

# **Individuelle Förderung im Unterricht: Eine Möglichkeit, soziale Herkunft und Schulerfolg zu entkoppeln?**

Eine Re-Analyse aus IGLU-E 2011

Dissertation

zur Erlangung des akademischen Grades einer  
Doktorin der Philosophie (Dr. phil.)

an der Fakultät für Erziehungswissenschaft der Universität Hamburg

vorgelegt von  
Katrin Schulz-Heidorf

Hamburg, 08.08.2016



Erstgutachter: Prof. Dr. Knut Schwippert  
(Universität Hamburg)

Zweitgutachter: Prof. Dr. Johannes König  
(Universität zu Köln)

Mündlicher Prüfer: Prof. Dr. Dietmar Höttecke  
(Universität Hamburg)

Disputation: Universität Hamburg, 05. Oktober 2016



*Für Christina, Wolfgang, Sophie und Charlotte.*

*Für meine Familie.*



## Inhaltsverzeichnis

1. Einleitung.....	1
2. Heterogenität im Schulalltag.....	4
3. Differenzierung im Unterricht.....	15
4. Individuelle Förderung: Voraussetzungen und Umsetzung.....	21
5. Individuelle Förderung: Implementierung in Deutschland.....	26
6. Effekte individueller Förderung .....	27
7. Herkunftseffekt .....	51
8. Schulerfolg.....	58
8.1 Leseleistung und ihre Konzeption in IGLU.....	59
8.2 Leseselbstkonzept und seine Konzeption in IGLU .....	62
8.3 Rahmenmodelle des Schulerfolgs .....	72
9. Wahrnehmung der schulischen Lernumwelt.....	75
9.1 Die Entwicklung der Klimaforschung.....	77
9.2 Dimensionen des schulischen Klimas .....	80
9.3 Kritik an der Klimaforschung .....	82
9.4 Wahrnehmung der sozialen Lernumwelt durch Schülerinnen und Schüler .....	83
9.5 Schüler-Schüler-Beziehung.....	84
9.6 Schüler-Lehrer-Beziehung .....	86
9.7 Soziale Lernumwelt und Schulerfolg .....	88
9.8 Individuelle Förderung und soziale Merkmale der Lernumwelt .....	106
10. Zusammensetzung der Klasse .....	108
11. Fragestellung und Ableitung der Hypothesen .....	110
11.1 Herkunftseffekt.....	112
11.2 Effekte der Wahrnehmung der schulischen Lernumwelt.....	112
11.3 Schulerfolg .....	113

11.4	Effekte der Klassenzusammensetzung .....	113
11.5	Effekte der individuellen Förderung.....	114
12.	Empirische Analysen .....	116
12.1	Datengrundlage: IGLU-E 2011 .....	116
12.2	Analysemethoden und Kennwerte .....	119
12.3	Prüfung der Einzelmodelle .....	128
12.3.1	ESCS: Die sozioökonomische Stellung der Familie.....	129
12.3.2	HLE: Die häusliche Lesesozialisation .....	145
12.3.3	Leseleistung.....	146
12.3.4	Leseselbstkonzept.....	148
12.3.5	Schüler-Lehrer-Beziehung.....	149
12.3.6	Schüler-Schüler-Beziehung .....	152
12.3.7	Klassenzusammensetzung .....	154
12.3.8	Individuelle Förderung.....	156
12.4	Prüfung der einzelnen Zusammenhänge.....	162
12.4.1	Nullmodell.....	163
12.4.2	Herkunft und Schulerfolg.....	165
12.4.3	Soziale Beziehungen und Schulerfolg .....	165
12.4.4	Klassenzusammensetzung und individuelle Förderung.....	166
12.4.5	Individuelle Förderung und Herkunft .....	167
12.4.6	Individuelle Förderung und soziale Beziehungen .....	170
12.4.7	Individuelle Förderung und Schulerfolg.....	171
12.5	Prüfung des Gesamtmodells.....	172
13.	Zusammenfassung und Diskussion der Ergebnisse.....	177
13.1	Zusammenfassung der Ergebnisse .....	177
13.2	Diskussion der Ergebnisse .....	178

13.3 Grenzen der empirischen Untersuchung.....	191
14. Ausblick .....	194
14.1 Implikationen für zukünftige Forschungsarbeiten .....	194
14.2 Implikationen für die schulische Praxis und den bildungspolitischen Diskurs um individuelle Förderung im Unterricht .....	200
Verzeichnisse.....	205
Literaturverzeichnis.....	205
Tabellenverzeichnis.....	230
Abbildungsverzeichnis.....	234
Anhang.....	236
Anhang 1: Mehrebenenstruktur: Ein Beispiel.....	236
Anhang 2: ISCO-08-Kodierung.....	237
Anhang 3: Kodierung der Items ASBH20A und ASBH20B .....	238
Anhang 4: Explorative Faktorenanalysen.....	239
Anhang 4.1: HLE: Die häusliche Lesesozialisation .....	239
Anhang 4.2: Leseselbstkonzept .....	240
Anhang 4.3: Schüler-Lehrer-Beziehung .....	241
Anhang 4.4: Schüler-Schüler-Beziehung.....	242
Anhang 4.5: Individuelle Förderung .....	243
Anhang 5: Parameter des Gesamtmodells .....	245
Anhang 5.1: Parameter des Strukturmodells unter Berücksichtigung nicht signifikanter Pfade .....	245
Anhang 5.2: Faktorladungen der Messmodelle .....	246
Anhang 5.3: Parameter des Strukturmodells .....	247
Anhang 6: Zusammenfassung der Dissertation .....	248
Anhang 6.1: Deutschsprachige Zusammenfassung .....	248
Anhang 6.2: Englischsprachige Zusammenfassung .....	251



## 1. Einleitung

Vielfalt mit all ihren Implikationen gilt heute als eine der größten gesellschaftlichen Herausforderungen und ist gleichzeitig mit der Chance verbunden, einen „Nährboden neuer Ideen“ (Schönwälder, Petermann & Vertovec, o.J., S. 2) darzustellen, wenn dem Vorhandensein unterschiedlicher Perspektiven, Fähigkeiten und Potenziale Rechnung getragen wird. Gerade im Hinblick auf rasante technologische, arbeitsmarkt- und gesellschaftsbezogene Veränderungen kann dies von herausragender Bedeutung sein – wissen wir doch heute noch nicht, welche zukünftigen Bedürfnisse mit diesen Entwicklungen einhergehen. Aus gesellschaftlicher Perspektive erscheint es daher vor allem als eine Frage des ‚Sich-leisten-könnens‘, die Heterogenität, die individuelle Potenziale und Begabungen mit sich bringen, nicht zu berücksichtigen.

Übertragen lässt sich dieser Diskurs auf den ‚Mikrokosmos Grundschule‘, der als Abbild der heterogenen Zusammensetzung einer Gesellschaft ähnliche Bewertungstendenzen aufweist: Wurden bis zu dem durch den PISA-Schock eingeleiteten bildungspolitischen Umdenken noch vermehrt<sup>1</sup> Homogenisierungsstrategien in Schule und Unterricht proklamiert, findet das Konzept der ‚individuellen Förderung‘, welches die individuellen Fähigkeiten und Interessen der Schülerinnen und Schüler<sup>2</sup> in Lehr-Lern-Situationen berücksichtigt, seit Veröffentlichung der Ergebnisse der PISA-Studie starke Popularität (Klieme & Warwas, 2011). Individuelle Förderung wird dabei als „Schlüssel zur Lösung bildungspolitischer und pädagogischer Probleme“ (Klieme & Warwas, 2011, S. 805) verstanden, der es jedem Kind ermögliche, sein Potenzial auszuschöpfen. Diesen Ansprüchen im Klassenzimmer zu begegnen, halten auch Lehrende für wichtig (Abs, 2005; Solzbacher, 2008; Fischer, 2014), gleichzeitig fehlt es an elaborierten Unterrichtskonzepten (Altrichter, Trautmann, Wischer, Sommerauer & Doppler, 2009; Klieme & Warwas, 2011; Bräu, 2007 und andere): So wird die Umsetzung entsprechend als didaktische Herausforderung bewertet, wenngleich individuelle Förderung per se als Chance verstanden wird, den heterogenen Lernvoraussetzungen, Interessen und Fähigkeiten der Schülerinnen und Schüler gerecht zu werden.

---

<sup>1</sup> Dass auch vor 2002 vor allem reformpädagogische Ansätze der individuellen Förderung existierten, soll in Kapitel 2 ausführlicher erläutert werden.

<sup>2</sup> In der vorliegenden Arbeit wird gegendert. An einzelnen Stellen soll jedoch zu Gunsten einer besseren Lesbarkeit jeweils nur das generische Maskulinum aufgeführt werden, z.B. bei ‚Schülerpopulation‘.

Dabei bleibt bisher offen, ob individuelle Förderung ebenfalls eine Möglichkeit darstellt, dem sogenannten Herkunftseffekt ausgleichend zu begegnen. Letzterer beschreibt die prognostische Bedeutsamkeit vorrangig sozioökonomischer Merkmale der Familie, etwa dem Sozialstatus und den familiären Ressourcen, aber auch dem häuslichen Anregungsgehalt für die Kompetenzentwicklung und Bildungsbeteiligung der Schülerinnen und Schüler (Boudon, 1974; Wild & Lorenz, 2010; Becker, 2010). Eine Folge konnten u.a. Bos und Kollegen (2003, aktueller auch Wendt, Stubbe & Schwippert, 2012) nachweisen: Viertklässlerinnen und Viertklässler aus Elternhäusern der untersten Sozialschicht liegen im Leseverständnis etwa eine halbe Standardabweichung (2001: 56 Punkte, 2011: 53 Punkte) hinter ihren Mitschülerinnen und Mitschülern der obersten Sozialschicht, was in etwa dem Leistungszuwachs eines Schuljahres entspricht. Es muss also davon ausgegangen werden, dass sich Schulklassen aus Schülerinnen und Schülern unterschiedlicher Herkunft und dementsprechend unterschiedlicher Eingangs- und Lernvoraussetzungen zusammensetzen. Ob differenzierte Lernangebote im Konzept der individuellen Förderung, die sich an diesen Voraussetzungen der Kinder und Jugendlichen orientiert, nicht nur den Schulerfolg, sondern auch den Herkunftseffekt beeinflussen können, ist Gegenstand der vorliegenden Analysen. Diese erfolgen auf Grundlage der Daten der Internationalen Grundschul-Lese-Untersuchung (IGLU), so dass sich in vielen Aspekten auf Merkmale des Lesen Lernens bezogen wird.

In Anlehnung an bekannte Rahmenmodelle des Schulerfolgs, die konstatieren, dass sich Lehrangebote nicht direkt und linear in Lernergebnissen bei den Schülerinnen und Schülern niederschlagen (vgl. v.a. Helmke, 2006; auch Ditton, 2000; Slavin, 1996; Stringfield, 1994 u.a.), wird das Modell zudem um soziale Merkmale der Lernumwelt erweitert, zu denen die Integration in den Klassenkontext und die Beziehung zu der Lehrkraft gehören. Für diese wird nicht nur angenommen, dass sie stark durch individualisierte Unterrichtssettings beeinflusst werden, sondern auch, dass ihre positive Wahrnehmung in einem positiven Zusammenhang mit dem Schulerfolg von Kindern und Jugendlichen steht (vgl. etwa Saldern, 1987; Eder, 1996; Roorda, Koomen, Spilt & Oort, 2011 und Hattie, 2009). Überprüft werden soll dies im Rahmen eines Mehrebenenstrukturgleichungsansatzes.

