-innen pro Klasse ist es möglich, dass in einzelnen Klassen ein Zusammenhang zwischen Geschlecht und sozialer Herkunft besteht (obwohl in der Stichprobe insgesamt kein Zusammenhang besteht). Dies wäre der Fall, wenn in einer Klasse Mädchen häufiger aus statusniedrigen Familien kommen als Jungen. Da das Geschlecht einen Effekt auf die Mathematikkompetenz hat (Mädchen haben schlechtere Mathematikkompetenzen), würden in diesem Fall herkunftsbedingte Kompetenzunterschiede überschätzt werden. Durch die Kontrolle des Geschlechts wird dies verhindert.

Die Koeffizienten β_2 , β_3 und β_4 repräsentieren die Effekte der Variablen auf der Schülerebene (Mathematikkompetenz in der fünften Jahrgangsstufe, Geschlecht, Migrationshintergrund). Der Koeffizient β_{0j} gibt die klassenspezifischen Intercepts an, welche hier nicht weiter von Bedeutung sind. Primär von Interesse ist der Koeffizient β_{1j} , welcher die zwischen den Klassen variierenden Slopes des elterlichen Bildungsstandes beschreibt. Das heißt, für jede Klasse wird ein *eigener Slope* des Effekts des elterlichen Bildungsstandes auf die Mathematikkompetenz in der siebten Jahrgangsstufe geschätzt (unter Kontrolle der Mathematikkompetenz in der fünften Jahrgangsstufe, des Geschlechts und des Migrationshintergrundes). Der klassenspezifische Slope zeigt an, *wie stark der Lernerfolg in der jeweiligen Klasse von der sozialen Herkunft abhängt*.

Auch die klassenspezifischen Slopes können als Abweichung vom Stichprobenmittelwert der klassenspezifischen Slopes (γ_{10}) ausgedrückt werden. Die Abweichungen u_{1j} stellen klassenspezifische Slope-Residuen dar, welche als Effektivitätswerte für geringe herkunftsbedingte Unterschiede im Lernerfolg verwendet werden (Kyriakides et al., 2018b, S. 11f.). Ein *negativer Wert* von u_{1j} bedeutet, dass der Effekt des elterlichen Bildungsstandes auf die Mathematikkompetenz in Klasse j niedriger ausfällt als in anderen Klassen. Diese Klassen sind *relativ wirksam* bezüglich der Geringhaltung von herkunftsbedingten Unterschieden im Lernerfolg. Ein *positiver Wert* von u_{1j} bedeutet, dass der Effekt des elterlichen Bildungsstandes auf die Mathematikkompetenz in Klasse j höher ausfällt als in anderen Klassen. Diese Klassen sind *relativ unwirksam* bezüglich der Geringhaltung von herkunftsbedingten Unterschieden im Lernerfolg.

Analysemethoden. Die dargestellten Effektivitätswerte werden für alle drei Fragestellungen zu den Effekten von Klassen verwendet. Für die Untersuchung von *relativen Klasseneffekten auf den durchschnittlichen Lernerfolg* (Fragestellung F2.1) wird zunächst ein leeres Random-Intercept-Modell berechnet (d. h. ohne Prädiktoren auf der Schülerebene). Die Varianz der klassenspezifischen Intercepts zeigt an, wie stark die *unadjustierten Kompetenzmittelwerte* zwischen den Klassen variieren. Die Varianz der klassenspezifischen Intercepts wird auf Signifikanz getestet und es erfolgt eine Quantifizierung mithilfe der *unadjustierten Intraclass Correlation* (Snijders & Bosker, 2012, S. 18 ff.). Anschließend wird das zuvor dargestellte Random-Intercept-Modell berechnet, in welchem die Ausgangsbedingungen der Schüler/-innen kontrolliert werden (Gleichungen 1.1 und 1.2). Die Varianz der klassenspezifischen Intercepts (also die Varianz der Effektivitätswerte für einen hohen durchschnittlichen Lernerfolg) zeigt an, wie stark die *adjustierten Kompetenzmittelwerte* zwischen den Klassen variieren. Die Varianz der klassenspezifischen Intercepts wird auf Signifikanz

getestet und es erfolgt eine Quantifizierung mithilfe der *adjustierten Intraclass Correlation* (Snijders & Bosker, 2012, S. 52). Durch die Gegenüberstellung von unadjustierten und adjustierten Kompetenzmittelwerten kann überprüft werden, inwieweit Unterschiede im Lernerfolg zwischen den Klassen auf Ausgangsbedingungen der Schüler/-innen zurückzuführen sind (Opdenakker et al., 2002, S. 406 f.).

Um zu überprüfen, ob *relative Klasseneffekte auf den durchschnittlichen Lernerfolg bei Schüler/-innen verschiedener sozialer Herkunft* unterschiedlich ausfallen (Fragestellung F2.1), wird die Stichprobe anhand des elterlichen Bildungsstandes in zwei Gruppen aufgeteilt. Schüler/-innen von Eltern mit mindestens einem Hochschulabschluss (735 Schüler/-innen in 118 Klassen) werden Schüler/-innen von Eltern ohne Hochschulabschluss (788 Schüler/-innen in 120 Klassen) gegenübergestellt. Die Varianz der klassenspezifischen Intercepts (Gleichungen 1.1 und 1.2) wird für beide Gruppen getrennt berechnet und verglichen.

Für die Analyse von *relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg* (Fragestellung F2.2) werden Random-Intercept-Random-Slope-Modelle verwendet, wobei ausschließlich der Random Slope von Interesse ist. Zunächst wird ein einfaches Modell berechnet, in dem nur der elterliche Bildungsstand als Prädiktor der Mathematikkompetenz in der siebten Jahrgangsstufe modelliert wird. Die Varianz der klassenspezifischen Slopes zeigt an, wie stark die *unadjustierten Kompetenzunterschiede nach sozialer Herkunft* zwischen den Klassen variieren. Die Varianz der klassenspezifischen Slopes wird auf Signifikanz getestet (Raudenbush & Bryk, 2002, S. 63 f.). Anschließend wird das zuvor dargestellte Random-Intercept-Random-Slope-Modell berechnet, in welchem die Mathematikkompetenz in der fünften Jahrgangsstufe, das Geschlecht und der Migrationshintergrund kontrolliert werden (Gleichungen 2.1 bis 2.3). Die Varianz der klassenspezifischen Slopes zeigt an, wie stark die *adjustierten Kompetenzunterschiede nach sozialer Herkunft* zwischen den Klassen variieren. Die Varianz der klassenspezifischen Slopes wird auf Signifikanz getestet. Da für die Varianz von Random Slopes keine Effektstärke zur Verfügung steht, wird die Bedeutsamkeit der Unterschiede zwischen Klassen anhand von *inhaltlichen Überlegungen* beurteilt.

Um das *Verhältnis zwischen relativen Klasseneffekten auf den durchschnittlichen Lernerfolg und relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg* zu bestimmen (Fragestellung F2.3), wird das Random-Intercept-Random-Slope-Modell verwendet (Gleichungen 2.1 bis Gleichungen 2.3). In diesem Modell ist es möglich, die Kovarianz zwischen den klassenspezifischen Intercepts und Slopes zu schätzen. Aus der Kovarianz wird mithilfe

der Standardabweichungen der klassenspezifischen Intercepts und Slopes ein Korrelationskoeffizient berechnet (Snijders & Bosker, 2012, S. 79). Der Korrelationskoeffizient ist als standardisierte Effektstärke ein gut interpretierbarer Indikator für das Verhältnis zwischen den beiden Arten von Klasseneffekten.²²

7.3.4 Analysestrategie für Effekte von pädagogischen Merkmalen

Im Weiteren wird erläutert, mit welcher Analysestrategie die Effekte von pädagogischen Merkmalen untersucht werden (Fragestellungen F3.1 bis F3.6). Zunächst wird ein analytisches Modell vorgestellt, anschließend wird auf die Analysemethoden eingegangen.

Analytisches Modell. Bei dem nachfolgend dargestellten analytischen Modell handelt es sich um eine vereinfachte Variante des in Kapitel 5 dargestellten theoretischen Modells der Wirksamkeit von Schule, welche nur diejenigen Variablengruppen und Wirkzusammenhänge umfasst, die einer empirischen Überprüfung unterzogen werden. Abbildung 8 zeigt das analytische Modell zu Effekten von pädagogischen Merkmalen. Der direkte Pfeil auf den Lernerfolg repräsentiert Effekte auf den durchschnittlichen Lernerfolg, der Pfeil auf den Zusammenhang zwischen sozialer Herkunft und Lernerfolg repräsentiert Effekte auf herkunftsbedingte Unterschiede im Lernerfolg. Als Einflussfaktoren auf die zwei Zielkriterien werden die drei Variablengruppen Unterricht, Ganztagsangebote und Lerngruppe analysiert.

Für die empirische Überprüfung der im analytischen Modell dargestellten Wirkzusammenhänge werden verschiedene Zwei-Ebenen-Modelle verwendet. Diese Modelle bauen auf den zuvor dargestellten Effektivitätswerten auf. Im Weiteren werden die Analysemethoden für Effekte auf den durchschnittlichen Lernerfolg und für Effekte auf herkunftsbedingte Unterschiede im Lernerfolg erläutert.

Analysemethoden für Effekte auf den durchschnittlichen Lernerfolg. Die Grundlage für die Analysen zu den Effekten von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg bilden die Effektivitätswerte für einen hohen durchschnittlichen Lernerfolg. Das zuvor dargestellte Modell (Glei-

22 Es kann eingewendet werden, dass im Random-Intercept-Random-Slope-Modell nur ein Teil der Ausgangsbedingungen der Schüler/-innen kontrolliert wird, weshalb die Klasseneffekte auf den durchschnittlichen Lernerfolg verzerrt geschätzt werden. In tiefergehenden Analysen zeigt sich, dass diese Verzerrung nur sehr gering ausfällt: die Korrelation zwischen den klassenspezifischen Intercepts im Random-Intercept-Modell und den klassenspezifischen Intercepts im Random-Intercept-Random-Slope-Modell beträgt $r = 0.912$ ($p < 0.001$).

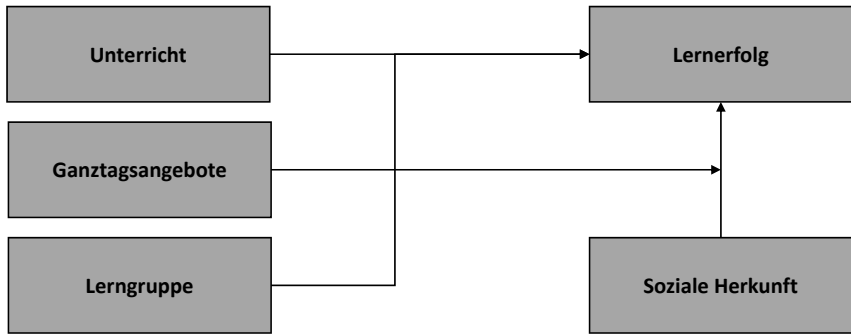


Abb. 8: Analytisches Modell zu Effekten von pädagogischen Merkmalen

chungen 1.1 und 1.2) wird erweitert, indem pädagogische Merkmale als Prädiktoren der klassenspezifischen Intercepts eingeführt werden. Diese Analysestrategie wird auch als *Intercept-As-Outcome-Modell* bezeichnet (Raudenbush & Bryk, 2002, S. 80 ff.). Dabei handelt es sich um integrative Modelle, in denen die Schätzung der Intercepts und die Schätzung der Zusammenhänge zwischen Intercepts und pädagogischen Merkmalen im gleichen Modell erfolgt. Durch integrative Modelle wird, im Vergleich zur Berechnung mehrerer Modelle nacheinander, die Genauigkeit der Schätzung erhöht (Ditton, 1998, S. 35 ff.; Raudenbush & Bryk, 2002, S. 90 ff.) Gleichungen 3.1 und 3.2 veranschaulichen ein Intercept-As-Outcome-Modell am Beispiel des pädagogischen Merkmals Unterrichtsumfang.

$$\begin{aligned}
 \text{L1: } \quad & \text{MATHEMATIKKOMPETENZ}_{ij} = \beta_{0j} + \beta_1 \text{ MATHEMATIKKOMPETENZ}_{ij} + \\
 & \beta_2 \text{ MATHEMATIKNOTE}_{ij} + \beta_3 \text{ KOGNITIVEGRUNDFÄHIGKEIT}_{ij} + \beta_4 \text{ LESEKOMPETENZ}_{ij} + \\
 & \beta_5 \text{ SELBSTKONZEPTMATHEMATIK}_{ij} + \beta_6 \text{ GESCHLECHT}_{ij} + \beta_7 \text{ LERNBEGLEITUNGSKONTROLLE}_{ij} + \\
 & \beta_8 \text{ BILDUNGSSTAND}_{ij} + \beta_9 \text{ NACHHILFEMATHEMATIK}_{ij} + \beta_{10} \text{ HÄUFIGKEITLESEN}_{ij} + \\
 & \beta_{11} \text{ HÄUFIGKEITLESEN}_{ij}^2 + \beta_{12} \text{ FREUNDINNENHAUPTSCHULASPIRATION}_{ij} + r_{ij} \quad (3.1)
 \end{aligned}$$

$$\text{L2: } \quad \beta_{0j} = \gamma_{00} + \gamma_{01} \text{ UNTERRICHTSUMFANG} + u_{0j} \quad (3.2)$$

Auf der Schülerebene ist das Modell identisch mit dem Random-Intercept-Modell (Gleichung 1.1). Demnach werden auf der Schülerebene die nicht schulischen Ausgangsbedingungen der Schüler/-innen kontrolliert. Von primärem Interesse ist der Koeffizient γ_{01} . Dieser beschreibt den Effekt des Unterrichtsumfangs auf die klassenspezifischen Intercepts und somit auf die adjustierten Kompetenzmittelwerte. Ein *positiver Wert* von γ_{01} bedeutet: Je höher der Unter-

richtsumfang in einer Klasse ist, desto höher fällt der durchschnittliche Lernerfolg in der Klasse aus (unter Konstanthaltung der Variablen auf der Schüler-ebene). Ein *negativer Wert* von γ_{01} bedeutet: Je höher der Unterrichtsumfang in einer Klasse ist, desto niedriger fällt der durchschnittliche Lernerfolg in der Klasse aus (unter Konstanthaltung der Variablen auf der Schülerebene).

Auch bei den Analysen zu Effekten von pädagogischen Merkmalen besteht die Vorgehensweise in einer *schrittweisen Modellbildung* (Helmke, 2014, S. 54; Kunter et al., 2006, S. 184 f.). Das heißt, zunächst wird jeweils ein einzelnes pädagogisches Merkmal als Prädiktor der klassenspezifischen Intercepts modelliert und anschließend werden mehrere pädagogische Merkmale als Prädiktoren der klassenspezifischen Intercepts modelliert. Diese Vorgehensweise ermöglicht eine *differenzierte Einschätzung* der vorliegenden Evidenz. Außerdem ist diese Vorgehensweise aus statistischen Gründen angezeigt. In den verwendeten Zwei-Ebenen-Modellen kann nur eine *begrenzte Anzahl von Prädiktoren* auf der Klassenebene geschätzt werden, wobei die Anzahl an schätzbaren Prädiktoren von der Anzahl an Klassen in der Stichprobe abhängt (Raudenbush & Bryk, 2002, S. 267).

Vor diesem Hintergrund wurde eine Vorgehensweise gewählt, die sich an den Empfehlungen von Raudenbush und Bryk (2002, S. 256 ff.) für die Entwicklung von Mehrebenenmodellen orientiert. Zunächst erfolgen Analysen zu pädagogischen Merkmalen *getrennt nach theoretischen Kategorien* (Unterricht, Ganztagsangebote, Lerngruppe). In jeder Kategorie werden relevante Variablen ermittelt, wobei sowohl *lineare* als auch *kurvilineare Effekte* überprüft werden. Auf der Grundlage dieser Analysen wird anschließend ein *integratives Gesamtmodell* gebildet, in das nur diejenigen Variablen aufgenommen werden, für die in den kategorienspezifischen Analysen signifikante Effekte festzustellen sind. Bei allen multivariaten Modellen werden nur diejenigen Prädiktoren im Modell belassen, die sich als signifikant erweisen (unter Kontrolle der restlichen Prädiktoren) (Raudenbush & Bryk, 2002, S. 126). Unter theoretischen Gesichtspunkten entspricht diese Vorgehensweise dem Sparsamkeitsprinzip für wissenschaftliche Erklärungen. Unter statistischen Gesichtspunkten kann hierdurch die Anzahl der zu schätzenden Prädiktoren geringgehalten werden. Weiterhin können statistische Suppressoreffekte vermieden werden (Lenz & Sahn, 2020).

Bei allen Analysen werden Zwei-Ebenen-Modelle berechnet (Schülerebene und Klassenebene). Merkmale auf der Schulebene (Ganztagsangebote) werden hierfür auf die Klassenebene *disaggregiert*. Durch diese Vorgehensweise ist es

möglich, Merkmale auf der Klassen- und Schulebene im gleichen Modell zu untersuchen (die gleichzeitige Modellierung der Klassen- und Schulebene ist aufgrund der Struktur der Stichprobe nicht möglich). Ein Nachteil von Disaggregation besteht darin, dass die Fallzahl künstlich erhöht wird (Hox, 2010, S. 3). Die Erhöhung der Fallzahl fällt in der vorliegenden Analyse allerdings relativ gering aus, da der Unterschied zwischen der Anzahl der Schulen ($N=71$) und der Anzahl der Klassen ($N=120$) nicht sehr groß ist. Insofern ist nicht anzunehmen, dass die Ergebnisse bedeutsam verzerrt werden. Diese Annahme wird durch Robustheitsanalysen bestätigt. Skalenbasierte pädagogische Merkmale werden mithilfe von Zwei-Ebenen-Strukturgleichungsmodellen *latent modelliert*.

Analysemethoden für Effekte auf herkunftsbedingte Unterschiede im Lernerfolg. Die Grundlage für die Analysen zu den Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg bilden die Effektivitätswerte für geringe herkunftsbedingte Unterschiede im Lernerfolg. Das zuvor dargestellte Modell (Gleichungen 2.1 bis 2.3) wird erweitert, indem pädagogische Merkmale als Prädiktoren der klassenspezifischen Intercepts und Slopes eingeführt werden. Diese Analysestrategie wird auch als *Intercept-And-Slope-As-Outcome-Modell* bezeichnet (Raudenbush & Bryk, 2002, S. 80 ff.). Gleichungen 4.1 bis 4.3 veranschaulichen ein Intercept-And-Slope-As-Outcome-Modell am Beispiel des pädagogischen Merkmals Unterrichtsumfang.

$$\begin{aligned}
 \text{L1: } \quad & \text{MATHEMATIKKOMPETENZ}_{ij} = \beta_{0j} + \beta_{1j} \text{ BILDUNGSSTAND}_{ij} + \\
 & \beta_2 \text{ MATHEMATIKKOMPETENZ}_{ij} + \beta_3 \text{ GESCHLECHT}_{ij} + \\
 & \beta_4 \text{ MIGRATIONSHINTERGRUND}_{ij} + r_{ij} \qquad \qquad \qquad (4.1)
 \end{aligned}$$

$$\text{L2: } \quad \beta_{0j} = \gamma_{00} + \gamma_{01} \text{ UNTERRICHTSUMFANG} + \gamma_{02} \text{ MITTELWERTBILDUNGSSTAND} + u_{0j} \qquad (4.2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \text{ UNTERRICHTSUMFANG} + u_{1j} \qquad \qquad \qquad (4.3)$$

Das Modell auf der Schülerebene entspricht dem Random-Slope-Modell (Gleichung 2.1). Der Effekt des Unterrichtsumfangs auf die klassenspezifischen Intercepts (γ_{01}) ist inhaltlich nicht weiter relevant und wird nur modelliert, um Schätzfehler zu vermeiden (Raudenbush & Bryk, 2002, S. 151 f.).²³ Der primär interessierende Koeffizient ist γ_{11} , in dem der Effekt des Unterrichtsumfangs

23 Der Koeffizient γ_{01} im Intercept-And-Slope-As-Outcome-Modell (Gleichung 4.1 bis 4.3) liefert eine ungenauere Schätzung des Effekts von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg als der Koeffizient γ_{01} im Intercept-As-Outcome-Modell (Gleichung 3.1 und 3.2), weil weniger nicht schulische Ausgangsbedingungen von Schüler/-innen kontrolliert werden.

auf die klassenspezifischen Slopes zum Ausdruck kommt. Bei diesem Koeffizienten handelt es sich um eine sogenannte Cross-Level-Interaktion (Aguinis et al., 2013). Bei der Interpretation ist zu beachten, dass geringe Unterschiede im Lernerfolg angestrebt werden, weshalb ein negativer Effekt als günstig anzusehen ist. Ein *negativer Wert* von γ_{11} bedeutet: Je höher der Unterrichtsumfang in einer Klasse ist, desto geringer fallen herkunftsbedingte Unterschiede im Lernerfolg in der Klasse aus. Ein *positiver Wert* von γ_{11} bedeutet: Je höher der Unterrichtsumfang in einer Klasse ist, desto stärker fallen herkunftsbedingte Unterschiede im Lernerfolg in der Klasse aus.

Die bei den Analysen zu Effekten von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg dargestellten Vorgehensweisen der *schrittweisen Modellbildung*, der *Disaggregation* von Merkmalen auf der Schulebene und der *latenten Modellierung* von skalenbasierten Merkmalen werden in gleicher Weise auch bei den Analysen zu Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg angewendet.

Zusätzlich zum Intercept-And-Slope-As-Outcome-Modell werden bei ausgewählten pädagogischen Merkmalen vertiefende Analysen mit *Mehrgruppenmodellen* durchgeführt (Urban & Mayerl, 2014, S. 222 ff.). Für die Mehrgruppenmodelle werden die Schüler/-innen anhand des elterlichen Bildungsstandes in zwei Gruppen eingeteilt (Schüler/-innen von Eltern mit mindestens einem Hochschulabschluss und Schüler/-innen von Eltern ohne Hochschulabschluss). Im Rahmen von Mehrgruppenmodellen wird das Intercept-As-Outcome-Modell (Gleichung 3.1 und 3.2) für beide Schülergruppen getrennt berechnet. Dadurch kann beispielweise der Effekt des Unterrichtsumfangs auf den durchschnittlichen Lernerfolg bei beiden Schülergruppen miteinander verglichen werden. Mithilfe der Mehrgruppenmodelle kann überprüft werden, inwieweit ein Effekt eines pädagogischen Merkmals auf Unterschiede im Lernerfolg entweder auf den Mechanismus der *Förderung sozial benachteiligter Schüler/-innen* oder auf den Mechanismus der *Beeinträchtigung sozial privilegierter Schüler/-innen* zurückzuführen ist.

7.3.5 Umsetzung der Analysen

Im Weiteren wird knapp die praktische Umsetzung der Analysen erläutert. Aus Datenschutzgründen müssen die Analysen innerhalb der Online-Plattform *RemoteNEPS* durchgeführt werden. Als Software wird *Mplus 8.2* (Muthén & Muthén, 1998–2018) verwendet. Im Rahmen der deskriptiven Analysen ist festzustellen, dass einige Variablen nicht normalverteilt sind, daher wird ein *robuster Maximum-Likelihood-Schätzer* verwendet (Syntax: ESTIMATOR=MLR).

Da nur die Schülerebene und die Klassenebene modelliert werden können, werden die Standardfehler außerdem *für die Schulebene korrigiert* (Syntax: TYPE = TWOLEVEL COMPLEX).

In Mplus ergeben sich bestimmte Einschränkungen durch die Modellierung von *zufälligen Effekten* (Syntax: TYPE = RANDOM). Zufällige Effekte werden in der vorliegenden Untersuchung für die Modellierung von Random Slopes und für die Modellierung von kurvilinearen Effekten von latenten Variablen benötigt. Werden zufällige Effekte modelliert, werden keine standardisierten Regressions- bzw. Pfadkoeffizienten zur Verfügung gestellt. Weiterhin werden die typischerweise verwendeten Indizes für die Bewertung der Modellpassung nicht ausgegeben (z. B. CFI, RMSEA).

7.3.6 Robustheitsanalysen

Abschließend wird ein knapper Überblick über die durchgeführten Robustheitsanalysen gegeben. *Robustheitsanalysen* zielen darauf ab, durch den Vergleich verschiedener Modelle zu überprüfen, inwieweit die Ergebnisse von den Modellspezifikationen abhängen. Ein hohes Maß an Robustheit ist gegeben, wenn verschiedene Modellspezifikationen zu vergleichbaren Ergebnissen führen, ein niedriges Maß an Robustheit dagegen, wenn verschiedenen Modellspezifikationen zu unterschiedlichen Ergebnissen führen.

Eine relevante Modellspezifikation ist die *Auswahl der Kontrollvariablen*. Diesbezüglich kann die durchgeführte schrittweise Modellbildung als Robustheitsanalyse angesehen werden, da ermittelt werden kann, ob sich die Effekte von Merkmalen auch unter Kontrolle anderer Merkmale zeigen. Da die untersuchte Stichprobe ganz Deutschland umfasst und somit Schulen in verschiedenen Bundesländern, erscheint es weiterhin sinnvoll, zu überprüfen, ob sich die Ergebnisse verändern, wenn die Zugehörigkeit zu Bundesländern kontrolliert wird. Effekte der Bundeslandzugehörigkeit sind denkbar, weil sich die Bundesländer hinsichtlich bildungspolitischer Rahmenbedingungen unterscheiden (z. B. Lehrpläne).

Eine weitere Modellspezifikation betrifft den *Umgang mit fehlenden Werten*. Angesichts der teilweise recht hohen Anteile an fehlenden Werten wird überprüft, inwieweit sich die Ergebnisse verändern, wenn Fälle mit fehlenden Werten ausgeschlossen werden, anstatt sie mit der FIML-Methode zu behandeln.

Bei den Analysen, bei denen die soziale Herkunft eine Rolle spielt, stellt die *Operationalisierung der sozialen Herkunft* eine wichtige Modellspezifikation dar. Daher wird untersucht, inwieweit sich die Ergebnisse verändern, wenn

anstelle des Bildungsstandes die Berufsposition als Indikator der sozialen Herkunft verwendet wird (empirische Hinweise auf unterschiedliche Ergebnisse finden sich bei Kyriakides et al., 2018a, S. 121).

7.4 Zusammenfassung des siebten Kapitels

Die Bearbeitung der Fragestellungen erfolgt mithilfe einer korrelativen Analyse auf der Grundlage der NEPS Startkohorte 3. Analysiert wird eine Stichprobe von Gymnasien im Zeitraum zwischen der fünften und der siebten Jahrgangsstufe. Die untersuchten Variablen werden mit Kompetenztests, Fragebögen für Schüler/-innen, Lehrkräfte und Schulleiter/-innen sowie Elterninterviews erhoben. Als Analyseebenen werden die Schülerebene und die Klassenebene betrachtet. Vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg werden mit Regressionsmodellen und Pfadmodellen analysiert. Relative Effekte von Klassen werden mithilfe von Effektivitätswerten untersucht, welche mit einem Random-Intercept-Modell und einem Random-Intercept-Random-Slope-Modell berechnet werden. Effekte von pädagogischen Merkmalen werden mit Intercept-As-Outcome-Modellen und Intercept-And-Slope-As-Outcome-Modellen analysiert.

8. Ergebnisse

Im vorliegenden Kapitel werden die Ergebnisse der empirischen Analysen berichtet. Nach einigen einleitenden Bemerkungen (Kapitel 8.1) wird zunächst knapp auf deskriptive Befunde eingegangen (Kapitel 8.2). Anschließend werden die Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg (Kapitel 8.3), zu den Effekten von Klassen (Kapitel 8.4) und zu Effekten von pädagogischen Merkmalen (Kapitel 8.5) präsentiert und diskutiert.

8.1 Einleitende Bemerkungen

Wie im zweiten Kapitel bereits ausgeführt wurde, stellen *Einschätzungen der vorliegenden Evidenz* einen unverzichtbaren Bestandteil des durch Wirkamkeitsstudien zur Verfügung gestellten Orientierungswissens dar. Diese Einschätzungen ermöglichen den Entscheidungsträger/-innen in der pädagogischen Praxis und der Bildungspolitik informierte Entscheidungen bei der Gestaltung von Schule und Unterricht. Die vorliegende Evidenz kann sowohl bei der Betrachtung einer einzelnen Studie als auch bei der Betrachtung mehrerer Studien eingeschätzt werden. Für die nachfolgende Ergebnisdarstellung ist der erstgenannte Fall relevant. Bei allen berichteten Befunden wird im Rahmen der Diskussion auch die vorliegende Evidenz eingeschätzt.

In der Literatur finden sich mehrere *Klassifikationen* zur Unterscheidung verschiedener Stufen von Evidenz (Borman et al., 2003, S. 154 ff.; Center for Data-Driven Reform in Education, 2020; Creemers, 1994, S. 92 ff.). In Anlehnung an diese Klassifikationen werden vier hierarchisch angeordnete *Stufen von Evidenz* unterschieden: keine Evidenz, begrenzte Evidenz, fundierte Evidenz und gesicherte Evidenz. Die Zuordnung von Ergebnissen zu den verschiedenen Stufen von Evidenz erfolgt anhand von empirischen und theoretischen Kriterien. *Empirische Kriterien* umfassen neben statistischen Kennwerten wie Effektstärke und p-Wert auch die Robustheit des Ergebnisses. *Theoretische Kriterien* verweisen auf die Frage, ob für einen Zusammenhang genau eine plausible theoretische Erklärung vorliegt oder nicht. Sowohl das Fehlen einer theoretischen Erklärung als auch das Vorhandensein alternativer theoretischer Erklärungen stellen Einschränkungen für Kausalinterpretationen dar.

Die vier Stufen der Evidenz sind folgendermaßen zu charakterisieren: *Keine Evidenz* liegt vor, wenn die Ergebnisse keinerlei Hinweise auf den an-

genommenen Effekt liefern. *Begrenzte Evidenz* bedeutet, dass die Ergebnisse gewisse Hinweise auf den angenommenen Effekt liefern. Es ist möglich, dass den Ergebnissen ein Kausaleffekt zugrunde liegt. Es besteht allerdings nur ein relativ niedriges Maß an Gewissheit, da die Ergebnisse empirisch nur wenig robust und/oder theoretisch nur bedingt plausibel sind. *Fundierte Evidenz* liegt vor, wenn die Ergebnisse überzeugende Belege für den angenommenen Effekt liefern. Ein Kausaleffekt ist wahrscheinlich. Es besteht ein relativ hohes Maß an Gewissheit, weil sich die Ergebnisse als statistisch robust und theoretisch plausibel erweisen. *Gesicherte Evidenz* liegt vor, wenn die angenommenen Effekte durch die Ergebnisse eindeutig belegt werden. Es besteht Gewissheit über einen Kausaleffekt. Diese Stufe der Evidenz kann nur mit experimentellen Studien erreicht werden (Scheerens & Bosker, 1997, S. 28) und wird daher nachfolgend nicht weiter berücksichtigt. Zusätzlich zu den hierarchisch angeordneten Stufen ist als weitere Kategorie *unklare Evidenz* zu nennen. Diese liegt vor, wenn die Interpretation nicht eindeutig ist, weil die empirischen Ergebnisse verschiedener Modelle widersprüchlich ausfallen, weil starke methodische Limitationen bestehen, weil keine plausible theoretische Erklärung vorliegt oder weil zwei gegensätzliche theoretische Erklärungen gleichermaßen plausibel sind.

Die analytische Strategie der vorliegenden Arbeit zielt auf eine *differenzierte Einschätzung* der vorliegenden Evidenz ab. Daher wird im Weiteren darauf verzichtet, explizit Hypothesen zu formulieren und anschließend anzunehmen oder zurückzuweisen. Die Vorgehensweise einer *binären Entscheidung* (Annahme der Hypothese vs. Zurückweisen der Hypothese) erscheint angemessen, wenn ein Effekt mit einem *einzelnen Modell* überprüft wird. In der vorliegenden Arbeit werden Effekte allerdings jeweils in *mehreren Modellen* überprüft (durch die schrittweise Modellbildung und die vertiefenden Robustheitsanalysen). Diese Strategie erscheint sinnvoll, um bei einer korrelativen Wirksamkeitsstudie aussagekräftige Ergebnisse erzielen zu können. Hierdurch können sich aber auch komplexe Befundlagen ergeben, wenn sich beispielsweise ein Effekt in einem Modell zeigt, in einem anderen Modell dagegen nicht. An die Stelle einer binären Entscheidung tritt daher eine differenziertere Einschätzung anhand der dargestellten Stufen von Evidenz (keine Evidenz, begrenzte Evidenz, fundierte Evidenz, unklare Evidenz).

Für ein klares Verständnis ist auf bestimmte *Begriffsverwendungen* in diesem Kapitel hinzuweisen. Erstens werden bei der Ergebnisdarstellung Begriffe wie „Effekt“, „fördern“ oder „verringern“ im *rein statistischen Sinne* gebraucht (d. h. im Sinne eines statistischen Zusammenhangs). Dies bedeutet nicht zwangsläufig auch, dass inhaltlich ein Kausaleffekt vorliegt. Theoretische Überlegungen

zur Plausibilität einer Kausalinterpretation werden explizit angestellt. Zweitens beziehen sich die Begriffe „unterschiedsverringern“ und „unterschiedsvergrößern“ auf die Gegenüberstellung verschiedener *Merkmalsausprägungen*. Wenn beispielsweise von einem „unterschiedsverringern Effekt“ gesprochen wird, bedeutet das, dass bei hohen Merkmalsausprägungen Kompetenzunterschiede geringer ausfallen als bei niedrigen Merkmalsausprägungen. Nicht gemeint ist, dass Kompetenzunterschiede im Zeitverlauf abnehmen.

Die verwendeten *statistischen Symbole* orientieren sich an der Nomenklatur von Raudenbush und Bryk (2002). Demnach werden Regressionskoeffizienten auf der Schülerebene mit β bezeichnet und Regressionskoeffizienten auf der Klassenebene mit γ . Hervorzuheben ist, dass im Laufe des Kapitels sowohl standardisierte als auch unstandardisierte Regressionskoeffizienten berichtet werden. Da bei den Analysen auf der Schülerebene *durchgängig standardisierte Regressionskoeffizienten* zur Verfügung stehen, werden ausschließlich diese berichtet. Bei den Analysen auf der Klassenebene sind dagegen in den meisten Fällen nur unstandardisierte Regressionskoeffizienten verfügbar, daher werden *durchgängig unstandardisierte Regressionskoeffizienten* berichtet. Um welche Art von Regressionskoeffizient es sich handelt, wird jeweils im Text bzw. in den Überschriften der Tabellen benannt.

8.2 Deskriptive Ergebnisse

Einleitend werden relevante deskriptive Befunde berichtet. Mittelwerte und Standardabweichungen der untersuchten Variablen finden sich in Tabelle 2 und Tabelle 3. Insgesamt sind bei den deskriptiven Statistiken keine Auffälligkeiten festzustellen. Die zentralen Variablen Bildungsstand und Mathematikkompetenzen verdienen allerdings eine tiefergehende Betrachtung.

Beim *Bildungsstand der Eltern* ist eine rechtssteile Verteilung festzustellen. Dies ist darauf zurückzuführen, dass Gymnasiast/-innen betrachtet werden. Abbildung 9 zeigt ein Balkendiagramm des elterlichen Bildungsstandes. Die größte Gruppe bilden mit 34.7% Eltern mit mindestens einem universitären Hochschulabschluss (18 Bildungsjahre). Die zweitgrößte Gruppe bilden mit 27.4% Eltern mit Realschulabschluss und Berufsausbildung oder mit Abitur ohne Berufsausbildung (beide äquivalent zu 13 Bildungsjahren). 20.6% der Eltern haben eine Hochschulreife mit Berufsausbildung (15 Bildungsjahre) und 13.5% Eltern einen Fachhochschulabschluss (16 Bildungsjahre). Bildungsferne Elternhäuser stellen mit 3.8% nur eine kleine Gruppe dar. An der Verteilung des Bildungsstandes wird deutlich, dass sich soziale Unterschiede in der Analysestichprobe vordergründig auf den Unterschied zwischen Eltern mit

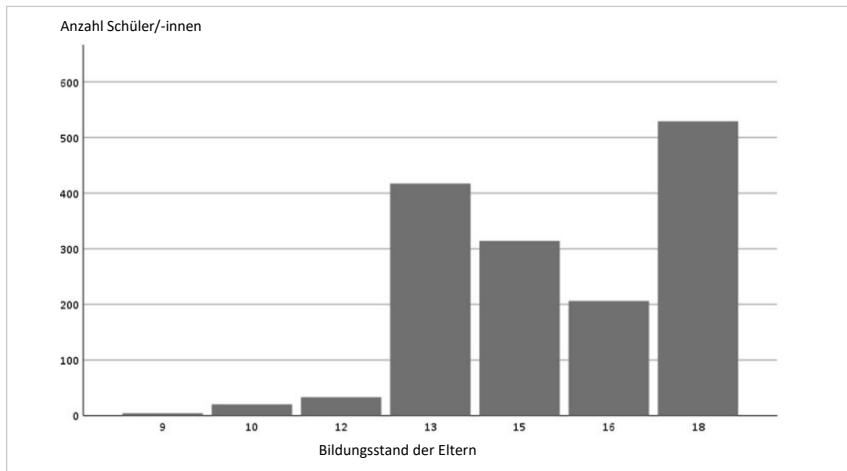


Abb. 9: Balkendiagramm der Variable Bildungsstand der Eltern

hohem Bildungsstand und Eltern mit mittlerem Bildungsstand beziehen. Bei der Betrachtung der Berufsposition zeigt sich eine vergleichbare Verteilung. Wenn im Weiteren von *sozial benachteiligten Schüler/-innen* gesprochen wird, bezieht sich dies also auf eine relative Benachteiligung innerhalb der Gruppe der Gymnasiast/-innen. Es handelt sich um Schüler/-innen, die im Vergleich zu ihren Mitschüler/-innen auf weniger familiäre Ressourcen zurückgreifen können.

Bei den *Mathematikkompetenzen* fallen die deskriptiven Statistiken in der fünften und siebten Jahrgangsstufe jeweils plausibel aus. An beiden Messzeitpunkten sind die Mathematikkompetenzen näherungsweise normalverteilt. Die vom NEPS zur Verfügung gestellten Kompetenzwerte werden mithilfe von Anchoritems einheitlich skaliert und können daher direkt miteinander verglichen werden. Am Anfang der fünften Jahrgangsstufe beträgt der Mittelwert der Mathematikkompetenzen $\bar{x}_{t1} = 0.742$ und am Anfang der siebten Jahrgangsstufe $\bar{x}_{t2} = 1.521$. Hieraus ergibt sich eine standardisierte Mittelwertsdifferenz von $d = 0.790$. Die Mathematikkompetenzen der Schüler/-innen nehmen somit pro Schuljahr um durchschnittlich $d = 0.395$ zu. Der Kompetenzzuwachs pro Schuljahr bewegt sich somit im oberen Bereich des Spektrums der typischerweise berichteten Effektstärken (Ehmke et al., 2006a; Hattie, 2013; Lehmann et al., 1999; Walter et al., 2006).

Weiterhin von Interesse ist die *Veränderung von Kompetenzunterschieden nach sozialer Herkunft* im Untersuchungszeitraum. Hierfür wird der Effekt des elterlichen Bildungsstandes auf die Mathematikkompetenz an beiden Mess-

zeitpunkten betrachtet. In der fünften Jahrgangsstufe beträgt der standardisierte Regressionskoeffizient des elterlichen Bildungsstandes auf die Mathematikkompetenz $\beta_{11} = 0.079$ ($p < 0.01$) und in der siebten Jahrgangsstufe $\beta_{12} = 0.126$ ($p < 0.01$). Kompetenzunterschiede nach sozialer Herkunft nehmen im betrachteten Zeitraum also leicht zu. Sozial privilegierte Schüler/-innen erreichen einen höheren Kompetenzzuwachs als sozial benachteiligte Schüler/-innen. Ein vergleichbares Bild zeigt sich bei Verwendung der Berufsposition als Indikator der sozialen Herkunft. Hinweise auf eine Zunahme von herkunftsbedingten Kompetenzunterschieden im Laufe der Schulzeit finden sich auch in mehreren anderen Studien (Neumann et al., 2014).

8.3 Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Nachfolgend werden die Ergebnisse zum ersten Bereich von Fragestellungen präsentiert, welcher sich auf vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg bezieht (Fragestellungen F1.1 bis F1.6). Zunächst werden die Ergebnisse in den Bereichen elterliche Lernbegleitung, Nutzung von Ganztagsangeboten, Nachhilfe, Freizeitaktivitäten, Gleichaltrigengruppe und Lernvoraussetzungen präsentiert. Mit Regressionsmodellen wird überprüft, inwieweit die untersuchten Variablen die Voraussetzungen für einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg erfüllen (nämlich ein Zusammenhang zur sozialen Herkunft und ein Zusammenhang zur Mathematikkompetenz, wobei die Zusammenhänge die gleiche Richtung aufweisen müssen). Bei Variablen, bei denen dies der Fall ist, werden einfache Pfadmodelle berechnet (d. h. mit jeweils einem Mediator). Abschließend wird ein bereichsübergreifendes, multivariates Pfadmodell berechnet (d. h. mit mehreren Mediatoren).

Die Ergebnisse werden in der Form von standardisierten Regressions- und Pfadkoeffizienten berichtet. Hinsichtlich der *Interpretation von Effektstärken* wird Referenzgrößen mit Bezug zum schulischen Lernen (Hattie, 2013, S. 21 ff.; Scheerens et al., 2007, S. 114 f.; Spreitzer et al., 2022, S. 714) der Vorzug gegenüber allgemeinen Referenzgrößen (Cohen, 1988) gegeben. Folgende Schwellenwerte werden verwendet: $\beta = 0.3$ kennzeichnet einen starken Effekt, $\beta = 0.2$ einen moderaten Effekt, $\beta = 0.1$ einen schwachen Effekt und $\beta = 0.05$ einen sehr schwachen Effekt.

In den Analysen zeigt sich, dass die Effekte der sozialen Herkunft auf potenzielle Mediatoren vergleichbar ausfallen, wenn der Bildungsstand oder die

Berufsposition der Eltern betrachtet wird. Daher werden im Weiteren nur die Ergebnisse zum Bildungsstand berichtet.

8.3.1 Ergebnisse zur Rolle der elterlichen Lernbegleitung

Im Weiteren werden die Ergebnisse zu der Frage präsentiert, inwieweit die elterliche Lernbegleitung (als Aspekt der familiären Lernumwelt) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt (Fragestellung F1.1). Auf der Grundlage vorhandener Befunde erscheint eine Rolle als vermittelnder Mechanismus vor allem bei einer unterstützenden Lernbegleitung plausibel. Bei allen Merkmalen der elterlichen Lernbegleitung ist außerdem anzunehmen, dass sie auch mit der Mathematikkompetenz in der fünften Jahrgangsstufe korrelieren, was sowohl bei der Ermittlung von Herkunftseffekten als auch bei der Ermittlung von lernförderlichen Effekten zu berücksichtigen ist.

Die Ergebnisse der bivariaten Zwei-Ebenen-Regressionsmodelle finden sich in Tabelle 4. Die drei Variablen Unterstützung, Erwartungen und Kontrolle werden jeweils latent modelliert. Die Indizes der Modellgüte fallen in allen Regressionsmodellen mindestens akzeptabel aus ($0.920 \leq CFI \leq 0.998$; $0.059 \leq RMSEA \leq 0.084$).

Der elterliche Bildungsstand hat keinen Effekt auf die *Häufigkeit der Lernbegleitung*. Die Häufigkeit der Lernbegleitung hat einen negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, der aber unter Kontrolle der Ausgangskompetenz verschwindet.

Der elterliche Bildungsstand hat ebenfalls keinen bivariaten Effekt auf *Unterstützung*. Wird die Ausgangskompetenz kontrolliert, zeigt sich allerdings ein sehr schwacher positiver Effekt des Bildungsstandes auf Unterstützung (nicht in der Tabelle, $\beta = 0.069$, $p < 0.10$). Bei gleicher Ausgangskompetenz erfahren sozial privilegierte Schüler/-innen also mehr Unterstützung durch ihre Eltern als sozial benachteiligte Schüler/-innen. Unterstützung hat einen negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, der aber unter Kontrolle der Ausgangskompetenz verschwindet (siehe Tabelle 4).

Weiterhin ist ein sehr schwacher positiver Effekt des elterlichen Bildungsstandes auf *Erwartungen* festzustellen (10%-Niveau), der unter Kontrolle der Ausgangskompetenz noch leicht zunimmt (nicht in der Tabelle, $\beta = 0.063$, $p < 0.05$). Bei gleicher Ausgangskompetenz erfahren sozial privilegierte Schüler/-innen also mehr Erwartungen von ihren Eltern als sozial benachteiligte Schüler/-innen. Erwartungen haben einen negativen Effekt auf die Ma-

thematischer Kompetenz in der siebten Jahrgangsstufe, auch unter Kontrolle der Ausgangskompetenz (dann nur noch auf dem 10%-Niveau).

Tab. 4: Zusammenhänge zwischen elterlichem Bildungsstand, elterlicher Lernbegleitung und Mathematikkompetenz in der siebten Jahrgangsstufe (standardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekt des Bildungsstandes auf die Variable	Effekt der Variable auf die Mathematikkompetenz_7	Effekt der Variable auf die Mathematikkompetenz_7 (unter Kontrolle von Mathematikkompetenz_5)
Lernbegleitung Häufigkeit	-0.016 (0.029)	-0.079* (0.028)	-0.035 (0.025)
Lernbegleitung Unterstützung	0.060 (0.041)	-0.072* (0.033)	-0.028 (0.028)
Lernbegleitung Erwartungen	0.055* (0.030)	-0.089* (0.038)	-0.051* (0.031)
Lernbegleitung Kontrolle	-0.001 (0.001)	-0.148*** (0.033)	-0.103** (0.031)

Anmerkungen: Jede Zelle stellt das Ergebnis eines eigenen Regressionsmodells dar. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Der elterliche Bildungsstand hat keinen Effekt auf *Kontrolle*. Kontrolle hat einen schwachen negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, auch wenn die Ausgangskompetenz konstant gehalten wird.

In tiefergehenden Analysen können lernförderliche Effekte der elterlichen Lernbegleitung auch dann nicht festgestellt werden, wenn *zusätzliche Kontrollvariablen* berücksichtigt werden (Mathematiknote, kognitive Grundfähigkeit, Lesekompetenz, Selbstkonzept Mathematik, Geschlecht).

Diskussion der Ergebnisse zur Rolle der elterlichen Lernbegleitung. Insgesamt betrachtet liegen bei keinem Merkmal der elterlichen Lernbegleitung die Bedingungen für einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg vor. Die Befunde sprechen zwar dafür, dass sozial privilegierte Schüler/-innen mehr *Unterstützung* und mehr *Erwartungen* von ihren Eltern erfahren als sozial benachteiligte Schüler/-innen, aber es können keine lernförderlichen Effekte dieser Merkmale festgestellt werden. Es bleibt unklar, ob Unterstützung und Erwartungen tatsächlich keine förderlichen Effekte auf den Lernerfolg haben, oder ob reziproke Effekte vorliegen (d. h., Schüler/-innen mit ungünstigen Lernvoraussetzungen werden beim Lernen intensiver begleitet), die durch die verwendeten Kontrollvariablen nicht angemessen modelliert werden können. Theoretisch plausibler erscheint die zweitgenannte Interpretation.

8.3.2 Ergebnisse zur Rolle der Nutzung von Ganztagsangeboten

Nachfolgend werden die Ergebnisse zu Fragestellung F1.2 berichtet, also inwieweit die Nutzung von Ganztagsangeboten (als Aspekt der schulischen Lernaktivitäten) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt. Basierend auf früheren Studien ist anzunehmen, dass sozial benachteiligte Schüler/-innen häufiger an Förderunterricht und Hausaufgabenbetreuungen teilnehmen als sozial privilegierte Schüler/-innen. Hierbei ist allerdings unklar, inwieweit sich dieses Bild ändert, wenn die anzunehmenden Effekte der Ausgangskompetenzen der Schüler/-innen auf die Nutzung dieser Ganztagsangebote berücksichtigt werden. Weiterhin in Betracht zu ziehen sind differenzielle Effekte der Nutzung von Ganztagsangeboten.

Tabelle 5 zeigt die Ergebnisse der bivariaten Zwei-Ebenen-Modelle. Der elterliche Bildungsstand hat einen schwachen negativen Effekt auf die Nutzung von *Förderunterricht*. Der Effekt des elterlichen Bildungsstandes verringert sich leicht, wenn die Ausgangskompetenz kontrolliert wird, bleibt aber signifikant (nicht in der Tabelle, $\beta = -0.096$, $p < 0.05$). Demnach nutzen sozial privilegierte Schüler/-innen auch bei gleicher Ausgangskompetenz seltener Förderunterricht als sozial benachteiligte Schüler/-innen. Die Nutzung von Förderunterricht hat einen schwachen negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, der aber verschwindet, wenn die Ausgangskompetenz kontrolliert wird. Weiterhin ist ein Interaktionseffekt zwischen elterlichem Bildungsstand und der Nutzung von Förderunterricht festzustellen (nicht in der Tabelle, $\beta = 0.053$, $p < 0.10$, Ausgangskompetenz kontrolliert). Der Interaktionseffekt ergibt sich daraus, dass die Nutzung von Förderunterricht bei statushohen Schüler/-innen keinen Effekt hat (nicht in der Tabelle, $\beta = 0.037$, $p > 0.10$, Ausgangskompetenz kontrolliert), bei statusniedrigen Schüler/-innen dagegen einen schwachen negativen Effekt (nicht in der Tabelle, $\beta = -0.079$, $p < 0.05$, Ausgangskompetenz kontrolliert). Anzumerken ist, dass der Interaktionseffekt knapp das Signifikanzniveau verfehlt, wenn statt dem Bildungsstand die Berufsposition der Eltern betrachtet wird.