Dabei wird zunächst auf die lern- und leistungsbezogene Heterogenität der Schülerinnen und Schüler im Klassenzimmer eingegangen, die trotz Homogenisierungsbestrebungen wie Klassenwiederholungen und schulformbezogene Selektionsmechanismen (Lehmann, Peek,

Pieper & Von Stritzky, 1995; Artelt, Stanat, Schneider & Schiefele, 2001; Arnold, Bos, Richert & Stubbe, 2007; Belser & Küsel, 1976; Tietze & Roßbach, 1998) kaum verringert wird und damit charakterisierendes Merkmal der Klassenzusammensetzung, vor allem in der hier untersuchten Grundschule ist (Kapitel 2). In diesem Kontext sind es vorrangig Maßnahmen der Differenzierung, und hier insbesondere individualisierte Unterrichtssettings, die sowohl von bildungspolitischer als auch lerntheoretischer und empirischer Seite für den Umgang mit unterschiedlichen Lernvoraussetzungen proklamiert werden (Kapitel 3). Deutlich wird dabei, dass dem hier interessierenden Konzept der Individualisierung höchst unterschiedliche Definitionen zugrunde liegen, was insbesondere Folgen für die Interpretation der zugehörigen Fragen im Fragebogen, aber auch der Zusammenfassung empirischer Arbeiten hat. Dass ihre Umsetzung gekoppelt ist an spezifische Voraussetzungen und Kompetenzen sowohl bei Lehrenden als auch Lernenden, ist Gegenstand des Kapitels 4. Dem schließt sich eine Situationsanalyse des aktuellen Stands ihrer Implementation in den deutschen Schulunterricht an (Kapitel 5), gefolgt von einer Darstellung des aktuellen Forschungsstands zu den Effekten individueller Förderung (Kapitel 6). In Verbindung gebracht wird das didaktische Konzept anschließend mit dem Herkunftseffekt (Kapitel 7), der durch individualisierte Lernsettings, in denen die schülerspezifischen Voraussetzungen, aber auch ihre Interessen und Neigungen explizit Berücksichtigung erfahren, abgeschwächt werden könnte. Dies könnte sich dann in einem herkunftsunabhängige(re)n Schulerfolg der Lernenden widerspiegeln, wobei unter ‚Schulerfolg‘ nicht nur kognitive Dimensionen wie die Leseleistung, sondern auch affektive Merkmale der Lernenden, hier ihr Selbstkonzept im Lesen, verstanden werden, da dies ebenfalls als wichtiger Ertrag von Schule gilt (Kapitel 8). Wie erwähnt werden diese zunächst theoretischen Betrachtungen – in Anlehnung an entsprechende Rahmenmodelle des Schulerfolgs – anschließend um die Facette der Wahrnehmung der sozialen Lernumwelt der Schülerinnen und Schüler erweitert. Diese können (neben Merkmalen des Unterrichts) dazu beitragen, den Schulerfolg der Lernenden zu erklären und gleichzeitig durch Maßnahmen individueller Förderung positiv beeinflusst werden (Kapitel 9). Bevor die Wirkzusammenhänge dann in Kapitel 11 abgeleitet werden, wird das Modell um eine weitere Facette ergänzt: So soll die herkunfts- und sprachbezogene Klassenzusammensetzung in die Analysen aufgenommen werden, da das Urteil über die Qualität von Unterricht gerechter ausfällt, wenn berücksichtigt wird, dass Lehrende unter Umständen in einer Klasse mit schwierigeren Lernausgangslagen unterrichten als ihre Kolleginnen und Kollegen und deswegen geringere *Leis-*

*tungswerte* am Ende eines Untersuchungszeitraums zeigen, wenngleich die *Leistungszuwächse* möglicherweise größer waren (Kapitel 10).

Die sich anschließende empirische Analyse auf Grundlage von Mehrebenenstrukturgleichungsmodellen gliedert sich zunächst auf in die Beschreibung der Datengrundlage IGLU-E 2011 (Kapitel 12.1) und die Darstellung der Analysemethode und der für die Modellbewertung relevanten Kennwerte (Kapitel 12.2). Dem folgt die separate Prüfung der einzelnen Messmodelle (Kapitel 12.3). In Kapitel 12.4 werden diese dann schrittweise miteinander in Beziehung gebracht, bis das Gesamtmodell (Kapitel 12.5) aufgebaut ist. Dem schließt sich die Zusammenfassung der Ergebnisse (Kapitel 13.1), ihre Diskussion (Kapitel 13.2) sowie eine Betrachtung der Grenzen dieser empirischen Untersuchung (Kapitel 13.3) an. Abgeschlossen wird die Arbeit mit den sich aus ihr ergebenden Implikationen für zukünftige Forschungsarbeiten (Kapitel 14.1) sowie für die schulische Praxis und den bildungspolitischen Diskurs um individuelle Förderung im Unterricht (Kapitel 14.2).

## **2. Heterogenität im Schulalltag**

Während in Deutschland Schülerinnen und Schüler ab der Sekundarstufe I im Zuge von Homogenisierungsbestrebungen in unterschiedlichen Schulformen unterrichtet werden, findet sich in der Grundschule aufgrund ihrer Konzeption als Gesamtschule eine nahezu unausgelesene Schülerpopulation wieder, in der sich die unterschiedlichen Ausprägungen etwa in Vorwissen, Lernvoraussetzungen, Leistungsvermögen, familiären Hintergrundmerkmalen sowie Sprachgebrauch der Grundgesamtheit spiegeln (Kluczniok, Große & Roßbach, 2014). Dies geht einher mit unterschiedlichen Klassenzusammensetzungen, die je nach Streuung bestimmter Schülermerkmale als ‚heterogen‘ oder ‚homogen‘ klassifiziert werden können. Dabei stellt ‚Heterogenität‘ per se keine objektive Norm dar: „Heterogenität ist also ein ‚relativer‘ Begriff, sie hängt vom Maßstab ab und ist nur zusammen mit Homogenität zu betrachten, wird erst durch Vergleichsoperationen ‚hergestellt‘ und ist wandelbar“ (Wenning, 2007, S. 24, Herv. i. O.). Haag und Streber (2014) verdeutlichen dies an einem Beispiel: Bezüglich der Vergleichskategorie ‚Alter‘ könne man Grundschul Kinder im Alter von sechs bis zehn Jahren – verglichen mit Erwachsenen – als relativ homogene Gruppe bezeichnen, während auf Klassenebene hingegen Standardabweichungen von nur zehn Monaten als Maß für *altersheterogene* und Standardabweichungen von drei Monaten als Indikator für besonders *alters-*

homogene Klassen identifiziert wurden (Roßbach & Tietze, 1996)<sup>3</sup>. Im Kontext der vorliegenden Arbeit sollen, in Anlehnung an Gröhlich, Scharenberg und Bos (2009) und Haag und Streber (2014), unter ‚Heterogenität‘ die kulturellen, ökonomischen und sprachlichen Unterschiede sowie die daraus und aus den unterschiedlichen häuslichen Anregungsgehalten resultierenden leistungs- und selbstbezogenen Ausgangsbedingungen für das schulische Lernen von Kindern und Jugendlichen verstanden werden. Damit grenzt sich diese Definition explizit von Merkmalen wie Interesse, Neigungen, Geschlecht, Alter, biographischen Erfahrungen und ähnlichem ab, die zwar durchaus für die Beschreibung von Heterogenität genutzt werden können (vgl. etwa Speck-Hamdan, 2009), allerdings nicht Bestandteil der nachstehenden Analysen<sup>4</sup> sind.

Dass der Begriff der Heterogenität im Sinne einer ‚Abwesenheit von Gleichartigkeit‘ und damit verbunden als etwas Defizitäres, als Belastung (vgl. Specht, 2009) bislang eher negativ konnotiert wurde, erscheint im Anbetracht der bisherigen schulischen Praxis der Lerngruppenhomogenisierung (Gröhlich et al., 2009) wenig verwunderlich. Zurückgeführt werden kann dies vor allem auf das Konzept der Altersgruppierung von Schülerinnen und Schülern in Jahrgangsklassen, welches seit dem 19. Jahrhundert maßgeblich das Schulwesen strukturiert (Rauin, 1987; Altrichter et al., 2009; Krammer, 2009). Dem liegt die Annahme zugrunde, dass ein gleiches biologisches Alter eine ‚Gleichartigkeit des Lernalters‘ (Altrichter et al., 2009) bedingt – und damit die Voraussetzungen für möglichst ökonomische Lehrbedingungen schafft. Weitere, vorrangig leistungsbezogene Homogenisierungsbemühungen finden sich in der Etablierung unterschiedlicher Sekundarschulformen (vorrangig Haupt- und Realschulen sowie Gymnasien<sup>5</sup>; ausgenommen Gesamtschulen, hier liegt gerade *kein* leistungsbezogenes Homogenisierungsbestreben vor) und in Klassenwiederholungen und Zurückstellungen. Diesen Maßnahmen „liegt eine regelrechte ‚Sehnsucht nach der homogenen

---

<sup>3</sup> Anhand von Analysen der BiKS-3-8-Studie stellen Kluczniok, Große und Roßbach (2014) hingegen klasseninterne Altersunterschiede zwischen dem jüngsten und ältesten Kind von einem halben Jahr (altershomogenste Klasse) bis hin zu zwei Jahren und acht Monaten (altersheterogenste Klasse) fest – verglichen mit der altersbezogenen Streuung zwischen Erwachsenen und Kindern (wie in dem Beispiel oben) ist dies allerdings immer noch als homogen zu bewerten, was verdeutlicht, dass es hierbei vor allem auf die Referenzkategorien ankommt.

<sup>4</sup> Einige der genannten Merkmale, die hier nicht für die Beschreibung von Heterogenität herangezogen werden, finden dennoch Berücksichtigung im theoretischen Teil dieser Arbeit. Dies liegt begründet in den Definitionen, die anderen Studien und Ausarbeitungen zugrunde liegen.

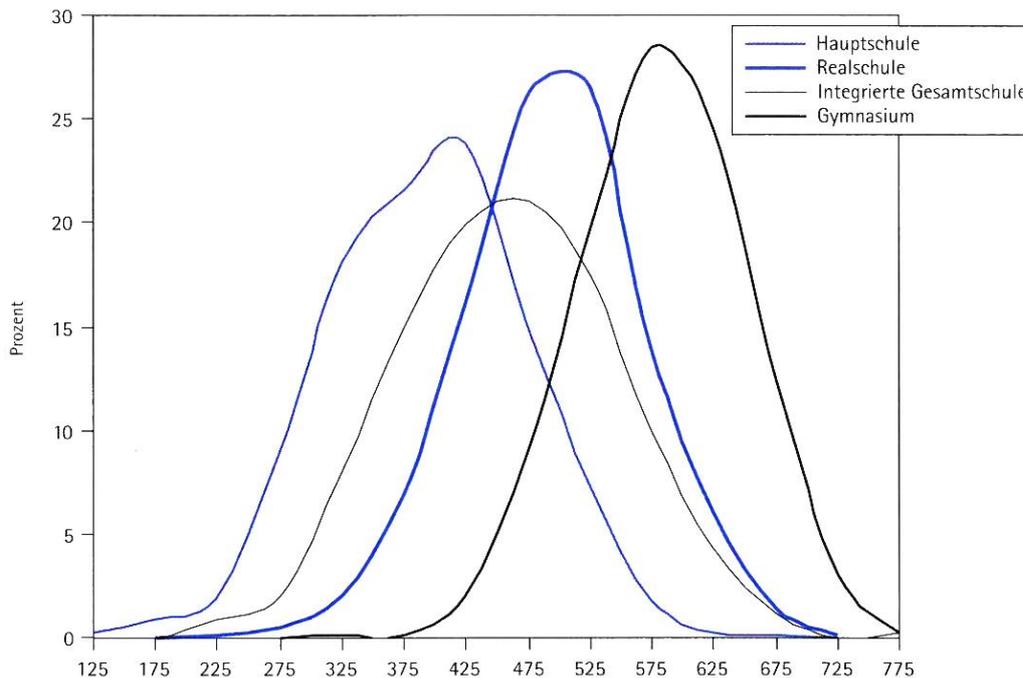
<sup>5</sup> Anzumerken sei jedoch, dass diese Bestrebungen in jüngster Zeit kontrovers diskutiert und gegenläufige Maßnahmen in einigen Bundesländern ergriffen wurden, beispielsweise durch die Integration von Haupt- und Realschulen in sog. Sekundarschulen, die teilweise über gymnasiale Oberstufen verfügen (s. etwa Hurrelmann, 2013).

Lerngruppe' [...] und die Überzeugung zugrunde, dass sich der Unterrichtsstoff in homogenen Lerngruppen leichter auf ein mittleres Niveau ausrichten lasse, um Überforderung bei den Leistungsschwächeren und Unterforderung bei den Leistungsstärkeren zu vermeiden“ (Bos & Scharenberg, 2010, S. 73f.). Dass eine leistungsbezogene Homogenisierung durch oben genannte Bemühungen nur unvollständig gelingt, konnte bereits mehrfach nachgewiesen werden:

So zeigen sich in der nach Bildungsgängen<sup>6</sup> differenzierten Darstellung der Ergebnisse des internationalen Lesekompetenztests (verstehender Umgang mit Texten) in PISA 2000 bei Artelt et al. (2001) zwar zunächst signifikante Unterschiede zwischen den mittleren Leistungsniveaus der Schülerinnen und Schüler in Hauptschulen, Realschulen, integrierten Gesamtschulen und Gymnasien. Gleichzeitig ist jedoch eine erhebliche Überlappung der Leistungsstreuung erkennbar (vgl. Abbildung 1). So erreichen ca. 12% der Schülerinnen und Schüler an Realschulen ein Leistungsniveau, welches oberhalb des gymnasialen Durchschnitts von 582 Punkten liegt. Etwa 10% der Hauptschülerinnen und -schüler weisen zudem bessere Lesekompetenzen auf als der Realschuldurchschnitt. Prüft man die Schulzuordnung von Jugendlichen, deren Leistungen dem Mittelwert der Lesekompetenz in allen OECD-Staaten entsprechen (M = 500 Punkte), befinden sich etwa 27% dieser Schülerinnen und Schüler an Realschulen, 15% an Gymnasien und 10% an Hauptschulen.