Weiterhin hat der elterliche Bildungsstand einen moderaten negativen Effekt auf die Nutzung von *Hausaufgabenbetreuungen*. Der Effekt des elterlichen Bildungsstandes verändert sich nicht, wenn die Ausgangskompetenz kontrolliert wird (nicht in der Tabelle, $\beta = -0.183$, $p < 0.01$). Demnach nutzen sozial privilegierte Schüler/-innen auch bei gleicher Ausgangskompetenz seltener Hausaufgabenbetreuungen als sozial benachteiligte Schüler/-innen. Die Nutzung von Hausaufgabenbetreuungen hat keinen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe. Ein Interaktionseffekt mit dem elterlichen Bildungsstand ist nicht festzustellen, aber ein Interaktionseffekt mit

Tab. 5: Zusammenhänge zwischen elterlichem Bildungsstand, Nutzung von Ganztagsangeboten und Mathematikkompetenz in der siebten Jahrgangsstufe (standardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekt des Bildungsstandes auf die Variable	Effekt der Variable auf die Mathematikkompetenz_7	Effekt der Variable auf die Mathematikkompetenz_7 (unter Kontrolle von Mathematikkompetenz_5)
Nutzung Förderunterricht	-0.120* (0.047)	-0.099* (0.034)	-0.043 (0.029)
Nutzung Hausaufgabenbetreuung	-0.189** (0.059)	-0.032 (0.038)	0.005 (0.030)
Nutzung fachbezogene Angebote	-0.002 (0.064)	-0.029 (0.036)	-0.016 (0.027)

Anmerkungen: Jede Zelle stellt das Ergebnis eines eigenen Regressionsmodells dar. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

der elterlichen Berufsposition (nicht in der Tabelle, $\beta = 0.050$, $p < 0.10$, Ausgangskompetenz kontrolliert). In tiefergehenden Analysen zeigt sich, wie die Interaktion zustande kommt. Bei statushohen Schüler/-innen hat die Nutzung von Hausaufgabenbetreuungen einen schwachen positiven Effekt (nicht in der Tabelle, $\beta = 0.067$, $p < 0.10$, Ausgangskompetenz kontrolliert). Bei statusniedrigen Schüler/-innen hat die Nutzung von Hausaufgabenbetreuungen dagegen einen schwachen negativen Effekt (nicht in der Tabelle, $\beta = -0.050$, $p < 0.05$, Ausgangskompetenz kontrolliert).

Bei der Nutzung von *fachbezogenen Angeboten* zeigt sich weder ein Zusammenhang zum elterlichen Bildungsstand noch ein Zusammenhang zur Mathematikkompetenz.

In tiefergehenden Analysen können lernförderliche (Haupt-)Effekte der Nutzung von Ganztagsangeboten auch dann nicht festgestellt werden, wenn *zusätzliche Kontrollvariablen* berücksichtigt werden (Mathematiknote, kognitive Grundfähigkeit, Lesekompetenz, Selbstkonzept Mathematik, Geschlecht).

Die Regressionsmodelle legen nur bei der Nutzung von Förderunterricht eine Mediation nahe. Daher wird diese Variable mithilfe eines Pfadmodells untersucht. Die Ergebnisse werden in Abbildung 10 berichtet. Das Pfadmodell ist ein gesättigtes Modell, daher fallen die Indizes der Modellpassung perfekt aus (CFI = 1; RMSEA = 0). Die Pfadkoeffizienten bestätigen eine Mediation: sowohl der Effekt des Bildungsstandes auf die Nutzung von Förderunterricht als auch der Effekt der Nutzung von Förderunterricht auf die Mathematikkompetenz fallen signifikant aus. Allerdings bleibt der direkte Effekt des Bildungsstandes auf die Mathematikkompetenz signifikant. Dies spricht für eine partielle Mediation. Der Mediationsanteil, also der Anteil des indirekten Effekts am Ge-

samteffekt (Ditlevsen et al., 2005, S. 116), beträgt 5.5%. Das bedeutet, 5.5% des Effekts des Bildungsstandes auf die Mathematikkompetenz wirken vermittelt über die Nutzung von Förderunterricht. Die Ergebnisse des Pfadmodells erweisen sich als statistisch robust.²⁴

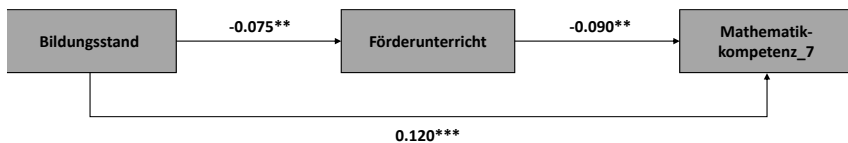


Abb. 10: Pfadmodell zu Förderunterricht (standardisierte Pfadkoeffizienten, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$)

Diskussion der Ergebnisse zur Rolle der Nutzung von Ganztagsangeboten. Die Befunde der Regressionsmodelle legen komplexe Wirkzusammenhänge zwischen der sozialen Herkunft, der Nutzung von Ganztagsangeboten und dem Lernerfolg nahe. Beim *Förderunterricht* ist zunächst festzuhalten, dass sozial benachteiligte Schüler/-innen häufiger Förderunterricht nutzen als sozial privilegierte Schüler/-innen. Weiterhin zeigen sich unterschiedliche Zusammenhänge zu den Mathematikkompetenzen: Bei sozial benachteiligten Schüler/-innen hat Förderunterricht einen negativen Effekt, bei sozial privilegierten Schüler/-innen dagegen keinen Effekt. Mögliche Erklärungen für den Interaktionseffekt sind differenzielle Wirksamkeit (d. h., Effekte auf den Lernerfolg variieren zwischen den beiden Schülergruppen) und differenzielle Selektion (d. h., bei statusniedrigen Schüler/-innen sind die Ausgangsunterschiede zwischen nutzenden und nicht nutzenden Schüler/-innen größer als bei statushohen Schüler/-innen).

Bei *Hausaufgabenbetreuungen* zeigt sich eine gegenläufige Tendenz: Sozial benachteiligte Schüler/-innen nutzen Hausaufgabenbetreuungen häufiger, aber bei sozial privilegierten Schüler/-innen fallen die Zusammenhänge zu den Mathematikkompetenzen günstiger aus. Dies kann als differenzielle Wirksamkeit interpretiert werden. Eine mögliche Erklärung für die stärkeren Effekte bei sozial privilegierten Schüler/-innen ist darin zu sehen, dass sie über günstigere Lernvoraussetzungen verfügen als sozial benachteiligte Schüler/-innen (Raudenbush & Eschmann, 2015, S. 453). Alternativ kann das Befundmuster auch als differenzielle Selektion interpretiert werden (siehe oben).

²⁴ Die Pfadkoeffizienten fallen auch dann signifikant aus, wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, wenn Fälle mit fehlenden Werten ausgeschlossen werden (statt sie mit der FIML-Methode zu behandeln) und wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird.

Die Ergebnisse des Pfadmodells zu Förderunterricht lassen zwei Interpretationen zu. Eine mögliche Interpretation ist, dass sozial benachteiligte Schüler/-innen häufiger Förderunterricht nutzen und dass Förderunterricht einen lernbeeinträchtigenden Effekt hat. In diesem Fall würde die Nutzung von Förderunterricht einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen. Die zweite Interpretation ist, dass dem negativen Zusammenhang zwischen der Nutzung von Förderunterricht und Mathematikkompetenz ein reziproker Effekt zugrunde liegt, d. h., leistungsschwächere Schüler/-innen nutzen häufiger Förderunterricht. In diesem Fall liegt kein vermittelnder Mechanismus vor. Für die zweitgenannte Interpretation spricht, dass der negative Zusammenhang verschwindet, wenn die mathematische Ausgangskompetenz kontrolliert wird. Außerdem erscheint die zweitgenannte Interpretation theoretisch plausibler, da für einen lernbeeinträchtigenden Effekt von Förderunterricht keine überzeugenden Argumente angebracht werden können.

8.3.3 Ergebnisse zur Rolle von Nachhilfe

Im Weiteren werden die Ergebnisse zu der Frage präsentiert, inwieweit die Nutzung von *Nachhilfe* (als Aspekt von schulbezogenen Einrichtungen) einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt (Fragestellung F1.3). Da die vorhandenen Befunde recht heterogen ausfallen, können keine klaren Annahmen darüber formuliert werden, ob die soziale Herkunft die Nutzung von Nachhilfe beeinflusst und ob Nachhilfe einen Effekt auf den Lernerfolg hat.

Die Ergebnisse der bivariaten Zwei-Ebenen-Regressionsmodelle finden sich in Tabelle 6. Der elterliche Bildungsstand hat einen schwachen negativen Effekt auf die Nutzung von Nachhilfe (10%-Niveau), der unter Kontrolle der Ausgangskompetenz knapp das Signifikanzniveau verfehlt (nicht in der Tabelle, $\beta = -0.094$, $p > 0.10$). Sozial privilegierte Schüler/-innen nutzen Nachhilfe also seltener als sozial benachteiligte Schüler/-innen, weil sie über bessere Ausgangskompetenzen verfügen. Nachhilfe hat einen moderaten negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe. Unter Kontrolle der Ausgangskompetenz fällt der negative Effekt nur noch schwach aus, bleibt aber hoch signifikant. Der negative Effekt von Nachhilfe bleibt auch dann fortbestehen, wenn zusätzliche Kontrollvariablen berücksichtigt werden (Mathematiknote, kognitive Grundfähigkeit, Lesekompetenz, Selbstkonzept Mathematik, Geschlecht).

Tab. 6: Zusammenhänge zwischen elterlichem Bildungsstand, Nutzung von Nachhilfe und Mathematikkompetenz in der siebten Jahrgangsstufe (standardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekt des Bildungsstandes auf die Variable	Effekt der Variable auf die Mathematikkompetenz_7	Effekt der Variable auf die Mathematikkompetenz_7 (unter Kontrolle von Mathematikkompetenz_5)
Nachhilfe Mathematik	-0.114 ⁺ (0.065)	-0.201*** (0.028)	-0.122*** (0.025)

Anmerkungen: Jede Zelle stellt das Ergebnis eines eigenen Regressionsmodells dar.

⁺p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

Da Nachhilfe in den Regressionsmodellen zwei gleichgerichtete Zusammenhänge zum Bildungsstand und zur Mathematikkompetenz aufweist, wird im nächsten Schritt ein Pfadmodell berechnet. Abbildung 11 zeigt die Ergebnisse. Da es sich um ein gesättigtes Modell handelt, weisen die Indizes der Modellpassung perfekte Werte auf (CFI = 1; RMSEA = 0). Das Pfadmodell passt zu der Annahme einer Mediation, da der Bildungsstand einen signifikanten Effekt auf die Nutzung von Nachhilfe hat und die Nutzung von Nachhilfe einen signifikanten Effekt auf die Mathematikkompetenz. Der direkte Effekt des Bildungsstandes fällt ebenfalls signifikant aus, was eine partielle Mediation nahelegt. Der Mediationsanteil beträgt 8.7%. Die Ergebnisse erweisen sich in tiefergehenden Analysen als statistisch robust.²⁵

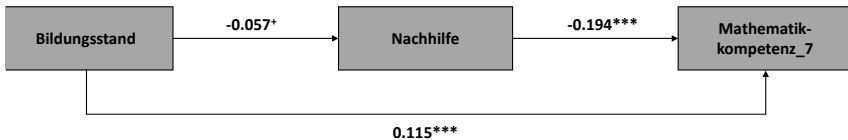


Abb. 11: Pfadmodell zu Nachhilfe (standardisierte Pfadkoeffizienten, ⁺p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001)

Diskussion der Ergebnisse zur Rolle von Nachhilfe. Die Ergebnisse der Regressions- und Pfadmodelle zu Nachhilfe lassen zwei mögliche Interpretationen zu (ähnlich wie bei der Nutzung von Förderunterricht). Erstens passen die Befunde zu der Annahme, dass sozial benachteiligte Schüler/-innen häufiger Nachhilfe nutzen und dass Nachhilfe den Lernerfolg beeinträchtigt. Bei dieser

²⁵ Die Pfadkoeffizienten fallen auch dann signifikant aus, wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, wenn Fälle mit fehlenden Werten ausgeschlossen werden und wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird.

Interpretation würde die Nutzung von Nachhilfe einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen. Eine alternative Interpretation ist, dass der negative Zusammenhang zwischen der Nutzung von Nachhilfe und der Mathematikkompetenz durch einen reziproken Effekt zustande kommt, d. h., leistungsschwächere Schüler/-innen nutzen häufiger Nachhilfe. Demnach kann der reziproke Effekt durch die verwendeten Kontrollvariablen nicht angemessen modelliert werden (ähnliche Interpretation bei Ömerogullari et al., 2020, S. 6). Theoretisch plausibler erscheint die zweitgenannte Interpretation, bei der Nachhilfe keinen vermittelnden Mechanismus darstellt.

8.3.4 Ergebnisse zur Rolle von Freizeitaktivitäten

Nachfolgend werden die Ergebnisse zu Fragestellung F1.4 berichtet, also inwieweit Freizeitaktivitäten einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen. Basierend auf früheren Studien ist anzunehmen, dass sozial privilegierte Schüler/-innen häufiger anregende Freizeitaktivitäten ausführen als sozial benachteiligte Schüler/-innen. Ebenso sind lernförderliche Effekte von Freizeitaktivitäten anzunehmen.

Tabelle 7 zeigt die Ergebnisse der Regressionsmodelle. Der elterliche Bildungsstand hat einen sehr schwachen positiven Effekt auf die Häufigkeit des *Lesens* (10%-Niveau). Sozial privilegierte Schüler/-innen lesen also häufiger als sozial benachteiligte Schüler/-innen. Die Häufigkeit des Lesens hat ihrerseits einen schwachen positiven Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe (siehe Tabelle 7).

Der Effekt wird durch die Kontrolle der Ausgangskompetenz verringert, bleibt aber signifikant. Weiterhin hat der elterliche Bildungsstand einen schwachen positiven Effekt auf die Häufigkeit des *Musizierens*. Demnach spielen sozial privilegierte Schüler/-innen häufiger Musikinstrumente als sozial benachteiligte Schüler/-innen. Die Häufigkeit des Musizierens hat einen sehr schwachen positiven Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, welcher allerdings unter Kontrolle der Ausgangskompetenz verschwindet. Die Häufigkeit von *Sport* korreliert weder mit der sozialen Herkunft noch mit der Mathematikkompetenz.

Tab. 7: Zusammenhänge zwischen elterlichem Bildungsstand, Freizeitaktivitäten und Mathematikkompetenz in der siebten Jahrgangsstufe (standardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekt des Bildungsstandes auf die Variable	Effekt der Variable auf die Mathematikkompetenz_7	Effekt der Variable auf die Mathematikkompetenz_7 (unter Kontrolle von Mathematikkompetenz_5)
Häufigkeit Lesen	0.057* (0.031)	0.131*** (0.027)	0.072** (0.022)
Häufigkeit Musizieren	0.140*** (0.030)	0.066* (0.029)	0.030 (0.026)
Häufigkeit Sport	0.005 (0.029)	0.000 (0.028)	0.018 (0.024)

Anmerkungen: Jede Zelle stellt das Ergebnis eines eigenen Regressionsmodells dar.
[†]p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

Bei der Häufigkeit des Lesens und der Häufigkeit des Musizierens legen die Ergebnisse der Regressionsmodelle eine Mediation nahe. Für diese zwei Variablen werden daher Pfadmodelle berechnet. Die Ergebnisse werden in den Abbildungen 12 und 13 berichtet. Beide Pfadmodelle sind gesättigt, daher fallen die Indizes der Modellpassung perfekt aus (CFI = 1; RMSEA = 0). Sowohl beim Lesen als auch beim Musizieren sprechen die Pfadkoeffizienten für eine Mediation: Der Bildungsstand hat jeweils positive Effekte auf die Freizeitaktivitäten und die Freizeitaktivitäten haben positive Effekte auf die Mathematikkompetenz. Allerdings bleiben die direkten Effekte des Bildungsstandes auf die Mathematikkompetenz jeweils signifikant, was partielle Mediationen nahelegt. Der Mediationsanteil beträgt bei der Häufigkeit des Lesens und bei der Häufigkeit des Musizierens jeweils 5.6%.

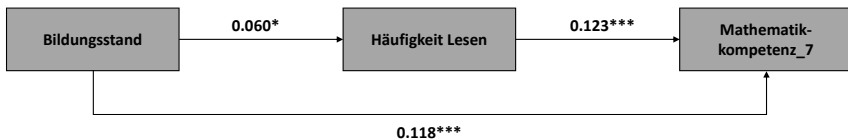


Abb. 12: Pfadmodell zur Häufigkeit des Lesens (standardisierte Pfadkoeffizienten, [†]p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001)

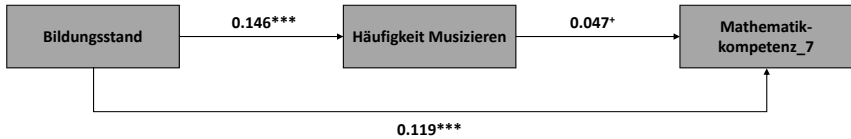


Abb. 13: Pfadmodell zur Häufigkeit des Musizierens (standardisierte Pfadkoeffizienten, $^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$)

Das Pfadmodell zur Häufigkeit des Lesens erweist sich als statistisch robust, das Pfadmodell zur Häufigkeit des Musizierens dagegen nur als teilweise statistisch robust.²⁶

Diskussion der Ergebnisse zur Rolle von Freizeitaktivitäten. Die Befunde der Regressions- und Pfadmodelle bestätigen bei der Häufigkeit des Lesens und bei der Häufigkeit des Musizierens eine Mediation. Auch unter theoretischen Gesichtspunkten erscheint es plausibel, dass diese Freizeitaktivitäten einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen. Hinsichtlich der Effekte auf den Lernerfolg erscheint es denkbar, dass Lesen und Musizieren den Erwerb von Mathematikkompetenz vermittelt über die Förderung von Konzentrationsfähigkeit beeinflussen. Die Interpretation von lernförderlichen Effekten ist zwar grundsätzlich plausibel, für die Entstehung der Zusammenhänge kommen aber auch andere Erklärungen in Betracht (z. B. Drittvariablen als gemeinsame Ursachen).

8.3.5 Ergebnisse zur Rolle der Gleichaltrigengruppe

Im Weiteren werden die Ergebnisse zu der Frage präsentiert, inwieweit die Gleichaltrigengruppe einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt (Fragestellung F1.5). Auf der Grundlage vorhandener Befunde ist anzunehmen, dass sozial privilegierte Schüler/-innen über

26 Bei der Häufigkeit des Lesens fallen die Pfadkoeffizienten auch dann signifikant aus, wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, wenn Fälle mit fehlenden Werten ausgeschlossen werden und wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird. Bei der Häufigkeit des Musizierens zeigen sich andere Ergebnisse. Wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, ist der Effekt des Bildungsstandes auf die Häufigkeit des Musizierens weiterhin signifikant, aber der Effekt der Häufigkeit des Musizierens auf die Mathematikkompetenz nicht mehr signifikant. Das gleiche Ergebnismuster ergibt sich, wenn Fälle mit fehlenden Werten ausgeschlossen werden. Dagegen bleiben beide Pfadkoeffizienten signifikant, wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird.

mehr Freund/-innen mit hohen Bildungsaspirationen verfügen und dass Freund/-innen mit hohen Bildungsaspirationen ihrerseits den Lernerfolg fördern.

Die Ergebnisse der bivariaten Zwei-Ebenen-Regressionsmodelle finden sich in Tabelle 8. Der elterliche Bildungsstand hat keinen Effekt auf den Anteil an *Freund/-innen mit Hauptschulaspiration*.

Tab. 8: Zusammenhänge zwischen elterlichem Bildungsstand, Gleichaltrigengruppe und Mathematikkompetenz in der siebten Jahrgangsstufe (standardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekt des Bildungsstandes auf die Variable	Effekt der Variable auf die Mathematikkompetenz_7	Effekt der Variable auf die Mathematikkompetenz_7 (unter Kontrolle von Mathematikkompetenz_5)
Freund/-innen Hauptschulaspiration	-0.026 (0.028)	-0.071* (0.028)	-0.054* (0.023)
Freund/-innen Realschulaspiration	-0.080** (0.029)	-0.066* (0.028)	-0.048* (0.021)
Freund/-innen Gymnasialaspiration	0.099*** (0.028)	-0.027 (0.031)	-0.001 (0.025)

Anmerkungen: Jede Zelle stellt das Ergebnis eines eigenen Regressionsmodells dar.

⁺p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

Der Anteil an Freund/-innen mit Hauptschulaspiration hat einen sehr schwachen negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, der auch unter Kontrolle der Ausgangskompetenz signifikant bleibt. Weiterhin besteht ein schwacher negativer Effekt des elterlichen Bildungsstandes auf den Anteil an *Freund/-innen mit Realschulaspiration*. Demnach haben sozial privilegierte Schüler/-innen weniger Freund/-innen, die einen Realschulabschluss anstreben, als sozial benachteiligte Schüler/-innen. Der Anteil an Freund/-innen mit Realschulaspiration hat seinerseits einen sehr schwachen negativen Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe. Der Effekt bleibt auch unter Kontrolle der Ausgangskompetenz signifikant. Schließlich zeigt sich ein schwacher positiver Effekt des elterlichen Bildungsstandes auf den Anteil an *Freund/-innen mit Gymnasialaspiration*. Der Anteil an Freund/-innen mit Gymnasialaspiration weist allerdings keinen Zusammenhang zur Mathematikkompetenz auf.

Die Regressionsmodelle legen nur beim Anteil an Freund/-innen mit Realschulaspiration eine Mediation nahe. Daher wird diese Variable mithilfe eines Pfadmodells untersucht. Die Ergebnisse werden in Abbildung 14 berichtet. Das

Pfadmodell ist ein gesättigtes Modell, daher fallen die Indizes der Modellpassung perfekt aus (CFI = 1; RMSEA = 0). Die Pfadkoeffizienten bestätigen eine Mediation: sowohl der Effekt des Bildungsstandes auf den Anteil an Freund/-innen mit Realschulaspisation als auch der Effekt des Anteils an Freund/-innen mit Realschulaspisation auf die Mathematikkompetenz fallen signifikant aus.

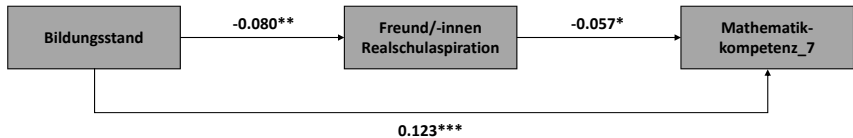


Abb. 14: Pfadmodell zum Anteil an Freund/-innen mit Realschulaspisation (standardisierte Pfadkoeffizienten, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$)

Allerdings bleibt der direkte Effekt des Bildungsstandes auf die Mathematikkompetenz signifikant, was für eine partielle Mediation spricht. Der Mediationsanteil beträgt 3.9%. Die Ergebnisse des Pfadmodells erweisen sich als statistisch robust.²⁷

Diskussion der Ergebnisse zur Rolle der Gleichaltrigengruppe. Von den untersuchten Merkmalen der Gleichaltrigengruppe kann die Mediationsannahme für den Anteil an Freund/-innen mit Realschulaspisation empirisch bestätigt werden. Auch unter theoretischen Gesichtspunkten erscheint eine Rolle als vermittelnder Mechanismus plausibel. Der negative Effekt des Anteils an Freund/-innen mit Realschulaspisation auf den Lernerfolg ist vermutlich darauf zurückzuführen, dass es sich um eine eher niedrige Bildungsaspisation handelt (gemessen an der Referenz des Gymnasiums). Als Wirkmechanismus von Freund/-innen mit niedrigen Bildungsaspitionen ist ein ungünstiger Einfluss auf die Lernmotivation denkbar.

8.3.6 Ergebnisse zur Rolle von Lernvoraussetzungen

Nachfolgend werden die Ergebnisse zu Fragestellung F1.6 berichtet, also inwieweit Lernvoraussetzungen von Schüler/-innen einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellen. Basierend auf früheren Studien ist anzunehmen, dass Zusammenhänge zwischen sozialer

²⁷ Die Pfadkoeffizienten fallen auch dann signifikant aus, wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, wenn Fälle mit fehlenden Werten ausgeschlossen werden und wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird.

Herkunft und kognitiven Merkmalen,²⁸ motivationalen Merkmalen und der Gewissenhaftigkeit bestehen. Weiterhin ist anzunehmen, dass sich diese Lernvoraussetzungen günstig auf den Lernerfolg auswirken.

Tabelle 9 zeigt die Ergebnisse der Regressionsmodelle. Zunächst wird auf kognitive Lernvoraussetzungen eingegangen. Der elterliche Bildungsstand hat einen schwachen positiven Effekt auf die *mathematische Ausgangskompetenz* in der fünften Jahrgangsstufe. Sozial privilegierte Schüler/-innen verfügen also bereits bei Eintritt in das Gymnasium über bessere Mathematikkompetenzen als sozial benachteiligte Schüler/-innen. Die Ausgangskompetenz hat ihrerseits einen starken positiven Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe. Ebenfalls ist ein sehr schwacher positiver Effekt des elterlichen Bildungsstandes auf die *metakognitive Kompetenz* festzustellen (10%-Niveau). Sozial privilegierte Schüler/-innen verfügen also über bessere metakognitive Kompetenzen als sozial benachteiligte Schüler/-innen.

Tab. 9: Zusammenhänge zwischen elterlichem Bildungsstand, Lernvoraussetzungen und Mathematikkompetenz in der siebten Jahrgangsstufe (standardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekt des Bildungsstandes auf die Variable	Effekt der Variable auf die Mathematikkompetenz_7	Effekt der Variable auf die Mathematikkompetenz_7 (unter Kontrolle von Mathematikkompetenz_5)
Mathematikkompetenz_5	0.079** (0.029)	0.516*** (0.022)	
Metakognitive Kompetenz	0.053* (0.028)	0.249*** (0.023)	-0.009 (0.025)
Lesekompetenz	0.135*** (0.026)	0.310*** (0.029)	0.145*** (0.023)
Interesse Mathematik	-0.014 (0.031)	0.285*** (0.036)	0.170*** (0.028)
Selbstkonzept Mathematik	-0.007 (0.028)	0.321*** (0.031)	0.176*** (0.029)
Gewissenhaftigkeit	-0.003 (0.010)	-0.002 (0.001)	0.001 (0.001)

Anmerkungen: Jede Zelle stellt das Ergebnis eines eigenen Regressionsmodells dar.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Die metakognitive Kompetenz hat einen moderaten Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, der allerdings unter Kontrolle der Ausgangskompetenz verschwindet. Weiterhin zeigt sich ein schwacher positiver Effekt des elterlichen Bildungsstandes auf die *Lesekompetenz*. Demnach

28 Bei den Analysen zu vermittelnden Mechanismen wird die kognitive Grundfähigkeit nicht berücksichtigt, da diese Variable mit nur einem Item erfasst wurde und daher für inhaltliche Analysen ungeeignet erscheint.

haben sozial privilegierte Schüler/-innen bessere Lesekompetenzen als sozial benachteiligte Schüler/-innen. Die Lesekompetenz hat einen starken positiven Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe, der unter Kontrolle der Ausgangskompetenz deutlich verringert wird, aber hoch signifikant bleibt.

Die drei Variablen Interesse an Mathematik, mathematikbezogenes Selbstkonzept und Gewissenhaftigkeit werden latent modelliert. Mehrheitlich fallen die Indizes der Modellpassung gut aus ($0.979 \leq CFI \leq 0.997$; $0.030 \leq RMSEA \leq 0.057$). Nur bei den zwei Modellen zum Effekt des Interesses auf die Mathematikkompetenz liegt der RMSEA geringfügig über dem üblichen Schwellenwert ($0.908 \leq CFI \leq 0.931$; $0.116 \leq RMSEA \leq 0.126$).

Unter inhaltlichen Gesichtspunkten ist festzustellen, dass das *Interesse an Mathematik* keinen Zusammenhang zum elterlichen Bildungsstand aufweist. Das Interesse hat aber einen starken Effekt auf die Mathematikkompetenz in der siebten Jahrgangsstufe. Der Effekt wird durch die Kontrolle der Ausgangskompetenz deutlich verringert, bleibt aber hoch signifikant. In ähnlicher Weise weist das *mathematikbezogene Selbstkonzept* keinen Zusammenhang zum elterlichen Bildungsstand auf, aber einen Zusammenhang zur Mathematikkompetenz in der siebten Jahrgangsstufe (auch unter Kontrolle der Ausgangskompetenz). Bei der *Gewissenhaftigkeit* zeigt sich schließlich weder ein Zusammenhang zum elterlichen Bildungsstand noch zur Mathematikkompetenz.

Basierend auf den Ergebnissen der Regressionsmodelle sind bei den kognitiven Lernvoraussetzungen der Schüler/-innen Mediationen in Betracht zu ziehen. Daher werden im nächsten Schritt für die mathematische Ausgangskompetenz, die metakognitive Kompetenz und die Lesekompetenz Pfadmodelle berechnet.

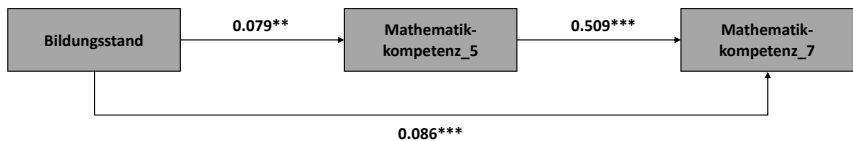


Abb. 15: Pfadmodell zur mathematischen Ausgangskompetenz (standardisierte Pfadkoeffizienten, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$)

Abbildungen 15 bis 17 zeigen die Ergebnisse. Alle drei Pfadmodelle sind gesättigt, daher fallen die Indizes der Modellpassung perfekt aus ($CFI=1$; $RMSEA=0$). Die drei Pfadmodelle bestätigen jeweils die Mediationsannahme:

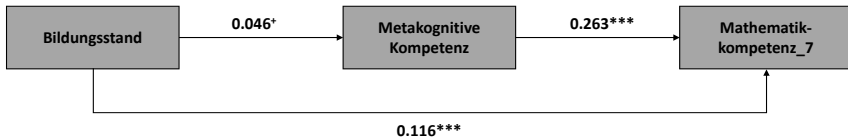


Abb. 16: Pfadmodell zur metakognitiven Kompetenz (standardisierte Pfadkoeffizienten, $^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$)

Sowohl die Effekte des Bildungsstandes auf die drei kognitiven Lernvoraussetzungen als auch die Effekte der drei kognitiven Lernvoraussetzungen auf die Mathematikkompetenz in der siebten Jahrgangsstufe fallen signifikant aus. In allen drei Pfadmodellen bleibt der direkte Effekt des Bildungsstandes auf die Mathematikkompetenz in der siebten Jahrgangsstufe signifikant, d. h., es werden partielle Mediationen nahegelegt. Es bestehen allerdings bedeutsame Unterschiede in den Mediationsanteilen. Bei der mathematischen Ausgangskompetenz beträgt der Mediationsanteil 31.7%, bei der metakognitiven Kompetenz 9.4% und bei der Lesekompetenz 36.9%. Die Ergebnisse der Pfadmodelle zur mathematischen Ausgangskompetenz und zur Lesekompetenz erweisen sich als statistisch robust, die Ergebnisse des Pfadmodells zur metakognitiven Kompetenz dagegen nur als teilweise statistisch robust.²⁹

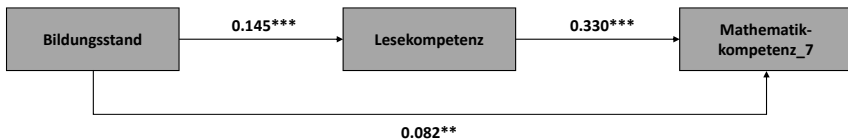


Abb. 17: Pfadmodell zur Lesekompetenz (standardisierte Pfadkoeffizienten, $^+p < 0.10$, $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$)

²⁹ Bei der mathematischen Ausgangskompetenz und der Lesekompetenz fallen die Pfadkoeffizienten auch dann signifikant aus, wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, wenn Fälle mit fehlenden Werten ausgeschlossen werden und wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird. Bei der metakognitiven Kompetenz zeigt sich ein anderes Ergebnismuster. Wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, fallen beide Pfadkoeffizienten signifikant aus. Wenn Fälle mit fehlenden Werten ausgeschlossen werden, ist der Effekt des Bildungsstandes auf die metakognitive Kompetenz nicht mehr signifikant, aber der Effekt der metakognitiven Kompetenz auf die Mathematikkompetenz bleibt signifikant. Wenn als Indikator für die soziale Herkunft die Berufsposition verwendet wird, fallen beide Pfadkoeffizienten signifikant aus.

Diskussion der Ergebnisse zur Rolle von Lernvoraussetzungen. Die Befunde sprechen dafür, dass die *mathematische Ausgangskompetenz* einen Mediator zwischen sozialer Herkunft und Lernerfolg darstellt. Es ist anzunehmen, dass die mathematische Ausgangskompetenz insbesondere die Qualität der fachlichen Lernaktivitäten beeinflusst (z. B. besseres Verstehen neuer Lerninhalte). Bei der *metakognitiven Kompetenz* sprechen die Befunde ebenfalls für eine Mediation. Als Wirkmechanismus der metakognitiven Kompetenz erscheint ein günstiger Einfluss auf die Qualität fachlicher Lernaktivitäten plausibel, weil Schüler/-innen eigene Fehler und Lernprobleme bemerken und korrigierende Maßnahmen ergreifen können. Auch bei der *Lesekompetenz* legen die Befunde eine Mediation nahe. Beim Effekt der Lesekompetenz auf die Mathematikkompetenz sind zwei Interpretationen plausibel. Erstens könnte eine hohe Lesekompetenz eine gute Leistung im Mathematiktest begünstigen, weil der im NEPS verwendete Test auf Textaufgaben basiert. Zweitens könnte eine gute Lesekompetenz den Erwerb von Mathematikkompetenz erleichtern, weil eine hohe Qualität fachlicher Lernaktivitäten begünstigt wird (z. B. haben Schüler/-innen mit guter Lesekompetenz keine Schwierigkeiten beim Lesen von Texten im Mathematikunterricht).

8.3.7 Ergebnisse des Gesamtmodells zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Als Zwischenergebnis der bisherigen Analysen kann festgehalten werden, dass bei acht Variablen die empirischen Befunde zu der Annahme einer Mediation passen. Dabei handelt es sich um die Nutzung von Förderunterricht, die Nutzung von Nachhilfe, die Häufigkeit des Lesens, die Häufigkeit des Musizierens, den Anteil an Freund/-innen mit Realschulaspiration, die mathematische Ausgangskompetenz, die metakognitive Kompetenz und die Lesekompetenz.

Aufbauend auf diesen Ergebnissen wird ein *multivariates Pfadmodell* berechnet. Dieses Pfadmodell erlaubt eine strengere Überprüfung von Mediationsannahmen, weil gleichzeitig die Zusammenhänge zwischen mehreren potenziellen Mediatoren und dem Bildungsstand bzw. der Mathematikkompetenz modelliert werden. Werden die zuvor genannten acht Variablen gemeinsam analysiert, kann die Mediationsannahme bei drei Variablen nicht mehr bestätigt werden: Bei multivariater Modellierung fallen die Effekte der Nutzung von Förderunterricht, der Häufigkeit des Musizierens und der metakognitiven Kompetenz auf die Mathematikkompetenz in der siebten Jahrgangsstufe nicht mehr signifikant aus. Diese drei Variablen werden daher aus dem Modell entfernt, um ein sparsames Modell zu entwickeln.

Abbildung 18 stellt die Ergebnisse des multivariaten Pfadmodells dar. Da alle Variablen manifest modelliert werden, ist das Modell gesättigt und die Indizes der Modellpassung fallen perfekt aus (CFI = 1; RMSEA = 0). Insgesamt erklären die modellierten Prädiktoren $R^2 = 32.2\%$ der Varianz der Mathematikkompetenz in der siebten Jahrgangsstufe (auf der Schülerebene).

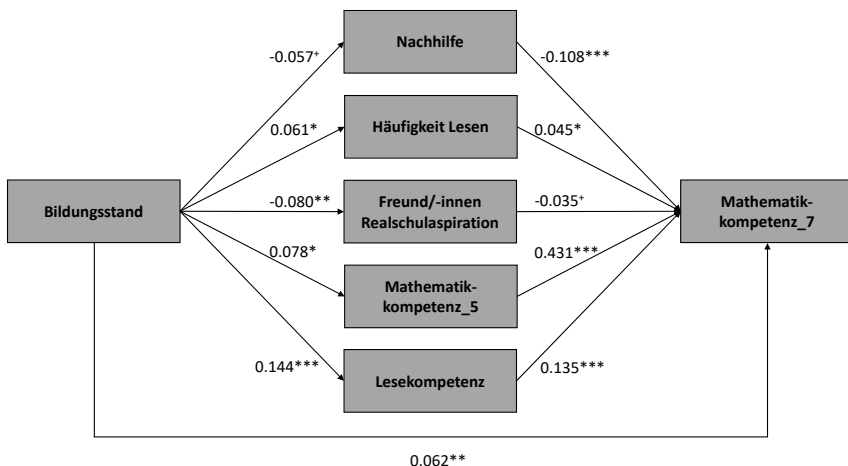


Abb. 18: Multivariates Pfadmodell (standardisierte Pfadkoeffizienten, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$)

Das Modell zeigt, dass die Nutzung von *Nachhilfe*, die Häufigkeit des *Lesens*, der Anteil von *Freund/-innen mit Realschulaspisation*, die *mathematische Ausgangskompetenz* und die *Lesekompetenz* als Mediatoren zwischen Bildungsstand und Mathematikkompetenz in der siebten Jahrgangsstufe bestätigt werden können. Der direkte Effekt des Bildungsstandes auf die Mathematikkompetenz in der siebten Jahrgangsstufe bleibt im multivariaten Pfadmodell signifikant. Der gemeinsame Mediationsanteil der fünf Mediatoren beträgt 51.2%. Das bedeutet, 51.2% des Effekts des Bildungsstandes auf die Mathematikkompetenz in der siebten Jahrgangsstufe wirken vermittelt über die Nutzung von *Nachhilfe*, die Häufigkeit des *Lesens*, den Anteil an *Freund/-innen mit Realschulaspisation*, die *mathematische Ausgangskompetenz* und die *Lesekompetenz*. Die Ergebnisse des multivariaten Pfadmodells erweisen sich in tiefergehenden Analysen als weitestgehend statistisch robust.³⁰

30 Wird die Bundeslandzugehörigkeit der Schule kontrolliert, bleiben alle Koeffizienten signifikant. Werden Fälle mit fehlenden Werten ausgeschlossen, bleiben alle Koeffizienten signifikant. Wird als Indikator für die soziale Herkunft die Berufsposition verwendet, bleiben alle Koeffizienten signifikant mit Ausnahme des Effekts des An-

8.3.8 Diskussion der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Nachfolgend werden die Analysen zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg abschließend diskutiert. Resümierend ist zunächst festzuhalten, dass nur für einen kleineren Teil der Untersuchungsvariablen eine Rolle als Mediator bestätigt werden kann. Von den 20 untersuchten Variablen finden sich nur für 8 Variablen empirische Evidenzen für eine Mediation. Mithilfe der Mediatoren kann etwa die Hälfte des Effekts der sozialen Herkunft auf die Mathematikkompetenz in der siebten Jahrgangsstufe erklärt werden. Hieraus kann geschlossen werden, dass weitere Variablen, die in der vorliegenden Untersuchung nicht berücksichtigt werden, bedeutsame vermittelnde Mechanismen darstellen. Weiterhin ist festzustellen, dass viele Pfadkoeffizienten bzw. Mediationsanteile nur gering ausfallen. Dies legt nahe, dass die Entstehung von herkunftsbedingten Kompetenzunterschieden auf eine große Anzahl von vermittelnden Mechanismen mit jeweils relativ kleinen Effekten zurückzuführen ist.

Auf der Grundlage der Ergebnisse der Regressions- und Pfadmodelle kann eine differenzierte Einschätzung der vorliegenden Evidenz vorgenommen werden. Es wird jeweils beurteilt, ob *keine Evidenz* (keine Hinweise), *begrenzte Evidenz* (gewisse Hinweise), *fundierte Evidenz* (überzeugende Belege) oder *unklare Evidenz* (mehrdeutige Interpretation) vorliegt.

Bei der elterlichen Lernbegleitung (Fragestellung F1.1) finden sich keine empirischen Hinweise auf eine Rolle als vermittelnder Mechanismus. Bei *Häufigkeit* und *Kontrolle* liegen *keine Evidenzen* vor, weil diese Variablen nicht mit der sozialen Herkunft korrelieren. Demgegenüber werden *Unterstützung* und *Erwartungen* durch die soziale Herkunft bedingt, aber es können keine förderlichen Effekte auf den Lernerfolg festgestellt werden. Es erscheint denkbar, dass dies auf methodische Limitationen des korrelativen Forschungsdesigns zurückzuführen ist. Eine abschließende Einschätzung ist nicht möglich, daher sind die Evidenzen als *unklar* anzusehen. Das Befundmuster der vorliegenden Analysen deckt sich mit anderen Studien, die ebenfalls keine lernförderlichen Effekte der elterlichen Lernbegleitung nachweisen können (Castro et al., 2015; Roßbach, 1995; Schauenberg, 2007).

Auch bei der Nutzung von Ganztagsangeboten (Fragestellung F1.2) fallen die Ergebnisse nicht eindeutig aus. Bei *Förderunterricht* und *Hausaufgabenbetreuung* sind einerseits Nutzungsunterschiede zugunsten von sozial benach-

teils der Freund/-innen mit Realschulaspiration auf die Mathematikkompetenz in der siebten Jahrgangsstufe.

teiligten Schüler/-innen festzustellen (wie bei Linberg et al., 2015; Züchner & Fischer, 2014) und andererseits Interaktionseffekte zugunsten von sozial privilegierten Schüler/-innen (wie bei Schüpbach et al., 2013). Bei der Nutzung von Förderunterricht liegen zusätzlich Ergebnisse eines Pfadmodells vor. Die Ergebnisse des Pfadmodells passen zwar zur Annahme einer Mediation, unter theoretischen Gesichtspunkten erscheint ein reziproker Effekt allerdings plausibler (d. h., leistungsschwächere Schüler/-innen nutzen häufiger Förderunterricht). Vor diesem Hintergrund sind die Evidenzen sowohl bei Förderunterricht als auch bei Hausaufgabenbetreuungen als *unklar* einzuschätzen. Demgegenüber finden sich bei der Nutzung von *fachbezogenen Angeboten keine Evidenzen* für eine Rolle als vermittelnder Mechanismus.

Die Analysen zur Rolle von *Nachhilfe* (Fragestellung F1.3) ergeben, dass sozial benachteiligte Schüler/-innen häufiger Nachhilfe nutzen (wie bei Able & Liebau, 1998; Guill & Wendt, 2016; Luplow & Schneider, 2014) und dass Schüler/-innen, die Nachhilfe nutzen, über schlechtere Kompetenzen verfügen (wie bei Guill & Bos, 2014; Guill et al., 2020; Ömerogullari et al., 2020). Die beiden negativen Zusammenhänge können auch im Rahmen eines Pfadmodells bestätigt werden. Unter theoretischen Gesichtspunkten erscheint allerdings wiederum ein reziproker Effekt plausibler (d. h., leistungsschwächere Schüler/-innen nutzen häufiger Nachhilfe). Die Evidenzen zur Nachhilfe sind daher als *unklar* anzusehen.

Bei den Freizeitaktivitäten (Fragestellung F1.4) liegen *keine Evidenzen* dafür vor, dass die Häufigkeit von *Sport* einen vermittelnden Mechanismus darstellt, aber *begrenzte Evidenzen* dafür, dass die Häufigkeit des *Lesens* und die Häufigkeit des *Musizierens* vermittelnde Mechanismen darstellen. Bei beiden Variablen kann eine Mediation mithilfe eines Pfadmodells bestätigt werden, wobei die Mediationsanteile mit jeweils 5.6% eher niedrig ausfallen. Hinweise auf Herkunftseffekte auf Freizeitaktivitäten und lernförderliche Effekte von Freizeitaktivitäten finden sich auch in anderen Studien (Autorengruppe Bildungsberichterstattung, 2014; Leven & Schneekloth, 2010; Pfof et al., 2010; Vaughn, 2000).

Bezüglich der Gleichaltrigengruppe (Fragestellung F1.5) finden sich *begrenzte Evidenzen* dafür, dass der Anteil an *Freund/-innen mit Realschulaspiration* einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt. Dieses Merkmal der Gleichaltrigengruppe kann mithilfe eines Pfadmodells als Mediator bestätigt werden, wenn auch nur mit einem geringen Mediationsanteil von 3.9%. Dieser Befund deckt sich mit den Ergebnissen anderer Studien (Behtui & Neergaard, 2016; Carolan, 2016; Johnson, 2000;

Schmitt, 2012). Demgegenüber finden sich beim Anteil an *Freund/-innen mit Hauptschulaspisation* und beim Anteil an *Freund/-innen mit Gymnasialaspisation keine Evidenzen* für eine Rolle als vermittelnder Mechanismus.

Innerhalb der Kategorie der Lernvoraussetzungen (Fragestellung F1.6) variieren die Evidenzen zwischen den verschiedenen Arten von Lernvoraussetzungen. Mediationseffekte der kognitiven Lernvoraussetzungen können mithilfe von Pfadmodellen klar bestätigt werden. Für die *mathematische Ausgangskompetenz* liegen angesichts des bedeutsamen Mediationsanteils von 31.7% und der hohen theoretischen Plausibilität *fundierte Evidenzen* vor. Bei der *metakognitiven Kompetenz* fällt der Mediationsanteil mit 9.4% niedriger aus. Die Evidenzen sind als *begrenzt* einzuschätzen, primär weil sich die Ergebnisse nur als bedingt robust erweisen. Bei der *Lesekompetenz* sind die Evidenzen trotz des Mediationsanteils von 36.9% ebenfalls als *begrenzt* einzuschätzen, da es theoretisch nicht eindeutig ist, ob der Effekt der Lesekompetenz auf die Mathematikkompetenz einen höheren Lernerfolg oder lediglich eine höhere Testleistung bei Textaufgaben widerspiegelt. Vergleichbare Befunde, wonach die kognitiven Lernvoraussetzungen von Schüler/-innen von ihrer sozialen Herkunft abhängen, finden sich in einer Reihe von Studien (Weinert & Ebert, 2013; Karlen et al., 2014; Richter et al., 2012).

Bei den motivationalen Lernvoraussetzungen *Interesse* und *Selbstkonzept* liefern die Analysen dagegen *keine Evidenzen* für einen Mediationseffekt. Diese Aspekte der Lernmotivation hängen in der untersuchten Gymnasialstichprobe nicht mit der sozialen Herkunft zusammen. Demgegenüber berichten Steinmayr und Kolleg/-innen (2012), ebenfalls für eine Gymnasialstichprobe, dass die Lernmotivation teilweise mit der sozialen Herkunft zusammenhängt. Eine mögliche Erklärung für die abweichenden Befunde ist darin zu sehen, dass Steinmayr und Kolleg/-innen die soziale Herkunft anders operationalisieren.

Bei der *Gewissenhaftigkeit* sind die Evidenzen schließlich als *unklar* einzuschätzen. Der Befund, dass keine Zusammenhänge zum Bildungsstand und zur Mathematikkompetenz bestehen, ist möglicherweise darauf zurückzuführen, dass die Messung von Gewissenhaftigkeit zu ungenau ist (nur zwei Items, niedrige interne Konsistenz). Eine weitere Schwierigkeit ist darin zu sehen, dass Gewissenhaftigkeit global erfasst wurde. Vorhandene empirische Befunde legen aber nahe, dass vor allem die Subfacette Selbstregulation mit der sozialen Herkunft korreliert (Kaiser & Diewald 2014; Miech et al., 2001; Sammons et al., 2014).

Zusammenfassend ist festzuhalten, dass in der untersuchten Gymnasialstichprobe die kognitiven Lernvoraussetzungen der Schüler/-innen, anregende Aktivitäten außerhalb der Schule und die Zusammensetzung des Freundeskreises empirische Bestätigung als vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg finden.

8.4 Ergebnisse zu Effekten von Klassen

Nachfolgend werden die Ergebnisse zum zweiten Bereich von Fragestellungen präsentiert. Diese Fragestellungen beziehen sich auf die Effekte von pädagogischen Einheiten. Es wird also betrachtet, wie stark sich Klassen in ihrer Wirksamkeit unterscheiden. Zunächst werden relative Klasseneffekte auf den durchschnittlichen Lernerfolg betrachtet (Fragestellung F2.1) und dann relative Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg (Fragestellung F2.2). Anschließend wird analysiert, in welchem Verhältnis die beiden Zielkriterien zueinanderstehen (Fragestellung F2.3).

Es ist darauf hinzuweisen, dass in den nachfolgenden Regressionstabellen durchgängig *unstandardisierte Regressionskoeffizienten* berichtet werden. Das liegt daran, dass bei Random-Slope-Modellen keine standardisierten Regressionskoeffizienten zur Verfügung stehen. Um eine einheitliche Darstellung zu gewährleisten, werden auch bei den Random-Intercept-Modellen unstandardisierte Regressionskoeffizienten berichtet.

8.4.1 Ergebnisse zu relativen Klasseneffekten auf den durchschnittlichen Lernerfolg

Nachfolgend werden die Ergebnisse zu relativen Klasseneffekten auf den durchschnittlichen Lernerfolg berichtet. Diese Analysen beziehen sich auf die Frage, wie stark die Kompetenzmittelwerte zwischen den Klassen variieren (Fragestellung F2.1). Hierdurch kann ermittelt werden, wie stark der Lernerfolg von Schüler/-innen davon abhängt, welche Klasse sie besuchen. Zunächst werden relative Klasseneffekte für die gesamte Stichprobe ermittelt, anschließend getrennt für sozial privilegierte und sozial benachteiligte Schüler/-innen.