---

<sup>6</sup> Die Differenzierung erfolgte nach Bildungsgängen und nicht nach Schulformen, da die Zuordnung von Schülerinnen und Schülern an Schulen mit mehreren Bildungsgängen ( $N_{SCL} = 19$ ) so eindeutiger vorgenommen werden konnte. Zu beachten ist, dass sich Schülerinnen und Schüler an integrierten Gesamtschulen nicht in leistungshomogenen Klassen (z.B. Hauptschul- oder Realschulklassen innerhalb der gleichen Schule) befinden und eine eindeutige Zuordnung dieser Jugendlichen zu den unterschiedlichen Schulformen nicht möglich ist. Die Interpretation des Gesamtmittelwerts für diesen Bildungsgang ist also weniger aussagekräftig (vgl. Artelt, Stanat, Schneider & Schiefele, 2001).

**Abbildung 1:** Verteilung der Leistungen auf der Gesamtskala Lesen nach Bildungsgang in PISA 2000

Quelle: Artelt et al., 2001, S. 121

Diese Ergebnisse replizieren die Befunde von Lehmann et al. (1995), die im Rahmen der Hamburger Lesestudie als deutsche Teiluntersuchung der Internationalen Lesestudie der IEA (ein Vorläufer der späteren Internationalen Grundschul-Lese-Untersuchung) bereits 1995 11.061 Schülerinnen und Schüler in allen west- und ostdeutschen Bundesländern in den Klassenstufen 3 und 8 bezüglich ihres Leseverständnisses untersuchten. Für den Schulformvergleich wurde das durchschnittliche Leseverständnis von Achtklässlern nach Schulformzugehörigkeit aufgeschlüsselt (vgl. Tabelle 1). Deutlich wird, analog zu den Ergebnissen von Artelt et al. (2001), dass das Leseverständnis von Hauptschülerinnen und -schülern, die sich eine Standardabweichung oberhalb des Hauptschulmittelwerts befinden, in etwa dem durchschnittlichen Niveau an Gesamt- und Realschulen entspricht. Auch überlappt sich die Leistungsverteilung von Hauptschulen und Gymnasien: Der Testwert 1,4 entspricht beispielsweise in etwa dem 68. Perzentil an Hauptschulen und liegt etwa eine Standardabweichung unterhalb des Mittelwerts an Gymnasien. Dieser Wert unterscheidet sich zudem nicht deutlich von der mittleren Leseleistung an Real- und Gesamtschulen.

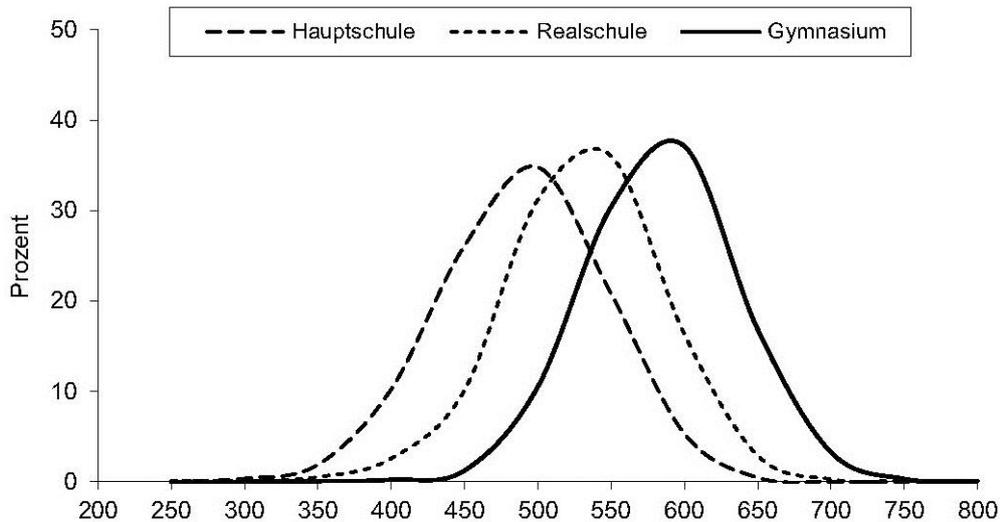
**Tabelle 1:** Leseverständnis von Achtklässlerinnen und Achtklässlern in west- und ost-deutschen Bundesländern im Schuljahr 1990/1991 nach Schulform

Schulform	arithm. Mittel	Standardabweichung	68. Perzentil
Hauptschule	0,55	0,86	-0,36 – 1,41
Gesamtschule	1,15	1,00	0,15 – 2,15
Realschule	1,46	0,76	0,70 – 2,22
Gymnasium	2,28	0,85	1,43 – 3,13

Quelle: Lehmann et al., 1995, S. 149, adaptiert

Und auch in aktuellen Untersuchungen zeigt sich ein ähnliches Bild: So wurde für entsprechende Analysen der IGLU-Daten aus 2006 und 2011 die Schullaufbahnpflichtung der Lehrenden<sup>7</sup> als Indikator für die (spätere) Schulformzugehörigkeit der Viertklässlerinnen und Viertklässler herangezogen und diese mit den Testleistungen der Lesekompetenz verglichen. Dass letztere einen guten Indikator für die gesamtschulischen Kompetenzen von Kindern darstellen, konnten Bos, Bonsen et al. (2007) nachweisen: Auch durch Hinzunahme weiterer Deutschfachanteile oder anderer Fächer führten Analysen zu der Schullaufbahnpräferenz von Lehrpersonen zu ähnlichen Ergebnissen wie unter alleiniger Betrachtung der Lesekompetenz (vgl. Arnold et al., 2007). Die differenzierte Betrachtung der Leistungen auf der Gesamtskala Lesen nach Schulformen zeigt auch hier deutliche Überlappungen (vgl. Abbildung 2). So erhalten etwa 37% der Schülerinnen und Schüler mit 535 Punkten im Lesekompetenztest eine Realschulempfehlung, während etwa 25% der Kinder mit gleicher Leistung eine Hauptschul- oder eine Gymnasialempfehlung erhalten. Wird – wie oben beschrieben – davon ausgegangen, dass die Lesekompetenz als Indikator für die schulische Gesamtkompetenz von Grundschulkindern herangezogen werden kann, lassen die hier und oben dargestellten Befunde darauf schließen, dass die angestrebte leistungsbezogene Homogenisierung durch die Gliederung in unterschiedliche Schulformen nicht vollständig gelingt.

<sup>7</sup> Anzumerken sei, dass nicht in allen Bundesländern die Schullaufbahnpflichtung durch die Lehrkraft bindend ist (verbindlich in Baden-Württemberg, Bayern, Brandenburg, Sachsen, Saarland, Thüringen, vgl. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2010). In IGLU 2011 berechnen Stubbe, Bos und Euen (2012) die Abweichungsquote der Elternpräferenz von denen der Lehrkräfte mit 17,3%.

**Abbildung 2:** Testleistungen im Leseverständnis (Gesamtskala) nach Schullaufbahnpräferenzen der Lehrkräfte differenziert, IGLU 2006

Quelle: Arnold et al., 2007, S. 216

Ein ähnliches Bild zeichnet sich für die Wirksamkeit der Leistungshomogenisierung durch Klassenwiederholungen ab: So finden nicht versetzte Schülerinnen und Schüler auch nach Klassenwiederholungen meist nicht den Anschluss an die mittleren Leistungen ihrer neuen Klassenkameraden – dieser Leistungsrückstand wird zudem noch größer, je öfter sie nicht versetzt werden (Ingenkamp, 1972). Belser und Küsel (1976) ergänzen, dass im Wiederholungsjahr zwar eine Verbesserung der Leistungen zu beobachten sei, diese aber im Folgejahr durch neue und höhere Anforderungen sinken würden.

Diese Befunde bestätigen sich auch in internationalen Studien. Nach Tietze und Roßbach (1998) zeigen sich in 60 analysierten Untersuchungen „im Durchschnitt Vorteile der versetzten Schüler im Vergleich zu den nicht versetzten [...]“. Besonders ausgeprägt sind die Unterschiede bei den Schulleistungen. Der Vergleich sitzengebliebener Kinder mit gleich leistungsschwachen, aber versetzten Schülern zum gleichen Alterszeitpunkt ergibt deutliche Leistungsunterschiede zuungunsten der Sitzenbleiber“ (Tietze & Roßbach, 1998, S. 467). Dies findet sich ebenfalls in einer Studie von Ehmke, Drechsel und Carstensen (2008), in welcher anhand der PISA-I-Plus-Daten zu zwei Messzeitpunkten überprüft wurde, wie sich Neuntklässlerinnen und Neuntklässler bezüglich ihrer mathematischen Kompetenz entwickelten. Hier zeigte sich, dass sich die Klassenwiederholer zwar um durchschnittlich 23 Punkte auf der Mathematikskala<sup>8</sup> verbesserten (das entspricht in etwa dem durchschnittlichen Leistungs-

<sup>8</sup> Die international verwendete PISA-Mathematikskala weist einen Mittelwert von 500 Punkten und eine Standardabweichung von 100 im OECD-Durchschnitt auf (Ehmke, Drechsel & Carstensen, 2008).

zuwachs innerhalb eines Schuljahres und unterscheidet sich nicht signifikant vom Leistungszuwachs der regulär versetzten Schülerinnen und Schüler). Ein mit 38 Prozent großer Anteil der Repetenten zeigte allerdings keinen bedeutsamen Lernzuwachs oder verschlechterte sich sogar (Ehmke et al., 2008). Auch Hong und Raudenbush (2005) reihen sich in diese Befunde ein: Bei einem Vergleich von Grundschulkindern in den USA zeigten die Schülerinnen und Schüler, die trotz schlechter Leistungen versetzt wurden, im Laufe ihres schulischen Werdegangs bessere Rechen- und Leseleistungen als die nicht versetzten Kinder der Vergleichsgruppe. Und auch Tietze und Roßbach (1998) konstatieren, dass „der Leistungsabstand [zwischen leistungsschwachen, aber versetzten Kindern und der nicht versetzten Vergleichsgruppe, Anm. KSH] im Verlaufe der folgenden Schuljahre noch zunimmt“ (Tietze & Roßbach, 1998, S. 467).

Bezüglich des vermeintlich homogenisierend wirkenden Konzepts der Klassenwiederholung kommt Klemm (2009) abschließend zu dem Urteil, dass weder ein „nachhaltiger Ertrag des Wiederholens einer Klasse zu beobachten [sei], noch [...] festzustellen [sei], dass die leistungstärkeren Schüler und Schülerinnen vom ‚Aussortieren‘ der Schwächeren profitieren“ (Klemm, 2009, S. 7f., Herv. i. O.). Im Mittel kann also keine leistungsbezogene Homogenisierung durch Klassenwiederholungen festgestellt werden: Zwar wird eine Klasse, deren leistungsschwächstes Mitglied sitzenbleibt, zunächst homogenere Leistungen aufweisen, allerdings wird die Klasse des nachfolgenden Jahrgangs langfristig heterogenere Leistungsstreuungen zeigen, da nicht (dauerhaft) davon auszugehen ist, dass der Repetent den Leistungsrückstand alleine durch die Klassenwiederholung<sup>9</sup> aufholt.<sup>10</sup> Zudem könnte auch die Klasse, in der Schülerinnen und Schüler nicht weiterversetzt werden, Repetentinnen und Repetenten der oberen Klassenstufen aufnehmen, so dass sich gewissermaßen ein ‚linearer Austausch‘ von oben nach unten einstellt.

Es zeigt sich also, dass insbesondere die Strategien der leistungsbezogenen Gliederung der Schülerinnen und Schülern in unterschiedlichen Schulformen und das Klassenwiederholen wenig zu einer heterogeneren Klassenzusammensetzung im deutschen Bildungssystem bei-

---

<sup>9</sup> Bisher lassen sich für Deutschland kaum weitere, die Klassenwiederholung begleitende Maßnahmen etwa in Form von zusätzlicher Unterstützung oder speziellen Unterrichtsprogrammen finden (Ehmke et al., 2008).

<sup>10</sup> Auf die negativen Wirkungen von Klassenwiederholungen auf das Selbstkonzept und motivationale Faktoren der Kinder und Jugendlichen soll hier – ebenfalls wie auf den damit verbundenen finanziellen Aufwand – nicht weiter eingegangen werden. Überblicke hierzu finden sich bspw. bei Liedtke-Schöbel (2012) und Klemm (2009).

tragen. Ergänzt werden muss dies um die Perspektive, dass Heterogenität sich – wie oben beschrieben – nicht ausschließlich auf Leistungsmerkmale wie Noten und Kompetenzen bezieht, sondern beispielsweise auch Aspekte der Lernvoraussetzungen wie Vorwissen, familiäres, häusliches Anregungspotenzial, Selbstwirksamkeit und Selbstkonzept, sprachliche Kenntnisse und ähnliches beinhaltet. Dabei handelt es sich um Merkmale, die kaum auf dem Unterricht übergeordneter Ebene ‚kontrolliert‘ (im Sinne einer Gruppierung in unterschiedliche, in sich möglichst homogene Lerngruppen) werden können. Es muss folglich davon ausgegangen werden, dass trotz aller Bemühungen, durch eine leistungsbezogene Homogenisierung die schulischen Lehrbedingungen möglichst ökonomisch zu gestalten, die Schülerzusammensetzung in Klassen als heterogen zu bezeichnen ist.