Der erste Schritt besteht darin, zu überprüfen, wie stark die *unadjustierten Kompetenzmittelwerte* zwischen den Klassen variieren. Die Ergebnisse eines Random-Intercept-Modells ohne Kontrollvariablen finden sich in der mittleren Spalte von Tabelle 10. Die Varianz der klassenspezifischen Intercepts beträgt $\tau_{00} = 0.141$ und weicht signifikant von Null ab. Die unadjustierten

Tab. 10: Ergebnisse der Random-Intercept-Modelle (unstandardisierte Regressionskoeffizienten bzw. Varianzen mit Standardfehlern in Klammern)

Variable	Random-Intercept-Modell ohne Kontrollvariablen	Random-Intercept-Modell mit Kontrollvariablen
Intercept (Mittelwert der klassenspezifischen Intercepts)	1.520*** (0.044)	1.555*** (0.034)
<i>Prädiktoren Schülerebene</i>		
Mathematikkompetenz_5		0.350*** (0.029)
Mathematiknote		0.180*** (0.038)
Kognitive Grundfähigkeit		0.221*** (0.037)
Lesekompetenz		0.151*** (0.022)
Selbstkonzept Mathematik		0.122*** (0.030)
Geschlecht		-0.216*** (0.047)
Lernbegleitung Kontrolle		-0.073* (0.034)
Bildungsstand		0.034*** (0.009)
Nachhilfe Mathematik		-0.231** (0.082)
Häufigkeit Lesen		0.040* (0.018)
Häufigkeit Lesen ²		-0.038** (0.013)
Freund/-innen Hauptschulaspiration		-0.038[†] (0.020)
<i>Varianzkomponenten</i>		
Varianz Mathematikkompetenz_7 Schülerebene	0.891*** (0.040)	0.543*** (0.026)
Varianz Mathematikkompetenz_7 Klassenebene (Varianz der klassenspezifischen Intercepts)	0.141** (0.053)	0.049* (0.021)

Anmerkung: Die Korrelationen zwischen den Prädiktoren auf der Schülerebene bewegen sich zwischen $r = -0.292$ und $r = 0.456$. [†] $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Kompetenzmittelwerte variieren also signifikant zwischen den Klassen. Ausgedrückt als Anteil an der Gesamtvarianz ergibt sich eine unadjustierte Intraclass Correlation von $ICC = 13.7\%$. Somit können 13.7% der Varianz der Mathematikkompetenz in der siebten Jahrgangsstufe durch die Klassenzugehörigkeit erklärt werden. Hierbei ist zu berücksichtigen, dass in die unadjustierte Intraclass Correlation auch Unterschiede in den Ausgangsbedingungen der Schüler/-innen miteinfließen.

Im nächsten Schritt wird daher überprüft, wie stark die *adjustierten Kompetenzmittelwerte* (also die Effektivitätswerte für einen hohen durchschnittlichen Lernerfolg) zwischen den Klassen variieren. Hierfür wird das im letzten Ka-

pitel eingeführte Random-Intercept-Modell verwendet (Gleichungen 1.1 und 1.2). Es handelt sich um ein Mehrwertmodell, bei dem relevante Ausgangsbedingungen der Schüler/-innen kontrolliert werden (z. B. Mathematikkompetenz in der fünften Jahrgangsstufe, Mathematiknote, kognitive Grundfähigkeit etc.). Die Ergebnisse werden in der rechten Spalte von Tabelle 10 berichtet. Die Kontrollvariablen auf der Schülerebene werden manifest modelliert, daher fallen die Indizes der Modellpassung perfekt aus ($CFI=1$; $RMSEA=0$). Die Kontrollvariablen erklären zusammen $R^2=39.1\%$ der Varianz der Mathematikkompetenz in der siebten Jahrgangsstufe (auf der Schülerebene).

Für die hier untersuchte Fragestellung ist primär die Varianz der klassenspezifischen Intercepts von Interesse. Durch die Kontrolle der Ausgangsbedingungen der Schüler/-innen wird die Varianz der Intercepts deutlich verringert. Sie beträgt nun $\tau_{00}=0.049$, ist aber weiterhin signifikant. Demnach variieren auch die adjustierten Kompetenzmittelwerte signifikant zwischen den Klassen. Es ergibt sich eine adjustierte Intraclass Correlation von $ICC=8.3\%$. Das bedeutet, bei gleichen Ausgangsbedingungen der Schüler/-innen können 8.3% der Varianz der Mathematikkompetenz in der siebten Jahrgangsstufe durch die Klassenzugehörigkeit erklärt werden. Die Klassen unterscheiden sich also bedeutsam in ihren Effekten auf den durchschnittlichen Lernerfolg.

Der zweite Teil von Fragestellung F2.1 bezieht sich darauf, ob relative Klasseneffekte auf den durchschnittlichen Lernerfolg bei sozial privilegierten und sozial benachteiligten Schüler/-innen *unterschiedlich stark* ausfallen. Es soll geklärt werden, ob bei einer Schülergruppe der Lernerfolg stärker davon abhängt, welche Klasse sie besucht, als bei der anderen Schülergruppe. Hierfür wird die adjustierte Intraclass Correlation getrennt für sozial privilegierte Schüler/-innen und sozial benachteiligte Schüler/-innen berechnet. Da die Vorgehensweise für die Berechnung der adjustierten ICC zuvor bereits veranschaulicht wurde, wird auf die Darstellung von Tabellen verzichtet und stattdessen nur die ICC berichtet.

Die Ergebnisse variieren in Abhängigkeit davon, ob die Gruppen anhand des elterlichen Bildungsstandes oder anhand der elterlichen Berufsposition gebildet werden. Werden Schüler/-innen von Eltern mit *verschiedenem Bildungsstand* verglichen, ist kein Unterschied festzustellen. Bei Schüler/-innen von Eltern mit mindestens einem Hochschulabschluss beträgt die adjustierte Intraclass Correlation $ICC=6.3\%$ und bei Schüler/-innen von Eltern ohne Hochschulabschluss beträgt die adjustierte Intraclass Correlation $ICC=6.2\%$. Ein anderes Bild ergibt sich, wenn Schüler/-innen von Eltern mit *verschie-*

dener Berufsposition betrachtet werden. Bei Schüler/-innen von *Eltern mit überdurchschnittlicher Berufsposition* fällt die adjustierte Intraclass Correlation mit $ICC=4.6\%$ deutlich niedriger aus als bei Schüler/-innen von *Eltern mit unterdurchschnittlicher Berufsposition* mit $ICC=10.4\%$. Das Befundmuster ist somit nicht eindeutig: Wird der Bildungsstand betrachtet, fallen relative Klasseneffekte bei beiden Schülergruppen sehr ähnlich aus; wird die Berufsposition betrachtet, fallen relative Klasseneffekte bei sozial benachteiligten Schüler/-innen größer aus als bei sozial privilegierten Schüler/-innen. Demnach hängt der Lernerfolg bei Schüler/-innen von Eltern mit niedrigerer Berufsposition stärker davon ab, welche Klasse sie besuchen, als bei Schüler/-innen mit höherer Berufsposition.

8.4.2 Ergebnisse zu relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg

Im Weiteren werden die Ergebnisse zu *relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg* vorgestellt. Diese Analysen beziehen sich auf die Frage, wie stark Kompetenzunterschiede nach sozialer Herkunft zwischen den Klassen variieren (Fragestellung F2.2). Hierdurch kann ermittelt werden, welche Bedeutung es für die Kompetenzunterschiede innerhalb einer Gruppe von Schüler/-innen hat, welche Klasse diese Gruppe von Schüler/-innen besucht. Da die Ergebnisse bei Betrachtung des Bildungsstandes und der Berufsposition der Eltern vergleichbar ausfallen, werden nur die Ergebnisse zum Bildungsstand berichtet.

Für die Analysen werden *Random-Intercept-Random-Slope-Modelle* verwendet, wobei inhaltlich nur der Random Slope von Interesse ist. Anzumerken ist, dass bei dieser Art von Modell (Syntax: TYPE= RANDOM) die üblichen Indizes der Modellpassung nicht zur Verfügung stehen (z. B. CFI, RMSEA). Zunächst wird betrachtet, wie stark *unadjustierte Kompetenzunterschiede nach sozialer Herkunft* zwischen den Klassen variieren. Hierfür wird ein einfaches Random-Intercept-Random-Slope-Modell verwendet, in welchem der Bildungsstand als Prädiktor der Mathematikkompetenz in der siebten Jahrgangsstufe modelliert wird und der Regressionskoeffizient zwischen Klassen variiert. Die Ergebnisse finden sich in der mittleren Spalte von Tabelle 11. Für die Analyse von Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg ist primär die Varianz der klassenspezifischen Slopes von Interesse. Diese beträgt $\tau_{11} = 0.006$ und fällt signifikant aus. Demnach variiert das Ausmaß an unadjustierten

tierten Kompetenzunterschieden nach sozialer Herkunft signifikant zwischen den Klassen. Die Aussagekraft dieses Modells ist insofern begrenzt, als dass es sich um ein querschnittliches Modell ohne Kontrollvariablen handelt.

Daher wird als nächstes überprüft, wie stark *adjustierte Kompetenzunterschiede nach sozialer Herkunft* zwischen den Klassen variieren. Hierfür wird ein Random-Intercept-Random-Slope-Modell berechnet, in dem zusätzlich die Mathematikkompetenz in der fünften Jahrgangsstufe, das Geschlecht und der Migrationshintergrund als Kontrollvariablen auf der Schülerebene berücksichtigt werden (Gleichungen 2.1 bis 2.3). Mit diesem Modell kann ermittelt werden, wie stark die Effektivitätswerte für geringe herkunftsbedingte Unterschiede im Lernerfolg zwischen den Klassen variieren. Die Ergebnisse finden sich in der rechten Spalte von Tabelle 11. Es ist festzustellen, dass die Varianz der klassenspezifischen Slopes mit $\tau_{11} = 0.001$ deutlich geringer ausfällt und nicht mehr signifikant von Null abweicht. Ob die Varianz der klassenspezifischen Slopes signifikant ausfällt oder nicht, hängt also davon ab, ob ein bivariates Modell oder ein multivariates Modell verwendet wird.

Eine wesentliche Einschränkung der durchgeführten Signifikanztests besteht darin, dass keine Aussage darüber getroffen wird, *wie groß* die Unterschiede zwischen den Klassen ausfallen. Während Unterschiede zwischen Klassen bezüglich des durchschnittlichen Lernerfolgs mithilfe der Intraclass Correlation beurteilt werden können, steht für die Einschätzung von Unterschieden zwischen Klassen bezüglich herkunftsbedingter Unterschiede im Lernerfolg keine standardisierte Effektstärke zur Verfügung. Um trotzdem eine Einschätzung der Größe von relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg vornehmen zu können, werden *inhaltliche Überlegungen* angestellt. Hierfür wird verglichen, wie stark Kompetenzunterschiede nach sozialer Herkunft ausfallen, wenn 1) *alle Klassen* betrachtet werden, wenn 2) nur *relativ wirksame Klassen* betrachtet werden und wenn 3) nur *relativ unwirksame Klassen* betrachtet werden. Als Grundlage dient das Random-Intercept-Random-Slope-Modell mit Kontrollvariablen (Gleichungen 2.1 bis 2.3). Betrachtet werden adjustierte Kompetenzunterschiede zwischen Schüler/-innen von Eltern mit mindestens einem Hochschulabschluss und Schüler/-innen von Eltern ohne Hochschulabschluss.

Im *Durchschnitt aller Klassen* betragen die Kompetenzunterschiede zwischen sozial privilegierten und sozial benachteiligten Schüler/-innen

Tab. 11: Ergebnisse der Random-Intercept-Random-Slope-Modelle (unstandardisierte Regressionskoeffizienten bzw. Varianzen mit Standardfehlern in Klammern)

Variable	Random-Intercept-Random-Slope-Modell ohne Kontrollvariablen	Random-Intercept-Random-Slope-Modell mit Kontrollvariablen
Intercept (Mittelwert der klassenspezifischen Intercepts)	1.519*** (0.045)	1.509*** (0.033)
<i>Prädiktoren Schülerebene</i>		
Bildungsstand (Mittelwert der klassenspezifischen Slopes)	0.061*** (0.015)	0.040*** (0.011)
Mathematikkompetenz_5		0.539*** (0.026)
Geschlecht		-0.231*** (0.047)
Migrationshintergrund		-0.016 (0.064)
<i>Prädiktoren Klassenebene: Effekt auf Intercept</i>		
Mittelwert Bildungsstand	0.089* (0.053)	0.047 (0.040)
<i>Varianzkomponenten</i>		
Varianz Mathematikkompetenz_7 Schülerebene	0.852*** (0.038)	0.620*** (0.028)
Varianz Mathematikkompetenz_7 Klassenebene (Varianz der klassenspezifischen Intercepts)	0.108** (0.033)	0.049* (0.023)
Varianz des Regressionskoeffizienten des Bildungsstandes (Varianz der klassenspezifischen Slopes)	0.006* (0.003)	0.001 (0.621)
Kovarianz der klassenspezifischen Intercepts und Slopes	-0.009 (0.006)	-0.002 (0.470)

Anmerkung: Die Korrelationen zwischen den Prädiktoren auf der Schülerebene bewegen sich zwischen $r = -0.161$ und $r = 0.117$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

$d = 0.191$.³¹ Demnach sind die Mathematikkompetenzen von sozial privilegierten Schüler/-innen durchschnittlich 0.191 Standardabweichungen höher als die Mathematikkompetenzen von sozial benachteiligten Schüler/-innen. Als nächstes werden nur *relativ wirksame Klassen* betrachtet. Dabei handelt es sich um die 25% Klassen, in denen der Slope des elterlichen Bildungsstandes auf die Mathematikkompetenz am niedrigsten ausfällt. In der Gruppe der relativ wirksamen Klassen betragen die Kompetenzunterschiede zwischen sozial privilegierten und sozial benachteiligten Schüler/-innen $d = 0.130$. Schließlich werden nur *relativ unwirksame Klassen* betrachtet. Dabei handelt es sich um die 25% Klassen, in denen der Slope des elterlichen Bildungsstandes auf die Mathematikkompetenz am höchsten ausfällt. In der Gruppe der relativ unwirksamen Klassen betragen die Kompetenzunterschiede zwischen sozial privilegierten und sozial benachteiligten Schüler/-innen $d = 0.255$.

Diese Gegenüberstellung zeigt, dass herkunftsbedingte Kompetenzunterschiede in einem Viertel der Klassen ($d = 0.255$) doppelt so stark ausfallen wie im anderen Viertel der Klassen ($d = 0.130$). Das Ausmaß, in dem Unterschiede im Lernerfolg verringert werden können, fällt absolut gesehen zwar gering aus ($\Delta d = 0.125$), kann aber als inhaltlich bedeutsam angesehen werden. Dies gilt insbesondere, da es sich um eine konservative Schätzung von Klasseneffekten handelt. Die Klassen unterscheiden sich also bedeutsam in ihren Effekten auf herkunftsbedingte Unterschiede im Lernerfolg. Hervorzuheben ist, dass auch innerhalb der Gruppe der relativ wirksamen Klassen Kompetenzunterschiede nach sozialer Herkunft vorhanden sind. Demnach gelingt es auch in den relativ wirksamen Klassen nicht, den Einfluss der sozialen Herkunft vollständig zu neutralisieren.

31 Eltern mit Hochschulabschluss haben durchschnittlich 17.44 Bildungsjahre, Eltern ohne Hochschulabschluss durchschnittlich 13.66 Bildungsjahre. Werden alle Klassen betrachtet, beträgt der durchschnittliche Slope des Bildungsstandes auf die Mathematikkompetenz $\gamma_{10} = 0.040$. Die Differenz zwischen den Kompetenzmittelwerten der beiden Schülergruppen beträgt somit $\Delta \bar{x} = (17.44 - 13.66) * 0.040 = 0.151$. Für die Standardisierung wird die unerklärte Varianz der Mathematikkompetenz in der siebten Jahrgangsstufe verwendet ($\sigma^2 = 0.620$), woraus sich eine Standardabweichung von $\sigma = 0.787$ ergibt. Die standardisierte Mittelwertsdifferenz beträgt somit $d = 0.151 / 0.787 = 0.191$.

8.4.3 Ergebnisse zum Verhältnis zwischen durchschnittlichem Lernerfolg und herkunftsbedingten Unterschieden im Lernerfolg

In den vorausgehenden Analysen wurde gezeigt, dass sowohl der durchschnittliche Lernerfolg als auch herkunftsbedingte Unterschiede im Lernerfolg bedeutsam zwischen Klassen variieren. Nachfolgend werden Analysen präsentiert, mit denen das Verhältnis zwischen den beiden Zielkriterien untersucht wird (Fragestellung F2.3). Hierfür wird das Random-Intercept-Random-Slope-Modell mit Kontrollvariablen verwendet (Gleichungen 2.1 bis 2.3). In diesem Modell wird die Kovarianz zwischen den klassenspezifischen Intercepts und Slopes ausgegeben (siehe rechte Spalte von Tabelle 11). Die Kovarianz beträgt $Cov = -0.002$, was einem Korrelationskoeffizienten von $r = -0.286$ entspricht. Das bedeutet, dass geringe herkunftsbedingte Unterschiede im Lernerfolg tendenziell mit einem hohen durchschnittlichen Lernerfolg einhergehen. Allerdings weicht die Kovarianz nicht signifikant von Null ab, d. h., der negative Zusammenhang kann nicht inferenzstatistisch abgesichert werden. Vergleichbare Ergebnisse zeigen sich, wenn die Berufsposition als Indikator der sozialen Herkunft betrachtet wird.

Aus diesem Ergebnis können zwei Schlussfolgerungen gezogen werden. Erstens besteht kein grundsätzlicher Konflikt zwischen den Zielsetzungen eines hohen durchschnittlichen Lernerfolgs und geringer herkunftsbedingter Unterschiede im Lernerfolg. Demnach können geringe Unterschiede im Lernerfolg erreicht werden, ohne dass der durchschnittliche Lernerfolg beeinträchtigt wird. Die Erreichung der beiden Ziele scheint relativ unabhängig voneinander zu sein (da die Korrelation nicht signifikant ausfällt). In der Tendenz zeichnet sich ab, dass sich die Erreichung des einen Ziels günstig auf die Erreichung des anderen Ziels auswirkt. Hieraus kann zweitens geschlossen werden, dass geringe herkunftsbedingte Unterschiede im Lernerfolg eher durch die Förderung von sozial benachteiligten Schüler/-innen erreicht werden und weniger durch die Beeinträchtigung von sozial privilegierten Schüler/-innen (wäre der primäre Mechanismus die Beeinträchtigung von sozial privilegierten Schüler/-innen, müssten geringe herkunftsbedingte Unterschiede im Lernerfolg mit einem niedrigen durchschnittlichen Lernerfolg einhergehen).

8.4.4 Diskussion der Ergebnisse zu Effekten von Klassen

Abschließend werden die Ergebnisse zu den Effekten von Klassen diskutiert. Die Ergebnisse zu *relativen Klasseneffekten auf den durchschnittlichen Lerner-*

folg (Fragestellung F2.1) zeigen konsistent, dass die Kompetenzmittelwerte zwischen den Klassen variieren. Die Evidenz für Klasseneffekte auf den durchschnittlichen Lernerfolg ist als *fundiert* einzuschätzen. Demnach hängt der Lernerfolg von Schüler/-innen wesentlich davon ab, welche Klasse sie besuchen. Die pädagogische Ausgestaltung von Lernbedingungen innerhalb von Klassen hat also einen maßgeblichen Einfluss auf den Lernerfolg der Schüler/-innen. Die Größe der Klasseneffekte fällt in der untersuchten Stichprobe im Vergleich zu anderen Studien eher klein aus, bewegt sich aber in einer plausiblen Größenordnung (Luyten, 2003; Marzano, 2000; Scheerens et al., 1989). Eine mögliche Erklärung für die vergleichsweise kleinen Effekte ist darin zu sehen, dass es sich bei Gymnasiast/-innen um eine Schülergruppe handelt, die über günstige Lernvoraussetzungen verfügt. Aufgrund ihrer Lernvoraussetzungen können Gymnasiast/-innen ungünstige schulische Lernangebote bis zu einem gewissen Grad ausgleichen und sind somit weniger von der pädagogischen Ausgestaltung abhängig als Schüler/-innen mit ungünstigen Lernvoraussetzungen.

Die Ergebnisse zur Größe *relativer Klasseneffekte auf den durchschnittlichen Lernerfolg bei Schüler/-innen verschiedener sozialer Herkunft* (Fragestellung F2.1) sind weniger konsistent. Wird der elterliche Bildungsstand betrachtet, fallen relative Klasseneffekte bei sozial privilegierten und sozial benachteiligten Schüler/-innen ähnlich aus. Wird dagegen die elterliche Berufsposition betrachtet, fallen relative Klasseneffekte bei sozial benachteiligten Schüler/-innen stärker aus als bei sozial privilegierten Schüler/-innen. Eine plausible theoretische Erklärung für die abweichenden Befunde bei Verwendung verschiedener Indikatoren der sozialen Herkunft kann nicht ermittelt werden. Insgesamt betrachtet liegen *begrenzte Evidenzen* für unterschiedlich starke Klasseneffekte vor. Demnach hängt der Lernerfolg bei sozial benachteiligten Schüler/-innen stärker davon ab, welche Klasse sie besuchen, als bei sozial privilegierten Schüler/-innen. Somit reagieren sozial benachteiligte Schüler/-innen sensitiver auf die Ausgestaltung von Lernbedingungen innerhalb von Klassen als sozial privilegierte Schüler/-innen. Die Befunde passen zu der bereits erwähnten Interpretation, dass günstige Lernvoraussetzungen von Schüler/-innen ungünstige schulische Lernangebote bis zu einem gewissen Grad ausgleichen können. Weitere Hinweise auf unterschiedlich starke Klasseneffekte liefern die Studien von Pfeifer (2011) und Nye und Kolleg/-innen (2004).

Auch bei den Ergebnissen zu *relativen Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg* (Fragestellung F2.2) sind die Befunde nicht ganz konsistent. Kompetenzunterschiede nach sozialer Herkunft variieren nur dann signifikant zwischen den Klassen, wenn keine weiteren Kontrollvariablen be-

rücksichtigt werden. Allerdings sind die Unterschiede zwischen Klassen auch bei Berücksichtigung von Kontrollvariablen als inhaltlich bedeutsam anzusehen. Vergleichbare Befundmuster finden sich in internationalen Studien: Es bestehen inhaltlich bedeutsame Unterschiede zwischen Schulen, obwohl die Varianz des Random Slopes nicht signifikant ausfällt (Strand, 2016, S. 130; Young & Fraser, 1993, S. 277, 284). Insgesamt betrachtet liegen *begrenzte Evidenzen* dafür vor, dass Klassen bedeutsame Effekte auf herkunftsbedingte Kompetenzunterschiede haben. Demnach können herkunftsbedingte Kompetenzunterschiede durch eine angemessene Ausgestaltung von Lernbedingungen innerhalb von Klassen gering gehalten werden. Die vollständige Beseitigung von Kompetenzunterschieden erscheint allerdings nicht möglich.

Bezüglich des *Verhältnisses zwischen durchschnittlichem Lernerfolg und herkunftsbedingten Unterschieden im Lernerfolg* (Fragestellung F2.3) finden sich keine Hinweise darauf, dass zwischen den Zielsetzungen eines hohen durchschnittlichen Lernerfolgs und geringer herkunftsbedingter Unterschiede im Lernerfolg ein Zielkonflikt besteht. Die Erreichung der beiden Zielsetzungen scheint unabhängig voneinander zu sein, in der Tendenz sprechen die Ergebnisse eher für die Vereinbarkeit der beiden Zielsetzungen. Ähnliche Befunde berichten Schwippert (2001) und Kyriakides und Kolleg/-innen (2018b).

8.5 Ergebnisse zu Effekten von pädagogischen Merkmalen

Der letzte Teil der Ergebnisdarstellung bezieht sich auf die Analysen zu den Effekten von pädagogischen Merkmalen in den drei Kategorien Unterrichtsmerkmale, Ganztagsangebote und Merkmale der Lerngruppe (Fragestellungen F3.1 bis F3.6). Es werden jeweils die Effekte von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg und auf herkunftsbedingte Unterschiede im Lernerfolg analysiert. Basierend auf den Analysen in den einzelnen Kategorien werden abschließend integrative Gesamtmodelle entwickelt.

Für eine *übersichtliche Darstellung* werden bei den kategorienspezifischen Analysen jeweils nur die Regressionskoeffizienten der *Variablen auf der Klassenebene* berichtet. Die Regressionskoeffizienten der *Variablen auf der Schülerebene* werden nicht berichtet, da sie für die formulierten Fragestellungen nicht weiter von Interesse sind (es handelt sich um reine Kontrollvariablen). Nur bei der abschließenden Darstellung der Gesamtmodelle werden die Regressionskoeffizienten der Variablen auf der Schülerebene berichtet. In ähnlicher Weise werden bei den kategorienspezifischen Analysen zu den Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg (mithilfe des Intercept-And-Slope-As-Outcome-Modell, Gleichungen

4.1 bis 4.3) nur die Regressionskoeffizienten der *Effekte auf den Random Slope* berichtet. Die Regressionskoeffizienten der *Effekte auf den Random Intercept* werden nicht berichtet, weil sie ebenfalls nicht von Interesse sind (siehe Fußnote 23 auf Seite 189). Erst im Rahmen der Gesamtmodelle werden diese Regressionskoeffizienten berichtet.³²

Bei der Interpretation der Ergebnisse tritt die Schwierigkeit auf, dass für viele Modelle *keine standardisierten Regressionskoeffizienten* ausgegeben werden können (weil zufällige Effekte mithilfe von TYPE = RANDOM modelliert werden). Um eine einheitliche Darstellung zu gewährleisten, werden nachfolgend durchgängig *unstandardisierte Regressionskoeffizienten* berichtet. Unstandardisierte Regressionskoeffizienten sind mit zwei Nachteilen verbunden. Erstens können die Koeffizienten verschiedener Variablen *nicht direkt miteinander verglichen* werden (d. h., wenn Variable A einen höheren Koeffizienten hat als Variable B, bedeutet dies nicht automatisch auch, dass Variable A einen stärkeren Effekt hat als Variable B). Zweitens kann bei der *Interpretation der Stärke von Zusammenhängen* nicht auf Schwellenwerte für standardisierte Effektstärken zurückgegriffen werden. Der letztgenannte Punkt ist auch deshalb relevant, weil es bei großen Stichproben theoretisch möglich ist, dass Koeffizienten signifikant werden, obwohl die Stärke des Zusammenhangs unter inhaltlichen Gesichtspunkten als bedeutungslos anzusehen ist. In explorativen Analysen fanden sich keine Hinweise darauf, dass dieser Fall bei der gegebenen Stichprobengröße eintreten kann.³³ Folglich werden alle signifikanten Koeffizienten von Variablen auf der Klassenebene als inhaltlich bedeutsam interpretiert.

32 Die Regressionskoeffizienten der Variablen auf der Schülerebene und die Regressionskoeffizienten der Effekte auf den Random Intercept (im Intercept-And-Slope-As-Outcome-Modell) *in den kategorienspezifischen Analysen* weichen jeweils nicht bedeutend von den Regressionskoeffizienten *in den Gesamtmodellen* ab. Insofern stellen die Regressionskoeffizienten in den Gesamtmodellen ein repräsentatives Abbild dar.

33 Um dies zu überprüfen, werden die standardisierten Regressionskoeffizienten zweier Intercept-As-Outcome-Modelle betrachtet, für welche der Befehl TYPE = RANDOM nicht benötigt wird. Der standardisierte Regressionskoeffizient des Effekts von Förderunterricht für leistungsstarke Schüler/-innen auf den durchschnittlichen Lernerfolg beträgt $\gamma = -0.249$ und erreicht knapp das 10%-Signifikanzniveau ($p = 0.091$). Gemessen an den zuvor gesetzten Schwellenwerten handelt es sich um einen moderaten bis starken Effekt. Hieraus kann geschlossen werden, dass bei der vorliegenden Stichprobengröße Regressionskoeffizienten von Variablen auf der Klassenebene nur bei relativ starken Effekten signifikant werden. Demgegenüber beträgt der standardisierte Regressionskoeffizient des Effekts von Aufgaben mit Hilfe auf den durchschnittlichen Lernerfolg $\gamma = 0.150$ und stellt somit einen schwachen bis moderaten Effekt dar. Der Regressionskoeffizient ist deutlich nicht signifikant ($p = 0.466$). Hieran

8.5.1 Ergebnisse zu Effekten von Unterrichtsmerkmalen

Im Weiteren werden die Ergebnisse zu Effekten von Unterrichtsmerkmalen berichtet. Zunächst werden Effekte auf den durchschnittlichen Lernerfolg behandelt (Fragestellung F3.1), anschließend Effekte auf herkunftsbedingte Unterschiede im Lernerfolg (Fragestellung F3.2).

Bei den Analysen zu Unterrichtsmerkmalen werden die drei Variablen Partner-/Gruppenarbeit, kognitive Aktivierung und Individualisierung latent modelliert. Es werden nicht in allen Modellen Indizes der Modellpassung ausgegeben (weil die Syntax TYPE = RANDOM verwendet wird). In den Modellen, in denen Indizes der Modellpassung zur Verfügung stehen, fallen diese gut aus ($0.998 \leq CFI \leq 1$; $0 \leq RMSEA \leq 0.022$).

Effekte von Unterrichtsmerkmalen auf den durchschnittlichen Lernerfolg. Auf der Grundlage theoretischer Überlegungen und vorhandener Studien sind förderliche Effekte auf den durchschnittlichen Lernerfolg insbesondere bei einem hohen Unterrichtsumfang, beim häufigen Einsatz von Klassenunterricht und Partner/Gruppenarbeiten, einem hohen Maß an Klassenführung, kognitiver Aktivierung und Unterstützung sowie einem mittleren Maß an Übung anzunehmen.

Die Ergebnisse zu den Effekten von Unterrichtsmerkmalen finden sich in Tabelle 12. In der zweiten Spalte finden sich die Regressionskoeffizienten von Unterrichtsmerkmalen, wenn jeweils nur ein Unterrichtsmerkmal als Prädiktor des Random Intercept modelliert wird. Diese Art von Modell wird nachfolgend als *bivariates Intercept-As-Outcome-Modell* bezeichnet. In den bivariaten Modellen erweisen sich drei Unterrichtsmerkmale als signifikant. *Partner/Gruppenarbeiten* haben einen kurvilinearen Effekt auf den durchschnittlichen Lernerfolg. Demnach fällt der Lernerfolg der Schüler/-innen bei einer mittleren Häufigkeit von Partner/Gruppenarbeiten am höchsten aus (siehe Abbildung 21 in Anhang C). Die *Klassenführung* hat einen positiven Effekt auf den durchschnittlichen Lernerfolg, d. h., je effizienter die Klassenführung, desto höher der durchschnittliche Lernerfolg. Bei der *kognitiven Aktivierung* liegt sowohl ein positiver linearer als auch ein kurvilinearere Effekt vor. Das bedeutet, der Lernerfolg der Schüler/-innen ist bei einem hohen, aber nicht maximalem Ausmaß an kognitiver Aktivierung am größten (siehe Abbildung 22 in Anhang C).

zeigt sich, dass bei der vorliegenden Stichprobengröße auch inhaltlich bedeutsame Zusammenhänge nicht signifikant werden. Es finden sich somit keine Hinweise darauf, dass Regressionskoeffizienten signifikant werden können, obwohl die Stärke des Zusammenhangs inhaltlich nicht bedeutsam ist.

Tab. 12: Ergebnisse der Intercept-As-Outcome-Modelle und Slope-As-Outcome-Modelle zu Unterrichtsmerkmalen (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekte auf den durchschnittlichen Lernerfolg		Effekte auf herkunftsbedingte Unterschiede im Lernerfolg	
	Intercept-As-Outcome bivariat	Intercept-As-Outcome multivariat	Slope-As-Outcome bivariat	Slope-As-Outcome multivariat
Unterrichtsumfang	-0.052 (0.102)		-0.018 (0.026)	
Klassenunterricht	0.039 (0.054)	-0.061 (0.080)	-0.017 (0.020)	
Klassenunterricht ²		-0.048* (0.027)		
Einzelarbeit	-0.066 (0.041)		-0.011 (0.011)	
Partner-/Gruppenarbeit	0.041 (0.073)		-0.025* (0.013)	
Partner-/Gruppenarbeit ²	-0.133** (0.043)			
Leistungsheterogene Gruppen	0.014 (0.054)		-0.032** (0.011)	-0.032** (0.011)
Leistungshomogene Gruppen	0.067 (0.064)		-0.004 (0.019)	
Klassenführung	0.034** (0.011)	0.038*** (0.081)	-0.002 (0.005)	
Kognitive Aktivierung	0.111* (0.041)	0.222** (0.081)	-0.019 (0.015)	
Kognitive Aktivierung ²	-0.075* (0.030)	-0.076* (0.032)		
Übung und Wiederholung	0.001 (0.007)		-0.003 (0.003)	
Aufgaben mit Hilfe	0.005 (0.007)		-0.004* (0.002)	-0.004* (0.002)
Aufgaben ohne Hilfe	0.002 (0.005)		0.002 (0.001)	
Besprechung Hausaufgaben	-0.009 (0.008)	-0.015* (0.007)	-0.003 (0.002)	-0.004* (0.002)
Individualisierung	-0.106 (0.239)	-0.200* (0.093)	-0.003 (0.026)	

Anmerkungen: In der zweiten und vierten Spalte stellt jede Zelle das Ergebnis eines bivariaten Mehrebenenmodells dar. In der dritten und fünften Spalte wird jeweils das Ergebnis eines multivariaten Mehrebenenmodells dargestellt. Die Korrelationen zwischen den Unterrichtsmerkmalen bewegen sich zwischen $r = -0.473$ und $r = 0.723$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

In der dritten Spalte von Tabelle 12 finden sich die Regressionskoeffizienten von Unterrichtsmerkmalen, wenn in einem Modell gleichzeitig mehrere Unterrichtsmerkmale als Prädiktoren des Random Intercept modelliert werden. Diese Art von Modell wird nachfolgend als *multivariates Intercept-As-Outcome-Modell* bezeichnet. Bei allen multivariaten Modellen werden jeweils nur Variablen mit signifikanten Effekten berücksichtigt (um sparsame Modelle zu entwickeln). Bei den Unterrichtsmerkmalen zeigt sich, dass die Effekte von *Klassenführung* und *kognitiver Aktivierung* auch im multivariaten Modell bestehen bleiben. Der Effekt von Partner-/Gruppenarbeiten ist dagegen nicht mehr signifikant. Bei drei weiteren Unterrichtsmerkmalen zeigen sich signifikante Effekte nur im multivariaten Modell. Bei *Klassenunterricht* ist ein kurvilinearereffekt festzustellen (10%-Niveau). Eine grafische Analyse ergibt, dass sich ein sehr seltener Einsatz von Klassenunterricht ungünstig auf den Lernerfolg auswirkt (siehe Abbildung 23 in Anhang C). Sowohl der Zeitumfang der *Hausaufgabenbesprechung* als auch das Ausmaß an *Individualisierung* haben negative Effekte auf den durchschnittlichen Lernerfolg, d. h., je stärker diese Variablen ausgeprägt sind, desto niedriger fällt der durchschnittliche Lernerfolg aus. Beim negativen Effekt der Hausaufgabenbesprechung auf den durchschnittlichen Lernerfolg liegen allerdings Hinweise auf einen reziproken Effekt vor (d. h., Lehrkräfte wenden in leistungsschwachen Klassen mehr Zeit für die Hausaufgabenbesprechung auf).³⁴ Die Ergebnisse des multivariaten Modells erweisen sich als weitestgehend statistisch robust.³⁵

34 Das mittlere mathematische Ausgangsniveau der Klasse hat einen negativen Effekt auf den Zeitanteil für die Hausaufgabenbesprechung ($\gamma = -0.218$, $p < 0.10$). Wird der Effekt des mathematischen Ausgangsniveaus auf die Hausaufgabenbesprechung mit einem Mehrebenen-Pfadmodell modelliert, ist der negative Effekt der Hausaufgabenbesprechung auf den durchschnittlichen Lernerfolg nicht mehr signifikant.

35 Wird die Bundeslandzugehörigkeit der Schule kontrolliert, sind die Effekte von Klassenführung und kognitiver Aktivierung weiterhin signifikant, während die Effekte von Hausaufgabenbesprechung und Individualisierung nicht mehr signifikant ausfallen. Werden Fälle mit fehlenden Werten ausgeschlossen (anstatt sie mit FIML zu behandeln), bleiben die Effekte aller Unterrichtsmerkmale außer Klassenunterricht signifikant.

Effekte von Unterrichtsmerkmalen auf herkunftsbedingte Unterschiede im Lernerfolg. Basierend auf theoretischen Überlegungen und vorhandenen Befunden ist anzunehmen, dass geringe herkunftsbedingte Unterschiede im Lernerfolg durch einen hohen Unterrichtsumfang, die häufige Verwendung von Partner/-Gruppenarbeiten (insbesondere bei leistungsheterogener Paar- bzw. Gruppenbildung) sowie ein hohes Maß an Klassenführung, Unterstützung und Übung erreicht werden können. Denkbar sind weiterhin unterschiedsvergrößernde Effekte von leistungshomogener Gruppenbildung und kognitiver Aktivierung. Bei Individualisierung ist ein Effekt auf herkunftsbedingte Unterschiede im Lernerfolg denkbar, aber die Richtung des Effekts kann theoretisch nicht bestimmt werden.

In der vierten Spalte von Tabelle 12 finden sich die Regressionskoeffizienten von Unterrichtsmerkmalen, wenn jeweils nur ein Unterrichtsmerkmal als Prädiktor des Random Slope modelliert wird. Diese Art von Modell wird nachfolgend als *bivariates Slope-As-Outcome-Modell* bezeichnet (es handelt sich zwar um Intercept-And-Slope-As-Outcome-Modelle, aber es wird nur der Slope betrachtet). In den bivariaten Analysen haben drei Unterrichtsmerkmale signifikante Effekte auf den Slope. Die Häufigkeit von *Partner-/Gruppenarbeiten* (10%-Niveau), die Bildung von *leistungsheterogenen Gruppen* und der Zeitan teil für *Aufgaben mit Hilfe* (10%-Niveau) haben jeweils negative Effekte auf den Slope. Das bedeutet, je stärker diese Merkmale im Unterricht ausgeprägt sind, desto geringer fallen herkunftsbedingte Unterschiede im Lernerfolg aus.

In der fünften Spalte von Tabelle 12 finden sich die Regressionskoeffizienten von Unterrichtsmerkmalen, wenn in einem Modell gleichzeitig mehrere Unterrichtsmerkmale als Prädiktoren des Random Slope modelliert werden. Diese Art von Modell wird nachfolgend als *multivariates Slope-As-Outcome-Modell* bezeichnet. Im multivariaten Modell bleiben die Effekte von *leistungsheterogenen Gruppen* und *Aufgaben mit Hilfe* signifikant, während der Effekt von *Partner-/Gruppenarbeiten* das Signifikanzniveau verfehlt. Allerdings zeigt sich im multivariaten Modell nun auch ein signifikanter negativer Effekt der *Besprechung der Hausaufgaben*. Die Ergebnisse des multivariaten Slope-As-Outcome-Modells erweisen sich als weitestgehend statistisch robust.³⁶

36 Wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird und wenn Fälle mit fehlenden Werten ausgeschlossen werden, bleiben die Effekte aller drei Unterrichtsmerkmale signifikant. Wird als Indikator der sozialen Herkunft die Berufsposition verwendet, sind die Effekte von *Aufgaben mit Hilfe* und *Besprechung der Hausaufgaben* weiterhin signifikant, aber der Effekt von *leistungsheterogenen Gruppen* ist nicht mehr signifikant.

In weiterführenden Analysen wird untersucht, inwieweit die Effekte der Unterrichtsmerkmale auf herkunftsbedingte Unterschiede im Lernerfolg auf die Förderung sozial benachteiligter Schüler/innen oder auf die Beeinträchtigung sozial privilegierter Schüler/-innen zurückzuführen sind. Hierfür werden *Mehrgruppenmodelle* verwendet, d. h., das Intercept-As-Outcome-Modell wird für beide Schülergruppen getrennt berechnet (Schüler/-innen von Eltern mit mindestens einem Hochschulabschluss vs. Schüler/-innen von Eltern ohne Hochschulabschluss). Es handelt sich um bivariate Mehrgruppenmodelle, d. h., es wird jeweils nur ein Unterrichtsmerkmal als Prädiktor des Random Intercepts modelliert.

Die Ergebnisse finden sich in Tabelle 13. Ein signifikantes Ergebnis ist nur bei der *Besprechung der Hausaufgaben* festzustellen. Demnach hat die Hausaufgabenbesprechung bei sozial privilegierten Schüler/-innen einen negativen Effekt auf den Lernerfolg, während bei sozial benachteiligten Schüler/-innen kein Zusammenhang mit dem Lernerfolg besteht. Bei den restlichen Unterrichtsmerkmalen zeichnen sich anhand der Vorzeichen der Regressionskoeffizienten zumindest gewisse Tendenzen ab. *Partner-/Gruppenarbeiten* und *Aufgaben mit Hilfe* haben jeweils bei sozial benachteiligten Schüler/-innen tendenziell positive Effekte (der Effekt von Aufgaben mit Hilfe verfehlt mit $p = 0.113$ knapp das 10%-Niveau), während sie bei sozial privilegierten Schüler/-innen tendenziell keine Effekte haben. Die Bildung *leistungsheterogener Gruppen* hat bei sozial benachteiligten Schüler/-innen tendenziell einen positiven Effekt und bei sozial privilegierten Schüler/-innen tendenziell einen negativen Effekt.

Tab. 13: Ergebnisse der bivariaten Mehrgruppenmodelle zu Unterrichtsmerkmalen (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Intercept-As-Outcome bivariat (nur sozial benachteiligte Schüler/-innen)	Intercept-As-Outcome bivariat (nur sozial privilegierte Schüler/-innen)
Partner-/Gruppenarbeit	0.078 (0.054)	-0.008 (0.085)
Leistungsheterogene Gruppen	0.060 (0.051)	-0.045 (0.064)
Aufgaben mit Hilfe	0.012 (0.008)	-0.003 (0.006)
Besprechung Hausaufgaben	-0.001 (0.009)	-0.017* (0.007)

Anmerkungen: Jede Zeile stellt das Ergebnis eines bivariaten Mehrebenenmodells dar.
⁺ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

8.5.2 Diskussion der Ergebnisse zu Effekten von Unterrichtsmerkmalen

Im Weiteren werden die Analysen zu Unterrichtsmerkmalen diskutiert. Die Befunde zu Effekten von Unterrichtsmerkmalen werden jeweils dahingehend eingeschätzt, ob *keine Evidenz* (keine Hinweise), *begrenzte Evidenz* (gewisse Hinweise), *fundierte Evidenz* (überzeugende Belege) oder *unklare Evidenz* (mehrdeutige Interpretation) vorliegt.

Beim *Unterrichtsumfang* liegen, im Kontrast zu den theoretischen Erwartungen, *keine Evidenzen* für Effekte auf den durchschnittlichen Lernerfolg oder auf herkunftsbedingte Unterschiede im Lernerfolg vor. Möglicherweise ist die Erfassung des Unterrichtsumfangs zu ungenau. Die Variable bezieht sich auf die vorgesehene Unterrichtszeit, welche nur einen groben Indikator für das tatsächliche Ausmaß an Lernaktivitäten darstellt.

Die kurvilinearen Effekte von *Klassenunterricht* und *Partner-/Gruppenarbeiten* auf den durchschnittlichen Lernerfolg wurden zwar theoretisch nicht erwartet, erscheinen aber plausibel. Dass der zu häufige Einsatz bestimmter Unterrichtsmethoden den Lernerfolg der Schüler/-innen einschränken kann, deckt sich mit Befunden zur Wirksamkeit von Methodenvielfalt (Brophy & Good, 1986; Helmke, 2014). Einschränkend ist anzumerken, dass sich die kurvilinearen Effekte von Klassenunterricht und Partner-/Gruppenarbeiten jeweils als nur wenig robust erweisen. Die Evidenz fällt daher *begrenzt* aus.

Es finden sich weiterhin Hinweise darauf, dass soziale Interaktionen zwischen Schüler/-innen einen unterschiedsverringernenden Effekt haben. Hierfür sprechen die negativen Effekte von *Partner-/Gruppenarbeiten* und *leistungsheterogenen Gruppen* auf herkunftsbedingte Unterschiede im Lernerfolg. Theoretisch sind diese Befunde plausibel, da durch die Interaktion mit (leistungsstärkeren) Mitschüler/-innen defizitäre Lernvoraussetzungen kompensiert werden können. Ähnliche Ergebnisse berichtet eine Meta-Analyse zum kooperativen Lernen (Rohrbeck et al., 2003). Einschränkend ist anzumerken, dass bei diesen Befunden nur ein moderates Maß an statistischer Robustheit gegeben ist. Vor diesem Hintergrund liefern die Analysen nur *begrenzte Evidenzen* für unterschiedsverringernende Effekte. Demgegenüber finden sich bei der Bildung von *leistungshomogenen Gruppen* *keine Evidenzen* für unterschiedsvergrößernde Effekte.

Der positive Effekt von *Klassenführung* auf den durchschnittlichen Lernerfolg ist theoretisch plausibel (höhere Quantität fachlicher Lernaktivitäten), statistisch robust und deckt sich mit dem bestehenden Forschungsstand (Helmke & Helmke, 2014). Die Evidenz ist als *fundiert* einzuschätzen. Demgegenüber

kann der theoretisch erwartete Effekt von Klassenführung auf herkunftsbedingte Unterschiede im Lernerfolg nicht nachgewiesen werden. Möglicherweise ist dies darauf zurückzuführen, dass Klassenführung global erfasst wurde. Im theoretischen Teil wurde argumentiert, dass primär bei denjenigen Aspekten von Klassenführung, die einen Bezug zur Verhaltensregulation aufweisen (z. B. klare Regeln im Unterricht), ein unterschiedsverringender Effekt zu erwarten ist.

Die Befunde zur *kognitiven Aktivierung* können ebenfalls als *fundiert* beurteilt werden. Als theoretischer Mechanismus ist insbesondere die Förderung der Qualität fachlicher Lernaktivitäten anzunehmen. Hervorzuheben ist, dass die vorliegende Analyse die erste deutsche Studie darstellt, in der ein kurvilinearere Effekt von kognitiver Aktivierung ermittelt wird. Vergleichbare Ergebnisse werden in der international vergleichenden Studie von Caro und Kolleg/-innen (2016) berichtet. Eine mögliche theoretische Erklärung ist darin zu sehen, dass eine zu einseitige Ausrichtung auf offene und herausfordernde Aufgaben dazu führt, dass nicht genug eng geführte Routineaufgaben behandelt werden, welche für die Automatisierung grundlegender kognitiver Prozesse benötigt werden. In diesem Sinne wäre die Kombination von offenen Aufgaben und eng geführten Aufgaben als günstig anzusehen. Ähnliche Befunde liegen zur Wirksamkeit verschiedener Arten von Fragen vor (Muijs et al., 2014, S. 234). Frühere Studien hatten gezeigt, dass Lernprozesse auch durch Fragen auf niedrigem kognitivem Niveau gefördert werden und nicht nur durch Fragen auf hohem kognitivem Niveau (Brophy & Good, 1986, S. 363). Beachtenswert ist, dass kognitive Aktivierung in der vorliegenden Untersuchung keinen unterschiedsvergrößernden Effekt hat, anders als in der Studie von Atlay und Kolleg/-innen (2019). Als Erklärung kommen Unterschiede bei den Stichproben (nur Gymnasien vs. mehrere Schulformen) und bei den Erhebungsmethoden (Lehrkräftefragebogen vs. Schülerfragebogen) in Betracht.

Für die theoretisch erwarteten Effekte von *Übung und Wiederholung* auf den durchschnittlichen Lernerfolg und auf herkunftsbedingte Unterschiede im Lernerfolg liefern die Analysen *keine Evidenzen*. Möglicherweise ist die Erfassung von Übung zu ungenau, da lediglich das quantitative Ausmaß betrachtet wird, während qualitative Aspekte unberücksichtigt bleiben.

Hinsichtlich der Wirksamkeit von *Aufgaben mit Hilfe* finden sich *keine Evidenzen* für einen globalen Effekt auf den durchschnittlichen Lernerfolg. Allerdings geht ein höherer Zeiteanteil für Aufgaben mit Hilfe mit geringeren herkunftsbedingten Unterschieden im Lernerfolg einher. Die Evidenz ist aufgrund der theoretischen Plausibilität als *fundiert* einzuschätzen. Unterstützung sollte die Anforderung an die Lernvoraussetzungen der Schüler/-innen und an das

Erziehungshandeln der Eltern gering halten. Weiterhin erscheint es denkbar, dass sozial benachteiligte Schüler/-innen im Unterricht mehr Unterstützung erfahren als sozial privilegierte Schüler/-innen (unterschiedliche Lerngelegenheiten). Eine Überprüfung der letztgenannten Annahme ist mit den vorliegenden Daten allerdings nicht möglich.

Bei den Effekten der *Besprechung der Hausaufgaben* auf den durchschnittlichen Lernerfolg fallen die Befunde *unklar* aus. Dass der negative Zusammenhang zwischen Hausaufgabenbesprechung und durchschnittlichem Lernerfolg verschwindet, wenn ein reziproker Effekt modelliert wird, legt nahe, dass Lehrkräfte den Unterricht an die Klassenkomposition anpassen. Nicht zu dieser Interpretation passt allerdings der Befund, dass bei sozial privilegierten Schüler/-innen ein negativer Zusammenhang zwischen Hausaufgabenbesprechung und Lernerfolg besteht und bei sozial benachteiligten Schüler/-innen kein Zusammenhang. Der letztgenannte Befund liefert eine Erklärung für den ebenfalls festgestellten unterschiedsverringernenden Effekt der Hausaufgabenbesprechung. Demnach verringert eine ausführliche Besprechung der Hausaufgaben den Lernerfolg von sozial privilegierten Schüler/-innen. Aufgrund von Einschränkungen bei der statistischen Robustheit ist die Evidenz für einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg allerdings nur als *begrenzt* einzuschätzen.