Doch selbst wenn es gelänge, Schülerinnen und Schüler aufgrund ihrer unterschiedlichen Leistungen in unterschiedliche, in sich homogene Gruppen aufzuteilen, kann nicht davon ausgegangen werden, dass dies das Lehren im Unterricht erleichtert: Wie aus lernpsychologischer, reformpädagogischer und neurowissenschaftlicher Perspektive bekannt ist, zeigen Schülerinnen und Schüler unter denselben (äußerlichen) Rahmenbedingungen unterschiedliche Lernerfolge.

Untersucht wurde dies beispielsweise im Rahmen des lernpsychologischen Konzepts der Aptitude-Treatment-Interaktion (ATI, vgl. vor allem Corno & Snow, 1986). Gegenstand der Analysen war dabei die leistungsbezogene Wirksamkeit unterschiedlicher *treatments* (im Sinne unterschiedlicher Lehrinstruktionen) in Abhängigkeit von Schülermerkmalen (*aptitudes*). Corno und Snow (1986) definieren *aptitudes* dabei wie folgt:

„The term *aptitude* signifies some aspect of the present state of an individual that is propaedeutic to some future achievement in some particular situation. As such, it incorporates conative and affective as well as cognitive attributes of persons that predict success in specified endeavours. It includes also prior achievement differences among persons that serve as such predictors.” (Corno & Snow, 1986, S. 605, Herv. i. O.)

Diesen weitgefassten Leistungsmerkmalen und ihren Voraussetzungen werden Lehrmethoden (*treatments*) gegenübergestellt, die sich möglichst an den Fähigkeiten (*aptitudes*) der Schülerinnen und Schüler orientieren sollen. Corno und Snow (1986) unterscheiden dabei zwei Pole der Anpassung: Die Makro-Adaption beinhaltet Maßnahmen, die typischerweise für einen längeren Zeitraum getroffen werden, häufig das Curriculum betreffen und dabei

auch Methoden, Medien, Sozialformen etc. beinhalten („month-to-month decisions“, Corno & Snow, 1986, S. 607). Die Mikro-Adaption fasst Handlungen zusammen, die in der klassenspezifischen, alltäglichen Unterrichtssituation der Anpassung dienen („moment-to-moment decisions“, ebd.)<sup>11</sup> und sich vorrangig auf die Interaktion zwischen Lehrperson und Lernendem beziehen. Dieser Strukturierung folgen die empirischen Analysen, bei denen als prominentestes Ergebnis jedoch lediglich der Zusammenhang zwischen der Strukturierung der Lernumgebung und den Leistungen der Lernenden nachgewiesen (und repliziert) werden konnte: So profitieren leistungsstarke Schülerinnen und Schüler vorrangig von offenen Lernumgebungen, während leistungsschwächeren Lernern stark strukturierte Lernumgebungen zu Gute kommen (Cronbach & Snow, 1977). In diese Befundlage reiht sich eine Untersuchung von Weinert und Kollegen (1989) ein, in der im Rahmen der Münchner Longitudinalstudie als Teil der IEA Classroom Environment Study: Teaching for Learning für Schulkinder in der fünften und später sechsten Klasse nachgewiesen werden konnte, dass Lernende mit schlechteren Lernvoraussetzungen (geringe Sprachkompetenz, geringeres Vorwissen u.a.) vor allem von kognitiv strukturierenden, nicht jedoch von offenen Unterrichtsangeboten profitieren – im Gegensatz zu ihren voraussetzungsstarken Mitschülerinnen und Mitschülern (vgl. auch Bruns, 2014). Wenngleich die Forschung im Rahmen des ATI-Ansatzes vorrangig in den 1970er und 1980er Jahren stattfand, konnten ähnliche Befunde in aktuelleren Untersuchungen repliziert werden (vgl. etwa Kirschhock, 2003; Hanke, 2005; Weinhold, 2006). So stellt Kirschhock (2003) bei einer Untersuchung von 135 Erstklässlerinnen und Erstklässlern bezüglich der Entwicklung ihrer schriftsprachlichen Kompetenzen in Abhängigkeit unterschiedlicher Lehrmethoden fest, dass eine „zu starke Strukturierung [...] Kinder mit bereits gut ausdifferenzierten schriftsprachlichen Fähigkeiten behindern“ (Kirschhock, 2003, S. 252) kann. Gleichzeitig benötigen Risikokinder, also solche, die ungünstige Voraussetzungen für den Schriftspracherwerb aufweisen, insbesondere im offenen Schriftspracherwerb strukturierte Unterstützung durch die Lehrkraft, um nicht den Anschluss zu ihren leistungsstärkeren Mitschülerinnen und Mitschülern zu verlieren.

Dass jegliches Lernen individuell ist, ist zudem auch Kerngedanke konstruktivistischer Lerntheorien. Ihnen zufolge seien Wissen und Erkenntnis „nicht der Niederschlag eines passiven Empfangs von Sinneseindrücken oder von Instruktionen, sondern Ergebnisse von

---

<sup>11</sup> Dass diese grobe Dichotomisierung eines Kontinuums (Formen der Adaption finden sich auch zwischen den genannten beiden Polen) etwas kurz gegriffen ist, merken die Autoren bereits selber an – sie dient der Veranschaulichung (vgl. Corno & Snow, 1986).

Handlungen eines aktiven Subjekts“ (Bräu, 2007, S. 174). Dies konstatierten Reformpädagoginnen und -pädagogen bereits seit Ende des 18. Jahrhunderts. So bemerkt etwa Christian Gotthilf Salzmann „einen [...] Hauptmangel“ des schulischen Lernens darin, „dass die Kinder bey dem Lernen mehr fremde, als ihre eignen Kraefte gebrauchen. Es ist noch sehr wenige Anleitung zum eignen Beobachten, eigener Erforschung, eigener Erwerbung der Kenntnisse, sondern der Lehrer arbeitet den Kindern vor, unterrichtet sie von dem, was er durch seine mühsamen Arbeiten herausgebracht hat, und das Kind verhaelt sich dabey mehrentheils ganz leidend“ (Salzmann, 1784, S. 48). Dass Lernprozesse nicht durch reine ‚Übertragung von Wissen‘ initiiert werden, sondern durch die aktive und individuell unterschiedliche Auseinandersetzung mit dem Lerngegenstand, hebt auch Herder (1970) hervor: „Was heißt lernen? [...] Was tun wir, wenn wir gehen, sprechen, zeichnen, tanzen lernen? Nicht wahr? wir [sic] üben und vollführen ein Werk; wir machens nach, bis wirs können, bis es gelingt, mit unsern Kräften, mit unsern Gliedern. So bei sichtbar in die Augen fallenden Künsten; bei unsichtbaren und bei der unsichtbarsten von allen, dem Denken, findet das Lernen auf keine andre Weise statt. Seine Gedanken kann mir der Lehrer nicht eingeben, eintrichtern; meine Gedanken kann, will und muß er durch Worte wecken; also daß sie meine, nicht seine Gedanken sind“ (Herder, 1970, S. 223–225). Faszinierenderweise kann das, was die Reformpädagogik schon vor über 200 Jahren proklamierte, mittlerweile durch neurowissenschaftliche Befunde untermauert werden: Natürliches Lernen vollzieht sich immer durch Beobachten und Nachmachen; Wissen kann nicht übertragen werden, es muss im Gehirn jedes Lernenden neu generiert werden (Roth, 2009; Herrmann, 2010). Eine zentrale Funktion nimmt dabei das limbische System des Gehirns ein: Als zentrales Bewertungssystem vermittelt es Affekte, Gefühle und Motivationen. So wird jede Situation danach beurteilt, ob sie „gut/vorteilhaft/lustvoll war und entsprechend wiederholt werden sollte, oder schlecht/nachteilig/schmerzhaft und entsprechend zu meiden ist“ (Roth, 2009, S. 61). Diese Bewertungen werden im emotionalen Gedächtnis gespeichert<sup>12</sup> und dienen allen zukünftigen Situationen als Referenz: Überprüft wird jedes Mal, wie ähnliche Erlebnisse bewertet und welche Erfahrungen in diesen Situationen gemacht wurden. Roth (2009) leitet daraus die zentrale Rolle des limbischen Systems für den individuellen Lernerfolg ab: So wird in jeder Lernsituation geprüft, was dafür spreche, „dass Hinhören, Lernen, Üben usw. sich tatsächlich lohnen“

---

<sup>12</sup> Zu unterscheiden sind diese Bewertungen von der Speicherung der *Einzelheiten* der Geschehnisse – diese werden in das deklarative Gedächtnis im Hippocampus aufgenommen und von dort der Erinnerung an Situationen ‚zugefüttert‘ (Roth, 2009).

(Roth, 2009, S. 62). Dem schließt sich die bewusste Organisation bzw. Kontrolle der kognitiven Leistung, des deklarativen Gedächtnisses und der emotionalen Konditionierung an. Neuromodulatorische Systeme steuern Prozesse von allgemeiner und gezielter Aufmerksamkeit bis hin zu Motivation und Lernfähigkeit, die durch hirneigene Opiate belohnt werden (Herrmann, 2010). Im Rückschluss bedeutet dies, dass Lernsituationen dem Lernenden attraktiv erscheinen müssen (wobei ‚Attraktivität‘ hochindividuelle Ausprägungen annimmt) und mit der Aussicht auf Erfolg verknüpft sein sollten, damit die Motivation für fortgesetzte Anstrengungen aufrechterhalten wird. In der Regel ergibt sich dies, wenn das neu zu Lernende an bestehende Wissensbestände angeknüpft wird und dem Lernenden von Bedeutung ist: So korreliert die Stärke des emotionalen Zustandes (Aktivierungsgrad des noradrenergen Systems (allgemeine Aufmerksamkeit)), des dopaminergen Systems (Neugier, Belohnungserwartung) und des cholinergen Systems (gezielte Aufmerksamkeit, Konzentration, empfunden als Interesse, Begeisterung) positiv mit der Gedächtnisleistung (Roth, 2009). Roth (2009) zieht daraus Schlüsse für das schulische Lernen: Bereits erworbenes Wissen wird in „Gedächtnisnetzen“ gespeichert, die entstehen, wenn unterschiedliche Lerninhalte miteinander verknüpft und so größere Bedeutungsfelder gebildet werden. Je mehr Wissen in diesen Clustern vorhanden ist, desto höher ist die Anschlussfähigkeit. Umgekehrt bedeutet dies, dass nicht Anschlussfähiges durch diese Strukturen ‚hindurchfällt‘, also nicht oder nur mühsam aufgenommen werden kann, da nirgendwo Verbindungen zu bereits Gelerntem gebildet werden können. Reines Auswendiglernen (hier verstanden als Lernen von Unbedeutendem) führt damit nur zu einer ‚mechanischen‘ Anwendung und steht im Gegensatz zu einem selbstständigen Durchdringen des Lerngegenstands als aktiver Prozess der Bedeutungserzeugung, der einen zeitlich stabileren Lernerfolg verspricht. Da diese Prozesse nicht nur hochkomplex, sondern auch hochindividuell und häufig unbewusst ablaufen, müsste eine Lehrperson die nahezu undenkbare Fähigkeit besitzen, „den Lern- und Gedächtnisstil eines jeden seiner Schüler genau [zu] kennen [...], um seine Tätigkeit daran optimal anzupassen“ (Roth, 2009, S. 64). Eine Brücke zu bereits erprobten und durchführbaren Lehrmethoden schlagen Herrmann (2010) und Roth (2009) dennoch: Formen der inneren Differenzierung, verstanden als „die Beachtung und Nutzung des trivialen Sachverhalts, dass Lernen und Verstehen per se intraneuronal individuell und selbstorganisiert sind“ (Herrmann, 2010, S. 156) gäben den Lernenden die Möglichkeit, nach ihren eigenen Neigungen, Kompetenzen, Potenzialen und Interessen sich mit individuell adäquaten Lernformen eigenständig einen

Lerngegenstand anzueignen. Aus neurowissenschaftlicher Perspektive proklamieren die Autoren damit den Übergang von einem angebotsorientierten, lehrerstrukturierten Unterricht zu einem nachfrageorientierten Selbstlernen der Schülerinnen und Schüler (Herrmann, 2010). Für die Rolle der Lehrperson würde dies einen Umschwung bedeuten: Statt Wissensvermittlung stünde dann verstärkt die individuelle Begleitung des Lernprozesses der Schülerinnen und Schüler und hier insbesondere die Vermittlung von ‚weichen‘ Kompetenzen wie Lerntechniken und Strategien der Selbstorganisation im Fokus.