Der negative Effekt von *Individualisierung* auf den durchschnittlichen Lernerfolg deckt sich mit den Befunden anderer deutscher Studien (Clausen, 2002; Gruehn, 1998; Schulz-Heidorf, 2016). Obwohl überzeugende theoretische Argumente für lernförderliche Effekte von Individualisierung sprechen, können diese erneut empirisch nicht bestätigt werden. Mögliche Erklärungen sind eine problematische Umsetzung von individualisierenden Maßnahmen oder sich aus der Individualisierung ergebende Schwierigkeiten bei der Unterrichtsorganisation. Die Evidenz ist als *begrenzt* einzuschätzen.

8.5.3 Ergebnisse zu Effekten von Ganztagsangeboten

Nachfolgend werden die Ergebnisse zu den Effekten von Ganztagsangeboten vorgestellt. Es werden Effekte auf den durchschnittlichen Lernerfolg (Fragestellung F3.3) und Effekte auf herkunftsbedingte Unterschiede im Lernerfolg (Fragestellung F3.4) behandelt. Bei den nachfolgenden Analysen werden alle Variablen manifest modelliert, daher fallen die Indizes der Modellpassung, wenn sie ausgegeben werden, jeweils perfekt aus (CFI = 1; RMSEA = 0).

Tab. 14: Ergebnisse der Intercept-As-Outcome-Modelle und Slope-As-Outcome-Modelle zu Ganztagsangeboten (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekte auf den durchschnittlichen Lernerfolg		Effekte auf herkunftsbedingte Unterschiede im Lernerfolg	
	Intercept-As-Outcome bivariat	Intercept-As-Outcome multivariat	Slope-As-Outcome bivariat	Slope-As-Outcome multivariat
Förderunterricht Mathematik	0.022 (0.062)		-0.046** (0.016)	-0.040** (0.013)
Förderunterricht schwach	-0.022 (0.019)		0.001 (0.011)	
Förderunterricht stark	-0.028⁺ (0.017)		0.013* (0.006)	0.013** (0.004)
Hausaufgabenbetreuung	-0.059** (0.019)		0.025* (0.012)	0.022* (0.011)
Mathematikangebote	-0.023 (0.017)		-0.015* (0.018)	-0.016** (0.005)

Anmerkungen: In der zweiten und vierten Spalte stellt jede Zelle das Ergebnis eines bivariaten Mehrebenenmodells dar. In der fünften Spalte wird das Ergebnis eines multivariaten Mehrebenenmodells dargestellt. Die Korrelationen zwischen den einzelnen Ganztagsangeboten bewegen sich zwischen $r = -0.096$ und $r = 0.498$. ⁺ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Effekte von Ganztagsangeboten auf den durchschnittlichen Lernerfolg. Unter theoretischen Gesichtspunkten erscheinen Effekte von Förderunterricht, Hausaufgabenbetreuung und Mathematikangeboten auf den durchschnittlichen Lernerfolg plausibel, weil die Quantität fachlicher Lernaktivitäten erhöht wird. Bei der Hausaufgabenbetreuung könnte sich die Unterstützung durch schulisches Personal außerdem förderlich auf die Qualität fachlicher Lernaktivitäten auswirken (im Vergleich zum Erledigen der Hausaufgaben zuhause). Im Kontrast zu diesen theoretischen Überlegungen können lernförderliche Effekte von Ganztagsangeboten bisher nur vereinzelt belegt werden.

Die Ergebnisse der empirischen Analysen zu Ganztagsangeboten finden sich in Tabelle 14. In den bivariaten Intercept-As-Outcome-Modellen haben zwei Merkmale signifikante Effekte. *Förderunterricht für leistungsstarke Schüler/-innen* (10%-Niveau) und *Hausaufgabenbetreuungen* weisen jeweils negative Zusammenhänge zum Intercept auf. Das heißt, je häufiger Förder-

unterricht für leistungsstarke Schüler/-innen und Hausaufgabenbetreuungen angeboten werden, desto niedriger fällt der durchschnittliche Lernerfolg der Schüler/-innen aus. Werden die Merkmale in einem gemeinsamen Modell analysiert, bleibt nur der Effekt der Hausaufgabenbetreuung signifikant. Vor diesem Hintergrund wird kein multivariates Intercept-As-Outcome-Modell berichtet. Beim negativen Effekt der Hausaufgabenbetreuung ist ein reziproker Effekt denkbar, kann empirisch aber nicht bestätigt werden (d. h., das mathematische Ausgangsniveau der Schüler/-innen hat keinen Effekt auf die Häufigkeit der Hausaufgabenbetreuung). Der negative Effekt der Hausaufgabenbetreuung erweist sich auch in weiterführenden Analysen als statistisch robust.³⁷

Effekte von Ganztagsangeboten auf herkunftsbedingte Unterschiede im Lernerfolg. Im Weiteren werden die Ergebnisse zu Effekten von Ganztagsangeboten auf herkunftsbedingte Unterschiede im Lernerfolg vorgestellt. Auf der Grundlage theoretischer Überlegungen sind bei Förderunterricht für leistungsschwache Schüler/-innen und Hausaufgabenbetreuung unterschiedsverringende Effekte zu erwarten, während bei Förderunterricht für leistungsstarke Schüler/-innen unterschiedsvergrößernde Effekte denkbar sind. Es liegen bisher aber keine empirischen Studien vor, die diese Annahmen stützen.

In den bivariaten Slope-As-Outcome-Modellen in Tabelle 14 haben vier der fünf Ganztagsangebote signifikante Effekte. Je häufiger *Förderunterricht in Mathematik* und *Mathematikangebote* erfolgen, desto niedriger fallen herkunftsbedingte Unterschiede im Lernerfolg aus. Je häufiger *Förderunterricht für leistungsstarke Schüler/-innen* und *Hausaufgabenbetreuungen* angeboten werden, desto stärker fallen herkunftsbedingte Unterschiede im Lernerfolg aus. Die Effekte dieser vier Merkmale zeigen sich auch in einem multivariaten Slope-As-Outcome-Modell. Die Ergebnisse des multivariaten Modells erweisen sich als weitestgehend statistisch robust.³⁸

Erkenntnisse über die Bedeutung von Ganztagsangeboten für verschiedene Schülergruppen liefern *Mehrgruppenmodelle* (siehe Tabelle 15). Drei der vier Ganztagsangebote haben nur bei sozial benachteiligten Schüler/-innen signi-

37 Der negative Effekt zeigt sich auch, wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, wenn Fälle mit fehlenden Werten ausgeschlossen werden und wenn die Schulebene statt die Klassenebene modelliert wird.

38 Wird die Bundeslandzugehörigkeit der Schulen kontrolliert, bleiben alle Effekte signifikant. Werden Fälle mit fehlenden Werten ausgeschlossen, fällt der Effekt der Häufigkeit von Mathematikangeboten nicht mehr signifikant aus, aber die Effekte der anderen drei Merkmale bleiben signifikant. Wird als Indikator für die soziale Herkunft die Berufsposition verwendet, bleiben alle Effekte signifikant. Wenn die Schulebene statt die Klassenebene modelliert wird, bleiben alle Effekte signifikant.

fikante Effekte. Bei dieser Schülergruppe hat *Förderunterricht in Mathematik* positive Effekte auf den Lernerfolg, während *Förderunterricht für leistungsstarke Schüler/-innen* sowie *Hausaufgabenbetreuungen* negative Effekte haben. Die Effekte von Mathematikangeboten fallen bei beiden Schülergruppen nicht signifikant aus, aber tendenziell gehen Mathematikangebote mit einem niedrigeren Lernerfolg bei sozial privilegierten Schüler/-innen einher.

8.5.4 Diskussion der Ergebnisse zu Effekten von Ganztagsangeboten

Die Ergebnisse zu den Effekten von Ganztagsangeboten fallen teilweise deutlich erwartungswidrig aus. Im Weiteren erfolgt eine Diskussion unter theoretischen und methodischen Gesichtspunkten.

Beim *Förderunterricht in Mathematik* ergibt sich ein plausibles Befundmuster. Wie theoretisch erwartet, geht mehr Förderunterricht in Mathematik mit geringeren herkunftsbedingten Unterschieden im Lernerfolg einher. Im Mehrgruppenmodell zeigt sich, dass der unterschiedsverringende Effekt darauf zurückzuführen ist, dass sozial benachteiligte Schüler/-innen stärker von Förderunterricht in Mathematik profitieren als sozial privilegierte Schüler/-innen. Die in Kapitel 8.3.2 berichteten Befunde zur Nutzung von Ganztagsangeboten liefern Hinweise auf den zugrundeliegenden Mechanismus. Der unterschiedsverringende Effekt von Förderunterricht in Mathematik ist demnach darauf zurückzuführen, dass sozial benachteiligte Schüler/-innen häufiger Förderunterricht nutzen als sozial privilegierte Schüler/-innen (unterschiedliche Lerngelegenheiten). Weiterhin sprechen die Befunde des Mehrgruppenmodells dafür, dass dem in Kapitel 8.3.2 festgestellten Interaktionseffekt zwischen sozialer Herkunft und Förderunterricht eine differenzielle Selektion zugrunde liegt. Insgesamt betrachtet ist die Evidenz für unterschiedsverringende Effekte von Förderunterricht in Mathematik als *fundierte* einzuschätzen.³⁹

Im Gegensatz liefern die Ergebnisse beim *Förderunterricht für leistungsschwache Schüler/-innen* keine Evidenzen für Effekte auf den durchschnittlichen Lernerfolg oder auf herkunftsbedingte Unterschiede im Lernerfolg. Mög-

39 Anzumerken ist allerdings, dass die Formulierung „Förderunterricht“ im Lehrkräfte- und Schülerfragebogen des NEPS nicht ganz eindeutig ist, da nicht spezifiziert wird, ob Förderunterricht für leistungsschwache oder leistungsstarke Schüler/-innen gemeint ist (anders als im Fragebogen für Schulleiter/-innen). Es ist anzunehmen, dass mit dem Begriff „Förderunterricht“ typischerweise Förderunterricht für leistungsschwache Schüler/-innen gemeint ist. Die berichteten Befunde stützen diese Annahme.

Tab. 15: Ergebnisse der bivariaten Mehrgruppenmodelle zu Ganztagsangeboten (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Intercept-As-Outcome bivariat (nur sozial benachteiligte Schüler/-innen)	Intercept-As-Outcome bivariat (nur sozial privilegierte Schüler/-innen)
Förderunterricht Mathematik	0.156* (0.075)	-0.049 (0.056)
Förderunterricht stark	-0.038* (0.018)	-0.012 (0.020)
Hausaufgabenbetreuung	-0.081*** (0.013)	-0.031 (0.029)
Mathematikangebote	-0.004 (0.015)	-0.038 (0.024)

Anmerkungen: Jede Zeile stellt das Ergebnis eines bivariaten Mehrebenenmodells dar. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

licherweise ist dies darauf zurückzuführen, dass bei der Erhebung der Variable im Fragebogen für Schulleiter/-innen nicht konkretisiert wird, auf welche Fächer sich der Förderunterricht bezieht.

Der negative Effekt von *Förderunterricht für leistungsstarke Schüler/-innen* auf den durchschnittlichen Lernerfolg deckt sich mit den Ergebnissen einer anderen Studie (Schmid et al., 2020). Der unterschiedsvergrößernde Effekt war theoretisch zu erwarten. Nicht zu erwarten war allerdings der Befund des Mehrgruppenmodells, wonach der unterschiedsvergrößernde Effekt darauf zurückzuführen ist, dass Förderunterricht für leistungsstarke Schüler/-innen den Lernerfolg von sozial benachteiligten Schüler/-innen verringert. Ein direkter lernbeeinträchtigender Effekt dieses Ganztagsangebots erscheint wenig überzeugend. Plausibler erscheint es, dass in der Häufigkeit von Förderunterricht für leistungsstarke Schüler/-innen die Prioritäten der Schule bezüglich der Förderung verschiedener Schülergruppen zum Ausdruck kommen. Die Priorisierung von leistungsstarken Schüler/-innen äußert sich darin, dass ein hoher Anteil der verfügbaren schulischen Ressourcen in diese Schülergruppe investiert wird. Dies geht damit einher, dass für leistungsschwächere Schüler/-innen weniger Ressourcen zur Verfügung stehen, was deren Lernerfolg einschränkt. Diese Interpretation liefert auch eine mögliche Erklärung für den negativen Effekt von Förderunterricht für leistungsstarke Schüler/-innen auf den durchschnittlichen Lernerfolg: Schulische Ressourcen werden in Schüler/-innen investiert, die auf diese Ressourcen weniger angewiesen sind, und fehlen dann bei Schüler/-innen, die auf diese Ressourcen stärker angewiesen sind. Insgesamt betrachtet kann das Befundmuster als *begrenzte Evidenz* für lernbeeinträchtigende und unterschiedsvergrößernde Effekte von Förderunterricht für leistungsstarke Schüler/-innen angesehen werden.

Bei *Hausaufgabenbetreuungen* fallen die Ergebnisse deutlich erwartungswidrig aus. Von Interesse ist zunächst einmal der negative Zusammenhang zwischen der Häufigkeit von *Hausaufgabenbetreuungen* und den Mathematikkompetenzen der Schüler/-innen. Es finden sich keine Hinweise darauf, dass Hausaufgabenbetreuungen reaktive schulische Maßnahmen bei ungünstiger Schülerkomposition darstellen. Vor diesem Hintergrund gilt es zu überprüfen, ob theoretische Argumente für einen lernbeeinträchtigenden Effekt von Hausaufgabenbetreuungen sprechen. Hierbei erscheint es sinnvoll, mögliche Vorteile des Bearbeitens von Hausaufgaben zuhause zu überdenken. Ein möglicher Vorteil ist, dass beim selbstständigen Bearbeiten von Hausaufgaben zuhause die Selbstregulation stärker gefördert wird als in der Hausaufgabenbetreuung (Trautwein & Lüdtke, 2008). Weiterhin möglich erscheint, dass die Qualität der Lernaktivitäten beim Bearbeiten der Hausaufgaben zuhause höher ausfällt, weil Lernen in unterschiedlichen Kontexten (in der Schule und in der Familie) anregender ist als Lernen in nur einem Kontext (nur in der Schule) (Smith et al., 1978). Die Interpretation eines lernbeeinträchtigenden Effekts von Hausaufgabenbetreuungen wird allerdings dadurch eingeschränkt, dass auf der Schülerebene kein negativer Zusammenhang zwischen der Nutzung einer Hausaufgabenbetreuung und der Mathematikkompetenz festzustellen ist (siehe Kapitel 8.3.2 „Ergebnisse zur Rolle der Nutzung von Ganztagsangeboten“). Insgesamt betrachtet ergeben die Befunde kein schlüssiges Gesamtbild. Die Evidenz zu den Effekten von Hausaufgabenbetreuungen auf den durchschnittlichen Lernerfolg ist daher als *unklar* einzuschätzen.

Dass Hausaufgabenbetreuungen mit stärkeren herkunftsbedingten Unterschieden im Lernerfolg einhergehen, steht ebenfalls im Kontrast zu den theoretischen Erwartungen. Hinweise auf den zugrunde liegenden Mechanismus bieten die in Kapitel 8.3.2 berichteten Befunde. Der dort festgestellte Interaktionseffekt lässt sich so interpretieren, dass sozial privilegierte Schüler/-innen von der Nutzung von Hausaufgabenbetreuungen profitieren, sozial benachteiligte Schüler/-innen dagegen nicht. Dass Hausaufgabenbetreuungen differenziell wirksam sind, stellt einen möglichen Mechanismus für den unterschiedsvergrößernden Effekt von Hausaufgabenbetreuungen dar. Obwohl somit ein möglicher Mechanismus identifiziert werden kann, weicht das Befundmuster insgesamt doch deutlich von den theoretisch begründeten Erwartungen ab. Vor diesem Hintergrund werden die Evidenzen für unterschiedsvergrößernde Effekte von Hausaufgabenbetreuungen als *begrenzt* eingeschätzt.

Hinsichtlich des Effekts von *Mathematikangeboten* auf herkunftsbedingte Unterschiede im Lernerfolg finden sich Hinweise mit gewissen Einschränkun-

gen. Im Slope-As-Outcome-Modell zeigt sich ein unterschiedsverringender Effekt von Mathematikangeboten. Die Ergebnisse des Mehrgruppenmodells sprechen aber in der Tendenz dafür, dass der unterschiedsverringende Effekt darauf zurückzuführen ist, dass sozial privilegierte Schüler/-innen in ihrem Lernerfolg beeinträchtigt werden. Dies erscheint theoretisch wenig plausibel. Außerdem erweist sich der unterschiedsverringende Effekt von Mathematikangeboten als nur teilweise robust. Die vorliegenden Evidenzen sind somit nur als *begrenzt* einzuschätzen.

8.5.5 Ergebnisse zu Effekten von Merkmalen der Lerngruppe

Die letzte Kategorie des analytischen Modells bilden Merkmale der Lerngruppe. Nachfolgend werden die Ergebnisse zu den Effekten der Klassengröße und der Klassenkomposition auf den durchschnittlichen Lernerfolg (Fragestellung F3.5) und auf herkunftsbedingte Unterschiede im Lernerfolg präsentiert (Fragestellung F3.6). Bei den nachfolgenden Analysen werden alle Variablen manifest modelliert, daher fallen die Indizes der Modellpassung, wenn sie ausgegeben werden, jeweils perfekt aus (CFI=1; RMSEA=0).

Effekte von Merkmalen der Lerngruppe auf den durchschnittlichen Lernerfolg. Auf der Grundlage vorhandener Studien sind Effekte auf den durchschnittlichen Lernerfolg bei der Klassenkomposition zu erwarten und hierbei insbesondere beim mittleren mathematischen Ausgangsniveau und beim mittleren Bildungsstand der Klasse.

Die Ergebnisse zu den Effekten von Merkmalen der Lerngruppe finden sich in Tabelle 16. Im Einklang mit den theoretischen Erwartungen fällt in den bivariaten Intercept-As-Outcome-Modellen der Effekt des *mittleren mathematischen Ausgangsniveaus der Klasse* (Mittelwert Mathematik_5) signifikant und positiv aus. Das bedeutet, je höher die durchschnittlichen mathematischen Ausgangskompetenzen der Klasse sind, desto höher fällt der durchschnittliche Lernerfolg der Schüler/-innen aus (unter Kontrolle der individuellen mathematischen Ausgangskompetenz). Da alle Variablen auf der Schülerebene um den Gesamtmittelwert zentriert wurden, gibt der Regressionskoeffizient auf der Klassenebene ($\gamma=0.206$, $p<0.05$) direkt den Kompositionseffekt wieder (Ditton, 1998, S. 92f.). Die Befunde zu den Effekten von Merkmalen der Lerngruppe fallen vergleichbar aus, wenn alle vier Merkmale gemeinsam modelliert werden, weshalb kein multivariates Modell berichtet wird. Der Effekt des ma-

thematischen Ausgangsniveaus der Klasse erweist sich teilweise als statistisch robust.⁴⁰

Tab. 16: Ergebnisse der Intercept-As-Outcome-Modelle und Slope-As-Outcome-Modelle zu Merkmalen der Lerngruppe (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Effekte auf den durchschnittlichen Lernerfolg		Effekte auf herkunftsbedingte Unterschiede im Lernerfolg	
	Intercept-As-Outcome bivariat	Intercept-As-Outcome multivariat	Slope-As-Outcome bivariat	Slope-As-Outcome multivariat
Klassengröße	-0.021 (0.017)		-0.002 (0.004)	
Mittelwert Mathematik_5	0.206* (0.085)		-0.061** (0.021)	-0.053* (0.021)
Mittelwert Bildungsstand	0.012 (0.036)		-0.009 (0.014)	
Anteil Migrationshintergrund	-0.370 (0.280)		0.147* (0.063)	0.101* (0.061)

Anmerkungen: In der zweiten und vierten Spalte stellt jede Zelle das Ergebnis eines bivariaten Mehrebenenmodells dar. In der fünften Spalte wird das Ergebnis eines multivariaten Mehrebenenmodells dargestellt. Die Korrelationen zwischen den einzelnen Merkmalen von Lerngruppen bewegen sich zwischen $r = -0.256$ und $r = 0.364$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Effekte von Merkmalen der Lerngruppe auf herkunftsbedingte Unterschiede im Lernerfolg. Basierend auf vorhandenen Studien sind Effekte der Klassengröße auf herkunftsbedingte Unterschiede im Lernerfolg denkbar. Weiterhin sind Effekte der Klassenkomposition zu erwarten, wobei allerdings die Richtungen der Effekte aufgrund des sehr heterogenen Forschungsstandes nicht eindeutig vorherzusagen sind.

In den bivariaten Slope-As-Outcome-Modellen in Tabelle 16 hat das *mittlere mathematische Ausgangsniveau der Klasse* einen signifikanten negativen Effekt und der *Anteil an Schüler/-innen mit Migrationshintergrund* einen signifikanten positiven Effekt auf den Random Slope. Demnach geht ein höheres mathematisches Ausgangsniveau der Klasse und ein niedrigerer Anteil an

40 Wenn die Bundeslandzugehörigkeit der Schule kontrolliert, bleibt der Effekt des mittleren Ausgangsniveaus der Klasse auf den Intercept signifikant. Wenn Fälle mit fehlenden Werten ausgeschlossen werden, fällt der Effekt nicht mehr signifikant aus.

Schüler/-innen mit Migrationshintergrund mit geringeren herkunftsbedingten Unterschieden im Lernerfolg einher. Diese zwei Effekte können auch in einem multivariaten Slope-As-Outcome-Modell bestätigt werden (wobei der Effekt des Anteils an Schüler/-innen mit Migrationshintergrund dann nur noch auf dem 10%-Niveau signifikant ist).

Die Befunde zur Klassenkomposition werfen die Frage auf, ob die zwei festgestellten Cross-Level-Interaktionen auf (im Modell nicht berücksichtigte) *Interaktionseffekte auf der Schülerebene* zurückzuführen sind. In vertiefenden Analysen zeigt sich keine Interaktion zwischen individueller mathematischer Ausgangskompetenz und individuellem Bildungsstand. Allerdings fällt die Interaktion zwischen individuellem Migrationshintergrund und individuellem Bildungsstand signifikant negativ aus (nicht in der Tabelle, $\beta = 0.068$, $p < 0.05$). Das bedeutet, der Effekt des Bildungsstandes ist bei Schüler/-innen ohne Migrationshintergrund stärker als bei Schüler/innen mit Migrationshintergrund. Die Cross-Level-Interaktion ist aber weiterhin signifikant (nicht in der Tabelle, $\gamma = 0.228$, $p < 0.05$), wenn die Interaktion auf der Schülerebene modelliert wird. Diese Ergebnisse bestätigen, dass es sich bei den ermittelten Zusammenhängen tatsächlich um Cross-Level-Interaktionen handelt.

Bezüglich des Effekts des mittleren mathematischen Ausgangsniveaus der Klasse wird außerdem überprüft, ob dieser Effekt mit der klassenspezifischen *Varianz der mathematischen Ausgangskompetenzen* zusammenhängt (d. h., in Klassen mit hohem mittlerem Ausgangsniveau fällt die Varianz der Ausgangskompetenzen geringer aus und dadurch fallen herkunftsbedingte Unterschiede im Lernerfolg geringer aus). Diese Annahme findet keine Bestätigung, da die Varianz der mathematischen Ausgangskompetenzen nicht mit herkunftsbedingten Unterschieden im Lernerfolg korreliert. In weiteren Analysen erweist sich der Effekt des mittleren mathematischen Ausgangsniveaus der Klasse als statistisch robust, der Effekt des Anteils an Schüler/-innen mit Migrationshintergrund dagegen nur als teilweise statistisch robust.⁴¹

Die Effekte von Merkmalen der Lerngruppe werden weiterhin mithilfe von Mehrgruppenmodellen analysiert (siehe Tabelle 17). Die Ergebnisse der

41 Wenn die Bundeslandzugehörigkeit der Schule kontrolliert wird, sind beide Effekte weiterhin signifikant. Wenn Fälle mit fehlenden Werten ausgeschlossen werden, ist der Effekt des mittleren Ausgangsniveaus der Klasse weiterhin signifikant, aber der Effekt des Anteils an Schüler/-innen mit Migrationshintergrund ist nicht mehr signifikant. Wenn als Indikator der sozialen Herkunft die Berufsposition verwendet wird, ist der Effekt des mittleren Ausgangsniveaus der Klasse weiterhin signifikant, aber der Effekt des Anteils an Schüler/-innen mit Migrationshintergrund ist nicht mehr signifikant.

Mehrgruppenmodelle zeigen, dass bei der Gruppe der sozial benachteiligten Schüler/-innen das mathematische Ausgangsniveau der Klasse einen signifikant positiven Effekt auf den Lernerfolg hat und der Anteil an Schüler/-innen mit Migrationshintergrund einen signifikant negativen Effekt. Bei der Gruppe der sozial privilegierten Schüler/-innen weisen diese zwei Merkmale der Lerngruppe dagegen keine signifikanten Zusammenhänge zum Lernerfolg der Schüler/-innen auf.

8.5.6 Diskussion der Ergebnisse zu Effekten von Merkmalen der Lerngruppe

Nachfolgend werden die Ergebnisse zu den Effekten von Merkmalen der Lerngruppe diskutiert. Die Befunde zur *Klassengröße* sprechen dafür, dass dieses Merkmal in der untersuchten Gymnasialstichprobe keine wesentliche Rolle spielt. Es finden sich *keine Evidenzen* für Effekte auf den durchschnittlichen Lernerfolg oder für Effekte auf herkunftsbedingte Unterschiede im Lernerfolg. Eine mögliche Erklärung für die genannten Befunde ist darin zu sehen, dass das Spektrum der Klassengrößen in der untersuchten Stichprobe zu gering ausfällt. Die große Mehrheit der Klassen bewegt sich zwischen 20 und 30 Schüler/-innen. Frühere Studien legen aber nahe, dass bedeutende lernförderliche Effekte erst bei Klassengrößen von weniger als 15 Schüler/-innen auftreten (Hattie, 2002).

Das *mathematische Ausgangsniveau* ist das einzige Kompositionsmerkmal, für das ein Effekt auf den durchschnittlichen Lernerfolg festgestellt werden kann. Als mögliche Mechanismen kommen ein günstiger Einfluss auf die Unterrichtsgestaltung und anregende Schüler-Schüler-Interaktionen (z. B. gegenseitige Unterstützung) in Betracht. Angesichts der theoretischen Plausibilität ist die Evidenz als *fundierte* einzuschätzen, was sich auch mit den Ergebnissen anderer Studien deckt (Kyriakides et al., 2018b; Reynolds et al., 2014).

Tab. 17: Ergebnisse der bivariaten Mehrgruppenmodelle zu Merkmalen der Lerngruppe (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Intercept-As-Outcome bivariat (nur sozial benachteiligte Schüler/-innen)	Intercept-As-Outcome bivariat (nur sozial privilegierte Schüler/-innen)
Mittelwert Mathematik_5	0.276** (0.100)	0.116 (0.087)
Anteil Migrationshintergrund	-0.656* (0.258)	-0.150 (0.231)

Anmerkungen: Jede Zeile stellt das Ergebnis eines bivariaten Mehrebenenmodells dar.
⁺p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

Beim *mittleren Bildungsstand* der Klasse finden sich dagegen *keine Evidenzen* für Effekte auf den durchschnittlichen Lernerfolg oder für Effekte auf herkunftsbedingte Unterschiede im Lernerfolg. Möglicherweise hängt dies mit der untersuchten Gymnasialstichprobe zusammen. Während in vielen Studien günstige Effekte der sozialen Schülerkomposition festzustellen sind (Van Ewijk & Slegers, 2010b), hatte bereits die Studie von Baumert und Kolleg/-innen (2006) ergeben, dass Effekte der Schülerkomposition im Gymnasium deutlich schwächer ausfallen als in anderen Schulformen.

Die Befunde aus den Slope-As-Outcome-Modellen und Mehrgruppenmodellen sprechen weiterhin mit großer Konsistenz dafür, dass sozial benachteiligte Schüler/-innen stärker von einer günstigen Schülerkomposition profitieren als sozial privilegierte Schüler/-innen. Dies gilt sowohl für das *mathematische Ausgangsniveau der Klasse* als auch für den *Anteil an Schüler/-innen mit Migrationshintergrund*, wobei der Effekt des letztgenannten Kompositionsmerkmals weniger robust ausfällt. Vor diesem Hintergrund liegt beim mathematischen Ausgangsniveau der Klasse *fundierte Evidenz* für einen unterschiedsverringenden Effekt vor, während beim Anteil der Schüler/-innen mit Migrationshintergrund nur *begrenzte Evidenz* vorliegt.

8.5.7 Ergebnisse der Gesamtmodelle zu Effekten von pädagogischen Merkmalen

Abschließend werden Analysen mithilfe integrativer Gesamtmodelle vorgestellt. Während sich die zuvor berichteten Analysen auf jeweils eine Kategorie pädagogischer Merkmale beschränken, leisten die nachfolgenden Analysen eine integrative Modellierung unter gleichzeitiger Berücksichtigung von Unterrichtsmerkmalen, Ganztagsangeboten und Merkmalen der Lerngruppe. Mithilfe dieser Analysen kann überprüft werden, inwieweit pädagogische

Tab. 18: Ergebnisse des Intercept-As-Outcome-Gesamtmodells (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Regressionskoeffizient
Intercept (Mittelwert der klassenspezifischen Intercepts)	1.533*** (0.069)
<i>Prädiktoren Schülerebene</i>	
Mathematikkompetenz_5	0.331*** (0.030)
Mathematiknote	0.174*** (0.038)
Kognitive Grundfähigkeit	0.232*** (0.038)
Lesekompetenz	0.145*** (0.021)
Selbstkonzept Mathematik	0.128*** (0.029)
Geschlecht	-0.223*** (0.048)
Lernbegleitung Kontrolle	-0.070* (0.032)
Bildungsstand	0.035*** (0.009)
Nachhilfe Mathematik	-0.225** (0.085)
Häufigkeit Lesen	0.037* (0.019)
Häufigkeit Lesen ²	-0.037*** (0.014)
Freund/-innen Hauptschulaspiration	-0.039* (0.019)
<i>Prädiktoren Klassenebene: Effekt auf Intercept</i>	
Klassenführung	0.029** (0.009)
Kognitive Aktivierung	0.166* (0.067)
Kognitive Aktivierung ²	-0.082* (0.032)
Individualisierung	-0.112* (0.054)
Hausaufgabenbetreuung	-0.042** (0.013)
Mittelwert Mathematik_5	0.138* (0.069)
<i>Varianzkomponenten</i>	
Varianz Mathematikkompetenz_7 Schülerebene	0.540*** (0.025)
Varianz Mathematikkompetenz_7 Klassenebene (Varianz der klassenspezifischen Intercepts)	0.004 (0.009)

Anmerkungen: Die Korrelationen zwischen den Prädiktoren auf der Schülerebene bewegen sich zwischen $r = -0.292$ und $r = 0.456$. Die Korrelationen zwischen den Prädiktoren auf der Klassenebene bewegen sich zwischen $r = -0.138$ und $r = 0.715$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Merkmale auch unter Kontrolle von pädagogischen Merkmalen aus anderen Kategorien Effekte auf den Lernerfolg der Schüler/-innen haben. Insofern leisten diese Analysen eine besonders strenge Überprüfung der Annahme der Wirksamkeit von pädagogischen Merkmalen. Es wird jeweils ein Intercept-As-Outcome-Gesamtmodell und ein Slope-As-Outcome-Gesamtmodell entwickelt.

Für die Entwicklung des *Intercept-As-Outcome-Gesamtmodells* werden alle pädagogischen Merkmale, bei denen in den kategoriespezifischen Analysen signifikante Effekte auf den Random Intercept festzustellen sind, gemeinsam modelliert.⁴² Anschließend werden nicht signifikante Prädiktoren entfernt, um ein sparsames Modell zu entwickeln. Die Ergebnisse des Intercept-As-Outcome-Gesamtmodells finden sich in Tabelle 18. Indizes der Modellpassung stehen in diesem Modell nicht zur Verfügung (aufgrund der Syntax TYPE = RANDOM). Als robuste Prädiktoren des durchschnittlichen Lernerfolgs erweisen sich Klassenführung (positiver Effekt), kognitive Aktivierung (positiver und kurvilinearereffekt), Individualisierung (negativer Effekt), Hausaufgabenbetreuung (negativer Effekt) und das mathematische Ausgangsniveau der Klasse (positiver Effekt). Somit hat aus allen drei Kategorien jeweils mindestens ein Merkmal einen signifikanten Effekt auf den durchschnittlichen Lernerfolg. Zusammen erklären die genannten Prädiktoren $R^2 = 91.8\%$ der Varianz der klassenspezifischen Intercepts.

Analog werden für die Entwicklung des *Slope-As-Outcome-Gesamtmodells* alle pädagogischen Merkmale mit signifikanten Effekten auf den Random Slope gemeinsam modelliert.⁴³ Nicht signifikante Prädiktoren werden ebenfalls entfernt. Tabelle 19 berichtet die Ergebnisse des Slope-As-Outcome-Gesamtmodells. Auch in diesem Modell werden keine Indizes der Modellpassung ausgegeben. Bezüglich herkunftsbedingter Unterschiede im Lernerfolg stellen leistungsheterogene Gruppen (negativer Effekt), Aufgaben mit Hilfe (negativer Effekt), Besprechung der Hausaufgaben (negativer Effekt), Förderunterricht in Mathematik (negativer Effekt), Förderunterricht für leistungsstarke Schüle-

42 Dabei handelt es sich um die folgenden Variablen: Klassenunterricht, Partner-/Gruppenarbeiten, Klassenführung, kognitive Aktivierung, Besprechung der Hausaufgaben, Individualisierung, Förderunterricht für leistungsstarke Schüler/-innen, Hausaufgabenbetreuung und mathematisches Ausgangsniveau der Klasse.

43 Dabei handelt es sich um die folgenden Variablen: Partner-/Gruppenarbeit, leistungsheterogene Gruppen, Aufgaben mit Hilfe, Besprechung der Hausaufgaben, Förderunterricht in Mathematik, Förderunterricht für leistungsstarke Schüler/-innen, Hausaufgabenbetreuung, Mathematikangebote, mathematisches Ausgangsniveau der Klasse und Anteil an Schüler/-innen mit Migrationshintergrund.

Tab. 19: Ergebnisse des Slope-As-Outcome-Gesamtmodells (unstandardisierte Regressionskoeffizienten mit Standardfehlern in Klammern)

Variable	Regressionskoeffizient
Intercept (Mittelwert der Random Intercepts)	1.523*** (0.029)
<i>Prädiktoren Schülerebene</i>	
Mathematikkompetenz_5	0.518*** (0.026)
Geschlecht	-0.217*** (0.048)
Migrationshintergrund	-0.013 (0.064)
Bildungsstand (Mittelwert der Random Slopes)	0.039*** (0.010)
<i>Prädiktoren Klassenebene: Effekt auf Intercept</i>	
Leistungsheterogene Gruppen	-0.006 (0.048)
Aufgaben mit Hilfe	-0.003 (0.006)
Besprechung Hausaufgaben	-0.005 (0.007)
Förderunterricht Mathematik	0.072 (0.066)
Förderunterricht stark	-0.023 (0.017)
Mathematikangebote	-0.032 (0.019)
Mittelwert Mathematik_5	0.232* (0.096)
Mittelwert Bildungsstand	0.006 (0.040)
<i>Prädiktoren Klassenebene: Effekt auf Slope</i>	
Leistungsheterogene Gruppen	-0.023** (0.009)
Aufgaben mit Hilfe	-0.003⁺ (0.001)
Besprechung Hausaufgaben	-0.004** (0.009)
Förderunterricht Mathematik	-0.028⁺ (0.016)
Förderunterricht stark	0.013* (0.005)
Mathematikangebote	-0.015* (0.006)
Mittelwert Mathematik_5	-0.040⁺ (0.024)
<i>Varianzkomponenten</i>	
Varianz Mathematikkompetenz_7 Schülerebene	0.608*** (0.028)
Varianz Mathematikkompetenz_7 Klassenebene (Varianz der klassenspezifischen Intercepts)	0.032⁺ (0.019)
Varianz des Regressionskoeffizienten des Bildungsstandes (Varianz der klassenspezifischen Slopes)	0.000 (0.005)
Kovarianz der klassenspezifischen Intercepts und Slopes	0.001 (0.005)

Anmerkungen: Die Korrelationen zwischen den Prädiktoren auf der Schülerebene bewegen sich zwischen $r = -0.161$ und $r = 0.117$. Die Korrelationen zwischen den Prädiktoren auf der Klassenebene bewegen sich zwischen $r = -0.239$ und $r = 0.364$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

r/-innen (positiver Effekt), Mathematikangebote (negativer Effekt) und das mathematische Ausgangsniveau der Klasse (negativer Effekt) robuste Prädiktoren dar. Auch bei herkunftsbedingten Unterschieden im Lernerfolg hat aus allen drei Kategorien jeweils mindestens ein Merkmal einen signifikanten Effekt. Die Prädiktoren erklären zusammen $R^2 = 78.2\%$ der Varianz der klassenspezifischen Slopes.

9. Abschließende Zusammenfassung und Diskussion

Im letzten Kapitel der vorliegenden Arbeit erfolgt eine abschließende Zusammenfassung und Diskussion. Einleitend sollen die zentralen Merkmale der vorliegenden Arbeit noch einmal kurz skizziert werden. Den Ausgangspunkt für die Untersuchung bildeten die wiederholt festgestellten Kompetenzunterschiede zwischen Schüler/-innen verschiedener sozialer Herkunft. Sowohl bezüglich der Ursachen dieser Kompetenzunterschiede als auch bezüglich geeigneter Maßnahmen zur Verringerung dieser Kompetenzunterschiede besteht Forschungsbedarf. Mit den empirischen Analysen wurde versucht, den Wissensstand in beiden Forschungsfeldern zu erweitern. Die Fragestellungen fokussierten auf die vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg, auf die Effekte von Klassen und auf die Effekte von pädagogischen Merkmalen. Zur Bearbeitung dieser Fragestellungen wurde eine Längsschnittanalyse auf der Grundlage des Nationalen Bildungspanels durchgeführt. Die Analysen beziehen sich auf eine für Deutschland repräsentative Stichprobe von Gymnasiast/-innen im Zeitraum zwischen der fünften und siebten Jahrgangsstufe.

Im nächsten Abschnitt wird zunächst die gewählte Forschungsstrategie bewertet (Kapitel 9.1). Danach werden die Ergebnisse zusammengefasst, in den vorhandenen Forschungsstand eingeordnet und bewertet. Dies erfolgt getrennt nach den drei Bereichen von Fragestellungen (Kapitel 9.2 bis 9.4). Nachfolgend werden ausgewählte Befunde einander vergleichend gegenübergestellt (Kapitel 9.5). Abschließend wird erörtert, welche Implikationen die berichteten Ergebnisse für die pädagogische Praxis und Bildungspolitik (Kapitel 9.6) sowie für die empirische Bildungsforschung (Kapitel 9.7) haben.

9.1 Bewertung der Forschungsstrategie

Im Weiteren wird die verwendete Forschungsstrategie einer systematischen Bewertung unterzogen. Die Bewertung orientiert sich an den im zweiten Kapitel formulierten Standards für Wirksamkeitsstudien. Auf dieser Grundlage werden die Stärken und Limitationen der Forschungsstrategie herausgearbeitet.

Eine Stärke der vorliegenden Arbeit ist darin zu sehen, dass die Durchführung der Untersuchung auf einer soliden *theoretischen Grundlage* erfolgte. Zentrale Begriffe wurden angemessen definiert und konzeptualisiert. Weiterhin wurden zwei theoretische Modelle entwickelt, die eine Ordnung von Variablen leisteten und eine theoriegeleitete Auswahl der Untersuchungsvariablen ermöglichten. Außerdem wurden sowohl für den durchschnittlichen Lernerfolg als auch für herkunftsbedingte Unterschiede im Lernerfolg theoretische Überlegungen zu Wirkmechanismen angestellt und bei der Interpretation der empirischen Befunde aufgegriffen.

Der Standard, *mehrere Dimensionen pädagogischer Wirksamkeit* gleichzeitig zu berücksichtigen, kann als mehrheitlich erfüllt angesehen werden. Pädagogische Effekte wurden bezüglich der Zielkriterien eines hohen durchschnittlichen Lernerfolgs und geringer herkunftsbedingter Unterschiede im Lernerfolg betrachtet. Einhergehend mit dem letztgenannten Zielkriterium wurden pädagogische Effekte bei sozial privilegierten und sozial benachteiligten Schüler/-innen einander vergleichend gegenübergestellt. Es kann eingewendet werden, dass die Berücksichtigung weiterer Dimensionen möglich und wünschenswert gewesen wäre (z. B. Berücksichtigung von motivationalen Lernzielen, Berücksichtigung verschiedener Schulformen). Der zusätzliche Arbeitsaufwand hat hierbei pragmatische Grenzen gesetzt.

Da es sich bei der vorliegenden Untersuchung um eine multivariate Längsschnittstudie handelt, liegt ein für die Analyse von pädagogischer Wirksamkeit *geeignetes Forschungsdesign* vor. Es wurden zahlreiche nicht schulische Einflussfaktoren auf Lernprozesse als Kontrollvariablen berücksichtigt. Insofern ist anzunehmen, dass die Schätzung der pädagogischen Effekte relativ präzise erfolgte. Auch die verwendeten *Auswertungsmethoden* sind als geeignet einzuschätzen. Mehrebenenmodelle, Pfadmodelle und Strukturgleichungsmodelle sind angemessene statistische Verfahren für die Analyse der verwendeten Daten. Positiv ist außerdem hervorzuheben, dass die genannten Verfahren in vielfältiger Weise für die *Überprüfung komplexer Wirkzusammenhänge* genutzt wurden. Neben linearen Effekten wurden auch kurvilineare Effekte und Interaktionseffekte überprüft. Kritisch anzumerken ist, dass nur direkte Effekte von pädagogischen Merkmalen überprüft wurden, während indirekte Effekte von pädagogischen Merkmalen außer Acht gelassen wurden (z. B. inwieweit Merkmale der Lerngruppe vermittelt über die Unterrichtsgestaltung wirken).

Eine zentrale Limitation betrifft die verwendeten *Erhebungsmethoden*. Während die Erfassung von Schülermerkmalen mithilfe von Tests bzw. Fragebögen sowie die Erfassung von Familienmerkmalen mithilfe von Telefoninterviews mit den Eltern angemessen erscheinen, ist die Erfassung von Un-

terrichtsmerkmalen und Ganztagsangeboten ausschließlich durch Befragung von Lehrkräften und Schulleiter/-innen als ungünstig anzusehen. Für die Erfassung von pädagogischen Prozessmerkmalen kommen als Datenquellen neben Befragungen von pädagogischen Fachkräften auch Schülerbefragungen und Beobachtungen in Frage. Bei der direkten Gegenüberstellung der drei Datenquellen werden insbesondere Schülerbefragungen und Beobachtungen als geeignete Erhebungsmethoden angesehen (Creemers et al., 2010, S. 32; Göllner et al., 2016, S. 63; Helmke, 2014, S. 292). Demgegenüber wird Befragungen von pädagogischen Fachkräften nur eine begrenzte Aussagekraft zugeschrieben, weil die Gefahr besteht, dass pädagogische Fachkräfte ihre Angaben aus Gründen des Selbstwertschutzes oder der sozialen Erwünschtheit verzerren (Göllner et al., 2016, S. 65). Somit wird in der vorliegenden Arbeit nur eine einseitige und enggeführte Erfassung von Unterricht und Ganztagsangeboten geleistet.

Neben der grundsätzlichen Einschränkung, dass nur bestimmte Datenquellen und Erhebungsmethoden zur Verfügung stehen, ist auch die konkrete *Ausgestaltung der Erhebungen* kritisch zu sehen. Ein großer Teil der pädagogischen Merkmale wird durch Einzelitems erfasst und nur ein kleinerer Teil durch Skalen. Bei den Unterrichtsmerkmalen dominieren in inhaltlicher Hinsicht quantitative Aspekte der Unterrichtsgestaltung (Häufigkeit und Dauer von Unterrichtsmethoden), während sich nur wenige Unterrichtsmerkmale auf qualitative Aspekte der Unterrichtsgestaltung beziehen (Art und Weise der Umsetzung von Unterrichtsmethoden). Bei den Ganztagsangeboten werden lediglich die Häufigkeiten erfasst und andere relevante Aspekte bleiben unberücksichtigt (z. B. methodische Ausgestaltung, Zusammensetzung der Lerngruppe).

9.2 Diskussion der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Nachfolgend werden die Ergebnisse der Untersuchung diskutiert. Begonnen wird mit dem ersten Bereich von Fragestellungen, welcher sich auf vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg bezieht (Fragestellungen Fl.1 bis Fl.6). Zunächst werden die Ergebnisse zusammengefasst. Danach werden die Ergebnisse in den vorhandenen Forschungsstand eingeordnet. Anschließend werden die Ergebnisse bewertet, wobei der Neuheitswert und die Generalisierbarkeit der Befunde im Vordergrund stehen.

9.2.1 Zusammenfassung der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Für die Zusammenfassung der Ergebnisse werden die untersuchten Variablen danach kategorisiert, ob *fundierte Evidenz* (überzeugende Belege), *begrenzte Evidenz* (gewisse Hinweise), *keine Evidenz* (keine Hinweise) oder *unklare Evidenz* (mehrdeutige Interpretation) für einen Mediationseffekt vorliegt. Eine übersichtliche Darstellung bietet Tabelle 20.

Fundierte Evidenzen für einen Mediationseffekt liegen nur für eine Variable vor, nämlich die mathematische Ausgangskompetenz in der fünften Jahrgangsstufe. Die mathematische Ausgangskompetenz korreliert jeweils positiv mit der sozialen Herkunft und der Mathematikkompetenz in der siebten Jahrgangsstufe.

Begrenzte Evidenzen für einen Mediationseffekt liegen bei der Häufigkeit des Lesens, der Häufigkeit des Musizierens, dem Anteil an Freund/-innen mit Realschulaspiration, der metakognitiven Kompetenz und der Lesekompetenz vor. Vier der fünf Variablen weisen jeweils eine positive Korrelation zur sozialen Herkunft und zur Mathematikkompetenz in der siebten Jahrgangsstufe auf. Nur der Anteil an Freund/-innen mit Realschulaspiration korreliert jeweils negativ mit der sozialen Herkunft und der Mathematikkompetenz in der siebten Jahrgangsstufe.

Keine Evidenzen für einen Mediationseffekt liegen bei der Häufigkeit der elterlichen Lernbegleitung, einer kontrollierenden Lernbegleitung, der Nutzung von fachbezogenen Angeboten, der Häufigkeit von Sport, den Anteilen von Freund/-innen mit Hauptschul- und Gymnasialaspiration, dem Interesse an Mathematik und dem mathematikbezogenen Selbstkonzept vor.

Unklare Evidenzen liegen bei Unterstützung und Erwartungen durch die Eltern, bei der Nutzung von Förderunterricht, Hausaufgabenbetreuungen und Nachhilfe sowie der Gewissenhaftigkeit vor.

Zusammenfassend ist festzuhalten, dass in der untersuchten Gymnasialstichprobe die kognitiven Lernvoraussetzungen der Schüler/-innen, anregende Aktivitäten außerhalb der Schule und die Zusammensetzung des Freundeskreises empirische Bestätigung als vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg finden.

Tab. 20: Evidenzen für Mediationseffekte zwischen sozialer Herkunft und Lernerfolg

Stufe der Evidenz	Variable	Art des Effekts
<i>Fundierte Evidenz</i>	Mathematikkompetenz_5	Beide Korrelationen positiv
<i>Begrenzte Evidenz</i>	Häufigkeit Lesen	Beide Korrelationen positiv
	Häufigkeit Musizieren	Beide Korrelationen positiv
	Freund/-innen Realschulaspisation	Beide Korrelationen negativ
	Metakognitive Kompetenz	Beide Korrelationen positiv
	Lesekompetenz	Beide Korrelationen positiv
<i>Keine Evidenz</i>	Lernbegleitung Häufigkeit	
	Lernbegleitung Kontrolle	
	Nutzung fachbezogene Angebote	
	Häufigkeit Sport	
	Freund/-innen Hauptschulaspisation	
	Freund/-innen Gymnasialaspisation	
	Interesse Mathematik	
	Selbstkonzept Mathematik	
<i>Unklare Evidenz</i>	Lernbegleitung Unterstützung	
	Lernbegleitung Erwartungen	
	Nutzung Förderunterricht	
	Nutzung Hausaufgabenbetreuung	
	Nachhilfe Mathematik	
	Gewissenhaftigkeit	

9.2.2 Einordnung der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg in den Forschungsstand

Nachfolgend werden die Ergebnisse der vorliegenden Untersuchung den Ergebnissen bereits vorhandener Studien gegenübergestellt. Hierbei wird auf die Studien Bezug genommen, welche im Forschungsüberblick im vierten Kapitel bereits dargestellt wurden.

Zu der Frage, ob die *elterliche Lernbegleitung* einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt, fallen die vorhandenen Befunde recht unterschiedlich aus. Sowohl bei der Häufigkeit der elterlichen Lernbegleitung (Autorengruppe Bildungsberichterstattung, 2014; Dumont, 2012; Ehmke et al., 2006b; Schauenberg, 2007; Süßlin, 2015) als auch bei einer kontrollierenden Lernbegleitung durch die Eltern (Dumont, 2012; Schauenberg, 2007; Trautwein et al., 2001) fallen die Ergebnisse inkonsistent

aus, d. h., ob und welche Zusammenhänge zum sozialen Status bestehen, variiert zwischen den Studien. Die Ergebnisse der vorliegenden Untersuchung verstärken die Zweifel an der Annahme, dass diese Merkmale systematisch vom sozialen Status bedingt werden. Konsistentere Befunde liegen beim Zusammenhang zwischen einer unterstützenden Lernbegleitung der Eltern und ihrem sozialen Status vor (Cooper et al., 2000; Dumont, 2012). Hinweise auf dieses Muster finden sich auch in der vorliegenden Untersuchung. Lernförderliche Effekte einer unterstützenden Lernbegleitung, wie sie andere Studien ermitteln (Dumont, 2012; Patall et al., 2008), können in der vorliegenden Untersuchung allerdings nicht bestätigt werden. Es erscheint denkbar, dass dies auf methodische Limitationen zurückzuführen ist (d. h., mit den verwendeten Kontrollvariablen können reziproke Effekte vermutlich nicht angemessen modelliert werden).

Bezüglich der *Nutzung von Ganztagsangeboten* fügen sich die Ergebnisse der vorliegenden Untersuchung gut in den bestehenden Forschungsstand ein. Die Befunde vorhandener Studien, wonach leistungsschwächere bzw. sozial benachteiligte Schüler/-innen häufiger Förderunterricht und Hausaufgabenbetreuung nutzen als leistungsstärkere bzw. sozial privilegierte Schüler/-innen (Linberg et al., 2015; Züchner & Fischer, 2014), zeigen sich auch in der vorliegenden Untersuchung. Lernförderliche Effekte der Nutzung dieser Ganztagsangebote können in den vorhandenen Studien teilweise ermittelt werden (Bellin & Wegner, 2010), teilweise dagegen nicht (Linberg et al., 2018; Steinmann et al., 2019). Der zweitgenannte Fall liegt auch in der vorliegenden Untersuchung vor. Dies sollte allerdings nicht als zunehmende Evidenz für die Unwirksamkeit dieser Angebote interpretiert werden, sondern als erneuter Hinweis auf methodische Herausforderungen bei der Modellierung von reziproken Effekten.

Die bisher vorhandenen Studien zu *Nachhilfe* berichten ausgesprochen heterogene Befunde bezüglich der Frage, ob Nutzungsunterschiede nach sozialer Herkunft bestehen (Abele & Liebau, 1998; Guill & Wendt, 2016; Hille et al., 2016; Park et al., 2016; Luplow & Schneider, 2014; Schneider, 2005). Die vorliegende Untersuchung stärkt die Einschätzung, dass sozial benachteiligte Schüler/-innen häufiger Nachhilfe nutzen als sozial privilegierte Schüler/-innen. Für eine abschließende Einschätzung werden allerdings weitere Studien und insbesondere Analysen zu moderierenden Variablen benötigt. Auch in Bezug auf lernförderliche Effekte fallen die vorhandenen Befunde sehr heterogen aus (Guill et al., 2020; Guill & Bos, 2014; Haag, 2001; Hosenfeld, 2011; Luplow & Schneider, 2014; Mischo & Haag, 2002; Ömerogullari et al., 2020). Dass in der vorliegenden Untersuchung keine lernförderlichen Effekte festgestellt werden

können, verweist wiederum auf die Möglichkeit reziproker Effekte und daraus resultierende methodische Schwierigkeiten.

Im Bereich der *Freizeitaktivitäten* zeigt sich in einer Reihe von Studien, dass sozial privilegierte Schüler/-innen häufiger anregende Freizeitaktivitäten ausüben als sozial benachteiligte Schüler/-innen (Autorengruppe Bildungsberichterstattung, 2014; DeGarmo et al., 1999; Leven & Schneekloth, 2010; Schauenberg, 2007). Weiterhin liefern vorhandene Studien Hinweise auf lernförderliche Effekte von Lesen und Musizieren (Pfost et al., 2010; Stutz et al., 2016; Vaughn, 2000). Die pfadanalytischen Ergebnisse der vorliegenden Untersuchung fügen sich gut in diese Befundlage ein. Das Ergebnis, dass kein Zusammenhang zwischen Sport und Lernerfolg besteht, verweist angesichts der bisher eher heterogenen Befundlage (Singh et al., 2012; Taras, 2005) auf weiteren Forschungsbedarf.

In vorhandenen Studien zur *Gleichaltrigen* finden sich mehrfach Hinweise darauf, dass sozial privilegierte Schüler/-innen mehr Freund/-innen mit positiven Einstellungen zur Schule bzw. hohen Bildungsaspirationen haben (Behtui & Neergaard, 2016; Carolan, 2016) und dass sich diese Art von Freund/-innen förderlich auf das Lernen auswirkt (Behtui & Neergaard, 2016; Johnson, 2000; Schmitt, 2012). Die Ergebnisse der vorliegenden Untersuchung passen grundsätzlich zu dieser Befundlage. Allerdings können Mediationseffekte nur für ein Merkmal der Gleichaltrigen bestätigt werden und hierbei scheint es eher eine Rolle zu spielen, dass sozial privilegierte Schüler/-innen weniger Freund/-innen mit niedrigen Bildungsaspirationen haben. Gleichwohl nehmen die Evidenzen dafür zu, dass die Gleichaltrigen einen vermittelnden Mechanismus zwischen sozialer Herkunft und Lernerfolg darstellt.

Im Bereich der Lernvoraussetzungen der Schüler/-innen zeigt eine umfangreiche und konsistente Befundlage, dass die *kognitiven Lernvoraussetzungen* von Schüler/-innen mit ihrer sozialen Herkunft zusammenhängen (Karlen et al., 2014; Schneider et al., 2017; Stahl, 2009; Steinmayr et al., 2010; Steinmayr et al., 2012; Weinert & Ebert, 2013) und ihrerseits den Lernerfolg bedingen (Alloway & Passolunghi, 2011; Hornung et al., 2014; Marzano, 2000; Ohtani & Hisaki, 2018; Steinmayr et al., 2010; Steinmayr et al., 2012). Die vorliegende Untersuchung bestätigt dieses Muster erneut mithilfe von Pfadmodellen.

In Bezug auf *motivationale Lernvoraussetzungen* der Schüler/-innen sind Herkunftseffekte in mehreren Studien (Dotterer et al., 2009; Karlen et al., 2014; Kriegaum & Spinath, 2016; Steinmayr et al., 2012), aber nicht in allen Studien (Caro, 2018; Yerdelen-Demar & Pesman, 2013) festzustellen. Die Ergebnisse der vorliegenden Untersuchung verstärken die Zweifel an einem generellen Zusammenhang zwischen Lernmotivation und sozialer Herkunft. Die Befund-

lage wird durch die vorliegende Untersuchung eher noch inkonsistenter, was die Frage nach moderierenden Variablen aufwirft. Interessant ist hierbei die Gegenüberstellung der vorliegenden Untersuchung mit der Untersuchung von Steinmayr und Kolleg/-innen (2012), da die Stichproben in beiden Fällen aus Gymnasiast/-innen bestehen. Hieran wird deutlich, dass die inkonsistente Befundlage nicht alleine durch die Betrachtung unterschiedlicher Schulformen erklärt werden kann.

Hinsichtlich der *Gewissenhaftigkeit* der Schüler/-innen sind in vorhandenen Studien Effekte der sozialen Herkunft insbesondere bei der Subfacette Selbstregulation festzustellen (Kaiser & Diewald, 2014; Miech et al., 2001; Sammons et al., 2014; Steinmayr et al., 2010). Weiterhin wurden mehrfach günstige Effekte von Gewissenhaftigkeit auf den Lernerfolg ermittelt (O'Connor & Pannonen, 2007; Poropat, 2009). Die entsprechenden Zusammenhänge können in der vorliegenden Untersuchung nicht bestätigt werden. Da die Erfassung von Gewissenhaftigkeit in der vorliegenden Untersuchung als ungünstig einzuschätzen ist, erscheint es denkbar, dass methodische Limitationen hierfür ausschlaggebend sind.

9.2.3 Bewertung der Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Im Weiteren werden die Ergebnisse zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg einer kritischen Bewertung unterzogen. Es wird erörtert, inwieweit die Ergebnisse Neuheitswert haben, d.h., einen Erkenntnisfortschritt leisten. Danach wird die Generalisierbarkeit der Ergebnisse diskutiert.

Bei der Einschätzung des *Neuheitswerts* der Befunde ist zu berücksichtigen, dass von den bisher verfügbaren Studien nur ein kleinerer Teil als methodisch hochwertig angesehen werden kann. Die Ergebnisse der vorliegenden Untersuchung haben im Vergleich zu vielen früheren Studien eine höhere Aussagekraft, da ein längsschnittliches Forschungsdesign und angemessene Analysemethoden verwendet werden. Das Verständnis für die Ursachen von herkunftsbedingten Kompetenzunterschieden wird sowohl dadurch erweitert, dass bestimmte Variablen als Mediatoren bestätigt werden können (z. B. kognitive Lernvoraussetzungen, Freizeitaktivitäten, Gleichaltrigengruppe) als auch dadurch, dass bestimmte Variablen nicht als Mediatoren bestätigt werden können (z. B. Häufigkeit der elterlichen Lernbegleitung, Lernmotivation).

Während diese Befunde durchaus einen gewissen Erkenntnisfortschritt leisten, sind eine Reihe von *Limitationen* festzustellen. Bei mehreren Variablen

fällt die Evidenz aufgrund von methodischen Limitationen unklar aus (z. B. Unterstützung und Erwartungen der Eltern, Nachhilfe). Das erreichte Niveau an Evidenz ist mehrheitlich nur als begrenzt einzuschätzen. Zudem handelt es sich bei drei der sechs Variablen, bei denen Evidenz für einen vermittelnden Mechanismus vorliegt, um kognitive Lernvoraussetzungen von Schüler/-innen. Die hieraus resultierende Frage, auf welche Ursachen die herkunftsbedingten Unterschiede in den kognitiven Lernvoraussetzungen zurückzuführen sind, vermag die vorliegende Untersuchung nicht zu beantworten.

Im Rahmen der Bewertung ist auch die *Generalisierbarkeit* der Befunde zu betrachten. Generalisierbarkeit bezieht sich darauf, inwieweit die für eine deutsche Gymnasialstichprobe ermittelten Befunde auf andere Länder, Schulformen und Schülergruppen übertragen werden können. In Bezug auf vermittelnde Mechanismen stellt sich insbesondere die Frage, inwieweit die Befunde, die sich primär auf die Gegenüberstellung von Schüler/-innen aus Familien mit mittlerem und hohem Sozialstatus beziehen, auch Gültigkeit für die Gegenüberstellung von Schüler/-innen aus Familien mit niedrigem und mittlerem Sozialstatus haben. Die Generalisierbarkeit ist bei empirisch bestätigten und empirisch nicht bestätigten Mediatoren unterschiedlich einzuschätzen. Bei den hier *bestätigten Mediatoren* erscheint es plausibel, dass sie auch für andere Kontexte gültig sind (d. h., die kognitiven Lernvoraussetzungen, die Freizeitaktivitäten und die Gleichaltrigengruppe stellen auch in anderen Kontexten vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg dar). Demgegenüber erscheint es denkbar, dass die hier *nicht bestätigten Mediatoren* in anderen Kontexten durchaus Mediatoren darstellen (d. h., die Häufigkeit der elterlichen Lernbegleitung und motivationale Lernvoraussetzungen stellen in anderen Kontexten möglicherweise vermittelnde Mechanismen zwischen sozialer Herkunft und Lernerfolg dar). Die Befunde sind daher nur als begrenzt generalisierbar einzuschätzen.

9.3 Diskussion der Ergebnisse zu Effekten von Klassen

Im Weiteren werden die Ergebnisse zu Effekten von Klassen diskutiert (Fragestellungen F2.1 bis F2.3). Wie im vorausgehenden Abschnitt werden die Ergebnisse zusammengefasst, in den bestehenden Forschungsstand eingeordnet und bewertet.

9.3.1 Zusammenfassung der Ergebnisse zu Effekten von Klassen

Zunächst werden die Ergebnisse zu den Effekten von Klassen zusammengefasst. Es liegen *fundierte Evidenzen* für bedeutsame *relative Klasseneffekte auf den durchschnittlichen Lernerfolg* vor (Fragestellung F2.1). Demnach hängt der Lernerfolg von Schüler/-innen wesentlich davon ab, welche Klasse sie besuchen. Weiterhin liegen *begrenzte Evidenzen* dafür vor, dass relative Klasseneffekte bei sozial benachteiligten Schüler/-innen *stärker* ausfallen als bei sozial privilegierten Schüler/-innen. Bei der zweitgenannten Frage variieren die Ergebnisse in Abhängigkeit davon, ob der Bildungsstand oder die Berufsposition der Eltern betrachtet wird.

In Bezug auf *relative Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg* (Fragestellung F2.2) liefern die Analysen *begrenzte Evidenzen* dafür, dass Kompetenzunterschiede nach sozialer Herkunft bedeutsam zwischen Klassen variieren. Inferenzstatistisch lassen sich Unterschiede zwischen Klassen nur bei Betrachtung von unadjustierten Kompetenzunterschieden absichern, nicht aber bei Betrachtung von adjustierten Kompetenzunterschieden. Unter inhaltlichen Gesichtspunkten sind die Unterschiede zwischen Klassen allerdings auch im zweitgenannten Fall als relevant anzusehen. Demnach können herkunftsbedingte Kompetenzunterschiede durch eine angemessene Ausgestaltung von Schule und Unterricht bis zu einem gewissen Grad verringert werden. Die vollständige Beseitigung von Kompetenzunterschieden erscheint allerdings nicht möglich.

Bezüglich des *Verhältnisses zwischen durchschnittlichem Lernerfolg und herkunftsbedingten Unterschieden im Lernerfolg* (Fragestellung F2.3) finden sich keine Hinweise darauf, dass zwischen den Zielsetzungen eines hohen durchschnittlichen Lernerfolgs und geringer herkunftsbedingter Unterschiede im Lernerfolg ein Zielkonflikt besteht. Die Erreichung der beiden Ziele scheint relativ unabhängig voneinander zu sein. In der Tendenz zeichnet sich ab, dass sich die Erreichung des einen Ziels günstig auf die Erreichung des anderen Ziels auswirkt. Die Befunde legen nahe, dass geringe herkunftsbedingte Unterschiede im Lernerfolg eher durch die Förderung von sozial benachteiligten Schüler/-innen erreicht werden und weniger durch die Beeinträchtigung von sozial privilegierten Schüler/-innen.

9.3.2 Einordnung der Ergebnisse zu Effekten von Klassen in den Forschungsstand

Hinsichtlich *relativer Klasseneffekte auf den durchschnittlichen Lernerfolg* steht eine umfangreiche und konsistente Befundlage zur Verfügung. Demnach hängt der Lernerfolg von Schüler/-innen wesentlich davon ab, welche Klasse sie besuchen (Luyten, 2003; Marzano, 2000; Scheerens et al., 1989). Die Ergebnisse der vorliegenden Untersuchung fügen sich gut in dieses Bild ein. Die Größe der Klasseneffekte fällt in der untersuchten Stichprobe zwar eher klein, aber dennoch plausibel aus. Eine mögliche Erklärung für die vergleichsweise kleinen Effekte ist darin zu sehen, dass es sich bei Gymnasiast/-innen um eine Schülergruppe handelt, die über günstige Lernvoraussetzungen verfügt.

Während sich zahlreiche Studien mit Klasseneffekten auf den durchschnittlichen Lernerfolg befassen, konnten im Rahmen der Literaturrecherche nur zwei Studien ermittelt werden, die Klasseneffekte auf den durchschnittlichen Lernerfolg *bei sozial privilegierten Schüler/-innen und bei sozial benachteiligten Schüler/-innen vergleichend gegenüberstellen*. Sowohl in der Studie von Pfeifer (2011) als auch in der Studie von Nye und Kolleg/-innen (2004) fallen Klasseneffekte bei sozial benachteiligten Schüler/-innen größer aus als bei sozial privilegierten Schüler/-innen. Die Befunde der vorliegenden Untersuchung liefern ebenfalls Hinweise in diese Richtung. Somit nehmen die Evidenzen für unterschiedlich starke Klasseneffekte bei Schüler/-innen verschiedener sozialer Herkunft zu. Gleichwohl besteht angesichts der überschaubaren Anzahl an Studien weiterer Forschungsbedarf.

Relative Klasseneffekte auf herkunftsbedingte Unterschiede im Lernerfolg werden in einer Reihe von Studien untersucht. Die Mehrzahl der Studien ermittelt, dass das Ausmaß an Kompetenzunterschieden nach sozialer Herkunft signifikant zwischen Klassen variiert (Kyriakides et al., 2018b; Nye et al., 2004; Pfeifer, 2011; Schwippert, 2001; Stahl, 2009; Stancel-Piatak, 2017), während zwei Studien nicht signifikante Unterschiede berichten (Carigiet Reinhard, 2012; Schulz-Heidorf, 2016). Die durch die vorliegende Untersuchung gelieferten Hinweise auf bedeutsame Unterschiede zwischen Klassen verstärken die Einschätzung, dass Klassen einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg haben.

Hinsichtlich der *Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg* stehen bisher nur wenige Studien zur Verfügung. Drei Studien untersuchen das Verhältnis der beiden Zielsetzungen auf der Klassenebene. Die Studien von Schwippert (2001) und Kyriakides und Kolleg/-innen (2018b) legen die Vereinbarkeit der beiden

Zielsetzungen nahe, während die Studie von Carigiet Reinhard (2012) einen Zielkonflikt nahelegt. Die Ergebnisse der vorliegenden Studie sprechen für unabhängige Zielsetzungen, mit Hinweisen auf vereinbare Zielsetzungen. In den Befunden zeichnet sich somit eine Tendenz zu vereinbaren Zielsetzungen ab (oder zumindest zu nicht konfligierenden Zielsetzungen). Angesichts der geringen Anzahl an Studien erscheint eine abschließende Einschätzung allerdings nicht möglich, was den Bedarf nach weiteren Studien verdeutlicht.

9.3.3 Bewertung der Ergebnisse zu Effekten von Klassen

Angesichts des umfangreichen Forschungsstandes zu *Klasseneffekten auf den durchschnittlichen Lernerfolg* haben die Ergebnisse der vorliegenden Untersuchung kaum Neuheitswert, sondern stellen eher Replikationen bereits vorhandener Befunde dar. Neue Erkenntnisse liefert die vorliegende Arbeit vor allem zu Klasseneffekten mit Bezug zu *herkunftsbedingten Unterschieden im Lernerfolg*. Die Ergebnisse, dass Klasseneffekte auf den durchschnittlichen Lernerfolg bei Schüler/-innen verschiedener sozialer Herkunft unterschiedlich stark ausfallen, dass herkunftsbedingte Unterschiede im Lernerfolg bedeutsam zwischen Klassen variieren und dass kein Zielkonflikt zwischen den Zielsetzungen eines hohen durchschnittlichen Lernerfolgs und geringer herkunftsbedingter Unterschiede im Lernerfolg besteht, stellen jeweils wichtige Erweiterungen des empirischen Kenntnisstandes dar.

Unter dem Gesichtspunkt der *Generalisierbarkeit* ist darauf hinzuweisen, dass nicht mit Sicherheit davon ausgegangen werden kann, dass die hier berichteten Befunde auf andere Schulformen und Länder übertragen werden können. Das liegt weniger am Forschungsdesign der vorliegenden Untersuchung als vielmehr daran, dass die bisher vorhandenen Befunde zu Klasseneffekten auf herkunftsbedingte Unterschiede im Lernerfolg und zur Vereinbarkeit eines hohen durchschnittlichen Lernerfolgs mit geringen herkunftsbedingten Unterschieden im Lernerfolg nicht ganz konsistent ausfallen. Ob diese inkonsistente Befundlage auf methodische Unterschiede zwischen den Studien oder auf institutionelle Unterschiede zwischen den untersuchten Schulformen zurückzuführen ist, ist gegenwärtig unklar. Der zweitgenannte Fall verweist auf die Möglichkeit, dass institutionelle Rahmenbedingungen von Schulformen eine moderierende Funktion haben. Dies würde die Generalisierbarkeit der Befunde bedeutsam einschränken.

9.4 Diskussion der Ergebnisse zu Effekten von pädagogischen Merkmalen

Nachfolgend werden die Ergebnisse zu Effekten von pädagogischen Merkmalen diskutiert (Fragestellungen F3.1 bis F3.6), wiederum unterteilt in Zusammenfassung, Einordnung in den Forschungsstand und Bewertung.

9.4.1 Zusammenfassung der Ergebnisse zu Effekten von pädagogischen Merkmalen

Zunächst werden die Befunde zu den Effekten von pädagogischen Merkmalen zusammengefasst. Der nachfolgende Überblick kategorisiert die pädagogischen Merkmale differenziert nach Zielkriterien und Evidenzstufen.

Zunächst werden Effekte auf den *durchschnittlichen Lernerfolg* betrachtet (siehe Tabelle 21). *Fundierte Evidenzen* für einen Effekt auf den durchschnittlichen Lernerfolg liegen bei Klassenführung (positiver Effekt), kognitiver Aktivierung (positiver und kurvilinearereffekt) und dem mathematischen Ausgangsniveau der Klasse (positiver Effekt) vor.

Begrenzte Evidenzen für einen Effekt auf den durchschnittlichen Lernerfolg liegen bei Klassenunterricht (kurvilinearereffekt), Partner-/Gruppenarbeit (kurvilinearereffekt), Individualisierung (negativer Effekt) und Förderunterricht für leistungsstarke Schüler/-innen (negativer Effekt) vor.

Keine Evidenzen für einen Effekt auf den durchschnittlichen Lernerfolg liegen beim Unterrichtsumfang, Einzelarbeit, leistungsheterogenen und leistungshomogenen Gruppen, Übung und Wiederholung, Aufgaben mit und ohne Hilfe, Förderunterricht in Mathematik, Förderunterricht für leistungsstarke Schüler/-innen, Mathematikangeboten, der Klassengröße, dem mittleren Bildungsstand der Klasse und dem Anteil an Schüler/-innen mit Migrationshintergrund vor.

Unklare Evidenzen liegen bei Hausaufgabenbesprechung und Hausaufgabenbetreuung vor.

Die Evidenzen für Effekte von pädagogischen Merkmalen auf *herkunftsbedingte Unterschiede im Lernerfolg* werden in Tabelle 22 zusammengefasst. *Fundierte Evidenzen* für einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg liegen bei Aufgaben mit Hilfe (negativer Effekt), Förderunterricht in Mathematik (negativer Effekt) und dem mathematischen Ausgangsniveau der Klasse (negativer Effekt) vor.

Begrenzte Evidenzen für einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg liegen bei Partner-/Gruppenarbeit (negativer Effekt), leistungsheterogenen Gruppen (negativer Effekt), Besprechung der Hausaufgaben (ne-

Tab. 21: Evidenzen für Effekte von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg

Stufe der Evidenz	Variable	Art des Effekts
<i>Fundierte Evidenz</i>	Klassenführung	Positiv
	Kognitive Aktivierung	Positiv und kurvilinear
	Mittelwert Mathematik_5	Positiv
<i>Begrenzte Evidenz</i>	Klassenunterricht	Kurvilinear
	Partner-/Gruppenarbeit	Kurvilinear
	Individualisierung	Negativ
	Förderunterricht stark	Negativ
<i>Keine Evidenz</i>	Unterrichtsumfang	
	Einzelarbeit	
	Leistungsheterogene Gruppen	
	Leistungshomogene Gruppen	
	Übung und Wiederholung	
	Aufgaben mit Hilfe	
	Aufgaben ohne Hilfe	
	Förderunterricht Mathematik	
	Förderunterricht schwach	
	Mathematikangebote	
	Klassengröße	
Mittelwert Bildungsstand		
Anteil Migrationshintergrund		
<i>Unklare Evidenz</i>	Besprechung Hausaufgaben	
	Hausaufgabenbetreuung	

gativer Effekt), Förderunterricht für leistungsstarke Schüler/-innen (positiver Effekt), Hausaufgabenbetreuung (positiver Effekt), Mathematikangeboten (negativer Effekt) und dem Anteil an Schüler/-innen mit Migrationshintergrund (positiver Effekt) vor.

Keine Evidenzen für einen Effekt auf herkunftsbedingte Unterschiede im Lernerfolg liegen beim Unterrichtsumfang, Klassenunterricht, Einzelarbeit, leistungshomogenen Gruppen, Klassenführung, kognitiver Aktivierung, Übung und Wiederholung, Aufgaben ohne Hilfe, Individualisierung, Förderunterricht für leistungsschwache Schüler/-innen, Klassengröße und dem mittleren Bildungsstand der Klasse vor.

Tab. 22: Evidenzen für Effekte von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg

Stufe der Evidenz	Variable	Art des Effekts
<i>Fundierte Evidenz</i>	Aufgaben mit Hilfe	Negativ
	Förderunterricht Mathematik	Negativ (Förderung sozial benachteiligter SuS)
	Mittelwert Mathematik_5	Negativ (Förderung sozial benachteiligter SuS)
<i>Begrenzte Evidenz</i>	Partner-/Gruppenarbeit	Negativ
	Leistungsheterogene Gruppen	Negativ
	Besprechung Hausaufgaben	Negativ (Beeinträchtigung sozial privilegierter SuS)
	Förderunterricht stark	Positiv (Beeinträchtigung sozial benachteiligter SuS)
	Hausaufgabenbetreuung	Positiv (Beeinträchtigung sozial benachteiligter SuS)
	Mathematikangebote	Negativ
	Anteil Migrationshintergrund	Positiv (Beeinträchtigung sozial benachteiligter SuS)
<i>Keine Evidenz</i>	Unterrichtsumfang	
	Klassenunterricht	
	Einzelarbeit	
	Leistungshomogene Gruppen	
	Klassenführung	
	Kognitive Aktivierung	
	Übung und Wiederholung	
	Aufgaben ohne Hilfe	
	Individualisierung	
	Förderunterricht schwach	
	Klassengröße	
	Mittelwert Bildungsstand	

Anmerkungen: SuS= Schüler/-innen. In der Spalte „Art des Effekts“ wird nur bei den Variablen eine zusätzliche Erläuterung in Klammern gegeben, bei denen sich im Mehrgruppenmodell signifikante Effekte zeigen.

9.4.2 Einordnung der Ergebnisse zu Effekten von pädagogischen Merkmalen in den Forschungsstand

Die Effekte von *Unterrichtsmerkmalen* auf den *durchschnittlichen Lernerfolg* wurden in zahlreichen Studien untersucht. Dementsprechend umfangreich ist der empirische Kenntnisstand in diesem Bereich (Hattie, 2013). Eine Reihe von Merkmalen wirksamen Unterrichts, die im Rahmen früherer Studien ermittelt wurden, können in der vorliegenden Untersuchung erneut bestätigt werden. Für die mehrfach belegten lernförderlichen Effekte von *Klassenunterricht* (Brophy & Good, 1986; Driessen & Sleegers, 2000; Muijs & Reynolds, 2001; Schwerdt & Wuppermann, 2011), *Partner-/Gruppenarbeiten* (Johnson et al., 2000; Lou et al., 1996), *Klassenführung* (Kyriakides et al., 2013; Marzano et al., 2003) und *kognitiver Aktivierung* (Fauth et al., 2014; Kunter et al., 2006; Lipowsky et al., 2009) finden sich auch in der vorliegenden Untersuchung Evidenzen. Die kurvilinearen Effekte von Klassenunterricht und Partner-/Gruppenarbeit, denen zufolge sich die zu häufige Verwendung dieser Unterrichtsmethoden ungünstig auf den Lernerfolg auswirken kann, passen ihrerseits zu Befunden über die Wirksamkeit von *Methodenvielfalt* (Brophy & Good, 1986; Helmke, 2014).

Bezüglich des Unterrichtsprinzips *Individualisierung* fallen die verfügbaren Befunde zu Effekten auf den durchschnittlich Lernerfolg recht heterogen aus. Lernförderliche Effekte von Individualisierung können nicht durchgängig belegt werden (Graham et al., 2021; Lipowsky & Lotz, 2015). In mehreren Studien finden sich sogar Hinweise auf negative Effekte von Individualisierung (Clausen, 2002; Gruehn, 1998; Schulz-Heidorf, 2016). Die vorliegende Untersuchung reiht sich in die letztgenannten Studien ein. Somit mehren sich die Zweifel an der Annahme, dass es sich bei Individualisierung um eine generell sinnvolle Vorgehensweise handelt. Weitere Studien werden benötigt, um zu klären, unter welchen Bedingungen Individualisierung den Lernerfolg der Schüler/-innen fördern kann (und unter welchen nicht).

Hinsichtlich der Effekte von Unterrichtsmerkmalen auf *herkunftsbedingte Unterschiede im Lernerfolg* ist die Befundlage deutlich weniger umfangreich. Einen vergleichsweise gut untersuchten Themenbereich stellen *Partner-/Gruppenarbeiten* und die hierbei verwendeten Prinzipien der Paar-/Gruppenbildung dar. Konsistent zeigt sich, dass leistungsschwächere bzw. sozial benachteiligte Schüler/-innen in besonderem Maß von leistungsheterogenen Paaren/Gruppen profitieren (Lou et al., 1996; Rohrbeck et al., 2003). Die vorliegende Untersuchung liefert erneut Evidenzen für unterschiedsverringende Effekte von leistungsheterogenen Partner-/Gruppenarbeiten.

In Bezug auf die Effekte von *Unterstützung und Feedback* auf herkunftsbedingte Unterschiede im Lernerfolg stehen nur wenige Studien zur Verfügung. Zudem fallen die Befunde nicht ganz konsistent aus. Während mehrere Studien ermitteln, dass benachteiligte Schüler/-innen stärker von einer individuellen Lernbegleitung durch die Lehrkraft profitieren (Brophy & Good, 1986; Pfeifer, 2011; Seiz et al., 2016), zeigt sich in einer Studie ein gegenteiliger Befund, d. h., privilegierte Schüler/-innen profitieren stärker (Atlay et al., 2019). Die in der vorliegenden Untersuchung ermittelten Befunde, wonach Aufgaben mit Hilfe und Hausaufgabenbesprechungen jeweils mit geringeren Kompetenzunterschieden einhergehen, stärken die Evidenzen für unterschiedsverringernde Effekte. Die sich abzeichnende Tendenz in der Befundlage sollte durch weitere Studien untermauert werden.

Neben Unterricht befasst sich die vorliegende Untersuchung auch mit *Ganztagsangeboten*. In den bisher verfügbaren deutschsprachigen Studien dominieren Analysen zur schulischen Organisationsform, d. h. Halbtagsschulen und Ganztagschulen werden miteinander verglichen. Die Befunde zu den Effekten dieses Strukturmerkmals auf den *durchschnittlichen Lernerfolg* fallen ernüchternd aus: In mehreren Studien zeigen sich keine lernförderlichen Effekte von Ganztagschulen (Aktionsrat Bildung, 2013; Linberg et al., 2018; Radisch et al., 2006; Steinmann & Strietholt, 2018; Strietholt et al., 2015). In ähnlicher Weise können auch in der vorliegenden Untersuchung keine lernförderlichen Effekte von Ganztagsangeboten festgestellt werden. Es finden sich sogar Hinweise darauf, dass bestimmte Ganztagsangebote negative Zusammenhänge zum Lernerfolg der Schüler/-innen aufweisen. Angesichts dieser unbefriedigenden Ergebnisse besteht weiterer Forschungsbedarf.

Auch hinsichtlich *herkunftsbedingter Unterschiede im Lernerfolg* finden sich in den vorhandenen Studien kaum Hinweise darauf, dass die schulische Organisationsform Effekte auf Kompetenzunterschiede hat (Aktionsrat Bildung, 2013; Linberg et al., 2015; Strietholt et al., 2015; Steinmann & Strietholt, 2018; Züchner & Fischer, 2014). Demgegenüber werden in der vorliegenden Untersuchung bei der Betrachtung bestimmter Ganztagsangebote durchaus Effekte auf herkunftsbedingte Unterschiede im Lernerfolg festgestellt. In Abhängigkeit von der Art des Ganztagsangebots zeigen sich sowohl unterschiedsverringernde als auch unterschiedsvergrößernde Effekte. Diese differenzierten Befunde legen nahe, dass Analysen, die sich auf die schulische Organisationsform begrenzen, ein unzutreffendes Bild zeichnen. Folglich werden weitere Studien zu Prozessmerkmalen benötigt, damit das kompensatorische Potential von Ganztagschulen angemessen beurteilt werden kann. Von besonderem Interesse ist

hierbei, inwieweit die vorgelegten Befunde in anderen Studien repliziert werden können.

Der dritte Bereich von pädagogischen Merkmalen bezieht sich auf die *Lerngruppe*. Zu den Effekten der Lerngruppe auf den *durchschnittlichen Lernerfolg* stehen zahlreiche Studien zur Verfügung. Die Befunde legen nahe, dass die *Klassengröße* nur eine untergeordnete Rolle spielt (Hanushek, 1986; Hattie, 2002; Robinson, 1990), während der *Klassenkomposition* eine erhebliche Bedeutung zukommt (Kyriakides et al., 2018b; Van Ewijk & Slegers, 2010b). Die Ergebnisse der vorliegenden Untersuchung bestätigen dieses Befundmuster erneut. In den Analysen können Kompositionseffekte hierbei nur für das mathematische Ausgangsniveau der Klasse nachgewiesen werden. Mit diesem Ergebnis verhärteten sich die Hinweise, dass die leistungsbezogene Zusammensetzung der Schülerschaft das wichtigste Kompositionsmerkmal darstellt (Baumert et al., 2006; Kyriakides et al., 2018b).

Auch zu den Effekten der Lerngruppe auf *herkunftsbedingte Unterschiede im Lernerfolg* liegt bereits eine Reihe von Studien vor. Die Befundlage fällt allerdings sehr heterogen aus, insbesondere bei Effekten der Schülerkomposition. Einige Studien ermitteln, dass der Effekt einer günstigen Komposition bei benachteiligten Schüler/-innen größer ausfällt (Coleman et al., 1966; Veldman & Sanford, 1984; Zimmer & Toma, 2000), während andere Studien ermitteln, dass der Effekt einer günstigen Komposition bei privilegierten Schüler/-innen größer ausfällt (Baumert et al., 1987a; Ma et al., 2016; Opdenakker et al., 2002). Die vorliegende Untersuchung verstärkt die Evidenz, dass sozial benachteiligte Schüler/-innen stärker von ihren Mitschüler/-innen beeinflusst werden als sozial privilegierte Schüler/-innen. Für eine abschließende Einschätzung werden allerdings weitere Studien und insbesondere Analysen zu moderierenden Rahmenbedingungen benötigt.

9.4.3 Bewertung der Ergebnisse zu Effekten von pädagogischen Merkmalen

Angesichts des umfangreichen Kenntnisstandes zu *Effekten von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg* haben viele Ergebnisse der vorliegenden Untersuchung kaum Neuheitswert, sondern stellen Replikationen bereits vorhandener Befunde dar. Neue Erkenntnisse liefern insbesondere die Ergebnisse zu kurvilinearen Effekten von Unterrichtsmerkmalen. Im Bereich der *Effekte von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg* wird dagegen ein größerer Erkenntnisfortschritt geleistet. Die Ergebnisse zu den Effekten der Unterrichtsgestaltung erweitern den bisher

recht begrenzten empirischen Kenntnisstand bedeutsam. Auch die Ergebnisse zu Ganztagsangeboten leisten einen wichtigen Beitrag. Die vorgelegten Analysen zu Prozessmerkmalen erzielen aussagekräftigere Ergebnisse als die bisher verfügbaren Studien zur schulischen Organisationsform.

Den Ergebnissen zu Effekten von pädagogischen Merkmalen kommt unter dem Gesichtspunkt eines empirisch fundierten Orientierungswissens eine besondere Bedeutung zu. Insofern ist bei diesen Befunden das erreichte *Niveau an Evidenz* besonders zu beachten. Bei einem größeren Teil der Befunde ist die Evidenz als begrenzt einzuschätzen, bei einem kleineren Teil als fundiert. Auf der Grundlage der ermittelten empirischen Zusammenhänge können Kausaleffekte somit als möglich bzw. als wahrscheinlich angesehen werden. Einen eindeutigen Nachweis von Kausaleffekten kann die vorliegende Arbeit nicht leisten. Dies stellt eine wesentliche Limitation des verwendeten korrelativen Forschungsdesigns dar.

Unter dem Gesichtspunkt der *Generalisierbarkeit* ist anzunehmen, dass die große Mehrheit der Befunde auch Gültigkeit für andere Kontexte hat. Es erscheint plausibel, dass die lernförderlichen Effekte von Klassenführung und kognitiver Aktivierung sowie die unterschiedsverringernenden Effekte von Unterstützung und fachspezifischem Förderunterricht in vergleichbarer Weise auch in anderen Kontexten gegeben sind. Einige Einschränkungen sind allerdings zu formulieren. Bei der hier untersuchten Gymnasialstichprobe handelt es sich um eine homogenisierte Schülerschaft. Insofern erscheint es denkbar, dass bestimmte Effekte in Schulformen mit heterogenerer Schülerschaft (z. B. Grundschulen, Gesamtschulen) anders ausfallen würden. Dies betrifft insbesondere die hier ermittelten günstigen Effekten von leistungsheterogenen Gruppenarbeiten und die ungünstigen Effekte von Individualisierung.

Bezüglich der Effekte der Klassenkomposition bei Schüler/-innen verschiedener sozialer Herkunft ist auf die sehr inkonsistente Befundlage in den vorhandenen Studien hinzuweisen. Möglicherweise haben institutionelle Rahmenbedingungen einen moderierenden Einfluss darauf, welche Schülergruppe stärker von einer günstigen Klassenkomposition profitiert (Gustafsson et al., 2018, S. 26). Der Befund, dass sozial benachteiligte Schüler/-innen stärker von einer günstigen Klassenkomposition profitieren als sozial privilegierte Schüler/-innen, ist somit nicht ohne Weiteres auf andere Schulformen und Länder übertragbar.

9.5 Vergleichende Gegenüberstellung ausgewählter Befunde

Die Zusammenfassung der Ergebnisse wird nun mit einigen Überlegungen abgeschlossen, bei denen die Ergebnisse ausgewählter Fragestellungen zueinander ins Verhältnis gesetzt werden. Zunächst einmal ist von Interesse, die Ergebnisse zu *vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg* (Fragestellungen F1.1 bis F1.6) und die Ergebnisse zu *Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg* (Fragestellungen F3.2, 3.4, 3.6) zu verbinden. Hieraus ergibt sich ein relativ schlüssiges Gesamtbild bezüglich der Ursachen von herkunftsbedingten Kompetenzunterschieden und geeigneter Maßnahmen zur Verringerung dieser. Bei den Analysen zu vermittelnden Mechanismen konnten in den Pfadmodellen bedeutsame Mediationsanteile primär bei den *kognitiven Lernvoraussetzungen der Schüler/-innen* festgestellt werden (mathematisches Vorwissen, metakognitive Kompetenz, Lesekompetenz). Vor diesem Hintergrund ist zu betrachten, dass fundierte Evidenzen für unterschiedsverringende Effekte insbesondere bei den pädagogischen Merkmalen *Unterstützung durch die Lehrkraft im Unterricht* und *Förderunterricht in Mathematik* vorliegen. Bei beiden Merkmalen erscheint es plausibel, dass sie mit den kognitiven Lernvoraussetzungen von Schüler/-innen interagieren. Dies würde bedeuten, dass lernbezogene Unterstützungsangebote von pädagogischen Fachkräften ungünstige kognitive Lernvoraussetzungen von sozial benachteiligten Schüler/-innen bis zu einem gewissen Grad kompensieren können. Ein vergleichbarer Mechanismus erscheint bei den pädagogischen Merkmalen *Bildung von leistungsheterogenen Gruppen* und *mathematisches Ausgangsniveau der Klasse* plausibel, bei denen ebenfalls unterschiedsverringende Effekte festgestellt wurden. Bei diesen Merkmalen erscheint es denkbar, dass lernbezogene Unterstützungsangebote von (leistungsstärkeren) Schüler/-innen als Kompensation für ungünstige kognitive Lernvoraussetzungen von sozial benachteiligten Schüler/-innen wirken. Insgesamt betrachtet legt das Befundmuster nahe, dass ungleiche Ausgangsbedingungen in kognitiven Lernvoraussetzungen eine wesentliche Ursache für herkunftsbedingte Kompetenzunterschiede darstellen und dass folglich lernbezogene Unterstützungsangebote von Lehrkräften und Mitschüler/-innen als Maßnahmen für die Verringerung von Kompetenzunterschieden geeignet sind.

Weiterhin von Interesse ist, die Ergebnisse zu *Effekten von pädagogischen Merkmalen auf den durchschnittlichen Lernerfolg* (Fragestellungen F3.1, F3.3, F3.5) und die Ergebnisse zu *Effekten von pädagogischen Merkmalen auf herkunftsbedingte Unterschiede im Lernerfolg* (Fragestellungen F3.2, F3.4, F3.6) vergleichend gegenüberzustellen. Es ergibt sich folgendes Bild: Insgesamt

können für 14 pädagogische Merkmale Effekte nachgewiesen werden. *Vier Merkmale* haben nur Effekte auf den durchschnittlichen Lernerfolg (Klassenunterricht, Klassenführung, kognitive Aktivierung, Individualisierung), *fünf Merkmale* haben nur Effekte auf herkunftsbedingte Unterschiede im Lernerfolg (leistungsheterogene Gruppen, Aufgaben mit Hilfe, Förderunterricht in Mathematik, Mathematikangebote, Anteil an Schüler/-innen mit Migrationshintergrund) und *fünf Merkmale* haben Effekte auf beide Zielkriterien (Partner-/Gruppenarbeit, Besprechung der Hausaufgaben, Förderunterricht für leistungsstarke Schüler/-innen, Hausaufgabenbetreuungen, mathematisches Ausgangsniveau der Klasse). Innerhalb der letztgenannten Kategorie haben drei Merkmale gleichgerichtete Effekte (Förderunterricht für leistungsstarke Schüler/-innen, Hausaufgabenbetreuung, mathematisches Ausgangsniveau der Klasse), ein Merkmal hat unterschiedlich gerichtete Effekte (Besprechung der Hausaufgaben) und ein Merkmal ist neutral (Partner-/Gruppenarbeit).⁴⁴

Aus dieser Gegenüberstellung können zwei Schlussfolgerungen gezogen werden. Erstens legt das Befundmuster nahe, dass für die Erreichung der beiden Zielsetzungen zu einem kleineren Teil *global wirksame Maßnahmen* eine Rolle spielen (d. h. Maßnahmen, die für beide Zielsetzungen wirksam sind) und zu einem größeren Teil *spezifisch wirksame Maßnahmen* (d. h. Maßnahmen, die nur für eine Zielsetzung wirksam sind). Ähnliche Befunde werden auch in anderen Studien berichtet (Helmke, 1988, S. 62; Schwippert, 2001, S. 148). Demnach werden allgemeine Qualitätsverbesserungen, die auf eine Steigerung des durchschnittlichen Lernerfolgs abzielen, vermutlich nicht ausreichen, um auch das Problem der herkunftsbedingten Kompetenzunterschiede zu beheben. Stattdessen erscheinen spezifische Maßnahmen notwendig, die gezielt auf die Verringerung von herkunftsbedingten Kompetenzunterschieden fokussieren. Zweitens bestätigen diese Ergebnisse erneut, dass zwischen den Zielsetzungen eines hohen durchschnittlichen Lernerfolgs und geringer herkunftsbedingter Unterschiede im Lernerfolg kein Zielkonflikt besteht.

9.6 Implikationen für die pädagogische Praxis und Bildungspolitik

Im Weiteren wird erörtert, welche Implikationen sich aus den Befunden der vorliegenden Arbeit für die pädagogische Praxis und Bildungspolitik ergeben.

44 Bei Partner-/Gruppenarbeiten sind die günstigen Ausprägungen bezüglich der beiden Zielsetzungen weder eindeutig vereinbar noch eindeutig konfligierend, da der Effekt auf den durchschnittlichen Lernerfolg kurvilinear ausfällt, der Effekt auf herkunftsbedingte Unterschiede im Lernerfolg dagegen linear.

Im Sinne eines *empirisch fundierten Orientierungswissen* wird die Frage beleuchtet, welche Maßnahmen geeignet erscheinen, um herkunftsbedingte Unterschiede im Lernerfolg gering zu halten. Hierbei ist darauf hinzuweisen, dass eine einzelne Studie keine angemessene empirische Grundlage für evidenzbasierte Entscheidungen darstellt. Wenn im Weiteren praxisrelevante Schlussfolgerungen gezogen werden, soll damit nicht zum Ausdruck gebracht werden, dass die hier erbrachten Evidenzen als hinreichend einzuschätzen sind. Dementsprechend werden bei den nachfolgenden Überlegungen auch Bezüge zum vorhandenen Forschungsstand hergestellt.

Weiterhin ist klarzustellen, dass sich die nachfolgenden Ausführungen auf empirisch fundierte *Ziel-Mittel-Zusammenhänge* beziehen (nach dem Schema „Um geringe herkunftsbedingte Unterschiede im Lernerfolg zu erreichen, erscheint es als eine angemessene Maßnahme, ...“). Diese Schlussfolgerungen sind auf der Grundlage empirischer Analysen grundsätzlich zulässig (Brezinka, 1995, S. 31 ff.). Die empirischen Analysen haben allerdings keine Aussagekraft für die *normative Frage*, ob die Zielsetzung geringer herkunftsbedingter Unterschiede im Lernerfolg überhaupt angestrebt werden sollte.

Bezüglich *geeigneter Maßnahmen* zur Erreichung geringer herkunftsbedingter Unterschiede im Lernerfolg legen die Befunde grundsätzlich nahe, dass die Unterrichtsgestaltung einen sinnvollen Ansatzpunkt darstellt. Die Befunde zu unterschiedsverringernenden Effekten von Aufgaben mit Hilfe sprechen dafür, dass dem Unterrichtsprinzip *Unterstützung und Feedback* eine wesentliche Bedeutung zukommt. Auch in anderen Studien finden sich Hinweise darauf, dass Schüler/-innen mit ungünstigen Lernbedingungen in besonderem Maße von Unterstützung und Feedback profitieren (Brophy & Good, 1986; Pfeifer, 2011; Seiz et al., 2016; Treiber & Weinert, 1985). Es erscheint aussichtsreich, den *Lernerfolg der Schüler/-innen regelmäßig zu überprüfen*, um Lernprobleme frühzeitig identifizieren zu können (formative Evaluation, Hattie, 2013, S. 215). Wenn Schüler/-innen Aufgaben bearbeiten, ist ein *spezifisches Feedback* relevant, welches neben Lösungen auch auf Strategien eingeht und dabei konkrete Verbesserungsvorschläge aufzeigt (Shute, 2008). Werden Lernprobleme festgestellt, ist es angezeigt, die zugrundeliegenden *Wissenslücken und Denkfehler gezielt zu beseitigen*. Bei allen unterstützenden Maßnahmen sind insbesondere *leistungsschwächere Schüler/-innen zu fokussieren*. Gleichwohl ist darauf zu achten, dass unterstützende Maßnahmen *nicht das Lerntempo von leistungsstärkeren Schüler/-innen verringern*. Dies machen die Hinweise auf lernbeeinträchtigende Effekte der Hausaufgabenbesprechung bei sozial privilegierten Schüler/-innen deutlich.

Ein weiterer relevanter Aspekt der Unterrichtsgestaltung betrifft den Einsatz von *kooperativen Lernformen*. Die empirischen Analysen ergaben, dass sowohl der häufige Einsatz von Partner-/Gruppenarbeiten als auch die Bildung von leistungsheterogenen Gruppen unterschiedsverringernde Effekte hat. Dieser Befund deckt sich mit den Befunden zweier Meta-Analysen, wonach leistungsschwächere und sozial benachteiligte Schüler/-innen in besonderem Maße von leistungsheterogenen Partner- und Gruppenarbeiten profitieren (Lou et al., 1996; Rohrbeck et al., 2003). Es erscheint daher als eine aussichtsreiche Strategie, im Unterricht *regelmäßig leistungsheterogene Paare/Gruppen* zu bilden und die Schüler/-innen gemeinsam Aufgaben bearbeiten zu lassen. Damit sich alle Schüler/-innen beteiligen und es zu lernförderlichen Interaktionen zwischen den Schüler/-innen kommt, sind *strukturierte kooperative Lernmethoden* zu empfehlen, bei denen die Schüler/-innen die Zielsetzungen nur gemeinsam erreichen können und daher intensiv zusammenarbeiten müssen (Johnson et al., 2000). Im Blick zu behalten ist hierbei, dass den *leistungsstärkeren Schüler/-innen nicht nur die Rolle von Tutor/-innen* zukommt, sondern dass sie auch zum Lernen angeregt und gefordert werden.

Neben der Unterrichtsgestaltung stellen *Ganztagsangebote* einen aussichtsreichen Ansatzpunkt für Maßnahmen dar. Die Befunde zu den Effekten von Förderunterricht in Mathematik und Förderunterricht für leistungsstarke Schüler/-innen legen nahe, dass *fachbezogener Förderunterricht für leistungsschwächere Schüler/-innen* eine geeignete Maßnahme für die Geringhaltung von herkunftsbedingten Unterschieden im Lernerfolg darstellt. Direkte Evidenzen für unterschiedsverringernde Effekte eines solchen Förderunterrichts aus anderen Studien stehen bisher aus. Allerdings liefern meta-analytische Befunde zu Tutoring Programs (Pellegrini et al., 2020; Slavin et al., 2011) klare Evidenzen dafür, dass Förderunterricht grundsätzlich lernförderliche Effekte hat. Insofern ist von einem unterschiedsverringernenden Effekt auszugehen, wenn *sozial benachteiligte Schüler/-innen häufiger Förderunterricht* erhalten als sozial privilegierte Schüler/-innen.

Anzumerken ist allerdings, dass Förderunterricht unter dem Gesichtspunkt der *Effizienz* eine sehr aufwändige Maßnahme darstellt. Ob der hohe Aufwand mit einem angemessen hohen Nutzen einhergeht, kann auf der Grundlage der vorliegenden Untersuchung nicht abschließend beantwortet werden. Als Argument zugunsten des fachbezogenen Förderunterrichts für leistungsschwächere Schüler/-innen kann aber angebracht werden, dass es sich bei Förderunterricht um eine Maßnahme handelt, bei der eine *Beeinträchtigung von sozial privilegierten bzw. leistungsstärkeren Schüler/-innen ausgesprochen unwahrscheinlich* ist. Demgegenüber kann es beim Unterricht durchaus zu Konflikten kommen,

d. h., für verschiedene Schülergruppen können unterschiedliche Ausgestaltungen des Unterrichts günstig sein.