Aus den unterschiedlichen Feldzugängen wird deutlich, dass Lernen ein hochindividueller Prozess ist, der nicht durch reine Wissensvermittlung, sondern erst durch die aktive und für den Lernenden bedeutsame Auseinandersetzung mit dem Lerngegenstand initiiert werden kann. Zudem ist der Wissenserwerb abhängig von dem Vorwissen, den bisher gewonnenen Einsichten und Überzeugungen und ihrer (emotionalen) Bewertung des Lernenden – ebenso wie von strukturellen äußeren Faktoren wie etwa der Lernumgebung, des Lernmaterials und ähnlichem. Es scheint daher unabdingbar, schulische Lehrangebote (unter Berücksichtigung des Bildungs- und Erziehungsauftrags der Grundschule) so zu gestalten, dass sie anschlussfähig an diese unterschiedlichen Lernvoraussetzungen sind – und sich nicht an einem fiktiven Durchschnittsschüler (Weinert, 1997) orientieren und so die tatsächliche (vorrangig lernbezogene) Heterogenität innerhalb der Klasse unberücksichtigt lassen. Möglichkeiten der Adaption des Unterrichts an den individuellen Lerner finden sich in den unterschiedlichen Formen der Differenzierung, die im Folgekapitel 3 eingehender erläutert werden sollen.

### **3. Differenzierung im Unterricht**

Wurde in Kapitel 2 gezeigt, dass die – ursprünglich als organisatorische Erleichterung gedachte – Homogenisierung die lern- und leistungsbezogene Vielfalt in Klassen nicht einzuschränken vermag, sollen in diesem Kapitel Möglichkeiten des Umgangs mit ihr beschrieben werden. Weinert (1997) unterscheidet vier prinzipielle Reaktionsmöglichkeiten auf Heterogenität seitens der Lehrperson:

- (1) Das (bewusste oder unbewusste) Ausrichten des Unterrichts an einem (fiktiven) Durchschnittsschüler ignoriert Lern- und Leistungsunterschiede zwischen den Schülerinnen und Schülern einer Klasse (passive Reaktionsform). Dies hat unter anderem zur Folge, dass gute Lernende besser, schlechte hingegen schlechter würden,

was insbesondere für die Formen des offenen Unterrichts gelte, vor allem dann, wenn sich der Lehrende lediglich als Moderator versteht: Erfolgt das Lernen dann autonom, sei „der individuelle Lernfortschritt eine direkte Funktion der persönlichen Lernvoraussetzungen“ (Weinert, 1997, S. 52).

- (2) Als substitutive Reaktionsform gilt die Anpassung des Schülers an die Anforderungen des Unterrichts. Gemeint sind hier zum einen die oben bereits beschriebenen (schulorganisatorischen) Maßnahmen der Homogenisierung, aber auch psychologische Programme zur Motivationssteigerung, Intelligenzförderung, Gedächtnisschulung, Lernstrategien etc. Dabei bliebe ihr Erfolg zur leistungswirksamen Verbesserung der Lernvoraussetzungen schwacher Schülerinnen und Schüler weitestgehend aus (Weinert, 1997). Reduziert werden könnten die Leistungsdifferenzen zwischen den Schulkindern eher mit der „systematische[n] Verbesserung der lernrelevanten Vorkenntnisse, [dem] gezielte[n] Schließen von Wissenslücken, [durch] die damit verbundenen Möglichkeiten der Vermittlung wirksamer Lernstrategien [...] und [durch] die Beeinflussung der Lernmotivation (durch attraktive Lernanreize, durch differenzielle Bekräftigungen und durch ein angstfreies, stimulierendes und aufgabenorientiertes Klassenklima)“ (Weinert, 1997, S. 52).
- (3) Als aktive Reaktionsform beschreibt Weinert die Anpassung des Unterrichts an die lernrelevanten Unterschiede zwischen den Schülerinnen und Schülern und meint damit vorrangig das Konzept des adaptiven Unterrichts als Überwindung der illusionären Hoffnung, durch eine einheitliche Lehrstrategie den gleichen Lernerfolg bei allen Schulkindern zu erzielen. Adaptiver Unterricht könne mithilfe differenzieller Lehrstrategien das Optimum an Lernfortschritten bei möglichst vielen Schülern bewirken und so auch leistungsschwächeren Lernern die Überzeugung der eigenen Selbstwirksamkeit vermitteln.
- (4) Die proaktive Reaktionsform baut auf dem Konzept der adaptiven Gestaltung des Unterrichts als gezielte Förderung der einzelnen Schülerinnen und Schüler auf. Als zentrale Anforderung an die Lehrperson gilt dabei das (rechtzeitige) Diagnostizieren und Interpretieren der Lernmöglichkeiten und -grenzen der Lernenden, da diese in pädagogisch günstigen Settings „mehr an Wissen und Können erwerben, als ihnen oft vor-schnell zugetraut wird“ (Weinert, 1997, S. 52). Neben einem adaptiven Lehrstil seien dabei vor allem differenzielle Lernziele Voraussetzung, zu denen etwa die Unter-

scheidung eines Basiscurriculums mit fundamentalen Lernzielen, die durch alle Schülerinnen und Schüler erreicht werden sollen, und eines Aufbaucurriculums, welchem Lernende nach ihren individuellen Lernvoraussetzungen und Interessen folgen können und welches ihnen möglichst umfangreiche geistige Entfaltungsmöglichkeiten bietet, gezählt werden kann.

Während im deutschen Schulsystem bislang vorrangig die ersten beiden Reaktionsformen vorherrschten, findet sich in den neueren Diskussionen um Chancengerechtigkeit, Heterogenität und Schulerfolg ein Umdenken hin zu einer besseren Passung von Lernausgangslagen und Lernangeboten, die sich in der vermehrten Forderung nach differenziellen Lehrformen ausdrückt (vgl. etwa Altrichter et al., 2009; Helmke, 2003; Meyer, 2005). Dabei ist das, was unter Differenzierung verstanden wird und wie ihre Umsetzung im Unterricht gelingen kann, nicht immer einheitlich.

Konsens besteht zunächst jedoch darüber, dass zwischen innerer und äußerer Differenzierung unterschieden werden kann. Letztere bezieht sich dabei auf die dem Unterricht übergeordnete, kriterienbezogene Bildung von Lerngruppen. Dabei werden die Schülerinnen und Schüler „in unterschiedliche, meist über längere und durch organisatorische Rhythmen vorgegebene Zeiträume in stabile Teilgruppen geteilt“ (Altrichter et al., 2009, S. 343). Hierbei kann es sich um schulorganisatorische Maßnahmen innerhalb einer Einzelschule (Jahrgangsklassen, Leistungsgruppen, Wahlpflichtangebote, Profilkurse etc.) oder um überschulische Maßnahmen (leistungsbezogene Mehrgliedrigkeit des Schulsystems, spezifische Schulabschlüsse, Sonderschulen etc.) handeln (vgl. Altrichter et al., 2009). Die innere Differenzierung hingegen bezieht sich auf didaktische Maßnahmen innerhalb bestehender Klassen bzw. Lerngruppen, die meist temporär sind und häufig das Ziel verfolgen, den einzelnen Schülerinnen und Schülern unterschiedliche Lernwege anzubieten (Altrichter et al., 2009). Häufig wird die innere Differenzierung daher auch als Binnendifferenzierung bezeichnet. Ein Modell, das die vielschichtigen Möglichkeiten der inneren Differenzierung aufzeigt, findet sich bei Klafki (2007, vgl. Abbildung 3). Dieser unterscheidet drei Hauptdimensionen, aus denen sich ein räumliches Raster ableiten lässt.

**Abbildung 3:** Dimensionen- und Kriterienraster (Ordnungs- und Suchraster) zur Inneren Differenzierung

A. Unterrichtsphasen	B. Differenzierungsaspekte						C. Aneignungs- bzw. Handlungsebenen		
	1. Stoffumfang/ Zeitaufwand	2. Komplexitäts- grad	3. Anzahl der notwendigen Durchgänge	4. Notwendigkeit direkter Hilfe/ Grad der Selbst- ständigkeit	5. Art der inhaltl. od. method. Zugänge/der Vorerfahrungen	6. Kooperations- fähigkeit	a) konkrete Aneignungs- bzw. Hand- lungsebene	b) explizit-sprach- liche Aneignungs- bzw. Handlungs- ebene	c) rein gedank- liche Aneignungs- bzw. Handlungs- ebene
I. Aufgaben- stellung, -entwicklung									
II. Erarbeitung									
III. Festlegung									
IV. Anwendung/ Transfer									

Quelle: Klafki, 2007, S. 188, adaptiert

Die Dimension A bezeichnet dabei die Phasen des Unterrichtsprozesses und folgt der Fragestellung „In welcher Phase des Unterrichts will ich oder kann ich Innere Differenzierung durchführen, bzw.: In welcher Phase haben Schüler oder Schülergruppen den Anschluß verloren?“ (Klafki, 2007, S. 187). Zu ihnen gehören (I) die Aufgabenstellung bzw. -entwicklung, (II) die Erarbeitung, (III) die Festigung und (IV) die Anwendung bzw. der Transfer.

In Dimension B werden schülerspezifische Differenzierungsaspekte zusammengefasst. Die konstituierende Fragestellung lautet „Unter welchen Gesichtspunkten kann es sinnvoll sein oder ist es notwendig, Schüler differenziert anzusprechen? Welches sind die spezifischen Möglichkeiten und Stärken bestimmter Schüler oder Schülergruppen oder aber ihre ‚Schwachstellen‘, die ihren weiteren Lernprozeß entscheidend behindern können?“ (ebd.). Zu den Differenzierungsaspekten gehören (1) der Stoffumfang bzw. der Zeitaufwand, (2) der Komplexitätsgrad, (3) die Anzahl der notwendigen Durchgänge, (4) die Notwendigkeit direkter Hilfe bzw. der Grad der Selbstständigkeit, (5) die Art der inhaltlichen oder methodischen Zugänge bzw. der Vorerfahrungen und (6) die Kooperationsfähigkeit.

Dimension C beinhaltet schließlich die unterschiedlichen Aneignungs- und Handlungsebenen, auf denen der Lerngegenstand zu eigen gemacht werden kann. Zu diesen zählen (a) die konkrete (dingliche Repräsentation), (b) die explizit-sprachliche (verlautbare Sprache, z.B. Analysieren, Begründen, Erörtern) und (c) die rein gedankliche (abstrakte) Aneignungs- bzw. Handlungsebene.

Die einzelnen ‚Bausteine‘ des räumlichen Rasters, welches sich durch die drei Dimensionen aufspannt, ergänzt Klafki durch unterrichtsnahe Beispiele, deren Ausführung hier allerdings zu weit führen würde. Deutlich wird hieraus jedoch die methodische Variationsbreite, mit der eine innere Differenzierung verwirklicht werden kann. Dabei sei explizit angemerkt, dass sich diese Binnendifferenzierung nicht an den jeweiligen Lernvoraussetzungen und Interessen des einzelnen Schülers bzw. der einzelnen Schülerin orientieren muss, sondern in erster Linie eine merkmalsbezogene Gruppierung innerhalb einer Lerngruppe bzw. Klasse darstellt (vgl. bspw. Bohl, Batzel & Richey, 2011). An dieser Stelle wird häufig zwischen Binnendifferenzierung und individueller Förderung unterschieden: So sei es die individuelle Förderung, welche explizit die individuellen Interessen, Neigungen, Vorkenntnisse, den Leistungsstand und anderen Lernbedingungen der Schülerinnen und Schüler berücksichtige. Begegnet werden kann diesen mit Mitteln der Binnendifferenzierung, die jedoch auch derlei Gestalt annehmen können, dass jede Schülerin und jeder Schüler an unterschiedlichen Lernmaterialien arbeitet. In einer Definition von Bräu (2007) wird dieses Verständnis von innerer Differenzierung und individueller Förderung deutlich. So heißt es dort zunächst, dass unter individualisierendem Unterricht „ein Unterricht verstanden werden [soll], in dem innere Differenzierung praktiziert wird, also Arrangements, in denen nicht alle gleichzeitig das Gleiche lernen sollen und müssen“ (Bräu, 2007, S. 173). Besonders im Fokus stehe in solch einem Unterricht, dass „die Schülerinnen und Schüler je nach Voraussetzungen, Möglichkeiten und Neigungen unterschiedliche Aufgaben bearbeiten, verschiedene Lernwege verfolgen oder Themen bearbeiten [können], manchmal auch verbunden mit unterschiedlichen Lehrzielen (ziendifferenter Unterricht). Im Extremfall arbeitet jedes Kind einer Klasse an einer eigenen Thematik, an einer eigenen Aufgabe oder an Aufgaben, die individuelle Ausprägungen und Bearbeitungsmodi erlauben. Ziel ist – wie in jedem Unterricht – das zumeist fachlich gebundene, inhaltliche Lernen, in das die individuelle Förderung jedes Einzelnen und gleichzeitig die Schulung selbstständigen Lernens eingebunden sind“ (ebd.). Individuelle Förderung wird dabei also als eine Unterform der Binnendifferenzierung verstanden, bei der differenzielle Materialien, Aufgabenstellungen und anderes mit dem Ziel eingesetzt werden, den unterschiedlichen Leistungsniveaus und Bedürfnissen der einzelnen Schülerinnen und Schüler gerecht zu werden.