Die Befunde zu differenziellen Effekten der *Schülerkomposition* verweisen darauf, dass Gruppenbildungsprozesse auf mehreren Ebenen von Schule relevant sind. Neben Gruppenbildung innerhalb von Klassen (Partner-/Gruppenarbeiten) gibt es Gruppenbildung innerhalb von Schulen (Verteilung von Schüler/-innen auf Klassen) und Gruppenbildung zwischen Schulen (Verteilung von Schüler/-innen auf Schulformen) (Creemers, 1994, S. 57). Der Befund, dass sozial benachteiligte Schüler/-innen stärker von einer günstigen Schülerkomposition profitieren als sozial privilegierte Schüler/-innen, legt nahe, dass die *Bildung von leistungsheterogenen Klassen* eine aussichtsreiche Maßnahme zur Geringhaltung von Kompetenzunterschieden darstellt. Diese Schlussfolgerung wird durch zahlreiche Befunde untermauert, wonach die Bildung von leistungsheterogenen Klassen den leistungsschwächeren Schüler/-innen zugutekommt, während die Bildung von leistungshomogenen Klassen den leistungsstärkeren Schüler/-innen zugutekommt (Creemers, 1994, S. 58 f.; Rui, 2009, S. 168 ff.). In ähnlicher Weise wird nahegelegt, dass *Gesamtschulsysteme* (mit leistungsheterogenen Schulen) eher geeignet sind, um Kompetenzunterschiede gering zu halten, als *gegliederte Schulsysteme* (mit leistungshomogenen Schulen). Diese Überlegung wird durch mehrere international vergleichende Studien untermauert, die zeigen, dass herkunftsbedingte Kompetenzunterschiede in Gesamtschulsystemen geringer ausfallen als in gegliederten Schulsystemen (van de Werfhorst & Mijs, 2010, S. 417 ff.). Vor diesem Hintergrund stellen der Zeitpunkt der Aufteilung auf verschiedene Bildungsgänge, die Anzahl der Bildungsgänge und das Verfahren zur Aufteilung wichtige Ansatzpunkte für bildungspolitische Maßnahmen dar.

Schließlich ermöglichen auch die Befunde zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg einige Überlegungen von praktischer Relevanz. Beachtenswert ist das Ergebnis, dass insbesondere die kognitiven Lernvoraussetzungen von Schüler/-innen einen vermittelnden Mechanismus darstellen. Dieser Befund liefert indirekte Evidenzen für Maßnahmen, die darauf abzielen, die *kognitiven Lernvoraussetzungen von Schüler/-innen verschiedener sozialer Herkunft schon vor Schuleintritt anzugleichen*. Von zentraler Bedeutung ist eine Verbesserung der vorschulischen kognitiven Anregung bei Kindern aus sozial benachteiligten Familien. Diese kann am ehesten durch Maßnahmen in den Bereichen *frühpädagogische Einrichtungen* und *aufsuchende Familienarbeit* erreicht werden. Aussichtsreich erscheinen insbesondere Interventionsprogramme, die sich mit qualitativ hochwertigen Angeboten ge-

zielt an benachteiligte Familien richten (Barnett, 2011, S. 975 ff.; Roßbach et al., 2008, S. 147 ff.).

In den zuvor angestellten Überlegungen zu geeigneten Maßnahmen für die Geringhaltung von herkunftsbedingten Unterschieden im Lernerfolg wurde mehrfach nahegelegt, dass die *Bevorzugung von sozial benachteiligten Kindern* bei der Verteilung von öffentlichen pädagogischen Ressourcen sinnvoll erscheint (z. B. mehr Förderunterricht, mehr frühpädagogische Förderung). Die zugrunde liegende Idee ist, dass Schüler/-innen, die über wenige familiäre Ressourcen verfügen (und dadurch auch über wenige persönliche Ressourcen in der Form von Lernvoraussetzungen), zusätzliche öffentliche pädagogische Ressourcen erhalten. Mehrfach findet sich in der Literatur die empirische Einschätzung, dass diese Vorgehensweise notwendig ist, um eine bedeutende Verringerung von Kompetenzunterschieden nach sozialer Herkunft zu erreichen (Expertenrat Herkunft und Bildungserfolg, 2011, S. 22; Fend, 2013, S. 138 f.). Allerdings wird mit dieser Einschätzung auch die normative Frage nach einer *gerechten Verteilung von öffentlichen pädagogischen Ressourcen* aufgeworfen. Eine tiefergehende Erörterung kann an dieser Stelle nicht geleistet werden. Pragmatisch soll aber darauf hingewiesen werden, dass eine Verteilung von öffentlichen pädagogischen Ressourcen ausschließlich anhand des *Kriteriums der sozialen Herkunft* problematisch erscheint, da die soziale Herkunft nur ein grober Indikator für Lernbenachteiligung ist (d. h., es gibt auch sozial benachteiligte Schüler/-innen, die günstige Lernvoraussetzungen aufweisen und viel Lernanregungen außerhalb der Schule erfahren). Im Extremfall würde diese Vorgehensweise dazu führen, dass leistungsstarke Kinder aus statusniedrigen Familien besonders gefördert werden, während leistungsschwache Kinder aus statushohen Familien nicht gefördert werden. Sinnvoller erscheint es, sich bei der Verteilung von öffentlichen pädagogischen Ressourcen an *direkt lernbezogenen Kriterien* zu orientieren (z. B. niedrige Kompetenzen, niedrige Lernmotivation, geringe Unterstützungsmöglichkeiten der Eltern). Bei dieser Vorgehensweise erhalten sozial benachteiligte Kinder ein überproportionales Maß an Förderung (weil die lernbezogenen Kriterien mit der sozialen Herkunft korrelieren), es werden aber nur diejenigen sozial benachteiligten Kinder besonders gefördert, die auch tatsächlich lernbenachteiligt sind.

Abschließend ist anzumerken, dass selbst mehrere Bereiche umfassende bildungspolitische Maßnahmen wahrscheinlich nur einen moderaten Beitrag zur Verringerung von Bildungsungleichheiten leisten können. Eine derart große gesellschaftliche Herausforderung wie der Abbau von Bildungsungleichheiten (oder gar von sozialen Ungleichheiten im Allgemeinen) kann nicht allein durch pädagogische und bildungspolitische Maßnahmen bewältigt werden.

Notwendig erscheinen umfangreiche Maßnahmen in weiteren Bereichen wie der Familien-, Sozial- und Finanzpolitik (Ditton, 2011, S. 261; Strand, 2016, S. 137).

9.7 Implikationen für die empirische Bildungsforschung

Abschließend werden Überlegungen angestellt, was sich aus der vorliegenden Arbeit für zukünftige Untersuchungen im Bereich der empirischen Bildungsforschung ergibt. Zunächst wird auf Studien zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg eingegangen und anschließend auf Wirksamkeitsstudien zu Unterschieden im Lernerfolg.

9.7.1 Implikationen für Studien zu vermittelnden Mechanismen zwischen sozialer Herkunft und Lernerfolg

Bezüglich vermittelnder Mechanismen zwischen sozialer Herkunft und Lernerfolg kann grundsätzlich festgehalten werden, dass weiterer Bedarf nach hochwertigen Untersuchungen besteht (Becker & Lauterbach, 2010, S. 17 f.; Maaz et al., 2010, S. 36; Solga & Dombrowski, 2009, S. 20). Aussagekräftige Forschungsdesigns sind hierbei insbesondere *multivariate Längsschnittstudien*. Wünschenswert ist ein möglichst *früher Untersuchungsbeginn*, so dass die Entstehung von herkunftsbedingten Kompetenzunterschieden von der frühesten Kindheit an nachgezeichnet werden kann. Weiterhin wünschenswert ist eine möglichst *lange Untersuchungsdauer* und die Berücksichtigung möglichst vieler potenzieller Mediatoren, so dass das komplexe Zusammenwirken *verschiedener Lernumwelten* analysiert werden kann (z. B. Familie, Kindergarten, Grundschule, Sekundarschule, schulbezogene Einrichtungen, Freizeiteinrichtungen). Ein besonderer Fokus ist auf den *Übergang* von der Grundschule auf die verschiedenen Schulformen der Sekundarstufe zu legen. Vielversprechende Datenquellen, die die skizzierten Anforderungen gut erfüllen, sind die Startkohorte 1 und Startkohorte 2 des NEPS. In Zukunft werden auf der Grundlage dieser Daten multivariate Analysen über einen Zeitraum von vielen Jahren hinweg möglich sein.

In forschungsmethodischer Hinsicht haben sich in der vorliegenden Arbeit *Pfadmodelle* als Analysemethoden für Mediationen grundsätzlich bewährt. Bei bestimmten Variablen sind neben einfachen Mediationsanalysen auch *tiefergehende Analysen* zu sekundären Herkunftseffekten (z. B. Nutzung von Nachhilfe) und zu differenzieller Wirksamkeit (z. B. Nutzung von Ganztagsangeboten) angezeigt. Die Ergebnisse machen aber auch gewisse Grenzen von

Pfadmodellen deutlich. Bei allen Variablen, die gezielte Unterstützungsangebote darstellen (z. B. Lernunterstützung durch die Eltern, Nutzung von Nachhilfe, Besuch von Förderunterricht), ist mit reziproken Effekten zu rechnen. Die Befunde der vorliegenden Arbeit legen nahe, dass reziproke Effekte durch die statistische Kontrolle von Lernvoraussetzungen der Schüler/-innen nicht angemessen modelliert werden können. Für Analysen zur Wirksamkeit von Unterstützungsangeboten erscheinen daher *spezifische Forschungsstrategien* notwendig. Aussichtsreich erscheinen insbesondere experimentelle Studien, oder, wenn diese nicht möglich sind, Analysen mithilfe des Propensity-Score-Matching-Verfahrens (Austin, 2011).

9.7.2 Implikationen für Wirksamkeitsstudien zu Unterschieden im Lernerfolg

Es können weiterhin einige Empfehlungen für Wirksamkeitsstudien mit dem Zielkriterium geringer Unterschiede im Lernerfolg formuliert werden. Grundsätzlich ist festzuhalten, dass der empirische Kenntnisstand in diesem Themenbereich bisher ungenügend ist und daher weiterer Forschungsbedarf besteht (Ditton, 2002, S. 206; Kyriakides & Creemers, 2011, S. 248; Pfeifer, 2011, S. 188; Strand, 2010, S. 310). Neben Studien zu Unterschieden zwischen Schüler/-innen verschiedener sozialer Herkunft sind hierbei auch Studien zu Unterschieden zwischen anderen Schülergruppen relevant, beispielsweise zwischen Schüler/-innen verschiedener Leistungsstärke, Schüler/-innen mit und ohne Migrationshintergrund oder Mädchen und Jungen.

In Bezug auf das *Forschungsdesign* erscheinen insbesondere drei Arten von Studien aussagekräftig. Ein geeignetes korrelatives Forschungsdesign ist eine *multivariate Längsschnittstudie*, wie es beispielsweise in der vorliegenden Arbeit verwendet wird. Die Stärken dieses Forschungsdesigns liegen darin, dass eine relativ große Anzahl an pädagogischen Merkmalen untersucht werden kann und dass mithilfe von Längsschnittanalysen zumindest ein fundiertes Maß an Evidenz erreicht werden kann. Ein weiteres geeignetes korrelatives Forschungsdesign ist eine *international vergleichende Querschnittsstudie* auf der Grundlage von Daten aus Schulleistungsstudien wie PISA, TIMSS oder IGLU. Als Beispiel kann die Untersuchung von Gustafsson und Kolleg/-innen (2018) genannt werden. Zwei Argumente sprechen für dieses Forschungsdesign. Erstens kann untersucht werden, inwieweit Effekte von pädagogischen Merkmalen auf Unterschiede im Lernerfolg über verschiedene Länder hinweg konsistent oder inkonsistent ausfallen. Zweitens können Effekte von Systemmerkmalen auf Unterschiede im Lernerfolg überprüft werden. Neben diesen

beiden korrelativen Forschungsdesigns sind weiterhin *experimentelle Forschungsdesigns* aussagekräftig. Auf der Grundlage theoretischer Überlegungen und vorhandener Befunde können spezifische Interventionen entwickelt und experimentell auf ihre Wirksamkeit geprüft werden. Das Potenzial von experimentellen Studien liegt darin, dass unterschiedsverringende Effekte von Interventionen mit gesicherter Evidenz belegt werden können. Ein Beispiel für eine experimentelle Studie zu Unterschieden im Lernerfolg ist die Untersuchung von Charalambous und Kolleg/-innen (2018).

Hinsichtlich geeigneter *Erhebungsmethoden* ist auf die Bedeutung von Schülerbefragungen (Göllner et al., 2016) und zielkindbezogenen Beobachtungen (Riedmeier, 2019) hinzuweisen. Beide Verfahren ermöglichen es, Merkmalsausprägungen bei einzelnen Schüler/-innen zu messen (z. B. individuell erfahrene Unterstützung durch die Lehrkraft im Unterricht, individuelle Lernaktivitäten im Rahmen von Ganztagsangeboten). Solche auf der Schülerebene verorteten Daten ermöglichen es, die Mechanismen *unterschiedlicher Lerngelegenheiten* und *differenziell wirksamer Lerngelegenheiten* empirisch zu überprüfen. Dies war in der vorliegenden Arbeit nur bei der Nutzung von Ganztagsangeboten möglich.

Weiterhin hat sich die *analytische Vorgehensweise* bewährt, parallel Effekte auf den durchschnittlichen Lernerfolg und Effekte auf Unterschiede im Lernerfolg zu untersuchen und für die Analysen zu Unterschieden im Lernerfolg sowohl Slope-As-Outcome-Modelle als auch Mehrgruppenmodelle zu verwenden. Durch diese Vorgehensweise kann ermittelt werden, ob unterschiedsverringende Effekte von pädagogischen Merkmalen auf den Mechanismus der *Förderung benachteiligter Schüler/-innen* oder auf den Mechanismus der *Beeinträchtigung privilegierter Schüler/-innen* zurückzuführen sind.

Schließlich legen die Ergebnisse der vorliegenden Arbeit nahe, dass bei der Verwendung von *Signifikanztests für die Varianz von Random Slopes* eine gewisse Vorsicht geboten ist. Obwohl die Varianz in der vorliegenden Arbeit (bei Kontrolle der Ausgangskompetenz) nicht signifikant ausfällt, sind die zwischen Klassen bestehenden Unterschiede im Random Slope als inhaltlich bedeutsam einzuschätzen. Weiterhin fallen zahlreiche Effekte von pädagogischen Merkmalen auf den Random Slope signifikant aus (ähnliche Befunde berichten Gustafsson und Kolleg/-innen, 2018, S. 21). Folglich sollte die Aussagekraft von Signifikanztests für die Varianz von Random Slopes nicht überwertet werden. Insbesondere sollten Cross-Level-Interaktionen auch dann geprüft werden, wenn die Varianz des Random Slopes nicht signifikant ausfällt (Aguinis et al., 2013, S. 1502; Creemers et al., 2010, S. 26; Raudenbush & Bryk, 2002, S. 258).

Schließlich können einige pädagogische Merkmale benannt werden, bei denen weiterführende Forschungsaktivitäten besonders aussichtsreich erscheinen. Relevant erscheint es, zu überprüfen, inwieweit bestimmte Unterrichtsmethoden einen Einfluss auf Unterschiede im Lernerfolg haben. *Kooperative Lernmethoden* haben hier ein großes Potential, insbesondere bei leistungsheterogenen Gruppen. Daneben sind die Effekte von *instruktionalen, konstruktivistischen und ausgewogenen Unterrichtsmethoden* bei verschiedenen Schülergruppen genauer zu untersuchen. Bei den Unterrichtsprinzipien sollten unterschiedsverringende Effekte von *Klassenführung* sowie von *Unterstützung und Feedback* geprüft werden. Auch mögliche unterschiedsvergrößernde Effekte von *kognitiver Aktivierung* verdienen Aufmerksamkeit. Bei den Ganztagsangeboten gilt es zu klären, welche *Arten von Ganztagsangeboten* lernförderlich sind und *wie diese Ganztagsangebote konkret auszugestalten* sind. Wünschenswert wären weitere Evidenzen für die Wirksamkeit von *fachspezifischem Förderunterricht* für leistungsschwächere Schüler/-innen. Außerdem sollte die Rolle von *Hausaufgabenbetreuungen* genauer untersucht werden, um zu ermitteln, ob sich der in der vorliegenden Arbeit ermittelte unterschiedsvergrößernde Effekt von Hausaufgabenbetreuungen auch in anderen Kontexten zeigt.

Abschließende Worte

Bildungsungleichheiten nach sozialer Herkunft stellen eine anhaltende gesellschaftliche Herausforderung dar. Die vorliegende Arbeit hat den bisher überschaubaren Kenntnisstand über die Ursachen von herkunftsbedingten Kompetenzunterschieden und mögliche schulische Maßnahmen durch einige relevante Befunde erweitert. Es werden weitere Forschungsaktivitäten benötigt, um den Entscheidungsträger/-innen in pädagogischer Praxis und Bildungspolitik ein evidenzbasiertes Orientierungswissen über wirksame Maßnahmen zur Verfügung zu stellen. Mit dieser Strategie erscheint es möglich, bedeutende Fortschritte bei der Verringerung von Bildungsungleichheiten zu erreichen.

Literatur

- Abele, A. & Liebau, E. (1998). Nachhilfeunterricht. Eine empirische Studie an bayerischen Gymnasien. *Die Deutsche Schule*, 90, 37–49.
- Adelantado-Renau, M., Moliner-Urdiales, D., Cavero-Redondo, I., Beltran-Valls, Martínez-Vizcaíno, V. & Álvarez-Bueno, C. (2019). Association Between Screen Media Use and Academic Performance Among Children and Adolescents: A Systematic Review and Meta-analysis. *JAMA pediatrics*, 173(11), 1058–1067. <https://doi.org/10.1001/jamapediatrics.2019.3176>
- Aguinis, H., Gottfredson, R. K. & Culpepper, S. A. (2013). Best-Practice Recommendations for Estimating Cross-Level Interaction Effects Using Multilevel Modeling. *Journal of Management*, 39(6), 1490–1528. <https://doi.org/10.1177/0149206313478188>
- Aktionsrat Bildung (2007). *Bildungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Aktionsrat Bildung (2013). *Zwischenbilanz Ganztagsgrundschulen: Betreuung oder Rhythmisierung?* Münster: Waxmann.
- Alexander, K. L., Entwisle, D. R. & Olson, L. S. (2001). Schools, Achievement, and Inequality: A Seasonal Perspective. *Educational Evaluation and Policy Analysis*, 23(2), 171–191. <https://doi.org/10.3102/01623737023002171>
- Alfieri, L., Brooks, P. J., Aldrich, N. J. & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103(1), 1–18. <https://doi.org/10.1037/a0021017>
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y. & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333(6045), 1034–1037. <https://doi.org/10.1126/science.1207998>
- Alloway, T. P. & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, 21(1), 133–137. <https://doi.org/10.1016/j.lindif.2010.09.013>
- Ammermüller, A. (2005). Educational Opportunities and the Role of Institutions. *ZEW Discussion Papers*. <https://doi.org/10.2139/ssrn.753366>
- Anders, Y. (2013). Stichwort: Auswirkungen frühkindlicher institutioneller Betreuung und Bildung. *Zeitschrift für Erziehungswissenschaft*, 16(2), 237–275. <https://doi.org/10.1007/s11618-013-0357-5>
- Anders, Y., Grosse, C., Rossbach, H.-G., Ebert, S. & Weinert, S. (2013). Preschool and primary school influences on the development of children's early numeracy skills between the ages of 3 and 7 years in Germany. *School Effectiveness and School Improvement*, 24(2), 195–211. <https://doi.org/10.1080/09243453.2012.749794>
- Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S. et al. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, 27(2), 231–244. <https://doi.org/10.1016/j.ecresq.2011.08.003>

- Andersen, I. G. & Andersen, S. C. (2017). Student-centered instruction and academic achievement: linking mechanisms of educational inequality to schools' instructional strategy. *British Journal of Sociology of Education*, 38(4), 533–550. <https://doi.org/10.1080/01425692.2015.1093409>
- Asendorpf, J. (2007). *Psychologie der Persönlichkeit*. Heidelberg: Springer Medizin. <https://doi.org/10.1007/978-3-540-71685-3>
- Aßmann, C., Steinhauer, H. W., Kiesl, H., Koch, S., Schönberger, B., Müller-Kuller, A. et al. (2011). Sampling designs of the National Educational Panel Study: challenges and solutions. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Hrsg.), *Education as a lifelong process. The German National Educational Panel Study (NEPS) (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 14, S. 51–65)*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/s11618-011-0181-8>
- Atlay, C., Tieben, N., Hillmert, S. & Fauth, B. (2019). Instructional quality and achievement inequality: How effective is teaching in closing the social achievement gap? *Learning and Instruction*, 63. <https://doi.org/10.1016/j.learninstruc.2019.05.008>
- Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46(3), 399–424. <https://doi.org/10.1080/00273171.2011.568786>
- Autorengruppe Bildungsberichterstattung (2008). *Bildung in Deutschland 2008. Ein indikatorengestützter Bericht mit einer Analyse zu Übergängen im Anschluss an den Sekundarbereich I*. Bielefeld: Bertelsmann.
- Autorengruppe Bildungsberichterstattung (2012). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann.
- Autorengruppe Bildungsberichterstattung (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: Bertelsmann.
- Autorengruppe Bildungsberichterstattung (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.
- Babad, E. (1993). Teachers' Differential Behavior. *Educational Psychology Review*, 5(4), 347–376. <https://doi.org/10.1007/BF01320223>
- Baier, F., Decker, A.-T., Voss, T., Kleickmann, T., Klusmann, U. & Kunter, M. (2019). What makes a good teacher? The relative importance of mathematics teachers' cognitive ability, personality, knowledge, beliefs, and motivation for instructional quality. *The British Journal of Educational Psychology*, 89(4), 767–786. <https://doi.org/10.1111/bjep.12256>
- Balbo, N., Billari, F. C. & Mills, M. (2013). Fertility in Advanced Societies: A Review of Research. *European Journal of Population*, 29(1), 1–38. <https://doi.org/10.1007/s10680-012-9277-y>
- Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, 333(6045), 975–978. <https://doi.org/10.1126/science.1204534>
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considera-

- tions. *Journal of Personality and Social Psychology*, 51(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bäumer, T., Preis, N., Roßbach, H.-G., Stecher, L. & Klieme, E. (2011). Education processes in life-course-specific learning environments. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Hrsg.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 14, S. 87–101). Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/s11618-011-0183-6>
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, P. et al. (Hrsg.). (2001a). *PISA 2000*. Opladen: Leske + Budrich. <https://doi.org/10.1007/978-3-322-83412-6>
- Baumert, J. & Köller, O. (1998). Nationale und internationale Schulleistungsstudien: was können sie leisten, wo sind ihre Grenzen? *Pädagogik*, 50(6), 12–18.
- 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>
- Baumert, J., Nagy, G. & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development*, 83(4), 1347–1367. <https://doi.org/10.1111/j.1467-8624.2012.01779.x>
- Baumert, J., Roeder, P.M., Sang, F. & Schmitz, B. (1986). Leistungsentwicklung und Ausgleich von Leistungsunterschieden in Gymnasialklassen. *Zeitschrift für Pädagogik*, 32, 639–660.
- Baumert, J., Roeder, P.M., Sang, F. & Schmitz, B. (1987). Chancenausgleich und Leistungsförderung in der Jahrgangsklasse: Ein Dilemma der Schulqualität? In U. Stefens & T. Barge (Hrsg.), *Untersuchungen zur Qualität des Unterrichts. Beiträge aus dem Arbeitskreis „Qualität von Schule“*, Heft 3 (S. 33–54). Wiesbaden: Hessisches Institut für Bildungsplanung und Schulentwicklung.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, P. Schneider et al. (Hrsg.), *PISA 2000* (S. 323–407). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_10
- Baumert, J., Stanat, P. & Demmrich, A. (2001b). *PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie*. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, P. Schneider et al. (Hrsg.), *PISA 2000* (S. 15–68). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_3
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90082-7_4
- Baumert, J., Watermann, R. & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. *Zeitschrift für Erziehungswissenschaft*, 6(1), 46–71. <https://doi.org/10.1007/s11618-003-0004-7>

- Baumrind, D. (1991). The Influence of Parenting Style on Adolescent Competence and Substance Use. *The Journal of Early Adolescence*, 11(1), 56–95. <https://doi.org/10.1177/0272431691111004>
- Bayerisches Landesamt für Schule (2018). *Bildungsbericht Bayern 2018*. Gunzenhausen: Bayerisches Landesamt für Schule.
- Becker, B. (2010a). Ethnische Unterschiede bei der Kindergartenselektion: Die Wahl von unterschiedlich stark segregierten Kindergärten in deutschen und türkischen Familien. In B. Becker & D. Reimer (Hrsg.), *Vom Kindergarten bis zur Hochschule. Generierung von ethnischen und sozialen Disparitäten in der Bildungsbio-graphie* (S. 17–47). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92105-1_2
- Becker, B. & Schober, P.S. (2017). Not Just Any Child Care Center? Social and Ethnic Disparities in the Use of Early Education Institutions With a Beneficial Learning Environment. *Early Education and Development*, 28(8), 1011–1034. <https://doi.org/10.1080/10409289.2017.1320900>
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik. *Zeitschrift für Pädagogische Psychologie*, 20(4), 233–242. <https://doi.org/10.1024/1010-0652.20.4.233>
- Becker, M., Stanat, P., Baumert, J. & Lehmann, R. (2008). Lernen ohne Schule: Differenzielle Entwicklung der Leseleistungen von Kindern mit und ohne Migrationshintergrund während der Sommerferien. In F. Kalter (Hrsg.), *Migration und Integration (Kölner Zeitschrift für Soziologie und Sozialpsychologie Sonderheft, Bd. 48, S. 252–276)*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Becker, R. (2010b). Bildungseffekte vorschulischer Erziehung und Elementarbildung – Bessere Bildungschancen für Arbeiter- und Migrantenkinder? In R. Becker (Hrsg.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (S. 129–160). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92484-7_5
- Becker, R. (2011). Entstehung und Reproduktion dauerhafter Bildungsungleichheiten. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 87–138). Wiesbaden: VS Verlag für Sozialwissenschaften/Springer Fachmedien Wiesbaden GmbH Wiesbaden. https://doi.org/10.1007/978-3-531-92759-6_4
- Becker, R. & Lauterbach, W. (2010). Bildung als Privileg – Ursachen, Mechanismen, Prozesse und Wirkungen. In R. Becker (Hrsg.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (S. 11–49). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92484-7_1
- Becker, S. & Kurz, K. (2011). Social Inequality in Early Childhood Health. Participation in the Preventive Health Care Program for Children. *Journal of Contextual Economics – Schmollers Jahrbuch*, 131(2), 381–394. <https://doi.org/10.3790/schm.131.2.381>
- Behoui, A. & Neergaard, A. (2016). Social capital and the educational achievement of young people in Sweden. *British Journal of Sociology of Education*, 37(7), 947–969. <https://doi.org/10.1080/01425692.2015.1013086>
- Beland, L.-P. & Murphy, R. (2015). *Ill Communication: Technology, Distraction & Student Performance*. CEP Discussion Paper No 1350. Zugriff am 11.04.2021. Verfügbar unter: <https://cep.lse.ac.uk/pubs/download/dp1350.pdf>

- Bellin, N. & Wegner, B. (2010). Forschungsfeld Ganztag: Lernförderliche Wirkung außerunterrichtlicher Angebote und Strategien der Sprachförderung. In H. Merkens & A. Schründer-Lenzen (Hrsg.), *Lernförderung unter den Bedingungen des Ganztags im Grundschulbereich* (S. 77–98). Münster: Waxmann.
- Blatchford, P., Bassett, P. & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, 21(6), 715–730. <https://doi.org/10.1016/j.learninstruc.2011.04.001>
- Blossfeld, H.-P., Maurice, J. von & Schneider, T. (2011). The National Educational Panel Study: need, main features, and research potential. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Hrsg.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 14, S. 5–17). Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/s11618-011-0178-3>
- Blossfeld, H.-P. & Roßbach, H.-G. (Hrsg.). (2019). *Education as a lifelong process: The German National Educational Panel Study (NEPS)*. Edition ZfE. Springer VS. https://doi.org/10.1007/978-3-658-23162-0_2
- Bohl, T., Batzel, A. & Richey, P. (2012). Öffnung – Differenzierung – Individualisierung – Adaptivität. In T. Bohl, M. Bönsch, M. Trautmann & B. Wischer (Hrsg.), *Didaktische Grundlagen und Forschungsergebnisse zur Binnendifferenzierung im Unterricht* (S. 40–69). Immenhausen bei Kassel: Prolog-Verlag. <https://doi.org/10.2307/j.ctvddzrnj.6>
- Borman, G.D., Hewes, G.M., Overman, L.T. & Brown, S. (2003). Comprehensive School Reform and Achievement: A Meta-Analysis. *Review of Educational Research*, 73(2), 125–230. <https://doi.org/10.3102/00346543073002125>
- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-12770-0>
- Boudon, R. (1974). *Education, Opportunity, and Social Inequality. Changing Prospects in Western Society*. New York, London, Sydney, Toronto: John Wiley & Sons.
- Bourdieu, P. (2012). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In U. Bauer, U.H. Bittlingmayer & A. Scherr (Hrsg.), *Handbuch Bildungs- und Erziehungssoziologie* (S. 229–242). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-18944-4_15
- Bourdieu, P. & Passeron, J.-C. (1977). *Reproduction in education, society and culture*. London: Sage Publications.
- Bradley, R. H., Caldwell, B. M., Rock, S. L., Hamrick, H. M. & Harris, P. (1988). Home Observation for Measurement of the Environment: Development of a Home Inventory for use with families having children 6 to 10 years old. *Contemporary Educational Psychology*, 13(1), 58–71. [https://doi.org/10.1016/0361-476X\(88\)90006-9](https://doi.org/10.1016/0361-476X(88)90006-9)
- Brake, A. & Büchner, P. (2012). *Bildung und soziale Ungleichheit. Eine Einführung*. Stuttgart: Verlag W. Kolhammer.
- Brandsma, H. P. & Knuver, J. W. M. (1989). Effects of school and classroom characteristics on pupil progress in language and arithmetic. *International Journal of Educational Research*, 13(7), 777–788. [https://doi.org/10.1016/0883-0355\(89\)90028-1](https://doi.org/10.1016/0883-0355(89)90028-1)

- Brezinka, W. (1995). *Erziehungsziele, Erziehungsmittel, Erziehungserfolg. Beiträge zu einem System der Erziehungswissenschaft*. München, Basel: Ernst Reinhardt Verlag. <https://doi.org/10.2378/9783497013425>
- Bromme, R., Prenzel, M. & Jäger, M. (2016). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. In J. Baumert & K.-J. Tillmann (Hrsg.), *Empirische Bildungsforschung. Der kritische Blick und die Antwort auf die Kritiker* (Zeitschrift für Erziehungswissenschaft – Sonderheft, v.31, S. 129–146). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-13785-4_8
- Brookover, W.B., Beady, C., Flood, P., Schweitzer, J. & Wisenbaker, J. (1979). *School social systems and student achievement. Schools can make a difference*. New York: Praeger.
- Brophy, J. (1981). Teacher Praise: A Functional Analysis. *Review of Educational Research*, 51(1), 5–32. <https://doi.org/10.3102/00346543051001005>
- Brophy, J. (2000). Teaching. Zugriff am 11.04.2021 Verfügbar unter: www.iaoed.org/downloads/prac01e.pdf
- Brophy, J. & Good, T.L. (1986). Teacher Behavior and Student Achievement. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (S. 328–375). New York: Macmillan.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Deutschland.
- Burger, K. (2016). Intergenerational transmission of education in Europe: Do more comprehensive education systems reduce social gradients in student achievement? *Research in Social Stratification and Mobility*, 44, 54–67. <https://doi.org/10.1016/j.rssm.2016.02.002>
- Burgess, S., Sanderson, E. & Umana-Aponte, M. (2011). School ties: An analysis of homophily in an adolescent friendship network. Zugriff am 11.04.2021. Verfügbar unter: <http://www.bristol.ac.uk/cmpo/publications/papers/2011/wp267.pdf>
- Burghardt, L. (2018). Relevance of child characteristics for the envisaged and real start of childcare attendance for the under-3s in Germany. *European Early Childhood Education Research Journal*, 26(6), 873–892. <https://doi.org/10.1080/1350293X.2018.1533706>
- Carigiet Reinhard, T. (2012). *Schulleistungen und Heterogenität. Eine mehrbenen-analytische Untersuchung der Bedingungsfaktoren der Schulleistungen am Ende der dritten Primarschulklasse*. Bern: Haupt.
- Caro, D.H. (2018). Socio-economic gaps in subject interest: the mediating role of parental cognitive involvement. *Large-scale Assessments in Education*, 6(1), 427. <https://doi.org/10.1186/s40536-018-0067-9>
- Caro, D.H., Lenkeit, J. & Kyriakides, L. (2016). Teaching strategies and differential effectiveness across learning contexts: Evidence from PISA 2012. *Studies in Educational Evaluation*, 49, 30–41. <https://doi.org/10.1016/j.stueduc.2016.03.005>
- Carolan, B.V. (2016). Unequal academic achievement in high school: the mediating roles of concerted cultivation and close friends. *British Journal of Sociology of Education*, 37(7), 1034–1055. <https://doi.org/10.1080/01425692.2015.1013085>
- Carroll, J.B. (1963). A model for school learning. *Teachers College Record*, 64, 723–733. <https://doi.org/10.1177/016146816306400801>

- Castro, M., Expósito-Casas, E., López-Martín, E., Lizasoain, L., Navarro-Asencio, E. & Gaviria, J. L. (2015). Parental involvement on student academic achievement: A meta-analysis. *Educational Research Review*, 14, 33–46. <https://doi.org/10.1016/j.edurev.2015.01.002>
- Center for Data-Driven Reform in Education. (2020). Best Evidence Encyclopedia. Zugriff am 28.09.2020. Verfügbar unter: <http://www.bestevidence.org/aboutbee.htm>
- Chapman, C., Muijs, D., Reynolds, D., Sammons, P. & Teddlie, C. (Hrsg.). (2016a). *The Routledge international handbook of educational effectiveness and improvement. Research, policy and practice* (Routledge international handbooks). London: Routledge. <https://doi.org/10.4324/9781315679488>
- Chapman, C., Reynolds, D., Muijs, D., Sammons, P., Stringfield, S. & Teddlie, C. (2016b). Educational effectiveness and improvement research and practice. The emergence of the discipline. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons & C. Teddlie (Hrsg.), *The Routledge international handbook of educational effectiveness and improvement. Research, policy and practice* (Routledge international handbooks, S. 1–24). London: Routledge. <https://doi.org/10.4324/9781315679488>
- Charalambous, E., Kyriakides, L. & Creemers, B. P. M. (2018). Promoting quality and equity in socially disadvantaged schools: A group-randomisation study. *Studies in Educational Evaluation*, 57, 42–52. <https://doi.org/10.1016/j.stueduc.2016.06.001>
- Christ, O. & Schlüter, E. (2012). *Strukturgleichungsmodelle mit Mplus. Eine praktische Einführung* (Sozialwissenschaften 10–2012). München: Oldenbourg. <https://doi.org/10.1524/9783486714807>
- Clausen, M. (2002). *Unterrichtsqualität. Eine Frage der Perspektive?* Münster: Waxmann.
- Coelen, H. & Siewert, J. (2008). Ferien und Ferieneffekte. In T. Coelen & H. U. Otto (Hrsg.), *Grundbegriffe Ganztagsbildung. Das Handbuch* (S. 432–441). Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91161-8_43
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hoboken: Taylor and Francis.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. et al. (1966). *Equality of Educational Opportunity*. Washington: U. S. Government Printing Office.
- Connor, C. M., Morrison, F. J. & Katch, L. E. (2004). Beyond the Reading Wars: Exploring the Effect of Child-Instruction Interactions on Growth in Early Reading. *Scientific Studies of Reading*, 8(4), 305–336. https://doi.org/10.1207/s1532799xssr0804_1
- Cooper, H., Lindsay, J. & Nye, B. (2000). Homework in the Home: How Student, Family, and Parenting-Style Differences Relate to the Homework Process. *Contemporary Educational Psychology*, 25(4), 464–487. <https://doi.org/10.1006/ceps.1999.1036>
- Cooper, H., Nye, B., Charlton, K., Lindsay, J. & Greathouse, S. (1996). The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-Analytic Review. *Review of Educational Research*, 66(3), 227–268. <https://doi.org/10.3102/00346543066003227>
- Craik, F. I. M. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)

- Creemers, B. (1994). *The effective classroom*. London: Cassell.
- Creemers, B. & Kyriakides, L. (2008). *The dynamics of educational effectiveness. A contribution to policy, practice and theory in contemporary schools*. London, New York: Routledge. <https://doi.org/10.4324/9780203939185>
- Creemers, B. & Kyriakides, L. (2012). *Improving quality in education. Dynamic approaches to school improvement*. Abingdon, Oxon: Routledge.
- Creemers, B., Kyriakides, L. & Sammons, P. (2010). *Methodological advances in educational effectiveness research*. London, New York: Routledge. <https://doi.org/10.4324/9780203851005>
- Dahmann, S. & Anger, S. (2014). *The Impact of Education on Personality – Evidence from a German High School Reform*. <https://doi.org/10.2139/ssrn.2432423>
- DeGarmo, D.S., Forgatch, M.S. & Martinez, C.R. (1999). Parenting of divorced mothers as a link between social status and boys' academic outcomes: unpacking the effects of socioeconomic status. *Child Development*, 70(5), 1231–1245. <https://doi.org/10.1111/1467-8624.00089>
- Degele, N. (2014). Modernisierung. In G. Endruweit, G. Trommsdorff & N. Burzan (Hrsg.), *Wörterbuch der Soziologie* (S. 326–329). Konstanz: UVK Lucius.
- Delgado, M. Y., Ettekal, A. V., Simpkins, S. D. & Schaefer, D. R. (2016). How Do My Friends Matter? Examining Latino Adolescents' Friendships, School Belonging, and Academic Achievement. *Journal of Youth and Adolescence*, 45(6), 1110–1125. <https://doi.org/10.1007/s10964-015-0341-x>
- Department for Education. (2019). *Find and compare schools in England*. Zugriff am 09.01.2020. Verfügbar unter: <https://www.gov.uk/school-performance-tables>
- Ditlevsen, S., Christensen, U., Lynch, J., Damsgaard, M. T. & Keiding, N. (2005). The Mediation Proportion: A Structural Equation Approach for Estimating the Proportion of Exposure Effect on Outcome Explained by an Intermediate Variable. *Epidemiology*, 16(1), 114–120. <https://doi.org/10.1097/01.ede.0000147107.76079.07>
- Ditton, H. (1998). *Mehrebenenanalyse. Grundlagen und Anwendungen des hierarchisch linearen Modells*. Weinheim: Juventa-Verl.
- Ditton, H. (2000). *Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung*. In A. Helmke, W. Hornstein & E. Terhart (Hrsg.), *Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule* (S. 73–92). Weinheim: Beltz.
- Ditton, H. (2002). *Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspektiven*. *Unterrichtswissenschaft*, (3), 197–212.
- Ditton, H. (2010). *Wie viel Ungleichheit durch Bildung verträgt eine Demokratie?* *Zeitschrift für Pädagogik*, 56, 53–68.
- Ditton, H. (2011). *Familie und Schule – eine Bestandsaufnahme der bildungssoziologischen Schuleffektforschung von James S. Coleman bis heute*. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 245–264). Wiesbaden: VS Verlag für Sozialwissenschaften/Springer Fachmedien Wiesbaden GmbH Wiesbaden. https://doi.org/10.1007/978-3-531-92759-6_9
- Ditton, H. & Krecker, L. (1995). *Qualität von Schule und Unterricht. Empirische Befunde zu Fragestellungen und Aufgaben der Forschung*. *Zeitschrift für Pädagogik*, 41, 507–529.

- Ditton, H. & Krüsken, J. (2009). Denn wer hat, dem wird gegeben werden? Eine Längsschnittstudie zur Entwicklung schulischer Leistungen und den Effekten der sozialen Herkunft in der Grundschulzeit. *Journal for Educational Research Online*, 1, 33–61.
- Ditton, H., Krüsken, J. & Schauenberg, M. (2005). Bildungsungleichheit – der Beitrag von Familie und Schule. *Zeitschrift für Erziehungswissenschaft*, 8(2), 285–304. <https://doi.org/10.1007/s11618-005-0138-x>
- Ditton, H. & Maaz, K. (2015). Sozioökonomischer Status und soziale Ungleichheit. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (S. 229–244). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-19994-8_17
- Dohmen, D., Erbes, A., Fuchs, K. & Günzel, J. (2008). Was wissen wir über Nachhilfe? – Sachstand und Auswertung der Forschungsliteratur zu Angebot, Nachfrage und Wirkungen. Berlin: Forschungsinstitut für Bildungs- und Sozialökonomie.
- Döring, N. & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-642-41089-5>
- Dotterer, A. M., McHale, S. M. & Crouter, A. C. (2009). The Development and Correlates of Academic Interests from Childhood through Adolescence. *Journal of Educational Psychology*, 101(2), 509–519. <https://doi.org/10.1037/a0013987>
- Downey, D. B. & Condrón, D. J. (2016). Fifty Years since the Coleman Report. *Sociology of Education*, 89(3), 207–220. <https://doi.org/10.1177/0038040716651676>
- Downey, D. B., Hippel, P. T. von & Broh, B. A. (2004). Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year. *American Sociological Review*, 69(5), 613–635. <https://doi.org/10.1177/000312240406900501>
- Driessen, G. & Slegers, P. (2000). Consistency of Teaching Approach and Student Achievement: An Empirical Test. *School Effectiveness and School Improvement*, 57–79. [https://doi.org/10.1076/0924-3453\(200003\)11:1;1-#;FT057](https://doi.org/10.1076/0924-3453(200003)11:1;1-#;FT057)
- Dronkers, J. & Avram, S. (2009). Choice and Effectiveness of Private and Public Schools in seven countries. A reanalysis of three PISA dat sets. *Zeitschrift für Pädagogik*, 55, 895–909.
- Dumont, H. (2012). Elterliche Hausaufgabenhilfe unter dem Blickwinkel sozialer Disparitäten. Eine Untersuchung pädagogisch-psychologischer Prozesse. Zugriff am 11.04.2021. Verfügbar unter: https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/47942/pdf/Dissertation_Hanna_Dumont.pdf?sequence=1&isAllowed=y
- Dumont, H., Maaz, K., Neumann, M. & Becker, M. (2014). Soziale Ungleichheiten beim Übergang von der Grundschule in die Sekundarstufe I: Theorie, Forschungsstand, Interventions- und Fördermöglichkeiten. In K. Maaz, M. Neumann & J. Baumert (Hrsg.), *Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter. Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive* (*Zeitschrift für Erziehungswissenschaft Sonderheft*, Bd. 24, S. 141–165). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-00454-5_7

- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. *Psychologie in Erziehung und Unterricht*, 60(3), 163–183. <https://doi.org/10.2378/peu2013.art14d>
- ECCE Study Group. (1999). School-age assessment of child development: Long-term impact of pre-school experiences on school success, and family-school relationships. Zugriff am 28.03.2021. Verfügbar unter: www.uni-bamberg.de/fileadmin/uni/fakultaeten/ppp_lehrstuehle/elementarpaedagogik/Team/Rossbach/Ecce_Study_Group.pdf
- Eccles, J. S. & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Ehmke, T., Blum, W., Neubrand, M., Jordan, A. & Ulfig, F. (2006a). Wie verändert sich die mathematische Kompetenz von der neunten zur zehnten Klassenstufe? In M. Prenzel (Hrsg.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 63–86). Münster: Waxmann.
- Ehmke, T. & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 231–254). Münster: Waxmann.
- Ehmke, T., Hohensee, F., Siegle, T. & Prenzel, M. (2006b). Soziale Herkunft, elterliche Unterstützungsprozesse und Kompetenzentwicklung. In M. Prenzel (Hrsg.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 225–248). Münster: Waxmann.
- Ehmke, T. & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. *Zeitschrift für Erziehungswissenschaft*, 8(4), 521–539. <https://doi.org/10.1007/s11618-005-0157-7>
- Enders, C. K. (2001). The Performance of the Full Information Maximum Likelihood Estimator in Multiple Regression Models with Missing Data. *Educational and Psychological Measurement*, 61(5), 713–740. <https://doi.org/10.1177/0013164401615001>
- 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>
- Erikson, R. & Goldthorpe, J. H. (1992). *The constant flux. A study of class mobility in industrial societies*. Oxford: Clarendon Press.
- Erikson, R. & Jonsson, J. O. (1996). Introduction: Explaining Class Inequality in Education: The Swedish Test Case. In R. Erikson (Hrsg.), *Can education be equalized? The Swedish case in comparative perspective* (S. 1–63). Boulder: Westview Press.
- Evans, G. W. & Schamberg, M. A. (2009). Childhood poverty, chronic stress, and adult working memory. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16), 6545–6549. <https://doi.org/10.1073/pnas.0811910106>
- Expertenrat Herkunft und Bildungserfolg (2011). *Empfehlungen für Bildungspolitische Weichenstellungen in der Perspektive auf das Jahr 2020*. Stuttgart: Ministerium für Kultus, Jugend und Sport.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E. & Büttner, G. (2014). Grundschulunterricht aus Schüler-, Lehrer- und Beobachterperspektive: Zusammenhänge und Vorhersage von Lernerfolg. *Zeitschrift für Pädagogische Psychologie*, 28(3), 127–137. <https://doi.org/10.1024/1010-0652/a000129>

- Fend, H. (2013). Starke und schwache Instrumente zur Beförderung von Chancengleichheit im Bildungswesen. In D. Deißner (Hrsg.), *Chancen bilden. Wege zu einer gerechteren Bildung – ein internationaler Erfahrungsaustausch* (S. 125–140). Wiesbaden: Springer. https://doi.org/10.1007/978-3-531-19393-9_8
- Fraser, B. J., Walberg, H. J., Welch, W. W. & Hattie, J. A. (1987). Syntheses of educational productivity research. *International Journal of Educational Research*, 11(2), 147–252. [https://doi.org/10.1016/0883-0355\(87\)90035-8](https://doi.org/10.1016/0883-0355(87)90035-8)
- Fuchs-Rechlin, K. & Bergmann, C. (2014). Der Abbau von Bildungsbenachteiligung durch Kindertagesbetreuung für unter 3-Jährige – zwischen Wunsch und Wirklichkeit. In K. Maaz, M. Neumann & J. Baumert (Hrsg.), *Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter. Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 24, S. 95–118). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-00454-5_5
- Gärtner, H., Thiel, F., Kellermann, C. & Nachbauer, M. (2021). Unterrichtsentwicklung durch Unterrichtsfeedback – Erste Entwicklungsschritte eines Beobachtungsbogens für Schulleitungen. *Zeitschrift für Bildungsforschung*, 11, 529–547. <https://doi.org/10.1007/s35834-021-00321-7>
- Gärtner, H., Thiel, F., Kellermann, C. & Nachbauer, M. (2022). Ein theoriebasierter Schülerfragebogen für Unterrichtsvaluation. *Journal for Educational Research Online*, 14, 147–173. <https://doi.org/10.31244/jero.2022.01.07>
- Ganzeboom, H. B. G. (2010). A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007. Zugriff am 14.07.2021. Verfügbar unter: https://www.researchgate.net/publication/242660186_A_new_International_Socio-Economic_Index_ISEI_of_occupational_status_for_the_International_Standard_Classification_of_Occupation_2008_ISCO-08_constructed_with_data_from_the_ISSP_2002-2007
- Garcia, E. (2015). Inequalities at the Starting Gate. Cognitive and Noncognitive Skills Gaps between 2010–2011 Kindergarten Classmates. Zugriff am 11.04.2021. Verfügbar unter: <https://files.eric.ed.gov/fulltext/ED560407.pdf>
- Geier, B. & Riedel, B. (2008). Ungleichheiten der Inanspruchnahme öffentlicher frühpädagogischer Angebote. Einflussfaktoren und Restriktionen elterlicher Betreuungentscheidungen. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen. Zeitschrift für Erziehungswissenschaft* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 11, S. 11–28). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH Wiesbaden. https://doi.org/10.1007/978-3-531-91452-7_2
- Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-93192-0>
- Geißler, R. & Weber-Menges, S. (2010). Bildungsungleichheit – Eine deutsche Altlast. Die bildungssoziologische Perspektive. In H. Barz (Hrsg.), *Handbuch Bildungsfinanzierung* (S. 155–165). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage. https://doi.org/10.1007/978-3-531-92520-2_13

- Glasgow, K. L., Dornbusch, S. M., Troyer, L., Steinberg, L. & Ritter, P. L. (1997). Parenting styles, adolescents' attributions, and educational outcomes in nine heterogeneous high schools. *Child Development*, 68(3), 507–529. <https://doi.org/10.2307/1131675>
- Gniewosz, B. & Gräsel, C. (2015). Überblick Soziale Ungleichheit. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (S. 195–199). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-19994-8_14
- Goebel, J. & Hoppe, L. (2015). Ausmaß und Trends sozialräumlicher Segregation in Deutschland. Zugriff am 06.05.2020. Verfügbar unter: https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/a-305-7-abschlussbericht-ausmass-trends-sozialraeumlicher-segregation.pdf?__blob=publicationFile&v=1
- Goldstein, H. (1997). *Methods in School Effectiveness Research*. *School Effectiveness and School Improvement*, 8(4), 369–395. <https://doi.org/10.1080/0924345970080401>
- Göllner, R., Wagner, W., Klieme, E., Lüdtke, O., Nagengast, B. & Trautwein (2016). Erfassung der Unterrichtsqualität mithilfe von Schülerurteilen: Chancen, Grenzen und Forschungsperspektiven. In Bundesministerium für Bildung und Forschung (Hrsg.), *Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments* (S. 63–82). Berlin: Bundesministerium für Bildung und Forschung.
- Goodman, A. & Gregg, P. (2010). Poorer children's educational attainment: how important are attitudes and behaviour? Zugriff am 04.11.2021. Verfügbar unter: <https://www.jrf.org.uk/report/poorer-children%E2%80%99s-educational-attainment-how-important-are-attitudes-and-behaviour>
- Görlitz, K., Spieß, C. K. & Ziege, E. (2018). Fast jedes zehnte Kind geht auf eine Privatschule – Nutzung hängt insbesondere in Ostdeutschland zunehmend vom Einkommen der Eltern ab. *DIW-Wochenbericht*, 1103–1111.
- Gräsel, C. (2015). Was ist Empirische Bildungsforschung? In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Strukturen und Methoden* (S. 15–30). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-531-19992-4_1
- Graham, L. J., de Bruin, K., Lassig, C. & Spandagou, I. (2021). A scoping review of 20 years of research on differentiation: investigating conceptualisation, characteristics, and methods used. *Review of Education*, 9(1), 161–198. <https://doi.org/10.1002/rev3.3238>
- Gruehn, S. (1995). Vereinbarkeit kognitiver und nichtkognitiver Ziele im Unterricht. *Zeitschrift für Pädagogik*, 41(4), 531–553.
- Gruehn, S. (1998). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Inaugural-Dissertation am Fachbereich Erziehungswissenschaften, Psychologie und Sportwissenschaften der Freien Universität Berlin. Berlin: Freie Universität Berlin.
- Guill, K. & Bos, W. (2014). Effectiveness of private tutoring in mathematics with regard to subjective and objective indicators of academic achievement. *Journal for Educational Research Online*, 6, 34–67.
- Guill, K., Gröhlich, C., Scharenberg, K., Wendt, H. & Bos, W. (2011). Die mathematischen Kompetenzen der Schülerinnen und Schüler. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern*

- lern am Ende der Jahrgangsstufe 8 (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 37–48). Münster: Waxmann.
- Guill, K., Lüdtke, O. & Köller, O. (2019). Assessing the instructional quality of private tutoring and its effects on student outcomes: Analyses from the German National Educational Panel Study. *The British Journal of Educational Psychology*, 90(2), 282–300. <https://doi.org/10.1111/bjep.12281>
- Guill, K. & Wendt, H. (2016). Außerschulischer Nachhilfunterricht am Ende der Grundschulzeit. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 247–256). Münster: Waxmann.
- Gülseven, Z., Kumru, A., Carlo, G., Palermo, F., Selçuk, B. & Sayıl, M. (2018). The mediational roles of harsh and responsive parenting in the longitudinal relations between socioeconomic status and Turkish children's emotional development. *International Journal of Behavioral Development*, 42(6), 563–573. <https://doi.org/10.1177/0165025418783279>
- Gustafsson, J.-E., Nilsen, T. & Hansen, K. Y. (2018). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*, 57, 16–30. <https://doi.org/10.1016/j.stueduc.2016.09.004>
- Haag, L. (2001). Hält bezahlter Nachhilfunterricht, was er verspricht? Eine Evaluationsstudie. *Zeitschrift für Pädagogische Psychologie*, 15(1), 38–44. <https://doi.org/10.1024//1010-0652.15.1.38>
- Hanushek, E. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature*, 24, 1141–1177.
- Hanushek, E. A. & Woessmann, L. (2006). Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries. *The Economic Journal*, 116(510), C63-C76. <https://doi.org/10.1111/j.1468-0297.2006.01076.x>
- Harnischfeger, A. & Wiley, D. E. (1978). Conceptual Issues in Models of School Learning. *Journal of Curriculum Studies*, 10(3), 215–231. <https://doi.org/10.1080/0022027780100304>
- Hattie, J. (2013). *Lernen sichtbar machen*. Baltmannsweiler: Schneider-Verl. Hohengehren.
- Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37(5), 449–481. [https://doi.org/10.1016/S0883-0355\(03\)00015-6](https://doi.org/10.1016/S0883-0355(03)00015-6)
- Hedges, L. V. & Hedberg, E. C. (2007). Intraclass Correlation Values for Planning Group-Randomized Trials in Education. *Educational Evaluation and Policy Analysis*, 29(1), 60–87. <https://doi.org/10.3102/0162373707299706>
- Helmke, A. (1988). Leistungssteigerung und Ausgleich von Leistungsunterschieden in Schulklassen: Unvereinbare Ziele? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 45–76.
- Helmke, A. (2007a). *Unterrichtsqualität erfassen, bewerten, verbessern*. Seelze: Klett Kallmeyer.

- Helmke, A. (2007b). Was wissen wir über guten Unterricht? Wissenschaftliche Erkenntnisse zur Unterrichtsforschung und Konsequenzen für die Unterrichtsentwicklung. Zugriff am 09.01.2020. Verfügbar unter: https://www.bildung.koeln.de/imperia/md/content/selbst_schule/downloads/andreas_helmke_.pdf
- Helmke, A. (2014). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts. Seelze-Velber: Klett/Kallmeyer.
- Helmke, A. & Helmke, T. (2014). Wie wirksam ist gute Klassenführung? *Lernende Schule*, 65, 9–12.
- Helmke, A. & Schrader, F.-W. (2008). Merkmale der Unterrichtsqualität: Potenzial, Reichweite und Grenzen. *SEMINAR – Lehrerbildung und Schule*, (3), 17–47.
- Helmke, A. & Schrader, F.-W. (2010). Determinanten der Schulleistung. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 90–102). Weinheim: Beltz.
- Helmke, A., Schrader, F.-W. & Lehneis-Klepper, G. (1991). Zur Rolle des Elternverhaltens für die Schulleistungsentwicklung ihrer Kinder. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 23(1), 1–22.
- Helmke, A. & Weinert, F.E. (1997a). Bedingungsfaktoren schulischer Leistungen. In F.E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule (Enzyklopädie der Psychologie Praxisgebiete Pädagogische Psychologie, Bd. 3, S. 71–176)*. Göttingen: Hogrefe Verl. für Psychologie.
- Helmke, A. & Weinert, F.E. (1997b). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F.E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241–251). Weinheim: Beltz.
- Hille, A., Spieß, C.K. & Staneva, M. (2016). Immer mehr Schülerinnen und Schüler nehmen Nachhilfe, besonders in Haushalten mit mittleren Einkommen. *DIW-Wochenbericht*.
- Hjalmarsson, S. & Mood, C. (2015). Do poorer youth have fewer friends? The role of household and child economic resources in adolescent school-class friendships. *Children and Youth Services Review*, 57, 201–211. <https://doi.org/10.1016/j.childyouth.2015.08.013>
- Hoff, E., Laursen, B. & Tardif, T. (2002). Socioeconomic Status and Parenting. In M. H. Bornstein (Hrsg.), *Handbook of parenting* (S. 231–252). Mahwah, NJ: Erlbaum.
- Hoffmann, L., Stanat, P., Maaz, K. & Klemm, K. (2019). Kompetenzen von Schülerinnen und Schülern an Schulen in privater und öffentlicher Trägerschaft. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 71(3), 385–408. <https://doi.org/10.1007/s11577-019-00638-2>
- Holtappels, H. G., Radisch, F., Rollett, W. & Kowoll, M.E. (2010). Bildungsangebot und Schülerkompetenzen in Ganztagsgrundschulen. In W. Bos, K. Schwippert & K.-H. Arnold (Hrsg.), *IGLU 2006 – die Grundschule auf dem Prüfstand. Vertiefende Analysen zu Rahmenbedingungen schulischen Lernens* (S. 165–198). Münster: Waxmann.
- Holzberger, D., Reinhold, S., Lütke, O. & Seidel, T. (2020). A meta-analysis on the relationship between school characteristics and student outcomes in science and maths – evidence from large-scale studies. *Studies in Science Education*, 56(1), 1–34. <https://doi.org/10.1080/03057267.2020.1735758>

- Hornung, C., Schiltz, C., Brunner, M. & Martin, R. (2014). Predicting first-grade mathematics achievement: the contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Frontiers in Psychology*, 5, 272. <https://doi.org/10.3389/fpsyg.2014.00272>
- Hosokawa, R. & Katsura, T. (2017). A longitudinal study of socioeconomic status, family processes, and child adjustment from preschool until early elementary school: the role of social competence. *Child and Adolescent Psychiatry and Mental Health*, 11, 62. <https://doi.org/10.1186/s13034-017-0206-z>
- Hox, J. J. (2010). *Multilevel Analysis. Techniques and Applications*. New York: Routledge. <https://doi.org/10.4324/9780203852279>
- Hulleman, C. S. & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, 326(5958), 1410–1412. <https://doi.org/10.1126/science.1177067>
- Hußmann, A., Stubbe, T. C. & Kasper, D. (2017). Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes et al. (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 195–218). Münster: Waxmann.
- Ivanov, S. (2011). Mathematische Kompetenz und Einstellungen zum Mathematikunterricht. In U. Vieluf, S. Ivanov & R. Nikolova (Hrsg.), *KESS 10/11. Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe* (S. 75–122). Hamburg: Freie und Hansestadt Hamburg Behörde für Schule und Berufsbildung.
- Ivanov, S. & Nikolova, R. (2010). Naturwissenschaftliche Kompetenzen. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8 (HANSE – Hamburger Schriften zur Qualität im Bildungswesen, Bd. 6, S. 79–99)*. Münster: Waxmann.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H. et al. (1972). *Inequality. A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jeynes, W. H. (2005). A Meta-Analysis of the Relation of Parental Involvement to Urban Elementary School Student Academic Achievement. *Urban Education*, 40(3), 237–269. <https://doi.org/10.1177/0042085905274540>
- Johnson, K. A. (2000). *The Peer Effect on Academic Achievement among Public Elementary School Students*. Zugriff am 11.04.2021. Verfügbar unter: <https://www.heritage.org/education/report/the-peer-effect-academic-achievement-among-public-elementary-school-students>
- Johnson, D. W., Johnson, R. T. & Stanne, M. B. (2000). *Cooperative Learning Methods: A Meta-Analysis*. Zugriff am 11.04.2021. Verfügbar unter: https://www.researchgate.net/profile/David_Johnson50/publication/220040324_Cooperative_learning_methods_A_meta-analysis/links/00b4952b39d258145c000000.pdf
- Jungbauer-Gans, M. (2004). Einfluss des sozialen und kulturellen Kapitals auf die Lesekompetenz. *Zeitschrift für Soziologie*, 33(5), 375–397. <https://doi.org/10.1515/zfsoz-2004-0502>

- Jungbauer-Gans, M., Lohmann, H. & Spieß, C. K. (2012). Bildungsungleichheiten und Privatschulen in Deutschland. In R. Becker & H. Solga (Hrsg.), *Soziologische Bildungsforschung (Kölner Zeitschrift für Soziologie und Sozialpsychologie: Sonderheft, Bd. 52, S. 64–85)*. Wiesbaden: Springer Fachmedien Wiesbaden; Imprint; Springer VS. https://doi.org/10.1007/978-3-658-00120-9_3
- Kaiser, T. & Diewald, M. (2014). Social origin, conscientiousness, and school grades: Does early socialization of the characteristics orderliness and focus contribute to the reproduction of social inequality? *Research in Social Stratification and Mobility*, 38, 93–105. <https://doi.org/10.1016/j.rssm.2014.07.001>
- Kalaycioglu, D. B. (2015). The Influence of Socioeconomic Status, Self-efficacy, and Anxiety on Mathematics Achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA. *Educational Sciences: Theory & Practice*, 15(5), 1391–1401. <https://doi.org/10.12738/estp.2015.5.2731>
- Karlen, Y., Merki, K. M. & Ramseier, E. (2014). The effect of individual differences in the development of metacognitive strategy knowledge. *Instructional Science*, 42(5), 777–794. <https://doi.org/10.1007/s11251-014-9314-9>
- Kawaguchi, D. (2013). Fewer School Days, More Inequality. Zugriff am 11.04.2021. Verfügbar unter: <https://pdfs.semanticscholar.org/2bce/5ab033aa482201acf8f333a4d188eab09354.pdf>
- Kellermann, C., Nachbauer, M., Gärtner, H. & Thiel, F. (2022). Effekte eines Unterrichtsfeedbacks von Schulleitungen auf die selbst eingeschätzte Entwicklung der Unterrichtskompetenz von Lehrkräften – Ergebnisse einer Interventionsstudie unter Berücksichtigung der Zielorientierungen von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 25, 1267–1291. <https://doi.org/10.1007/s11618-022-01087-y>
- Kellermann, C., Nachbauer, M., Gärtner, H. & Thiel, F. (2023). The effect of school leaders' feedback on teaching quality perceived by students – results of an intervention study. *School Effectiveness and School Improvement*, 34, 24–42. <https://doi.org/10.1080/09243453.2022.2075018>
- Kierner, K., Haag, N., Müller, K. & Ehmke, T. (2017). Einfluss sozialer und zuwanderungsbezogener Disparitäten, sowie der Klassenkomposition auf die Veränderung der mathematischen Kompetenz von der neunten zur zehnten Klassenstufe. *Zeitschrift für Erziehungswissenschaft*, 20, 125–149. <https://doi.org/10.1007/s11618-017-0753-3>
- Kim, J. (2016). Female education and its impact on fertility. *IZA World of Labor*, 1–10. <https://doi.org/10.15185/izawol.228>
- Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
- Klauer, K. J. & Leutner, D. (2012). *Lehren und Lernen. Einführung in die Instruktionspsychologie*. Weinheim: Beltz Verlagsgruppe.
- Kleckmann, T., Steffensky, M. & Praetorius, A.-K. (2020). Quality of Teaching in Science Education. More Than Three Basic Dimensions? *Zeitschrift für Pädagogik*, 37–53.

- Klieme, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen? *Pädagogik*, 56, 10–13.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel, L. Allolio-Näcke & Prenzel-Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127–146). Münster: Waxmann.
- Klieme, E., Neubrand, M. & Lüdtke, O. (2001a). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, P. Schneider et al. (Hrsg.), *PISA 2000* (S. 139–190). Opladen: Leske + Budrich. https://doi.org/10.1007/978-3-322-83412-6_5
- Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras Study: Investigating Effects of Teaching and Learning in Swiss and German Mathematics Classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–160). Münster: Waxmann.
- Klieme, E., Schümer, G. & Knoll, S. (2001b). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung. In E. Klieme & J. Baumert (Hrsg.), *TIMSS – Impulse für Schule und Unterricht* (S. 43–57). Bonn: Bundesministerium für Bildung und Forschung.
- Kline, R.B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Kluczniok, K., Lehl, S., Kuger, S. & Roszbach, H.-G. (2013). Quality of the home learning environment during preschool age – Domains and contextual conditions. *European Early Childhood Education Research Journal*, 21(3), 420–438. <https://doi.org/10.1080/1350293X.2013.814356>
- Köppe, S. (2012). Wahlfreiheit und Nutzerrollen im deutschen Bildungssystem. *WSI Mitteilungen*, 206–215. <https://doi.org/10.5771/0342-300X-2012-3-206>
- Kornhauser, A. W. (1918). The economic standing of parents and the intelligence of their children. *Journal of Educational Psychology*, 9, 159–164. <https://doi.org/10.1037/h0072624>
- Kraft, M. A., Blazar, D. & Hogan, D. (2018). The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Kramer, B. & Wagner, M. (2012). Die Bedeutung des kulturellen Kapitals für Freundschaften unter Schülern. *Soziale Welt*, 63(3), 213–231. <https://doi.org/10.5771/0038-6073-2012-3-213>
- Kriegbaum, K. & Spinath, B. (2016). Explaining Social Disparities in Mathematical Achievement: The Role of Motivation. *European Journal of Personality*, 30(1), 45–63. <https://doi.org/10.1002/per.2042>
- Kuhn, H. P. & Fischer, N. (2011). Entwicklung der Schulnoten in der Ganztagschule. In N. Fischer, H. G. Holtappels, E. Klieme, T. Rauschenbach, L. Stecher & I. Züchner (Hrsg.), *Ganztagschule: Entwicklung, Qualität, Wirkungen. Längsschnittliche Befunde der Studie zur Entwicklung von Ganztagschulen (StEG) (Studien zur ganztägigen Bildung*, S. 207–226). Weinheim und Basel: Beltz Juventa.

- Kultusministerkonferenz. (2002). PISA 2000 – Zentrale Handlungsfelder. Zugriff am 11.04.2021. Verfügbar unter: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf
- Kultusministerkonferenz. (2004). Bildungsstandards der Kultusministerkonferenz. Zugriff am 22.11.2019. Verfügbar unter: www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Konzeption-Entwicklung.pdf
- Kultusministerkonferenz. (2015). Ganztagschulen in Deutschland. Zugriff am 11.04.2012. Verfügbar unter: www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015-12-03-Ganztagsschulbericht.pdf
- Kunter, M. (2005). Multiple Ziele im Mathematikunterricht (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 51). Zugl.: Berlin, Freie Univ., Diss., 2004. Münster: Waxmann.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV. Münster: Waxmann. <https://doi.org/10.31244/9783830974338>
- Kunter, M., Dubberke, T., Baumert, J., Blum, W., Brunner, M., Jordan, A. et al. (2006). Mathematikunterricht in den PISA-Klassen 2004: Rahmenbedingungen, Formen und Lehr-Lernprozesse. In M. Prenzel (Hrsg.), PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres (S. 161–194). Münster: Waxmann.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820. <https://doi.org/10.1037/a0032583>
- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts*. Paderborn: Ferdinand Schöningh. <https://doi.org/10.36198/9783838538952>
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Münster: Waxmann. <https://doi.org/10.31244/9783830974338>
- Kyriakides, L. (2004). Differential School Effectiveness in Relation to Sex and Social Class: Some Implications for Policy Evaluation. *Educational Research and Evaluation*, 10(2), 141–161. <https://doi.org/10.1076/edre.10.2.141.27907>
- Kyriakides, L. (2007). Generic and Differentiated Models of Educational Effectiveness: Implications for the Improvement of Educational Practice. In T. Townsend (Hrsg.), *International Handbook of School Effectiveness and Improvement* (S. 41–56). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-5747-2_3
- Kyriakides, L., Christoforou, C. & Charalambous, C. Y. (2013). What matters for student learning outcomes: A meta-analysis of studies exploring factors of effective teaching. *Teaching and Teacher Education*, 36, 143–152. <https://doi.org/10.1016/j.tate.2013.07.010>
- Kyriakides, L. & Creemers, B. (2011). Can Schools Achieve Both Quality and Equity? Investigating the Two Dimensions of Educational Effectiveness. *Journal of Educa-*

- tion for Students Placed at Risk (JESPAR), 16(4), 237–254. <https://doi.org/10.1080/10824669.2011.610269>
- Kyriakides, L., Creemers, B., Antoniou, P. & Demetriou, D. (2010). A synthesis of studies searching for school factors: implications for theory and research. *British Educational Research Journal*, 36(5), 807–830. <https://doi.org/10.1080/01411920903165603>
- Kyriakides, L., Creemers, B. & Charalambous, E. (2018a). Equity and Quality Dimensions in Educational Effectiveness. Cham: Springer. <https://doi.org/10.1007/978-3-319-72066-1>
- Kyriakides, L., Creemers, B. & Charalambous, E. (2018b). Searching for differential teacher and school effectiveness in terms of student socioeconomic status and gender: implications for promoting equity. *School Effectiveness and School Improvement*, 30(3), 286–308. <https://doi.org/10.1080/09243453.2018.1511603>
- Lauer, P.A., Akiba, M., Wilkerson, S.B., Apthorp, H.S., Snow, D. & Martin-Glenn, M.L. (2006). Out-of-School-Time Programs: A Meta-Analysis of Effects for At-Risk Students. *Review of Educational Research*, 76(2), 275–313. <https://doi.org/10.3102/00346543076002275>
- Lee, V.E. & Bryk, A.S. (1989). A Multilevel Model of the Social Distribution of High School Achievement. *Sociology of Education*, 62(3), 172. <https://doi.org/10.2307/2112866>
- Lee, V.E. & Burkam, D.T. (2002). Inequality at the starting gate. Social background differences in achievement as children begin school. Washington, DC: Economic Policy Institute.
- Lehmann, R., Gänsfuß, R. & Peek, R. (1999). Aspekte der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern, die im Schuljahr 1996/97 eine fünfte Klasse in Hamburg besuchten. Bericht über die Erhebung im September 1998 (LAU7).
- Lehmann, R. & Lenkeit, J. (2008). ELEMENT Erhebung zum Lese- und Mathematikverständnis. Entwicklungen in den Jahrgangsstufen 4 bis 6 in Berlin.
- Lehrl, S., Ebert, S., Rossbach, H.-G. & Weinert, S. (2012). Die Bedeutung der familiären Lernumwelt für Vorläufer schriftsprachlicher Kompetenzen im Vorschulalter. *Zeitschrift für Familienforschung*, 24, 115–133. <https://doi.org/10.20377/jfr-181>
- Lehrl, S., Kuger, S. & Anders, Y. (2014). Soziale Disparitäten beim Zugang zu Kindergartenqualität und differenzielle Konsequenzen für die vorschulische mathematische Entwicklung. *Unterrichtswissenschaft*, 42, 132–151.
- Leibniz Institut für Bildungsverläufe. (2019). Studienübersicht NEPS Startkohorte 3 – Klasse 5. Zugriff am 11.04.2021. Verfügbar unter: https://www.neps-data.de/Portals/0/NEPS/Datenzentrum/Forschungsdaten/SC3/9-0-0/SC3_Studien_W1-9_de.pdf
- Lenz, G. & Sahn, A. (2020). Achieving Statistical Significance with Covariates and without Transparency. Zugriff am 04.01.2021. Verfügbar unter: <https://osf.io/preprints/metaarxiv/s42ba/>
- Letourneau, N.L., Duffett-Leger, L., Levac, L., Watson, B. & Young-Morris, C. (2013). Socioeconomic Status and Child Development. *Journal of Emotional and Behavioral Disorders*, 21(3), 211–224. <https://doi.org/10.1177/1063426611421007>

- Leven, I. & Schneekloth, U. (2010). Die Freizeit: Sozial getrennte Kinderwelten. In K. Hurrelmann & S. Andresen (Hrsg.), *Kinder in Deutschland 2010*. 2. World Vision Kinderstudie. Frankfurt am Main: Fischer.
- Linberg, A., Attig, M. & Weinert, S. (2020). Social disparities in the vocabulary of 2-year-old children and the mediating effect of language-stimulating interaction behavior. *Journal for Educational Research Online*, 12, 12–35.
- Linberg, T., Struck, O. & Bäumer, T. (2015). Analysen zum Zusammenhang von ganztägiger Beschulung mit den Kompetenzen von Schülern in Deutschland. Unveröffentlichter Forschungsbericht. <https://doi.org/10.1007/s11618-018-0830-2>
- Linberg, T., Struck, O. & Bäumer, T. (2018). Vorzug Ganztagschule? Zusammenhänge mit der Kompetenzentwicklung im Bereich Lesen und Mathematik. *Zeitschrift für Erziehungswissenschaft*, 21(6), 1205–1227.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (Zeitschrift für Pädagogik Beiheft, Bd. 51, S. 47–70). Weinheim: Beltz.
- Lipowsky, F. & Lotz, M. (2015). Ist Individualisierung der Königsweg zum erfolgreichen Lernen? Eine Auseinandersetzung mit Theorien, Konzepten und empirischen Befunden. In G. Mehlhorn, K. Schöppe & F. Schulz (Hrsg.), *Begabungen entwickeln & Kreativität fördern* (S. 155–219). München: kopaed.
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E. & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>
- Lochner, L. & Moretti, E. (2004). The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports. *American Economic Review*, 94(1), 155–189. <https://doi.org/10.1257/000282804322970751>
- Lou, Y., Abrami, P.C., Spence, J.C., Poulsen, C., Chambers, B. & d'Apollonia, S. (1996). Within-Class Grouping: A Meta-Analysis. *Review of Educational Research*, 66(4), 423–458. <https://doi.org/10.3102/00346543066004423>
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T. & Muthén, B. (2008). The multilevel latent covariate model: a new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13(3), 203–229. <https://doi.org/10.1037/a0012869>
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *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 modelling. *Contemporary Educational Psychology*, 34, 120–131. <https://doi.org/10.1016/j.cedpsych.2008.12.001>
- Lüdtke, O., Trautwein, U., Kunter, M. & Baumert, J. (2006). Analyse von Lernumwelten. *Zeitschrift für Pädagogische Psychologie*, 20, 85–96. <https://doi.org/10.1024/1010-0652.20.12.85>

- Luplow, N. & Schneider, T. (2014). Nutzung und Effektivität privat bezahlter Nachhilfe im Primarbereich. *Zeitschrift für Soziologie*, 43. <https://doi.org/10.1515/zfs-oz-2014-0104>
- Lüttinger, P. & König, W. (1988). Die Entwicklung einer international vergleichbaren Klassifikation für Bildungssysteme. *ZUMA Nachrichten*, 12, 1–14.
- Luyten, H. (2003). The Size of School Effects Compared to Teacher Effects: An Overview of the Research Literature. *School Effectiveness and School Improvement*, 14(1), 31–51. <https://doi.org/10.1076/sesi.14.1.31.13865>
- Luyten, H., Merrell, C. & Tymms, P. (2017). The contribution of schooling to learning gains of pupils in Years 1 to 6. *School Effectiveness and School Improvement*, 28(3), 374–405. <https://doi.org/10.1080/09243453.2017.1297312>
- Ma, X., Yuan, J. & Luo, X. (2016). Achievement-related within-school socioeconomic gaps in science subjects in China: evidence on existence, consistency, and compensation. *School Effectiveness and School Improvement*, 27(4), 511–533. <https://doi.org/10.1080/09243453.2016.1201122>
- Maaz, K., Baumert, J. & Trautwein, U. (2010). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 12). Wiesbaden: VS Verlag für Sozialwissenschaften/Springer Fachmedien Wiesbaden, Wiesbaden. https://doi.org/10.1007/978-3-531-92216-4_2
- Maehler, D. B., Massing, N. Helmschrott, S., Rammstedt, B. Staudinger, U. M. & Wolf, C. (2013). Grundlegende Kompetenzen in verschiedenen Bevölkerungsgruppen. In B. Rammstedt (Hrsg.), *Grundlegende Kompetenzen Erwachsener im internationalen Vergleich. Ergebnisse von PIAAC 2012* (S. 77–126). Münster: Waxmann.
- Marzano, R. J. (2000). *A New Era of School Reform: Going Where the Research Takes Us*. Aurora: McREL.
- Marzano, R. J., Gaddy, B. B. & Dean, C. (2000). *What Works in Classroom Instruction*. Zugriff am 11.04.2021. Verfügbar unter: https://www.researchgate.net/publication/265663591_What_Works_In_Classroom_Instruction
- Marzano, R. J., Marzano, J. S. & Pickering, D. J. (2003). *Classroom Management That Works. Research-Based Strategies for Every Teacher*. Zugriff am 11.04.2021. Verfügbar unter: [http://perino.pbworks.com/f/4%20CLM-Basic-Text\(M\).pdf](http://perino.pbworks.com/f/4%20CLM-Basic-Text(M).pdf)
- Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist*, 59(1), 14–19. <https://doi.org/10.1037/0003-066X.59.1.14>
- McCartney, K., Dearing, E., Taylor, B. A. & Bub, K. L. (2007). Quality Child Care Supports the Achievement of Low-Income Children: Direct and Indirect Pathways Through Caregiving and the Home Environment. *Journal of Applied Developmental Psychology*, 28(5–6), 411–426. <https://doi.org/10.1016/j.appdev.2007.06.010>
- McDonald, R. P. (1999). *Test theory. A unified treatment*. Mahwah, NJ: L. Erlbaum Associates.
- McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>

- Miech, R., Essex, M. J. & Goldsmith, H. H. (2001). Socioeconomic Status and the Adjustment to School: The Role of Self-Regulation during Early Childhood. *Sociology of Education*, 74(2), 102. <https://doi.org/10.2307/2673165>
- Mietzel, G. (2007). *Pädagogische Psychologie des Lernens und Lehrens*. Göttingen: Hogrefe.
- Mischo, C. & Haag, L. (2002). Expansion and effectiveness of private tutoring. *European Journal of Psychology of Education*, 17(3), 263–273. <https://doi.org/10.1007/BF03173536>
- Möller, K., Jonen, A., Hardy, I. & Stern, E. (2002). Die Förderung von naturwissenschaftlichem Verständnis bei Grundschulkindern durch Strukturierung der Lernumgebung. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule. Schulische und außerschulische Bedingungen mathematischer, naturwissenschaftlicher und einfacher Kompetenzen* (Zeitschrift für Pädagogik Beiheft, S. 176–191). Weinheim: Beltz.
- Mortimore, P. (1998). *The Road to Improvement. Reflections on School Effectiveness*. Lisse, Exton: Swets & Zeitlinger Publishers.
- Moser, U., Oostlander, J. & Tomasik, M. J. (2017). Soziale Ungleichheiten im Leistungszuwachs und bei Bildungsübergängen. In M. P. Neuenschwander & C. Näggle (Hrsg.), *Bildungsverläufe von der Einschulung bis in den ersten Arbeitsmarkt: Theoretische Ansätze, empirische Befunde und Beispiele* (S. 59–77). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-16981-7_4
- Muijs, D. (2012). Methodological change in educational effectiveness research. In C. Chapman, P. Armstrong, A. Harris, D. Muijs, D. Reynolds & P. Sammons (Hrsg.), *School effectiveness and improvement research, policy, and practice. Challenging the orthodoxy* (S. 58–66). New York, NY: Routledge.
- Muijs, D., Campbell, J., Kyriakides, L. & Robinson, W. (2005). Making the Case for Differentiated Teacher Effectiveness: An Overview of Research in Four Key Areas. *School Effectiveness and School Improvement*, 16(1), 51–70. <https://doi.org/10.1080/09243450500113985>
- Muijs, D., Kyriakides, L., van der Werf, G., Creemers, B., Timperley, H. & Earl, L. (2014). State of the art – teacher effectiveness and professional learning. *School Effectiveness and School Improvement*, 25(2), 231–256. <https://doi.org/10.1080/09243453.2014.885451>
- Muijs, D. & Reynolds, D. (2001). *Effective teaching. Evidence and practice*. London: Paul Chapman.
- Muijs, R. D. (1997). Predictors of academic achievement and academic self-concept; a longitudinal perspective. *The British Journal of Educational Psychology*, 67(3), 263–77. <https://doi.org/10.1111/j.2044-8279.1997.tb01243.x>
- Müller, A. & Ditton, H. (2014). Feedback: Begriff, Formen und Funktionen. In H. Ditton & A. Müller (Hrsg.), *Feedback und Rückmeldungen. Theoretische Grundlagen, empirische Befunde, praktische Anwendungsfelder* (S. 11–28). Münster: Waxmann.
- Müller, K. & Ehmke, T. (2016). Soziale Herkunft und Kompetenzerwerb. In K. Reiss, C. Sälzer, A. Schiepe-Tiska, E. Klieme & O. Köller (Hrsg.), *PISA 2015. Eine Studie zwischen Kontinuität und Innovation* (S. 285–316). Münster: Waxmann.

- Muthén, L. K. & Muthén, B. (1998–2018). *Mplus User's Guide*. Los Angeles: Muthén & Muthén.
- Nachbauer, M. & Kyriakides, L. (2020). A review and evaluation of approaches to measure equity in educational outcomes. *School Effectiveness and School Improvement*, 31(2), 306–331.
- NEPS-Netzwerk (2021). Nationales Bildungspanel, Scientific Use File der Startkohorte Klasse 5. Leibniz-Institut für Bildungsverläufe (LIfBi), Bamberg. <https://doi.org/10.5157/NEPS:SC3:11.0.1>
- Neumann, M., Becker, M. & Maaz, K. (2014). Soziale Ungleichheiten in der Kompetenzentwicklung in der Grundschule und der Sekundarstufe I. In K. Maaz, M. Neumann & J. Baumert (Hrsg.), *Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter. Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive* (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 24). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-00454-5_8
- Nye, B., Konstantopoulos, S. & Hedges, L. V. (2004). How Large Are Teacher Effects? *Educational Evaluation and Policy Analysis*, 26(3), 237–257. <https://doi.org/10.3102/01623737026003237>
- O'Connor, M. C. & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43(5), 971–990. <https://doi.org/10.1016/j.paid.2007.03.017>
- OECD. (2007). PISA 2006. Science Competencies for Tomorrow's World: Part 1 Analysis. <https://doi.org/10.1787/9789264040014-en>
- OECD. (2013). PISA 2012 Results: Excellence through Equity. Giving every Student the Chance to Succeed Volume II. Zugriff am 11.04.2021. Verfügbar unter: <https://www.oecd.org/pisa/keyfindings/pisa-2012-results-volume-II.pdf>
- OECD. (2016a). PISA 2015. Ergebnisse im Fokus. Zugriff am 09.01.2020. Verfügbar unter: https://www.oecd.org/berlin/themen/pisa-studie/PISA_2015_Zusammenfassung.pdf
- OECD. (2016b). Teaching Strategies for Instructional Quality. Zugriff am 11.04.2021. Verfügbar unter: http://www.oecd.org/education/school/TALIS-PISA-LINK-teaching_strategies_brochure.pdf
- OECD. (2016c). Ten Questions for Mathematics Teachers. Zugriff am 11.04.2021. Verfügbar unter: <https://www.oecd-ilibrary.org/docserver/9789264265387-en.pdf?expires=1594417417&id=id&accname=guest&checksum=7E547902F064DD585E572E8C5D8ADF7D>
- OECD. (2019). Under Pressure: The Squeezed Middle Class. Zugriff am 11.04.2021. Verfügbar unter: https://www.oecd-ilibrary.org/social-issues-migration-health/under-pressure-the-squeezed-middle-class_689afed1-en
- Ohtani, K. & Hisasaka, T. (2018). Beyond intelligence: a meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning*, 13(2), 179–212. <https://doi.org/10.1007/s11409-018-9183-8>
- Ömerogullari, M., Guill, K. & Köller, O. (2020). Effectiveness of private tutoring during secondary schooling in Germany: Do the duration of private tutoring and tutor qualification affect school achievement? *Learning and Instruction*, 66, 101306. <https://doi.org/10.1016/j.learninstruc.2020.101306>

- Opdenakker, M.-C. & van Damme, J. (2000). Effects of Schools, Teaching Staff and Classes on Achievement and Well-Being in Secondary Education: Similarities and Differences Between School Outcomes. *School Effectiveness and School Improvement*, 11(2), 165–196. [https://doi.org/10.1076/0924-3453\(200006\)11:2;1-Q;FT165](https://doi.org/10.1076/0924-3453(200006)11:2;1-Q;FT165)
- Opdenakker, M.-C., van Damme, J., Fraine, F. de, van Landeghem, G. & Onghena, P. (2002). The Effect of Schools and Classes on Mathematics Achievement. *School Effectiveness and School Improvement*, 13(4), 399–427. <https://doi.org/10.1076/sesi.13.4.399.10283>
- Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: a multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement*, 19(1), 21–49. <https://doi.org/10.1080/09243450801936845>
- Park, H., Buchmann, C., Choi, J. & Merry, J. J. (2016). Learning Beyond the School Walls: Trends and Implications. *Annual Review of Sociology*, 42(1), 231–252. <https://doi.org/10.1146/annurev-soc-081715-074341>
- Patall, E. A., Cooper, H. & Allen, A. B. (2010). Extending the School Day or School Year. *Review of Educational Research*, 80(3), 401–436. <https://doi.org/10.3102/0034654310377086>
- Patall, E. A., Cooper, H. & Robinson, J. C. (2008). Parent Involvement in Homework: A Research Synthesis. *Review of Educational Research*, 78(4), 1039–1101. <https://doi.org/10.3102/0034654308325185>
- Peisert, H. (1967). *Soziale Lage und Bildungschancen in Deutschland*. München: Piper.
- Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L. & Yazejian, N. (2001). The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. *Child development*, 72(5). <https://doi.org/10.1111/1467-8624.00364>
- Pellegrini, M., Lake, C., Neitzel, A. & Slavin, R. E. (2020). Effective Programs in Elementary Mathematics: A Meta-Analysis. Zugriff am 17.05.2020. Verfügbar unter: http://www.bestevidence.org/word/elem_math_May_26_2020_full.pdf
- Peter, F. & Storck, J. (2015). Persönlichkeitseigenschaften beeinflussen die Studienabsicht von Jugendlichen. DIW-Wochenbericht.
- Pfeifer, M. (2011). Bildungsbenachteiligung und das Potenzial von Schule und Unterricht. Lesekompetenz bei sozioökonomisch benachteiligten Schülern. Wiesbaden: VS Verlag für Sozialwissenschaften/Springer Fachmedien Wiesbaden GmbH Wiesbaden. <https://doi.org/10.1007/978-3-531-94014-4>
- Pfost, M., Dörfler, T. & Artelt, C. (2010). Der Zusammenhang zwischen außerschulischem Lesen und Lesekompetenz. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42(3), 167–176. <https://doi.org/10.1026/0049-8637/a000017>
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D. et al. (2005). Features of Pre-Kindergarten Programs, Classrooms, and Teachers: Do They Predict Observed Classroom Quality and Child-Teacher Interactions? *Applied Developmental Science*, 9(3), 144–159. https://doi.org/10.1207/s1532480xads0903_2

- Pinquart, M. (2016). Associations of Parenting Styles and Dimensions with Academic Achievement in Children and Adolescents: A Meta-analysis. *Educational Psychology Review*, 28(3), 475–493. <https://doi.org/10.1007/s10648-015-9338-y>
- Piwovar, V., Thiel, F. & Ophardt, D. (2013). Training inservice teachers' competencies in classroom management. A quasi-experimental study with teachers of secondary schools. *Teaching and Teacher Education*, 30, 1–12. <https://doi.org/10.1016/j.tate.2012.09.007>
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338. <https://doi.org/10.1037/a0014996>
- Praetorius, A.-K. (2013). Einschätzung von Unterrichtsqualität durch externe Beobachterinnen und Beobachter. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 31, 174–185. <https://doi.org/10.36950/bzl.31.2013.9650>
- Praetorius, A.-K., Klieme, E., Herbert, B. & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM Mathematics Education*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Protzko, J., Aronson, J. & Blair, C. (2013). How to Make a Young Child Smarter: Evidence From the Database of Raising Intelligence. *Perspectives on Psychological Science: a Journal of the Association for Psychological Science*, 8(1), 25–40. <https://doi.org/10.1177/1745691612462585>
- Quinn, D. M. (2015). Black-White Summer Learning Gaps. *Educational Evaluation and Policy Analysis*, 37(1), 50–69. <https://doi.org/10.3102/0162373714534522>
- Quinn, D. M., Cooc, N., McIntyre, J. & Gomez, C. J. (2016). Seasonal Dynamics of Academic Achievement Inequality by Socioeconomic Status and Race/Ethnicity. *Educational Researcher*, 45(8), 443–453. <https://doi.org/10.3102/0013189X16677965>
- Radisch, F., Klieme, E. & Bos, W. (2006). Gestaltungsmerkmale und Effekte ganztägiger Angebote im Grundschulbereich. *Zeitschrift für Erziehungswissenschaft*, 9(1), 30–50. <https://doi.org/10.1007/s11618-006-0003-6>
- Ramseier, E. & Brühwiler, C. (2003). Herkunft, Leistung und Bildungschancen im gegliederten Bildungssystem. Vertiefte PISA-Analyse unter Einbezug der kognitiven Grundfähigkeiten. *Schweizerische Zeitschrift für Bildungswissenschaften*, 25, 23–58. <https://doi.org/10.24452/sjer.25.1.4646>
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Thousand Oaks: Sage Publ.
- Raudenbush, S. W. & Eschmann, R. D. (2015). Does Schooling Increase or Reduce Social Inequality? *Annual Review of Sociology*, 41(1), 443–470. <https://doi.org/10.1146/annurev-soc-071913-043406>
- Reinders, H., Ditton, H., Gräsel, C. & Gniewosz, B. (Hrsg.). (2015). *Empirische Bildungsforschung. Gegenstandsbereiche*. Wiesbaden: Springer VS. <https://doi.org/10.1007/978-3-531-19994-8>
- Reynolds, D., Sammons, P., Fraine, B. de, van Damme, J., Townsend, T., Teddlie, C. et al. (2014). Educational effectiveness research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25(2), 197–230. <https://doi.org/10.1080/09243453.2014.885450>

- Reynolds, D. & Teddlie, C. (2000). The processes of school effectiveness. In C. Teddlie & D. Reynolds (Hrsg.), *The international handbook of school effectiveness research* (S. 134–159). London: Falmer Press.
- Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 191–208). Stuttgart: Kohlhammer.
- Riedmeier, M. (2019). Standardisierte Verfahren zur Erhebung zielkindbezogener Prozessqualität in der Frühpädagogik: Ein vergleichender Überblick. *Frühe Bildung*, 8(3), 144–152. <https://doi.org/10.1026/2191-9186/a000426>
- Rivkin, S. G., Hanushek, E. A. & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73(2), 417–458. <https://doi.org/10.1111/j.1468-0262.2005.00584.x>
- Robinson, G. E. (1990). Synthesis of Research on the Effects of Class Size. *Educational Leadership*, 47, 80–90.
- Robinson, V. M. J., Lloyd, C. A. & Rowe, K. J. (2008). The Impact of Leadership on Student Outcomes: An Analysis of the Differential Effects of Leadership Types. *Educational Administration Quarterly*, 44(5), 635–674. <https://doi.org/10.1177/0013161X08321509>
- Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W. & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240–257. <https://doi.org/10.1037/0022-0663.95.2.240>
- Roßbach, H.-G. (1995). Hausaufgaben in der Grundschule. Ergebnisse einer empirischen Untersuchung. *Die Deutsche Schule*, 87, 103–112.
- Roßbach, H.-G., Kluczniok, K. & Kuger, S. (2008). Auswirkungen eines Kindergartenbesuchs auf den kognitiv-leistungsbezogenen Entwicklungsstand von Kindern. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen. Zeitschrift für Erziehungswissenschaft (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 11, S. 139–158)*. Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH Wiesbaden. https://doi.org/10.1007/978-3-531-91452-7_10
- Roubinov, D. S. & Boyce, W. T. (2017). Parenting and SES: relative values or enduring principles? *Current Opinion in Psychology*, 15, 162–167. <https://doi.org/10.1016/j.copsyc.2017.03.001>
- Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, 35(1), 185–205. <https://doi.org/10.1017/S0305000907008343>
- Rui, N. (2009). Four decades of research on the effects of detracking reform: Where do we stand? A systematic review of the evidence. *Journal of Evidence-Based Medicine*, 2(3), 164–183. <https://doi.org/10.1111/j.1756-5391.2009.01032.x>
- Rutter, M., Maughan, B., Mortimore, P., Ouston, J. & Smith, A. *Fifteen Thousand Hours. Secondary Schools and Their Effects on Children*. Cambridge: Harvard University Press.
- Ryan, R. M. & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>

- Sammons, P. (1999). *School effectiveness. Coming of age in the twenty-first century*. Lisse: Swets & Zeitlinger.
- Sammons, P. (2010). Equity and Educational Effectiveness. In P. Peterson, E. Baker & B. McGraw (Hrsg.), *International encyclopedia of education* (S. 51–57). Oxford: Elsevier. <https://doi.org/10.1016/B978-0-08-044894-7.00419-X>
- Sammons, P., Anders, Y., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. et al. (2008). Children's Cognitive Attainment and Progress in English Primary Schools During Key Stage 2: Investigating the potential continuing influences of pre-school education. In H.-G. Roßbach & H.-P. Blossfeld (Hrsg.), *Frühpädagogische Förderung in Institutionen. Zeitschrift für Erziehungswissenschaft (Zeitschrift für Erziehungswissenschaft Sonderheft, Bd. 11, S. 179–198)*. Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH Wiesbaden. https://doi.org/10.1007/978-3-531-91452-7_12
- Sammons, P. & Bakkum, L. (2011). *Effective Schools, Equity and Teacher Effectiveness: A Review to the Literature*.
- Sammons, P., Davis, S. & Gray, J. (2016). Methodological and scientific properties of school effectiveness research. Exploring the underpinnings, evolution, and future directions of the field. In C. Chapman, D. Muijs, D. Reynolds, P. Sammons & C. Teddlie (Hrsg.), *The Routledge international handbook of educational effectiveness and improvement. Research, policy and practice* (S. 25–76). London: Routledge.
- Sammons, P., Hillman, J. & Mortimore, P. (1995). *Key characteristics of effective schools. A review of school effectiveness research*, London University Institute of Education. Zugriff am 09.01.2020. Zugriff am 11.04.2021. Verfügbar unter: <https://de.scribd.com/document/376201419/Sammons-P-Hillman-J-Mortimore-P-1995-Key-characteristics-of-effective-schools-a-review-of-school-effectiveness-research-London-Institute-of-Ed>
- Sammons, P., Nuttall, D. & Cuttance, P. (1993). Differential School Effectiveness: results from a reanalysis of the Inner London Education Authority's Junior School Project Data. *British Educational Research Journal*, 19(4), 381–405. <https://doi.org/10.1080/0141192930190407>
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B. & Elliot, K. (2002). *The Effective Provision of Pre-School Education (EPPE) Project: Measuring the Impact of Pre-School on Children's Cognitive Progress over the Pre-School Period*. Zugriff am 11.04.2021. Verfügbar unter: https://www.researchgate.net/publication/43982070_The_Effective_Provision_of_Pre-School_Education_EPPE_Project_Measuring_the_impact_of_pre-school_on_children%27s_cognitive_progress_over_the_pre-school_period
- Sammons, P., Sylva, K., Melhuish, E., Siraj-Blatchford, I., Taggart, B., Elliot, K. et al. (2014). *The Effective Provision of Pre-School Education [EPPE] Project. Technical Paper 9 Report on age 6 assessment*. Zugriff am 11.04.2021. Verfügbar unter: https://www.researchgate.net/publication/43982068_The_Effective_Provision_of_Pre-School_Education_EPPE_Project_Report_on_Age_6_Assessment_2004
- Schaffner, E., Schiefele, U. & Schmidt, M. (2013). Die Bedeutung des familiären Hintergrundes für die Lesemotivation und Lesehäufigkeit von Gymnasialschülern. *Zeit-*

- schrift für Entwicklungspsychologie und Pädagogische Psychologie, 45(3), 131–141. <https://doi.org/10.1026/0049-8637/a000085>
- Schauenberg, M. (2007). Übertrittsentscheidungen nach der Grundschule. Empirische Analysen zu familialen Lebensbedingungen und Rational-Choice. München: Utz.
- Scheerens, J. (2000). Improving school effectiveness (Fundamentals of educational planning, vol. 68). Paris: UNESCO International Institute for Educational Planning.
- Scheerens, J. (2014). School, teaching, and system effectiveness: some comments on three state-of-the-art reviews. *School Effectiveness and School Improvement*, 25(2), 282–290. <https://doi.org/10.1080/09243453.2014.885453>
- Scheerens, J. & Bosker, R. J. (1997). The foundations of educational effectiveness. Oxford: Pergamon.
- Scheerens, J., Hendriks, M., Luyten, H., Slegers, P. & Glas, C. (2013). Productive Time in Education. Zugriff am 11.07.2020. Verfügbar unter: https://ris.utwente.nl/ws/portalfiles/portal/5143337/Productive_time_in_education.pdf
- Scheerens, J., Luyten, H., Steen, R. & Luyten-de Thouars, Y. (2007). Review and Meta-Analyses of School and Teacher Effectiveness. Zugriff am 11.07.2020. Verfügbar unter: <https://www.utwente.nl/nl/bms/omd/Medewerkers/artikelen/scheerens/rev-meta-analysis2007.PDF>
- Scheerens, J., Vermeulen, C. J. A. J. & Pelgrum, W. J. (1989). Generalizability of instructional and school effectiveness indicators across nations. *International Journal of Educational Research*, 13(7), 789–799. [https://doi.org/10.1016/0883-0355\(89\)90029-3](https://doi.org/10.1016/0883-0355(89)90029-3)
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 120–148.
- Schmerse, D., Anders, Y., Flöter, M., Wieduwilt, N., Roßbach, H.-G. & Tietze, W. (2018). Differential effects of home and preschool learning environments on early language development. *British Educational Research Journal*, 44(2), 338–357. <https://doi.org/10.1002/berj.3332>
- Schmid, C., Trendtel, M., Bruneforth, M. & Hartig, J. (2020). Effectiveness of a governmental action to improve Austrian primary schools – results of multilevel analyses based on repeated cycles of educational standards assessments. *School Effectiveness and School Improvement*, 31(2), 149–171. <https://doi.org/10.1080/09243453.2019.1620294>
- Schmitt, M. (2012). Soziale Beziehungen und Schulerfolg. Die Bedeutung sozialer Beziehungen für Schulerfolg im Primar- und Sekundarbereich. Bamberg: Univ. of Bamberg Press.
- Schnabel, K. U., Alfeld, C., Eccles, J. S., Köller, O. & Baumert, J. (2002). Parental Influence on Students' Educational Choices in the United States and Germany: Different Ramifications – Same Effect? *Journal of Vocational Behavior*, 60(2), 178–198. <https://doi.org/10.1006/jvbe.2001.1863>
- Schneider, T. (2005). Nachhilfe als Strategie zur Verwirklichung von Bildungszielen. Eine empirische Untersuchung mit Daten des Sozio-oekonomischen Panels (SOEP). *Zeitschrift für Pädagogik*, 51(3), 363–379. <https://doi.org/10.25656/01:4760>