Lipowsky (2015) unterscheidet Binnendifferenzierung und individuelle Förderung anhand eines anderen Merkmals: „Während die Lehrperson in Phasen innerer Differenzierung

*Gruppen von Schülern* unterschiedliche Aufgaben, unterschiedliche Aufgabenmengen und/oder unterschiedliche Lernzeitkontingente zuweist und/oder ihnen unterschiedliche Unterstützungsangebote gewährt, verfolgen Maßnahmen der Individualisierung den Anspruch, die Lernangebote und -bedingungen an die Voraussetzungen *einzelner Schüler* anzupassen“ (Lipowsky, 2015, S. 93, Herv. i. O.). Beide Differenzierungsmaßnahmen würden dabei „gleichermaßen das Ziel [verfolgen], Schüler mit unterschiedlichen Lernvoraussetzungen in ihrem Lernprozess wirkungsvoll zu fördern und zu unterstützen und dazu beizutragen, dass alle Schüler bestimmte Mindestziele erreichen“ (ebd.). Demnach unterscheiden sich beide Differenzierungsformen nur in der Größe der Zielgruppe (Einzelschüler oder Schülergruppen), eine Adaption an ihre Lernvoraussetzungen findet in beiden Maßnahmen statt.

Altrichter et al. (2009) fokussieren bei der Unterscheidung von innerer Differenzierung und individueller Förderung eher die *Planung der Passung* des Unterrichtsangebots an die Lernvoraussetzungen der Schülerinnen und Schüler: So werden bei der Binnendifferenzierung „durch verschiedene didaktische Arrangements (z.B. differenzierte Aufgabenstellungen, unterschiedliche Arbeitsformen, Projekte) temporäre Lerngruppen für bestimmte Lernaufgaben gebildet. Dadurch sollen Schüler/inne/n – individuell oder in Gruppen – unterschiedliche Lernwege angeboten werden [...]. Auch der Begriff ‚Individualisierung‘ zielt auf die bestmögliche Passung von Unterrichtsangebot und Schülervoraussetzungen. Anders als die lehrer- und gruppenbezogene Sichtweise von ‚innerer Differenzierung‘ denkt ‚Individualisierung‘ stärker von den Lernvoraussetzungen, -wegen und -zielen individueller Schüler/innen her. Individualisierter Unterricht ist dabei nicht unbedingt ‚Einzelunterricht‘. Sein Charakteristikum ist vielmehr, Schüler/inne/n durch Aufgabenstellungen und flexible Unterrichtsmethoden solche Lernwege und Lernziele zu ermöglichen, die ihren individuellen Voraussetzungen in Hinblick auf Leistungsvermögen, Interessen usw. gut entsprechen, sie durch diese ‚Passung‘ zu optimaler Ausschöpfung ihrer Lernpotenziale motivieren und sie dabei auf ihrem Lernweg zu unterstützen“ (Altrichter et al., 2009, S. 346–347, Herv. i. O.). Diesem Verständnis nach stellt die Binnendifferenzierung also ein Konzept dar, bei welchem den Schülerinnen und Schülern unterschiedliche Lernwege angeboten werden, die sich an ihren individuellen Lernvoraussetzungen orientieren *können*. Letzteres könnte sich zum Beispiel bereits daraus ergeben, dass Schülerinnen und Schüler aus einer größeren Vielfalt an Lernmethoden die für sie am ehesten geeignete auswählen können, wenngleich diese nicht explizit für den individuellen Schüler zur Verfügung gestellt wurde. Anders bei der individuellen Förderung, hier-

unter verstehen Altrichter et al. (2009) eine bewusste Vorbereitung und Anpassung des Lernmaterials an die Bedürfnisse *einzelner* Schülerinnen und Schüler.

Alle drei Definitionen spiegeln das uneinheitliche Bild des Verständnisses von individueller Förderung in Abgrenzung zu der Binnendifferenzierung wider. Ihnen allen gemein jedoch ist, dass individuelle Förderung stets mit einer Anpassung der Lernmethoden und des Materials an die Lernvoraussetzungen, Bedürfnisse und Interessen *jedes einzelnen Schülers* einhergeht, während dies nicht explizites Ziel der internen Differenzierung sein muss: Hier kann zwar eine hohe Passung der Bedürfnisse und Lernangebote erreicht werden, allerdings beinhaltet dies einen größeren Zufallsmoment, da die Lernmethoden und Materialien nicht explizit für *einzelne* Schülerinnen und Schüler sondern eher für Schülergruppen unterschiedlicher Leistungsniveaus, unterschiedlicher Lerntypen und ähnlichem vorbereitet werden.

Was bedeutet dies für die folgenden Ausarbeitungen? In den oben dargestellten Definitionen zeigt sich zunächst, dass die Anpassung des Lehr- und Lernmaterials und der Unterrichtsmethoden an die Voraussetzungen, Bedürfnisse und Interessen der einzelnen Schülerinnen und Schüler zwar in jedem Fall der individuellen Förderung, nicht immer jedoch der Binnendifferenzierung zugeordnet wird. Es muss also angenommen werden, dass der inneren Differenzierung unterschiedliche Verständnisse zugrunde liegen. Dies ist gerade im Hinblick auf die Operationalisierung bei (quantitativen) Befragungen von Belang: So könnte eine unterschiedliche Auffassung darüber, was individuelle Förderung und Binnendifferenzierung ausmacht, zu höchst unterschiedlichen Angaben etwa von Lehrpersonen führen, die dann entsprechend nicht mehr als reliabel einzuschätzen wären. In einem Vergleich mit den Maßnahmen der inneren und äußeren Differenzierung kann zudem angenommen werden, dass individuelle Förderung die am ehesten geeignete didaktische Herangehensweise darstellt, mit der der Heterogenität im Klassenzimmer (s. Kapitel 2) in angemessenem Umfang begegnet werden kann. Somit steht diese folgend im Fokus der Betrachtungen.

#### **4. Individuelle Förderung: Voraussetzungen und Umsetzung**

Wie in dem vorangegangenen Kapitel gezeigt werden konnte, werden grundsätzlich drei Maßnahmen unterschieden, mit denen den heterogenen lern- und leistungsbezogenen Merkmalen der Kinder und Jugendlichen im deutschen Bildungssystem begegnet wird.

Neben den Formen der inneren und äußeren Differenzierung ist es vor allem die individuelle Förderung, durch welche Unterrichtsmethoden und -materialien an die unterschiedlichen Voraussetzungen und Interessen der Schülerinnen und Schüler angepasst werden, wodurch die größtmögliche Berücksichtigung der im Klassenzimmer vorherrschenden Vielfalt in Begabungen, Bedürfnissen und Vorlieben erreicht werden kann.

Für die Umsetzung individueller Förderung in der pädagogischen Praxis lässt sich bisweilen jedoch kein „theoretisch und empirisch plausible[s] Modell wesentlicher Elemente des Phänomens ‚Individualisierung‘ und seiner wichtigsten Rahmenbedingungen“ (Altrichter et al., 2009, S. 359, Herv. i. O.) finden (vgl. z.B. auch Bräu, 2007 und Fischer, 2014), so dass ihre Realisierung unterschiedliche Formen annehmen kann. Klieme und Warwas (2011) unterscheiden drei<sup>13</sup> grundsätzliche Varianten des pädagogischen Verständnisses individueller Förderung:

- (1) **Kompensatorische Trainings- und Zusatzangebote:** Auf Grundlage einer ausführlichen Differentialdiagnostik können Pädagogen und Psychologen kompensatorische Trainings- und Zusatzangebote indizieren, die einzeln oder in Kleingruppen durchgeführt werden. Hierzu zählen vor allem unterschiedliche, dem jeweiligen Fördergegenstand angepasste Formen der Instruktion, kooperative Lernformen sowie Problemlöse- und Lernstrategietrainings. In Anlehnung an Kroesbergen und Van Luit (2003) betonen Klieme und Warwas, dass sich die schulfachlichen Fähigkeiten „durch Trainings dann verbessern lassen, wenn die Maßnahmen von der Lehrkraft strategisch ausgerichtet, feedbackgebend, gut strukturiert und übungsbetont durchgeführt werden“ (Klieme & Warwas, 2011, S. 808).
- (2) **Vielfältige Lernwege durch offenen Unterricht:** Dem reformpädagogischen Verständnis individueller Förderung entlehnt sind die Formen offenen Unterrichts wie etwa Freiarbeit, Projektarbeit, Lernen nach Jahres- und Wochenplänen sowie Stationenlernen<sup>14</sup>. Mit diesen sollen den Schülerinnen und Schülern eine Vielzahl von Lernarrangements, -wegen und -methoden angeboten sowie ihre lernbezogene

---

<sup>13</sup> Eine weitere didaktische Herangehensweise, die häufig in Verbindung mit individueller Förderung gebracht wird, ist das Mastery Learning nach Bloom (1968, 1971), bei welchem der Lernstoff in kleinschrittige Lerneinheiten aufgeteilt wird, die jeder Schüler bewältigen (daher „Mastery“ aus dem Englischen „to master“) soll, bevor zu Neuem übergegangen wird. Dies beinhaltet jedoch nicht die Komponente der *Planung der Passung* der Unterrichtsangebote an die Voraussetzungen und Neigungen der Lernenden, wie sie Kernelement des hier dargestellten Verständnisses individueller Förderung ist.

<sup>14</sup> Erläuterungen dieser Formen offenen Unterrichts finden sich z.B. bei Haag und Streber (2014).

Selbstständigkeit gefördert werden. Eine Anpassung der Lerngelegenheiten an die individuellen Voraussetzungen und Neigungen der Lernenden ergibt sich dabei vor allem durch „die Auswahl, zeitliche Anordnung und Dauer der Lerntätigkeiten“ (Klieme & Warwas, 2011, S. 809), bei denen die Schülerinnen und Schüler Unterstützung und Beratung durch die Lehrkraft erfahren. Kritisiert werden muss jedoch, dass zum einen der Grad der Offenheit des Unterrichts häufig überschätzt wird, zum anderen eine Passung zwischen Unterrichtsangebot und Lernvoraussetzungen nicht zwangsläufig gegeben sein muss. Dementsprechend gehe eine Öffnung des Unterrichts nicht automatisch mit einer individuellen Förderung einher (Klieme & Warwas, 2011, in Anlehnung an Brügelmann, 2000; Hanke, 2001 und Graumann, 2008). Zudem würde offener Unterricht vorrangig das Selbstbestimmungserleben der Schülerinnen und Schüler stärken, nicht notwendigerweise jedoch ihr fachliches Verständnis (Hartinger, 2005; Bennett, 1976 und weitere, s. Klieme & Warwas, 2011). Bezugnehmend auf Hardy, Jonen, Möller und Stern (2006) sowie Lipowsky (2002) wirke sich offener Unterricht jedoch dann lernförderlich aus, wenn dieser durch die Lehrperson so strukturiert würde, dass „Schüler neue Inhalte mit bestehendem Wissen verknüpfen und ihre Konzepte erweitern und umstrukturieren können, dass es letztlich auf die Mikrostrukturen der Lernangebote ankommt“ (Klieme & Warwas, 2011, S. 809).

- (3) **Binnendifferenzierung durch adaptiven Unterricht:** Der Begriff des ‚adaptiven Unterrichts‘ wurde wesentlich durch die ATI-Forschung geprägt, die in Kapitel 2 bereits erläutert wurde. Im Gegensatz zu dem offenen Unterricht ist hierbei eine gezielte Steuerung durch die Lehrperson vorgesehen, die aus verschiedenen Instruktionen und Lerngelegenheiten diejenige auswählt, die am ehesten auf die (diagnostizierten) Lernvoraussetzungen eines Schülers bzw. einer Schülerin zugeschnitten ist. Explizites Ziel ist es dabei also, die Lernumgebung an die individuellen Unterschiede zwischen den Lernenden anzupassen, was durch Makro- (Curriculum etc.) und Mikro-Adaption (innerhalb der Lehrer-Schüler-Interaktion) gelingt (vgl. insb. Corno & Snow, 1986 und Kapitel 2 hier).

Aus dieser Zusammenschau unterschiedlicher Verständnisse dessen, wie individuelle Förderung im Unterricht realisiert werden kann, wird vor allem deutlich, dass ihre Umsetzung keinem einheitlichen Konzept folgt. Wie oben bereits angemerkt, fehlt ein solches Unter-

richtskonzept bisweilen. Allerdings lassen sich Elemente identifizieren, die in unterschiedlichen Ausprägungen in individualisierenden Unterrichtssettings auftreten (vgl. hierfür Altrichter et al., 2009, auch Weinert, 1997 und Hofmann, 2000). Hierzu gehören vor allem *Lern-diagnosen*, anhand derer die Stärken und Defizite der Schülerinnen und Schüler in bestimmten Teilbereichen eruiert werden können. Diese ermöglichen die Bereitstellung von *Lernaufgaben*, die eine möglichst hohe Passung von individuellen Lernvoraussetzungen und Lernzielen aufweisen. In Anlehnung an Herber (1983) differenzieren Altrichter et al. (2009) zudem *unterschiedliche Lernziele* im individualisierenden Unterricht. Im ‚Fundamentum‘ seien Lernziele festgehalten, die grundlegende Konzepte des jeweiligen Lerngegenstands vermitteln und entsprechend für alle Lernenden gelten, das ‚Additum‘ definiere „Zusatzaufgaben, interessante Anwendungen, Vertiefungen und Übertragungen für Schüler/innen, die z.B. von einer Erarbeitung im Durchschnittstempo unter- oder überfordert sind oder deren spezielle Übungsmöglichkeiten angeboten werden sollen“ (Altrichter et al., 2009, S. 348). Es seien also gerade diese vertiefenden Aufgaben, die eine Anpassung an die Begabungen, Lernstände und Interessen der Schülerinnen und Schüler ermöglichen. Unterstützung erfahren diese dabei durch die Lehrperson, die durch *Lernberatung* und *Instruktion* den Kindern und Jugendlichen bei der Erreichung der individuellen Lernziele behilflich ist und *Rückmeldungen* gibt, welche die Schülerinnen und Schüler für das weitere Lernen motivieren. Eingebettet ist dies in *flexible Unterrichtsformen*, die das Lernen in individuellen wie gruppenbezogenen Settings zulassen. Die Formen der *Leistungsbeurteilung* und *Leistungsdokumentation* müssen ebenso für einen individualisierenden Unterricht geeignet sein wie die *didaktischen Materialien* und – häufig unbedacht – die *Räumlichkeiten*. Auch geht individuelle Förderung mit spezifischen *Kompetenzen bei Lehrenden und Lernenden* einher, auf die im Folgenden vertiefend eingegangen werden soll.