- Schneider, T. & Pfost, M. (2013). Social and Immigration-Specific Differences in the Development of Reading Comprehension: A Longitudinal Analysis of Primary School Students in Germany. In M. Pfost, C. Artelt & S. Weinert (Hrsg.), *The Development of Reading Literacy from Early Childhood to Adolescence. Empirical Findings from the Bamberg BiKS Longitudinal Studies* (S. 151–187). Bamberg: University of Bamberg Press.
- Schneider, W., Lingel, K., Artelt, C. & Neuenhaus, N. (2017). Metacognitive Knowledge in Secondary School Students: Assessment, Structure, and Developmental Change. In D. Leutner, J. Fleischer, J. Grünkorn & E. Klieme (Hrsg.), *Competence Assessment in Education* (S. 285–302). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-50030-0_17
- Schröder, C., Spieß, C. K. & Storck, J. (2015). Private Bildungsausgaben für Kinder: Einkommensschwache Familien sind relativ stärker belast. DIW-Wochenbericht.
- Schuller, T. & Desjardins, R. (2007). *Understanding the social outcomes of learning*. Paris: Organisation for Economic Co-Operation and Development.
- Schulz-Heidorf, K. (2016). Individuelle Förderung im Unterricht: Eine Möglichkeit, soziale Herkunft und Schulerfolg zu entkoppeln? Eine Re-Analyse aus IGLU-E 2011. Zugriff am 11.04.2021. Verfügbar unter: <https://ediss.sub.uni-hamburg.de/volltexte/2016/8224/pdf/Dissertation.pdf>
- Schüpbach, M. (2014). Extended education and social inequality in Switzerland: Compensatory effects? *Journal for Educational Research Online*, 6, 95–114.
- Schüpbach, M., Herzog, W. & Ignaczewska, J. (2013). Entwicklung der Mathematikleistung von Ganztagschulkindern: kompensatorische Wirkung der Ganztagschule? *Zeitschrift für Pädagogische Psychologie*, 27(3), 157–167. <https://doi.org/10.1024/1010-0652/a000101>
- Schwab, J. F. & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *Wiley Interdisciplinary Reviews. Cognitive Science*, 7(4), 264–275. <https://doi.org/10.1002/wcs.1393>
- Schwerdt, G. & Wuppermann, A. C. (2011). Is traditional teaching really all that bad? A within-student between-subject approach. *Economics of Education Review*, 30(2), 365–379. <https://doi.org/10.1016/j.econedurev.2010.11.005>
- Schwippert, K. (2001). *Optimalklassen. Mehrebenenanalytische Untersuchungen*. Münster, New York, München, Berlin: Waxmann.
- Scott-Little, C., Hamann, M. S. & Jurs, S. G. (2002). Evaluations of After-School Programs: A Meta-Evaluation of Methodologies and Narrative Synthesis of Findings. *American Journal of Evaluation*, 23(4), 387–419. <https://doi.org/10.1177/109821400202300403>
- Seidel, T. & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (S. 253–276). Weinheim: Julius Beltz.
- Seifert, W. (2005). Bildungsmobilität: Wie weit fällt der Apfel vom Stamm? In Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (Hrsg.), *Statistische Analysen und Studien Nordrhein Westfalen Band 24* (S. 3–11). Düsseldorf: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen.
- Seiz, J., Decristan, J., Kunter, M. & Baumert, J. (2016). Differenzielle Effekte von Klassenführung und Unterstützung für Schülerinnen und Schüler mit Migrationshin-

- tergrund. *Zeitschrift für Pädagogische Psychologie*, 30(4), 237–249. <https://doi.org/10.1024/1010-0652/a000186>
- September, S. J., Rich, E. G. & Roman, N. V. (2016). The role of parenting styles and socio-economic status in parents' knowledge of child development. *Early Child Development and Care*, 186(7), 1060–1078. <https://doi.org/10.1080/03004430.2015.1076399>
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Singh, A., Uijtdewilligen, L., Twisk, J. W. R., van Mechelen, W. & Chinapaw, M. J. M. (2012). Physical activity and performance at school: a systematic review of the literature including a methodological quality assessment. *Archives of Pediatrics & Adolescent Medicine*, 166(1), 49–55. <https://doi.org/10.1001/archpediatrics.2011.716>
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Slavin, R. E. (1994). Quality, appropriateness, incentive, and time: A model of instructional effectiveness. *International Journal of Educational Research*, 21(2), 141–157. [https://doi.org/10.1016/0883-0355\(94\)90029-9](https://doi.org/10.1016/0883-0355(94)90029-9)
- Slavin, R. E., Lake, C., Davis, S. & Madden, N. A. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, 6(1), 1–26. <https://doi.org/10.1016/j.edurev.2010.07.002>
- Slavin, R. E., Madden, N. A., Chambers, B. & Haxby, B. (2009). *2 Million children. Success for all*. Thousand Oaks, Calif.: Corwin Press.
- Smith, S. M., Glenberg, A. & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, 6, 342–353. <https://doi.org/10.3758/BF03197465>
- Snow, R. E. (1989). Aptitude-Treatment Interaction as a Framework for Research on Individual Differences in Learning. In P. L. Ackerman, R. J. Sternberg & R. Glaser (Hrsg.), *Learning and individual differences. Advances in theory and research* (S. 13–59). New York: W. H. Freeman.
- Solga, H. & Dombrowski, R. (2009). Soziale Ungleichheiten in schulischer und außerschulischer Bildung. Stand der Forschung und Forschungsbedarf. Hans Böckler Stiftung Arbeitspapier 171. Zugriff am 11.04.2021. Verfügbar unter: https://www.boeckler.de/pdf/p_arbp_171.pdf
- Spinath, B. (2015). Lernmotivation. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (S. 55–67). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-531-19994-8_4
- Splendid Research. (2017). Soziale Kontakte. Eine repräsentative Umfrage unter 1.039 Deutschen zur Anzahl und Qualität ihrer sozialen Kontakte. Hamburg: Splendid Research.
- Spreitzer, C., Hafner, S., Krainer, K. & Vohns, A. (2022). Effects of Generic and Subject-Didactic Teaching Characteristics on Student Performance in Mathematics in Secondary School: A Scoping Review. *European Journal of Educational Research*, 11(2), 711–737. <https://doi.org/10.12973/eu-jer.11.2.711>

- Stahl, N. (2009). *Der institutionelle Beitrag zur Entstehung primärer und sekundärer Herkunftseffekte. Eine Untersuchung an bayerischen Grundschulen*. Berlin: Mensch und Buch-Verl.
- Stancel-Piatak, A. S. (2017). *Effektivität des Schulsystems beim Abbau sozialer Ungleichheit. Latentes Mehrebenenmodell individueller und institutioneller Faktoren der sozialen Reproduktion*. Münster: Waxmann.
- Statistisches Bundesamt. (2018). *Alleinerziehende in Deutschland 2017*. Zugriff am 11.04.2021. Verfügbar unter: https://www.destatis.de/DE/Presse/Pressekonferenzen/2018/Alleinerziehende/pressebroschuere-alleinerziehende.pdf?__blob=publicationFile
- Steinmann, I. & Strietholt, R. (2018). Student achievement and educational inequality in half- and all-day schools: Evidence from Germany. *International Journal for Research on Extended Education*, 6, 175–197. <https://doi.org/10.3224/ijree.v6i2.08>
- Steinmann, I., Strietholt, R. & Caro, D. (2019). Participation in extracurricular activities and student achievement: evidence from German all-day schools. *School Effectiveness and School Improvement*, 30(2), 155–176. <https://doi.org/10.1080/09243453.2018.1540435>
- Steinmayr, R., Dinger, F. C. & Spinath, B. (2010). Parents' education and children's achievement: The role of personality. *European Journal of Personality*, 24(6). <https://doi.org/10.1002/per.755>
- Steinmayr, R., Dinger, F. C. & Spinath, B. (2012). Motivation as a Mediator of Social Disparities in Academic Achievement. *European Journal of Personality*, 26(3), 335–349. <https://doi.org/10.1002/per.842>
- Stockard, J., Wood, T. W., Coughlin, C. & Rasplia Khoury, C. (2018). The Effectiveness of Direct Instruction Curricula: A Meta-Analysis of a Half Century of Research. *Review of Educational Research*, 88(4), 479–507. <https://doi.org/10.3102/0034654317751919>
- Stocké, V. (2007). Explaining Educational Decision and Effects of Families' Social Class Position: An Empirical Test of the Breen Goldthorpe Model of Educational Attainment. *European Sociological Review*, 23(4), 505–519. <https://doi.org/10.1093/esr/jcm014>
- Strand, S. (2010). Do some schools narrow the gap? Differential school effectiveness by ethnicity, gender, poverty, and prior achievement. *School Effectiveness and School Improvement*, 21(3), 289–314. <https://doi.org/10.1080/09243451003732651>
- Strand, S. (2014). School effects and ethnic, gender and socio-economic gaps in educational achievement at age 11. *Oxford Review of Education*, 40(2), 223–245. <https://doi.org/10.1080/03054985.2014.891980>
- Strand, S. (2016). Do some schools narrow the gap? Differential school effectiveness revisited. *Review of Education*, 4(2), 107–144. <https://doi.org/10.1002/rev3.3054>
- Strietholt, R., Manitius, V., Berkemeyer, N. & Bos, W. (2015). Bildung und Bildungsgleichheit an Halb- und Ganztagschulen. *Zeitschrift für Erziehungswissenschaft*, 18(4), 737–761. <https://doi.org/10.1007/s11618-015-0634-6>
- Stringfield, S. (1994). Outlier Studies of School Effectiveness. In D. Reynolds, B. Creemers, P. S. Nesselrodt, E. C. Shaffer, S. Stringfield & C. Teddlie (Hrsg.), *Advances in School Effectiveness Research and Practice* (S. 73–83). Burlington: Elsevier Science. <https://doi.org/10.1016/B978-0-08-042392-0.50009-7>

- Stubbe, T. C., Schwippert, K. & Wendt, H. (2016). Soziale Disparitäten der Schülerleistung in Mathematik und Naturwissenschaften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 299–316). Münster: Waxmann.
- Stutz, F., Schaffner, E. & Schiefele, U. (2016). Relations among reading motivation, reading amount, and reading comprehension in the early elementary grades. *Learning and Individual Differences*, 45, 101–113. <https://doi.org/10.1016/j.lindif.2015.11.022>
- Susperreguy, M. I., Di Lonardo Burr, S., Xu, C., Douglas, H. & LeFevre, J.-A. (2020). Children's Home Numeracy Environment Predicts Growth of their Early Mathematical Skills in Kindergarten. *Child Development*. <https://doi.org/10.1111/cdev.13353>
- Süßlin, W. (2015). Unterstützungsbedarf von Eltern bei der Erziehung und Förderung ihrer Kinder. In Vodafone Stiftung Deutschland (Hrsg.), Was Eltern wollen. Informations- und Unterstützungswünsche zu Bildung und Erziehung (S. 8–17). Düsseldorf: Vodafone Stiftung Deutschland.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- Taras, H. (2005). Physical activity and student performance at school. *The Journal of School Health*, 75(6), 214–218. <https://doi.org/10.1111/j.1746-1561.2005.00026.x>
- Tazouti, Y. & Jarlégan, A. (2019). The mediating effects of parental self-efficacy and parental involvement on the link between family socioeconomic status and children's academic achievement. *Journal of Family Studies*, 25(3), 250–266. <https://doi.org/10.1080/13229400.2016.1241185>
- Teodorovic, J. (2011). Classroom and school factors related to student achievement: what works for students? *School Effectiveness and School Improvement*, 22(2), 215–236. <https://doi.org/10.1080/09243453.2011.575650>
- Thiel, F. (2016). Interaktion im Unterricht. Ordnungsmechanismen und Störungsdynamiken. Opladen, Toronto: Verlag Barbara Budrich. <https://doi.org/10.36198/9783838545714>
- Thomas, S., Sammons, P., Mortimore, P. & Smees, R. (1997). Differential Secondary School Effectiveness: comparing the performance of different pupil groups. *British Educational Research Journal*, 23(4), 451–469. <https://doi.org/10.1080/0141192970230405>
- Tietze, W., Meischner, T., Gänsfuß, R., Grenner, K., Schuster, K.-M., Völkel, P. et al. (1998). Wie gut sind unsere Kindergärten? Eine Untersuchung zur pädagogischen Qualität in deutschen Kindergärten. Neuwied: Luchterhand.
- Tippelt, R. & Reich-Claassen, J. (2010). Stichwort: Evidenzbasierung. *DIE Zeitschrift für Erwachsenenbildung*, (4), 22–23.
- Titsworth, S., Mazer, J. P., Goodboy, A. K., Bolkan, S. & Myers, S. A. (2015). Two Meta-analyses Exploring the Relationship between Teacher Clarity and Student Learning. *Communication Education*, 64(4), 385–418. <https://doi.org/10.1080/03634523.2015.1041998>
- Torquati, J. C., Raikes, H. H., Huddleston-Casas, C. A., Bovaird, J. A. & Harris, B. A. (2011). Family income, parent education, and perceived constraints as predictors

- of observed program quality and parent rated program quality. *Early Childhood Research Quarterly*, 26(4), 453–464. <https://doi.org/10.1016/j.ecresq.2011.03.004>
- Trautwein, U., Köller, O. & Baumert, J. (2001). Lieber oft als viel: Hausaufgaben und die Entwicklung von Leistung und Interesse im Mathematik-Unterricht der 7. Jahrgangsstufe. *Zeitschrift für Pädagogik*, 47, 703–724.
- Trautwein, U. & Lüdtke, O. (2008). Die Förderung der Selbstregulation durch Hausaufgaben – Herausforderungen und Chancen. In C. Rohlf, M. Harring & C. Palentien (Hrsg.), *Kompetenz-Bildung. Soziale, emotionale und kommunikative Kompetenzen von Kindern und Jugendlichen* (S. 275–288). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH Wiesbaden. https://doi.org/10.1007/978-3-658-03441-2_12
- Treiber, B. & Schneider, W. (1980). Qualifizierende und varianzreduzierende Wirkungen von Schulklassenunterricht. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 7, 261–283.
- Treiber, B. & Weinert, F. E. (1985). *Gute Schulleistungen für alle? Psychologische Studien zu einer pädagogischen Hoffnung*. Münster: Aschendorff.
- Treinius, G. & Einsiedler, W. (1996). Zur Vereinbarkeit von Steigerung des Lernleistungsniveaus und Verringerung von Leistungsunterschieden in Grundschulklassen. *Unterrichtswissenschaft*, 24, 290–311.
- Tuovinen, J. E. & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91(2), 334–341. <https://doi.org/10.1037/0022-0663.91.2.334>
- UNESCO Institute for Statistics. (2012). *International standard classification of education. ISCED 2011*. Montreal, Quebec: UNESCO Institute for Statistics.
- Urban, D. & Mayerl, J. (2014). *Strukturgleichungsmodellierung*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-01919-8>
- Urke, H. B., Bull, T. & Mittelmark, M. B. (2011). Socioeconomic status and chronic child malnutrition: Wealth and maternal education matter more in the Peruvian Andes than nationally. *Nutrition Research*, 31(10), 741–747. <https://doi.org/10.1016/j.nutres.2011.09.007>
- Van de Werfhorst, H. G. & Mijs, J. J. B. (2010). Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective. *Annual Review of Sociology*, 36(1), 407–428. <https://doi.org/10.1146/annurev.soc.012809.102538>
- Van Ewijk, R. & Sleegers, P. (2010a). Peer ethnicity and achievement: a meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, 21(3), 237–265. <https://doi.org/10.1080/09243451003612671>
- Van Ewijk, R. & Sleegers, P. (2010b). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134–150. <https://doi.org/10.1016/j.edurev.2010.02.001>
- Vanlaar, G., Denies, K., Vandecandelaere, M., van Damme, J., Verhaeghe, J. P., Pinxten, M. et al. (2014). How to improve reading comprehension in high-risk students: effects of class practices in Grade 5. *School Effectiveness and School Improvement*, 25(3), 408–432. <https://doi.org/10.1080/09243453.2013.811088>
- Vanlaar, G., Kyriakides, L., Panayiotou, A., Vandecandelaere, M., McMahon, L., Fraine, B. de et al. (2016). Do the teacher and school factors of the dynamic model affect

- high- and low-achieving student groups to the same extent? a cross-country study. *Research Papers in Education*, 31(2), 183–211. <https://doi.org/10.1080/02671522.2015.1027724>
- Vaughn, K. (2000). Music and Mathematics: Modest Support for the Off-Claimed Relationship. *Journal of Aesthetic Education*, 34(3), 149. <https://doi.org/10.2307/3333641>
- Veldman, D. J. & Sanford, J. P. (1984). The Influence of Class Ability Level on Student Achievement and Classroom Behavior. *American Educational Research Journal*, 21(3), 629–644. <https://doi.org/10.3102/00028312021003629>
- Vygotsky, L. S. (1978). *Mind in Society: the Development of Higher Psychological Processes*. Cambridge: Harvard University Press.
- Walberg, H. J. (1984). Improving the Productivity of America's Schools. *Educational Leadership*, 41, 19–27.
- Walter, O., Senkbeil, M., Rost, J., Carstensen, C. H. & Prenzel, M. (2006). Die Entwicklung der naturwissenschaftlichen Kompetenz von der neunten zur zehnten Klassenstufe: Deskriptive Befunde. In M. Prenzel (Hrsg.), *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 87–118). Münster: Waxmann.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450. <https://doi.org/10.1007/BF02294627>
- Weinert, F.E. (2000). Lehren und Lernen für die Zukunft – Ansprüche an das Lernen in der Schule. Zugriff am 05.07.2019. Verfügbar unter: <http://www2.ibw.uni-heidelberg.de/~gerstner/WeinertLehren&Lernen.pdf>
- Weinert, F.E. & Helmke, A. (1987). Schulleistungen – Leistungen der Schule oder der Kinder? In U. Steffens & T. Bargel (Hrsg.), *Untersuchungen zur Qualität des Unterrichts. Beiträge aus dem Arbeitskreis „Qualität von Schule“*, Heft 3 (S. 17–31). Wiesbaden: Hessisches Institut für Bildungsplanung und Schulentwicklung.
- Weinert, F.E. & Helmke, A. (Hrsg.). (1997). *Entwicklung im Grundschulalter*. Weinheim: Beltz.
- Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T. & Carstensen, C. H. (2011). Development of competencies across the life span. In H.-P. Blossfeld, H.-G. Roßbach & J. von Maurice (Hrsg.), *Education as a lifelong process. The German National Educational Panel Study (NEPS)* (*Zeitschrift für Erziehungswissenschaft Sonderheft*, Bd. 14, S. 67–86). Wiesbaden: VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/s11618-011-0182-7>
- Weinert, S., Attig, M. & Roßbach, H.-G. (2017). The emergence of social disparities: Evidence on early mother-child interaction and infant development from the German National Educational Panel Study (NEPS). In H.-P. Blossfeld, N. Kulic, J. Skopek & M. Triventi (Hrsg.), *Childcare, Early Education and Social Inequality* (S. 89–108). Edward Elgar Publishing. <https://doi.org/10.4337/9781786432094.00014>
- Weinert, S. & Ebert, S. (2013). Spracherwerb im Vorschulalter. *Zeitschrift für Erziehungswissenschaft*, 16(2), 303–332. <https://doi.org/10.1007/s11618-013-0354-8>
- Wenglinsky, H. (1998). Finance Equalization and Within-School Equity: The Relationship between Education Spending and the Social Distribution of Achievement. *Educational Evaluation and Policy Analysis*, 20(4), 269–283. <https://doi.org/10.2307/1164325>

- Wentzel, K. R., Jablansky, S. & Scalise, N. R. (2018). Do Friendships Afford Academic Benefits? A Meta-analytic Study. *Educational Psychology Review*, 30(4), 1241–1267. <https://doi.org/10.1007/s10648-018-9447-5>
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461–481. <https://doi.org/10.1037/0033-2909.91.3.461>
- Wiechmann, J. & Wildhirt, S. (2015). Unterrichtsmethoden – vom Nutzen der Vielfalt. In J. Wiechmann & S. Wildhirt (Hrsg.), *Zwölf Unterrichtsmethoden. Vielfalt für die Praxis*. Weinheim: Beltz.
- Wilder, S. (2014). Effects of parental involvement on academic achievement: a meta-synthesis. *Educational Review*, 66(3), 377–397. <https://doi.org/10.1080/00131911.2013.780009>
- Williams, P. A., Haertel, E. H., Haertel, G. D. & Walberg, H. J. (1982). The Impact of Leisure-Time Television on School Learning: A Research Synthesis. *American Educational Research Journal*, 19(1), 19–50. <https://doi.org/10.3102/00028312019001019>
- Woessmann, L., Lergetporer, P., Grewenig, E., Kersten, S., Kugler, F. & Werner, K. (2019). Was die Deutschen über Bildungsungleichheit denken. Ergebnisse des ifo Bildungsbarometers 2019. *ifo Schnelldienst*, 72(17), 27–41.
- Yerdelen-Damar, S. & Peşman, H. (2013). Relations of Gender and Socioeconomic Status to Physics Through Metacognition and Self-Efficacy. *The Journal of Educational Research*, 106(4), 280–289. <https://doi.org/10.1080/00220671.2012.692729>
- Young, D. J. & Fraser, B. J. (1993). Socioeconomic and Gender Effects on Science Achievement: An Australian Perspective. *School Effectiveness and School Improvement*, 4(4), 265–289. <https://doi.org/10.1080/0924345930040403>
- Zief, S. G., Lauver, S. & Maynard, R. A. (2006). Impacts of after-school programs on student outcomes: A systematic review for the Campbell collaboration. Zugriff am 11.04.2021. Verfügbar unter: https://www.researchgate.net/publication/228630250_Impacts_of_after-school_programs_on_student_outcomes_A_systematic_review_for_the_Campbell_collaboration
- Ziegler, E. & Stern, E. (2014). Delayed benefits of learning elementary algebraic transformations through contrasted comparisons. *Learning and Instruction*, 33, 131–146. <https://doi.org/10.1016/j.learninstruc.2014.04.006>
- Zielonka, M., Beier, L. & Blossfeld, H.-P. (2014). Schulverläufe während der Sekundarstufe. In M. Mudiappa & C. Artelt (Hrsg.), *BiKS – Ergebnisse aus den Längsschnittstudien. Praxisrelevante Befunde aus dem Primar- und Sekundarschulbereich* (Bd. 15, S. 119–130). Bamberg: University of Bamberg Press.
- Zimmer, R. W. & Toma, E. F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management*, 19(1), 75–92. [https://doi.org/10.1002/\(SICI\)1520-6688\(200024\)19:1<75::AID-PAM5>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1520-6688(200024)19:1<75::AID-PAM5>3.0.CO;2-W)
- Züchner, I. & Fischer, N. (2014). Kompensatorische Wirkungen von Ganztagschulen – Ist die Ganztagschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? *Zeitschrift für Erziehungswissenschaft*, 17, 349–367. <https://doi.org/10.1007/s11618-013-0477-y>

Anhang A: Erhebungsinstrumente

In diesem Anhang wird erläutert, wie die in der vorliegenden Arbeit untersuchten Variablen erhoben wurden. Hierfür werden relevante Ausschnitte aus den im NEPS verwendeten Erhebungsinstrumenten dargestellt. Bei allen Variablen werden Datenquelle, Item(s) und Antwortkategorien samt resultierender Messwerte berichtet. Die originalen Erhebungsinstrumente können auf der Website des NEPS heruntergeladen werden.⁴⁵ Die verwendeten Kompetenztest werden nicht dargestellt, da diese vom NEPS nicht veröffentlicht werden. Es können aber einige Beispiel-Items auf der Website des NEPS eingesehen werden.

Erhebungsinstrumente für Variablen auf der Schülerebene

Tab. 23: Erhebungsinstrumente für Variablen auf der Schülerebene

Variable	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Schüler/-in</i>		
Mathematiknote	S	<p>Welche Note hattest du im letzten Jahreszeugnis in Mathematik?</p> <p><input type="checkbox"/> sehr gut (1) <input type="checkbox"/> gut (2) <input type="checkbox"/> befriedigend (3) <input type="checkbox"/> ausreichend (4) <input type="checkbox"/> mangelhaft (5) <input type="checkbox"/> ungenügend (6)</p> <p>(Messwerte wurden invertiert)</p>
Interesse Mathematik	S	<p>Wie sehr treffen die folgenden Aussagen auf dich zu?</p> <p>An einem mathematischen Problem zu knobeln, macht mir einfach Spaß. (Item 1)</p> <p>Wenn ich an einem mathematischen Problem sitze, kann es passieren, dass ich gar nicht merke, wie die Zeit verfliegt. (Item 2)</p> <p>Wenn ich in Mathematik etwas Neues dazulernen kann, bin ich bereit, auch Freizeit dafür zu verwenden. (Item 3)</p> <p>Mathematik gehört für mich persönlich zu den wichtigsten Dingen. (Item 4)</p> <p><input type="checkbox"/> trifft gar nicht zu (1) <input type="checkbox"/> trifft eher nicht zu (2) <input type="checkbox"/> trifft eher zu (3) <input type="checkbox"/> trifft völlig zu (4)</p>
Selbstkonzept Mathematik	S	<p>Wie schätzt du dich in der Schule ein?</p> <p>Im Fach Mathematik bekomme ich gute Noten. (Item 1)</p> <p>Mathematik ist eines meiner besten Fächer. (Item 2)</p> <p>Ich war schon immer gut in Mathematik. (Item 3)</p> <p><input type="checkbox"/> trifft gar nicht zu (1) <input type="checkbox"/> trifft eher nicht zu (2) <input type="checkbox"/> trifft eher zu (3) <input type="checkbox"/> trifft völlig zu (4)</p>

Anmerkungen: S = Fragebogen Schüler/-in, E = Telefon-Interview Eltern

45 www.neps-data.de/Datenzentrum/Daten-und-Dokumentation/Startkohorte-Klasse-5/Dokumentation

Tab. 23: Erhebungsinstrumente für Variablen auf der Schülerebene (Fortsetzung)

Variable	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Schüler/-in</i>		
Gewissenhaftigkeit	S	<p>Inwieweit treffen die folgenden Aussagen auf dich zu?</p> <p>Ich bin bequem, neige zur Faulheit. (Item 1) Ich erledige Aufgaben gründlich. (Item 2)</p> <p> <input type="checkbox"/> trifft gar nicht zu (1) <input type="checkbox"/> trifft eher nicht zu (2) <input type="checkbox"/> teils/teils (3) <input type="checkbox"/> trifft eher zu (4) <input type="checkbox"/> trifft völlig zu (5) </p> <p>(Bei Item 1 wurden die Messwerte invertiert)</p>
Geschlecht	E	Ist <Name des Zielkinds> ein Junge oder ein Mädchen? (1 = Mädchen, 0 = Junge)
<i>Familie</i>		
Lernbegleitung Häufigkeit	S	<p>Wie oft helfen dir die folgenden Personen bei den Hausaufgaben oder dem Lernen für die Schule? Deine Eltern</p> <p> <input type="checkbox"/> nie (1) <input type="checkbox"/> selten (2) <input type="checkbox"/> manchmal (3) <input type="checkbox"/> oft (4) <input type="checkbox"/> immer (5) </p>
Lernbegleitung Unterstützung	S	<p>Wie ist das bei euch zu Hause?</p> <p>Wenn ich allein nicht mit den Hausaufgaben klar komme, nehmen sich meine Eltern immer Zeit für mich. (Item 1) Wenn ich eine schlechte Note bekommen habe, fragen mich meine Eltern, wie sie mir helfen können. (Item 2) Meine Eltern reden mir aufmunternd zu, wenn ich eine Hausaufgabe verpatzt habe. (Item 3) Wenn ich im Unterricht etwas nicht verstehe, kann ich mit meinen Eltern darüber reden. (Item 4)</p> <p> <input type="checkbox"/> stimme gar nicht zu (1) <input type="checkbox"/> stimme eher nicht zu (2) <input type="checkbox"/> stimme eher zu (3) <input type="checkbox"/> stimme völlig zu (4) </p>
Lernbegleitung Erwartungen	S	<p>Wie ist das bei euch zu Hause?</p> <p>Meine Eltern erwarten, dass ich mich bei den Hausaufgaben wirklich mit den Aufgaben auseinandersetze. (Item 1) Meine Eltern wollen, dass ich Sachen nicht nur auswendig lerne, sondern sie auch wirklich verstehe. (Item 2) Wenn ich für eine Arbeit lerne, weiß ich ganz genau, wie viel Anstrengung meine Eltern von mir erwarten. (Item 3)</p> <p> <input type="checkbox"/> stimme gar nicht zu (1) <input type="checkbox"/> stimme eher nicht zu (2) <input type="checkbox"/> stimme eher zu (3) <input type="checkbox"/> stimme völlig zu (4) </p>

Anmerkungen: S = Fragebogen Schüler/-in, E = Telefon-Interview Eltern

Tab. 23: Erhebungsinstrumente für Variablen auf der Schülerebene (Fortsetzung)

Variablen	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Familie</i>		
Lernbegleitung Kontrolle	S	Wie ist das bei euch zu Hause? Meine Eltern kontrollieren oft, ob ich die Hausaufgaben ordentlich erledigt habe. (Item 1) Meine Eltern achten immer darauf, dass ich meine Hausaufgaben mache. (Item 2) <input type="checkbox"/> stimme gar nicht zu (1) <input type="checkbox"/> stimme eher nicht zu (2) <input type="checkbox"/> stimme eher zu (3) <input type="checkbox"/> stimme völlig zu (4)
Bildungsstand	E	Welchen höchsten allgemein bildenden Schulabschluss haben Sie? Welchen höchsten beruflichen Ausbildungsabschluss haben Sie? Welchen höchsten allgemein bildenden Schulabschluss hat Ihr/e Partner/-in? Welchen höchsten beruflichen Ausbildungsabschluss hat Ihr/e Partner/-in?
Berufsposition	E	Sagen Sie mir bitte, welche berufliche Tätigkeit Sie ausüben. Sagen Sie mir bitte, welche berufliche Tätigkeit er/sie [der/die Partner/-in] ausübt.
Migrationshintergrund	E	In welchem Land sind Sie geboren? In welchem Land ist Ihr/e Partner/-in geboren?
<i>Schulische Lernaktivitäten</i>		
		Nimmst du an den folgenden Angeboten an deiner Schule außerhalb des Unterrichts teil?
Nutzung Förderunterricht	S	Fördergruppen/Förderunterricht
Nutzung Hausaufgabenbetreuung	S	Hausaufgabenhilfe/Hausaufgabenbetreuung
Nutzung fachbezogene Angebote	S	Fachbezogene Lernangebote (z. B. Zusatz- oder Erweiterungskurse in Mathematik oder Deutsch) <input type="checkbox"/> ja (1) <input type="checkbox"/> nein (0)
<i>Schulbezogene Einrichtungen</i>		
Nachhilfe Mathematik	E	Erhält <Name des Zielkinds> zurzeit Nachhilfeunterricht? Und in welchem Fach oder welchen Fächern erhält <Name des Zielkinds> Nachhilfe? (1 = Nachhilfe in Mathematik, 0 = Keine Nachhilfe oder keine Nachhilfe in Mathematik)

Anmerkungen: S = Fragebogen Schüler/-in, E = Telefon-Interview Eltern

Tab. 23: Erhebungsinstrumente für Variablen auf der Schülerebene (Fortsetzung)

Variable	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Freizeitaktivitäten</i>		
Häufigkeit Lesen	S	Wie viel Zeit verbringst du normalerweise damit, außerhalb der Schule zu lesen? An einem normalen Schultag lese ich ... <input type="checkbox"/> gar nicht außerhalb der Schule (1) <input type="checkbox"/> bis zu einer halben Stunde (2) <input type="checkbox"/> zwischen einer halben und einer Stunde (3) <input type="checkbox"/> 1 bis 2 Stunden (4) <input type="checkbox"/> mehr als 2 Stunden (5)
Häufigkeit Musizieren	S	An wie vielen Tagen im vergangenen Monat hast du selbst musiziert, z. B. ein Musikinstrument gespielt oder im Chor gesungen? an ungefähr __ Tagen
Häufigkeit Sport	S	Wie häufig machst du Sport? Zähle dabei den Sportunterricht in der Schule nicht mit. <input type="checkbox"/> nie (1) <input type="checkbox"/> einmal im Monat oder seltener (2) <input type="checkbox"/> mehrmals pro Monat oder einmal pro Woche (3) <input type="checkbox"/> mehrmals pro Woche (4) <input type="checkbox"/> (fast) täglich (5)
<i>Gleichaltrige</i>		
Freund/-innen Hauptschulaspiration	S	Wie viele Personen aus deinem Freundeskreis haben vor, den Hauptschulabschluss zu machen?
Freund/-innen Realschulaspiration	S	... haben vor, die Mittlere Reife zu machen?
Freund/-innen Gymnasialaspiration	S	... haben vor, Abitur zu machen? <input type="checkbox"/> keine (1) <input type="checkbox"/> fast keine (2) <input type="checkbox"/> weniger als die Hälfte (3) <input type="checkbox"/> ungefähr die Hälfte (4) <input type="checkbox"/> mehr als die Hälfte (5) <input type="checkbox"/> fast alle (6) <input type="checkbox"/> alle (7)

Anmerkungen: S = Fragebogen Schüler/-in, E = Telefon-Interview Eltern

Erhebungsinstrumente für Variablen auf Klassen- und Schulebene

Tab. 24: Erhebungsinstrumente für Variablen auf Klassen- und Schulebene

Variable	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Unterricht</i>		
Unterrichtsumfang	ML	Wie viele Unterrichtsstunden pro Woche sind in diesem Halbjahr regulär für den Mathematikunterricht vorgesehen? __ Zahl der Unterrichtsstunden
Klassenunterricht	ML	Wie häufig setzen Sie im Mathematikunterricht dieser Klasse folgende Sozialformen des Lernens ein? Ich erkläre der ganzen Klasse etwas.
Einzelarbeit	ML	Die Schülerinnen und Schüler arbeiten alleine an Arbeitsblättern.
Partner-/Gruppenarbeit	ML	Partnerarbeit (Item 1) Arbeit mit kleinen Schülergruppen (Item 2) Diskussionsrunden (Item 3)
		<input type="checkbox"/> Nie (1) <input type="checkbox"/> Ein- bis zweimal im Schuljahr (2) <input type="checkbox"/> Alle paar Monate (3) <input type="checkbox"/> Alle zwei bis vier Wochen (4) <input type="checkbox"/> Einmal pro Woche (5) <input type="checkbox"/> (fast) in jeder Stunde (6)
Leistungsheterogene Gruppen	ML	Inwieweit treffen die folgenden Aussagen auf Ihren Mathematikunterricht in dieser Klasse zu? Ich bilde Gruppen von Schülerinnen und Schülern mit unterschiedlichen Fähigkeiten.
Leistungshomogene Gruppen	ML	Ich bilde Gruppen von Schülerinnen und Schülern mit ähnlichen Fähigkeiten.
		<input type="checkbox"/> trifft gar nicht zu (1) <input type="checkbox"/> trifft eher nicht zu (2) <input type="checkbox"/> teils/teils (3) <input type="checkbox"/> trifft eher zu (4) <input type="checkbox"/> trifft völlig zu (5)
Klassenführung	ML	Wie viel Prozent der Zeit verbringen die Schülerinnen und Schüler in einer typischen Schulwoche im Mathematikunterricht mit folgenden Aktivitäten? __ Teilnahme an Tätigkeiten, die das Klassenmanagement betreffen und die nichts mit dem __ Unterrichtsinhalt/-ziel zu tun haben (z. B. Unterbrechungen und Ordnung halten)
Übung und Wiederholung	ML	__ lehrergestütztes wiederholendes und übendes Lernen
Aufgaben mit Hilfe	ML	__ Aufgabenstellungen mit Ihrer Hilfe bearbeiten
Aufgaben ohne Hilfe	ML	__ Aufgabenstellungen alleine, ohne Ihre Hilfe bearbeiten
Besprechung Hausaufgaben	ML	__ Hausaufgaben besprechen
		(Das Item zur Klassenführung wurde invertiert: $x_{inv} = 1 - x$)

Anmerkungen: ML = Fragebogen Mathematiklehrkraft, KL = Fragebogen Klassenlehrkraft, SL = Fragebogen Schulleiter/-in.

Tab. 24: Erhebungsinstrumente für Variablen auf Klassen- und Schulebene (Fortsetzung)

Variable	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Unterricht</i>		
Kognitive Aktivierung	ML	<p>Inwieweit treffen die folgenden Aussagen zu den Aufgaben zu, die Sie Ihren Schülerinnen und Schülern im Mathematikunterricht dieser Klasse geben?</p> <p>Ich stelle Aufgaben, bei denen es nicht nur um Standardlösungswege geht, sondern darum, den richtigen Ansatz auszuwählen. (Item 1) Ich stelle Aufgaben, für deren Lösung sich die Schülerinnen und Schüler Zeit zum Nachdenken nehmen müssen. (Item 2) Ich stelle Aufgaben, bei denen die Schülerinnen und Schüler verschiedene Lösungswege aufzeigen müssen. (Item 3) Ich stelle Aufgaben, die Erläuterungen und vertiefte Ausführungen statt einfacher Lösungen erfordern. (Item 4)</p> <p> <input type="checkbox"/> trifft gar nicht zu (1) <input type="checkbox"/> trifft eher nicht zu (2) <input type="checkbox"/> teils/teils (3) <input type="checkbox"/> trifft eher zu (4) <input type="checkbox"/> trifft völlig zu (5) </p>
Individualisierung	ML	<p>Inwieweit treffen die folgenden Aussagen auf Ihren Mathematikunterricht in dieser Klasse zu?</p> <p>Ich gebe Schülerinnen und Schülern je nach Leistungsstärke unterschiedlich schwierige Hausaufgaben. (Item 1) Wenn Schülerinnen und Schüler etwas nicht verstanden haben, vergebe ich gezielte Zusatzaufgaben. (Item 2) Leistungsstarken Schülerinnen und Schülern gebe ich Extraaufgaben, durch die sie wirklich gefordert werden. (Item 3)</p> <p> <input type="checkbox"/> trifft gar nicht zu (1) <input type="checkbox"/> trifft eher nicht zu (2) <input type="checkbox"/> teils/teils (3) <input type="checkbox"/> trifft eher zu (4) <input type="checkbox"/> trifft völlig zu (5) </p>

Anmerkungen: ML = Fragebogen Mathematiklehrkraft, KL = Fragebogen Klassenlehrkraft, SL = Fragebogen Schulleiter/-in.

Tab. 24: Erhebungsinstrumente für Variablen auf Klassen- und Schulebene (Fortsetzung)

Variable	Daten- quelle	Item(s), Antwortkategorien und Messwerte (in Klammern)
<i>Ganztagsangebote</i>		
Förderunterricht Mathematik	ML	Wie viele Unterrichtsstunden pro Woche sind in diesem Halbjahr für zusätzlichen Mathematikförderunterricht vorgesehen? __ Zahl der Unterrichtsstunden
Förderunterricht schwach	SL	Werden die folgenden außerunterrichtlichen Ganztagelemente und -angebote an Ihrer Schule praktiziert und falls ja, wie oft? Förderunterricht für Schülerinnen und Schüler mit niedrigen Fachleistungen
Förderunterricht stark	SL	Förderunterricht für Schülerinnen und Schüler mit hohen Fachleistungen
Hausaufgabenbetreuung	SL	Hausaufgabenhilfe, Hausaufgabenbetreuung, Lernzeit
Mathematikangebote	SL	Fachbezogene Lernangebote (Projekte, AGs) in den folgenden Bereichen: Mathematik <input type="checkbox"/> nein (1) <input type="checkbox"/> halbjährlich und seltener (2) <input type="checkbox"/> vierteljährlich (3) <input type="checkbox"/> monatlich (4) <input type="checkbox"/> einmal pro Woche (5) <input type="checkbox"/> 2-3-mal pro Woche (6) <input type="checkbox"/> 4-5-mal pro Woche (7)
<i>Merkmale der Lerngruppe</i>		
Klassengröße	KL	Aus wie vielen Schülerinnen und Schülern setzt sich Ihre Klasse zusammen? __ Mädchen __ Jungen (Messwert = Anzahl Mädchen + Anzahl Jungen)

Anmerkungen: ML = Fragebogen Mathematiklehrkraft, KL = Fragebogen Klassenlehrkraft, SL = Fragebogen Schulleiter/-in.

Die drei Merkmale der Lerngruppe Mittelwert Mathematik_5, Mittelwert Bildungsstand und Anteil Migrationshintergrund wurden jeweils durch Aggregation der entsprechenden Schülervariablen berechnet (getrennt für einzelne Klassen).

Anhang B: Ergebnisse der konfirmatorischen Faktorenanalysen

In den nachfolgenden Abbildungen und Tabellen werden die Ergebnisse der konfirmatorischen Faktorenanalysen auf der Schülerebene und auf der Klassenebene berichtet.

Faktorenanalyse auf der Schülerebene

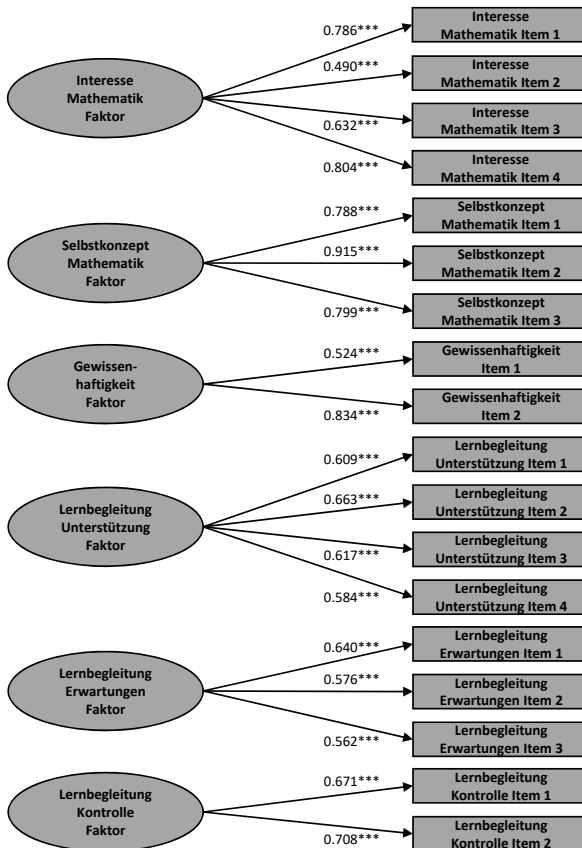


Abb. 19: Faktorladungen auf der Schülerebene (standardisierte Faktorladungen, *** $p < 0.001$)

Tab. 25: Faktorkorrelationen auf der Schülerebene (standardisierte Korrelationskoeffizienten)

	Interesse Mathema- tik	Selbstkon- zept Ma- thematik	Gewissen- haftigkeit	Lernbe- gleitung Unterstüt- zung	Lernbe- gleitung Erwartun- gen	Lernbeglei- tung Kont- rolle
Interesse Mathema- tik	1					
Selbstkon- zept Ma- thematik	0.521***	1				
Gewissen- haftigkeit	0.248***	0.117***	1			
Lernbe- gleitung Unterstüt- zung	-0.009	0.093**	0.252***	1		
Lernbe- gleitung Erwartun- gen	0.024	0.063⁺	0.184***	0.628***	1	
Lernbeglei- tung Kont- rolle	-0.037	0.057	0.121**	0.650***	0.678***	1

Anmerkung: ⁺p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001.

Faktorenanalyse auf der Klassenebene

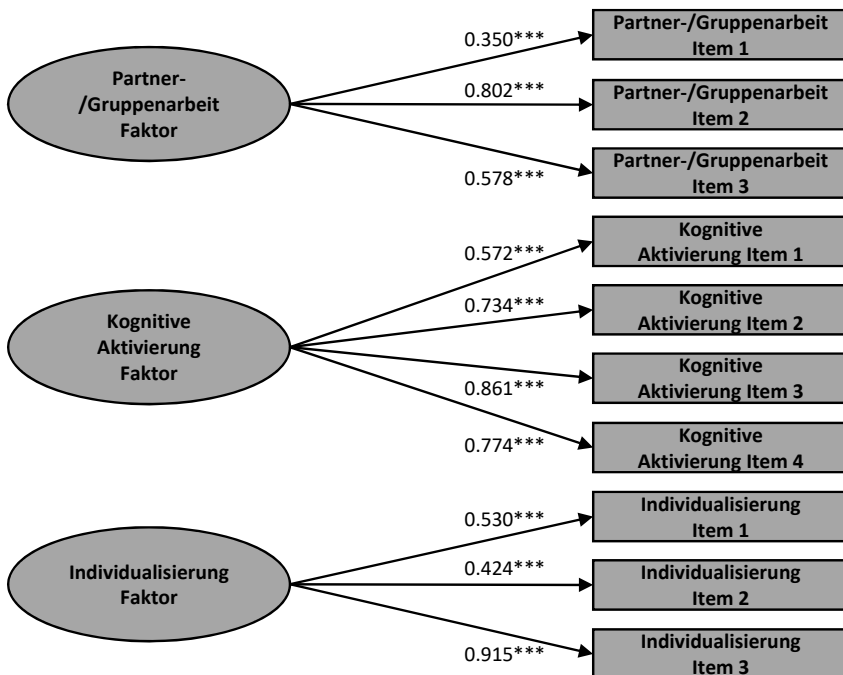


Abb. 20: Faktorladungen auf der Klassenebene (standardisierte Faktorladungen, *** $p < 0.001$)

Tab. 26: Faktorkorrelationen auf der Klassenebene (standardisierte Korrelationskoeffizienten)

	Partner-/ Gruppenarbeit	Kognitive Aktivierung	Individualisierung
Partner-/ Gruppenarbeit	1		
Kognitive Aktivierung	0.542***	1	
Individualisierung	0.580**	0.720***	1

Anmerkung: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$.

Anhang C: Abbildungen von Regressionslinien

In diesem Anhang finden sich Abbildungen von Regressionslinien von pädagogischen Merkmalen mit kurvilinearen Effekten. Diese Abbildungen sollen die Interpretation der im Ergebnisteil berichteten Regressionskoeffizienten erleichtern.

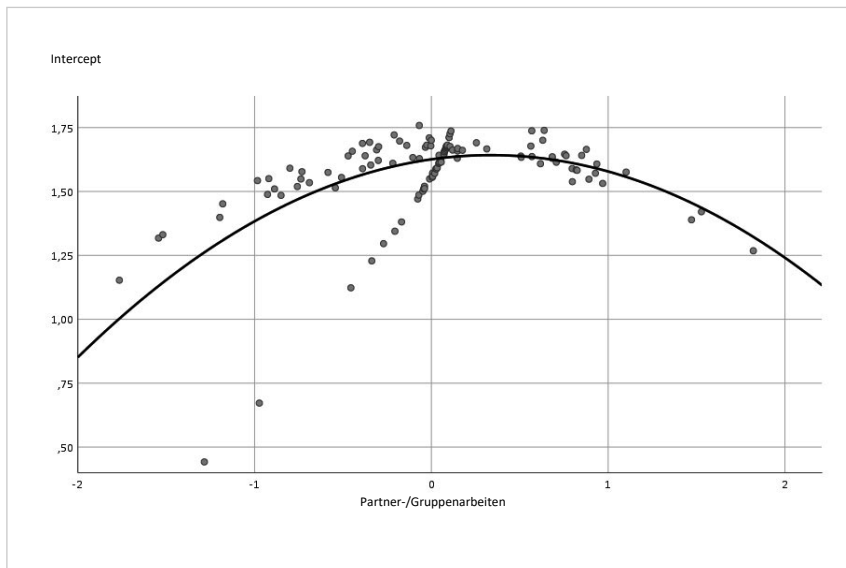


Abb. 21: Regressionslinie des Effekts von Partner-/Gruppenarbeiten auf den durchschnittlichen Lernerfolg (Partner-/Gruppenarbeiten latent modelliert)

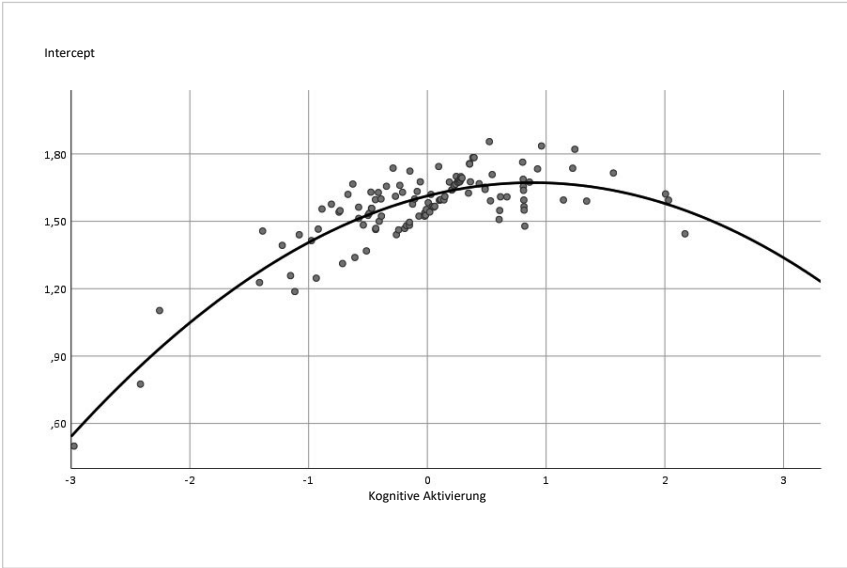


Abb. 22: Regressionslinie des Effekts von kognitiver Aktivierung auf den durchschnittlichen Lernerfolg (kognitive Aktivierung latent modelliert)

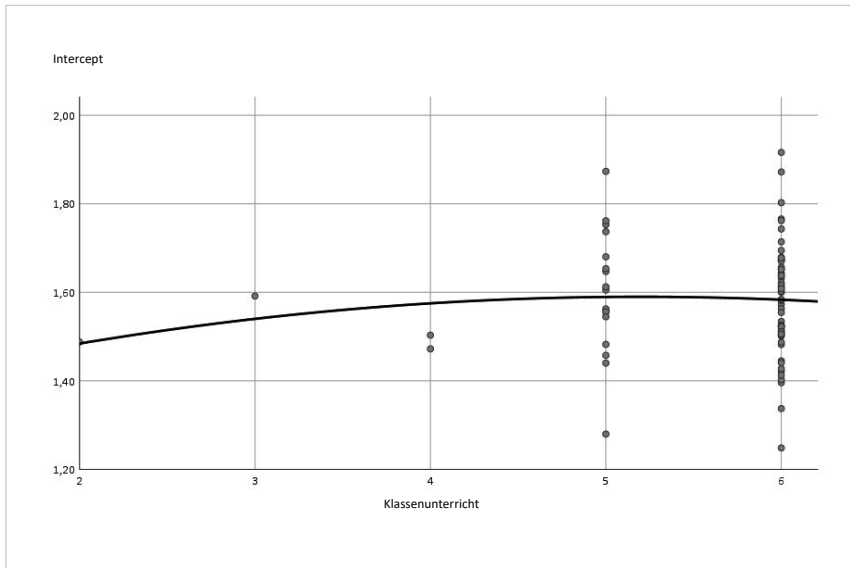


Abb. 23: Regressionslinie des Effekts von Klassenunterricht auf den durchschnittlichen Lernerfolg