Die unter ‚adaptiven Lehrkompetenzen‘ erfassten Fähigkeiten von Lehrerinnen und Lehrern, die zu einer gelungenen individuellen Förderung beitragen, lassen sich aus den oben dargestellten Elementen individualisierenden Unterrichts ableiten. Die Unterscheidung von vier Kernkompetenzen nach Beck et al. (2008) gehört dabei zu den am häufigsten zitierten. Zu diesen gehören ein

- „reichhaltiges, flexibel nutzbares eigenes Sachwissen, in dem sich die Lehrperson leicht und rasch geistig bewegen kann (Sachkompetenz),

- die Fähigkeit, bezogen auf den jeweiligen Unterrichtsgegenstand die Lernenden bezüglich ihrer Lernvoraussetzungen und -bedingungen (Vorwissen, Lernweisen, Lerntempo, Lernschwächen usw.) sowie ihrer Lernergebnisse zutreffend einschätzen zu können (diagnostische Kompetenz);
- reichhaltiges, methodisch-didaktisches Wissen und Können, wozu auch gehört, dass die Lehrperson die Vor- und Nachteile der einsetzbaren didaktischen Möglichkeiten und die Bedingungen kennt, unter denen diese Erfolg versprechend eingesetzt werden können (didaktische Kompetenz) sowie
- die Fähigkeit, eine Klasse so zu führen, dass sich die Lernenden – als Grundvoraussetzung für Lernfortschritt und Lernerfolg – aktiv, anhaltend und ohne ein Zuviel an störenden Nebenaktivitäten (hohe time-on-task-Werte) mit dem Unterrichtsgegenstand auseinandersetzen können (Klassenführungs-kompetenz)“ (Beck et al., 2008, S. 41)

Dabei konnten die Autoren in einem dreijährigen Forschungsprojekt mit 50 Schulklassen in Primar- und Sekundarstufe I zeigen, dass eine hohe adaptive Lehrkompetenz (gefördert durch eine Intervention) zu einem signifikant größeren Lernzuwachs bei den Schülerinnen und Schülern – verglichen mit der Kontrollgruppe Lehrender mit geringerer adaptiver Lehrkompetenz – führt. Dabei fielen die Lernzuwächse in leistungsheterogenen Gruppen sogar stärker aus als in leistungshomogenen Gruppen. Zudem stellen Beck et al. (2008) fest, dass adaptive Lehrkompetenz durch gezielte Fortbildungen deutlich steigerbar ist. Dies ist zwischenzeitlich auch in die ‚Standards für die Lehrerbildung‘ (vgl. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2014) aufgenommen worden. Auch ist die individuelle Förderung mittlerweile beständiger Teil von Darstellungen der Merkmale guten Unterrichts (z.B. bei Meyer, 2005; Helmke, 2006 und Terhart, 2006, auf eine ausführlichere Darstellung soll hier verzichtet werden).

Neben diesen lehrbezogenen Fähigkeiten bedarf es jedoch mitunter auch Kompetenzen bei den Schülerinnen und Schülern, welche für ein Gelingen individueller Förderung maßgeblich sein können. Zu diesen zählt vor allem eine lernbezogene Selbstregulation, welche als „ein aktiver, konstruktiver Prozess [verstanden wird], bei dem der Lernende sich Ziele für sein Lernen selbst setzt und zudem seine Kognitionen, seine Motivation und sein Verhalten in Abhängigkeit von diesen Zielen und den gegebenen äußeren Umständen beobachtet, reguliert und kontrolliert“ (Pintrich, 2000, S. 453). Aus dieser Definition wird deutlich, dass effektives

Lernen von mehreren Variablen abhängt, zu denen vor allem eine kognitive Dimension (konzeptionelles, strategisches Wissen und die Fähigkeit, kognitive Lernstrategien anzuwenden), eine motivationale Komponente (Aktivitäten, die die Initiierung und Aufrechterhaltung des Lernens unterstützen; handlungsförderliche Attributionen, Selbstwirksamkeitsüberzeugungen) und metakognitive Facetten (Planung, Selbstbeobachtung, Reflexion des eigenen Lernprozesses) gehören (vgl. Otto, Perels & Schmitz, 2011). Klieme und Warwas (2011) und Boekaerts (1999) beschreiben das Ineinandergreifen der unterschiedlichen Dimensionen. So werden auf kognitiver Ebene unterschiedliche Lernstrategien gewählt, die durch Prozesse auf der übergeordneten metakognitiven Ebene überwacht und reguliert werden, während auf der motivationalen Ebene die individuellen Bedürfnisse der Schülerinnen und Schüler Berücksichtigung finden. Dabei erlaubt das Erlernen der Selbstregulation, Lern- und Verstehensprozesse zunehmend unabhängig von Lehrenden und damit selbstständig zu vollziehen. Klieme und Warwas (2011) sowie Lipowsky (2002) bezeichnen dies als wichtiges Ziel individueller Förderung, zumal davon auszugehen sei, dass „individuelle Fördermaßnahmen von solchen Schülern besonders gut aufgenommen werden, die bereits über bestimmte selbstregulatorische Fähigkeiten verfügen“ (Klieme & Warwas, 2011, S. 813).

Es zeigt sich also, dass individuelle Förderung als didaktische Gestaltung des Unterrichts mit Kompetenzen sowohl auf Seiten der Lehrkraft als auch auf Seiten der Lernenden einhergeht – letzteres scheint gerade für die hier interessierende Kohorte von Viertklässlerinnen und Viertklässlern nicht ganz trivial, kann jedoch auf Grundlage der vorliegenden PIRLS-Daten nicht in die Analysen aufgenommen werden. Dies wird noch einmal in der Diskussion des Modells und der Befunde aufgegriffen. In welcher Form individuelle Förderung bisher in Deutschland angewendet wird, ist Gegenstand des folgenden Kapitels.

## **5. Individuelle Förderung: Implementierung in Deutschland**

Gerade im Zuge der Inklusions- und Heterogenitätsdebatte in Deutschland tritt individuelle Förderung immer stärker in den Fokus bildungspolitischer Entscheidungsträger. So fand sie etwa seit 2005 Eingang in fast alle Schulgesetze der sechzehn Bundesländer, wenngleich mit unterschiedlichem inhaltlichen Verständnis und unterschiedlichen Absichten. Fischer (2014) systematisiert die Gesetzestexte und kommt zu dem Schluss, dass individuelle Förderung fast ausschließlich für Schüler (1) mit Entwicklungsverzögerungen oder Behinderungen (15 Bundesländer), (2) für Schüler mit besonderen Begabungen (12 Bundesländer) und

(3) für Schüler, deren Muttersprache nicht Deutsch ist und die entsprechend Bedarf im Erwerb der deutschen Sprache haben (8 Bundesländer), Anwendung finden soll. Darüber hinaus werde die schulpraktische Realisierung der individuellen Förderung in fast allen Bundesländern durch sonderpädagogische Maßnahmen proklamiert. Dies spiegelt auch die Einschätzungen der Lehrkräfte wider, die in einer Untersuchung von Kunze und Solzbacher (2008) für die Sekundarstufe I aller Schulformen deutlich werden: Zwar wird individuelle Förderung durchweg als wichtiges und anzustrebendes Ziel betrachtet, allerdings beinhaltet dies vor allem eine defizitäre Perspektive der Lehrkräfte auf den Lernstand des einzelnen Schülers bzw. der einzelnen Schülerin, die darauf abzielt, Stärken aus- und Schwächen abzubauen. Selten findet sich ein Verständnis individueller Förderung als Berücksichtigung und Anpassung des Unterrichts an die individuellen Neigungen des Einzelschülers (vgl. Solzbacher, 2008). Bezüglich der Umsetzung individueller Förderung geben die meisten Lehrenden Instrumente des Förderunterrichts an oder berichten, dass sehr wenige bis gar keine Maßnahmen, Instrumente oder Konzepte individueller Förderung systematisch verankert seien. Damit einher geht eine große Unsicherheit der Lehrenden, entsprechende Maßnahmen im Unterricht zu ergreifen. So finden sich maximal Angaben, ‚gelegentlich‘ individuelle Förderung anzuwenden, die auch eher ‚nur einmal ausprobiert wird‘ und zudem keinem vorgegebenen Konzept folgt. Wischer (2008) weist darauf hin, dass Differenzierungsmaßnahmen vorrangig dazu genutzt werden, den Unterricht methodisch abwechslungsreich und schüleraktivierend zu gestalten, die Umsetzung organisatorisch komplexerer Formen der Anpassung des Unterrichts an die Neigungen, Interessen etc. der Schülerinnen und Schüler jedoch selten realisiert wird. Insgesamt zeichnet sich für Deutschland damit ein verhaltenes Bild der Nutzung individueller Förderung ab: Zwar wird ihre Bedeutung für das schulische Lernen und der Begegnung von Heterogenität im Klassenzimmer anerkannt – dass entsprechende Konzepte, Instrumente und Schulungen der Lehrkräfte bisher nur bedingt implementiert sind, scheint ihre konsequente Umsetzung jedoch zu erschweren. Welche Wirkung individueller Förderung in bisherigen empirischen Arbeiten nachgewiesen wurde, ist Gegenstand des folgenden Kapitels.

## **6. Effekte individueller Förderung**

Für die Darstellung empirischer Forschungsergebnisse zu der Wirkung individueller Förderung ergeben sich zwei Hürden: Zum einen liegen nur wenige empirische Arbeiten vor, zum

anderen unterliegen diese einem sehr heterogenen Verständnis dessen, was individuelle Förderung ausmacht. So finden sich in diesem Kontext ebenfalls Analysen zu der inneren Differenzierung oder zu Formen des offenen Unterrichts. Wie in Kapitel 3 dargestellt, muss es sich dabei nicht um Lehrmethoden handeln, die explizit auf die Begabungen, Leistungsvoraussetzungen und Interessen der Schülerinnen und Schüler eingehen. Für die folgende Darstellung soll sich daher lediglich auf Studien bezogen werden, denen ein solches Verständnis zugrunde liegt. Aufgeführt werden diese zunächst einzeln, damit auf die jeweiligen Verständnisse individueller Förderung und andere Besonderheiten eingegangen werden kann – eine Zusammenfassung gestaltet sich aus ebendiesen Gründen schwierig, soll aber dem Kapitel folgen.

Eine der ersten Studien in diesem Bereich im deutschsprachigen Raum wurde von Helmke und Weinert zwischen 1987 und 1997 durchgeführt (Helmke & Weinert, 1997a, auch Helmke & Schrader, 1998). Bei der SCHOLASTIK-Studie (Schulorganisierte Lernangebote und Sozialisation von Talenten, Interessen und Kompetenzen) wurden in einem Längsschnittdesign die lern- und leistungsbezogenen Entwicklungen von 1.150 Schülerinnen und Schülern (in 54 Klassen in und um München) in Abhängigkeit unterrichtlicher und schulischer Kontextfaktoren analysiert. Dabei wurden Unterrichtsmerkmale durch externe, geschulte Beobachter erhoben und anschließend in korrelative Zusammenhänge mit den Ergebnissen von Mathematik- und Rechtschreibtests gesetzt. Für das Unterrichtsmerkmal ‚Individuelle fachliche Unterstützung‘ ergab sich ein signifikanter Zusammenhang von  $r = 0,32$  ( $p < 0,05$ ) mit den Mathematikleistungen. Der Zusammenhang zwischen Individualisierung und Rechtschreiben zeigte sich nicht signifikant ( $r = 0,16$ ). Demnach stelle die Intensität individueller Förderung, neben anderen Unterrichtsmerkmalen, eine wesentliche Bedingung für die überdurchschnittlichen Lernzuwächse in Mathematik, nicht jedoch im Rechtschreiben dar (vgl. Helmke & Weinert, 1997a). In vertiefenden Analysen konnten Weinert et al. (1989) zudem zeigen, dass dieser Leistungszuwachs unter den Bedingungen individueller Förderung vor allem in Kombination mit einer hohen diagnostischen Kompetenz bei den Lehrenden einherging.

Letztere Befunde bestätigen auch die Ergebnisse der oben bereits erwähnten Schweizer Studie von Beck et al. (2008), die anhand eines elaborierten Experimental-Kontrollgruppendesigns ein durch die adaptive Lehrkompetenz der Lehrenden (Definition s. Kapitel 4)

bedingten Leistungszuwachs bei Primar- und Sekundarschülern nachweisen konnten. Befragt wurden 50 nicht-zufällig ausgewählte Lehrpersonen und ihre Klassen (N = 976 Schülerinnen und Schüler) der vierten, später fünften sowie siebten, später achten Klassenstufe. Dabei wurde die adaptive Handlungs- und Planungskompetenzen aller Lehrenden mittels Videotest und Vignetten in einem Pre- und Posttest erfasst. Die 32 Lehrenden der Treatmentgruppe erhielten zudem eine zweitägige Schulung zu adaptiven Lehrkompetenzen sowie ein unterrichtsbegleitendes fachspezifisch-pädagogisches Coaching. Anhand einer anschließend über zwei Monate durchgeführten, einheitlichen Unterrichtssequenz im Fach Naturwissenschaften, für die alle Lernziele vorab definiert wurden, und einem standardisierten, aus PISA und TIMSS adaptierten Leistungstest zu den naturwissenschaftlichen Kompetenzen der Schülerinnen und Schüler wurde dann der Leistungszuwachs der Lernenden in beiden Domänen erfasst. Dadurch, dass die Unterrichtssequenz in allen Klassen unter vergleichsweise standardisierten Bedingungen umgesetzt wurde, können „Unterschiede im Lernfortschritt der Klasse [...] somit hauptsächlich auf den konkreten Unterricht zurückgeführt werden“ (Beck et al., 2008, S. 108). Dabei zeigte sich ein signifikanter Leistungsvorsprung der Schülerinnen und Schüler, deren Lehrenden eine hohe adaptive Lehrkompetenz nachgewiesen wurde. Die Autoren weisen hier eine mittlere Effektstärke (partielles Eta-Quadrat<sup>15</sup>) von  $\eta^2 = 0,10$  nach. Für den allgemeinen naturwissenschaftlichen Leistungstest, der nicht explizit das im Rahmen der Unterrichtssequenz erworbene Wissen erfasst, wurde solch ein signifikanter Effekt nicht nachgewiesen. Die Autoren führen dies darauf zurück, dass möglicherweise Themen erfragt wurden, die im Unterricht nicht behandelt wurden. Nachgewiesen werden konnte diese Annahme aufgrund hierzu nicht vorliegender Daten nicht. In vertiefenden Analysen wurde untersucht, ob adaptives Unterrichten zudem eine Möglichkeit darstellt, der Heterogenität einer Klasse zu begegnen. Überprüft wurde dabei, ob adaptiv unterrichtende (und entsprechend kompetente) Lehrende größere Lernfortschritte in leistungsheterogenen Klassen erzielen als Lehrpersonen, die weniger adaptive Lehrformen anwenden und weniger qualifiziert sind. Dabei konnte ein starker Effekt<sup>16</sup> von  $d = 0,98$  bezüglich der vorgegebenen Unterrichtsreihe gefunden werden (der damit sogar stärker ausfällt als in leistungshomogenen Gruppen; dort ist der Unterschied nicht signifikant). Allerdings zeigt sich

<sup>15</sup> In Anlehnung an Cohen (1988) lassen sich Effektstärken ab  $\eta^2 = 0,01$  als schwach, ab  $\eta^2 = 0,06$  als mittelstark und ab  $\eta^2 = 0,14$  als stark interpretieren.

<sup>16</sup> Hierbei handelt es sich um das Effektstärkemaß Cohens d, welches nach Cohen (1969) bei Werten von 0,20 als klein, 0,50 als moderat und 0,80 und höher als stark bezeichnet wird.

solch ein Zusammenhang nicht für den Erwerb allgemeiner naturwissenschaftlicher Kompetenzen. Die Autoren führen letzteres darauf zurück, dass über den längeren Zeitraum die Lernleistungen der Schülerinnen und Schüler durch Drittvariablen beeinflusst worden sein könnten, möglicherweise aber auch die „kausale Distanz“ (Beck et al., 2008, S. 129) zwischen Lehrkompetenz, tatsächlichem Unterricht und später gemessenem Lernfortschritt zu groß sei, um solche Effekte nachzuweisen. Diese Befunde sind entsprechend vorsichtig zu interpretieren. Zusammenfassend lässt sich also festhalten, dass adaptives Unterrichten einen positiven Effekt auf den Lernzuwachs von Schülerinnen und Schülern sowohl der Primar- als auch der Sekundarstufe hat und darüber hinaus eine Möglichkeit darstellt, der leistungsbezogenen Heterogenität einer Schulklasse zu begegnen.

Der Implementation eines spezifischen Konzepts für den Leseunterricht, welches Aspekte individueller Förderung beinhaltet, folgen Reis, McCoach, Little, Muller und Kaniskan (2011) mit einem experimentellen Studiendesign: Das ‚Schoolwide Enrichment Model-Reading‘ (SEM-R) wurde an fünf US-amerikanischen Grundschulen in den Klassen zwei bis fünf von insgesamt 63 Lehrenden (1.192 Lernende) in einem randomisierten Kontroll-Treatment-Verfahren eingesetzt und seine Wirksamkeit bezüglich der Lesekompetenz und des Leseverständnisses überprüft. Die Umsetzung des Konzepts folgt dabei drei Stufen: „Phase I focuses on exposing students to books, Phase 2 incorporates differentiated instruction, including specific reading strategy instruction, applied to self-selected independent reading, and Phase 3 allows students to pursue self-selected enrichment activities and projects related to reading“ (Reis et al., 2011, S. 465). Damit verfolgt der SEM-R-Ansatz nicht nur das Ziel der Vermittlung von Lesestrategien, sondern auch, den Fokus auf die individuellen Interessen und Lernbedingungen der Lernenden zu legen und explizit in den Lernprozess einzubeziehen. Anzumerken sei, dass unterrichtliche Settings, in denen sich die Schülerinnen und Schüler den Lerngegenstand (in einem gesetzten Rahmen) aussuchen können, nicht beschreibendes Merkmal der individuellen Förderung, wie sie hier verstanden wird, sind. Dies ist bei der Verallgemeinerung der Befunde zu berücksichtigen.

In Mehrebenenanalysen zeigte sich unter Verwendung von standardisierten Tests und unter Kontrolle der Eingangsniveaus, dass alle Treatmentklassen mindestens gleich oder statistisch signifikant besser abschnitten als die Kontrollklassen der gleichen Schulen, die kein spezifisches Konzept im Leseunterricht verfolgten. Für den Leistungsvorsprung in der Lese-

kompetenz der SEM-R-Klassen in zwei Schulen konnten die Autoren eine Effektstärke von Cohens  $d = 0,33$  bzw.  $d = 0,10$  für die zweite Schule feststellen; in den übrigen drei Schulen zeigten sich keine signifikanten Effekte. Für das Leseverständnis wurde ein signifikanter Vorsprung der Schule festgestellt, die einem extrem armen städtischen Raum zugeordnet wurde ( $d = 0,27$ ). Unter Einbeziehung der qualitativen Unterrichtsbeobachtungen erklären Reis et al. (2011) diesen Befund damit, dass sich die Kontrollklassen derselben Schule vor allem durch das Erlernen von Sachkenntnissen, Testvorbereitungen und nicht-differenzierendem Leseunterricht auszeichneten, in denen sehr wenig Zeit für das eigenständige Lesen zur Verfügung stand. In den SEM-R-Klassen der Schule sei zudem eine größere Leselust und ein verstärktes Leseengagement festgestellt worden – möglicherweise deswegen, weil die Gelegenheit, eigenständig Bücher auszuwählen, die dem Interesse und Leseniveau der Schülerinnen und Schüler entsprechen, in ihren häuslichen und schulischen Umgebungen weniger gegeben ist. Wengleich eine Verallgemeinerung dieser Befunde aufgrund der Stichprobengröße sicherlich eingeschränkt ist, verweisen sie jedoch darauf, dass eine individuelle Förderung der Kompetenzen eines bzw. einer Lernenden unter Einbeziehung seiner bzw. ihrer Interessen und Neigungen zu einer positiven Leistungsentwicklung, aber auch zu mehr Engagement und Motivation führen kann; insbesondere auch dann, wenn die Eingangsvoraussetzungen – bedingt beispielsweise durch die häusliche Lernumgebung – ungünstig sind.

Für entgegengesetzte Befunde wird häufig eine Veröffentlichung von Gruehn (2000) zitiert, die auf Daten des BIJU-Projekts beruht und untersucht, ob individuelle Förderung einen systematischen Einfluss auf den Schulerfolg von Kindern und Jugendlichen hat. Gruehn verwendet – neben weiteren Indikatoren – drei Skalen, die in anderen Darstellungen mit individueller Förderung in Verbindung gebracht werden konnten. Zu diesen gehören die individuelle Bezugsnormorientierung, die diagnostische Kompetenz des Lehrers im Leistungsbereich und die Verwendung binnendifferenzierender Maßnahmen, wovon die ersten beiden (basierend auf den Ergebnissen von Faktorenanalysen) für die späteren Analysen zu einem Konstrukt („Adaptivität des Lehrerverhaltens“) zusammengefasst wurden. Dabei wurde für die Skala der Verwendung binnendifferenzierender Maßnahmen auf den Individualized Classroom Environment Questionnaire (ICEQ; Fraser, 1981) in gekürzter und modifizierter Form zurückgegriffen. Dieser „bezieht sich auf individualisierende Unterrichtsstrategien im Hinblick auf leistungsrelevante Aufgabenstellungen durch Vergabe unterschiedlicher Lernmaterialien sowie leistungsabhängige Anforderungsniveaus bei Lehrerfragen“ (Gruehn, 2000, S.

131), als Beispiel gibt Gruehn (2000, S. 129) das folgende Item an: „Im Mathematik-/Biologie-/Physikunterricht haben die Schüler oft verschiedene Aufgaben“. Hier könnte geurteilt werden, dass die Vergabe von unterschiedlich schweren Aufgaben für die verschiedenen Leistungsniveaus der Schülerinnen und Schüler noch nicht zwingend mit einer individuellen Anpassung an deren Begabungen, Lernvoraussetzungen und Interessen einhergehen müsse. Klärung würde die Einsicht in alle tatsächlich verwendeten Items geben – insbesondere, da Fraser (1981) die Teilskala ‚Differentiation‘ seines ICEQ wie folgt beschreibt: „Emphasis on the selective treatment of students on the basis of ability, learning style, interests, and rate of working“, dem allerdings das Beispielitem „Different students use different books, equipment, and materials“ (beides Fraser, 1981, S. 5) gegenüber steht, welches – im Gegensatz zu der Skalenbeschreibung – nur bedingt eine Passung von Unterrichtsmethodik und Lernvoraussetzungen etc. spezifiziert. Es bleibt daher unklar, ob die in BIJU verwendete Skala dem für die folgenden Analysen in Kapitel 3 spezifizierten Verständnis von individueller Förderung entspricht. Allerdings zeigen sich für die binnendifferenzierenden Maßnahmen erwartungswidrig negative Korrelationen mit dem klassendurchschnittlichen Lernzuwachs der über 5.000 Schülerinnen und Schüler in Mathematik ( $r = -0,20$  bei  $p \leq 0,05$ ), Biologie ( $r = -0,30$  bei  $p \leq 0,01$ ) und Physik ( $r = -0,28$  bei  $p \leq 0,01$ ), die möglicherweise auf einen kompensatorischen Charakter der Maßnahme verweisen, die entsprechend vermutlich eher in leistungsschwachen Klassen angewendet wird (vgl. Gruehn, 2000).

Eine Teildimension des Faktors ‚Adaptivität des Lehrerverhaltens‘ stellt die individuelle Bezugsnormorientierung dar, welche „die Tendenz des Lehrers, Leistungsveränderungen von Schülern anhand eines intraindividuellen Vergleichsmaßstabs zu bewerten“ (vgl. Gruehn, 2000, S. 130–131) misst. Als Beispielitem gibt Gruehn (2000, S. 129) Folgendes an: „Wenn ein Schüler seine Leistung verbessert, wird er vom Lehrer gelobt, auch dann, wenn er im Vergleich zur Klasse unter dem Durchschnitt liegt“. Die zweite Teildimension des Faktors ergibt sich aus der diagnostischen Kompetenz des Lehrers im Leistungsbereich. Hier wird „das auf diagnostischen Informationen beruhende Wissen des Lehrers um den aktuellen Leistungsstand jedes einzelnen Schülers [...]“ (Gruehn, 2000, S. 131) erfasst und unter anderem anhand des folgenden Items gemessen: „Unser Mathematik-/Biologie-/Physiklehrer merkt sofort, wenn man etwas nicht verstanden hat“ (Gruehn, 2000, S. 129). Für diesen gemeinsamen Faktor ergeben sich lediglich Korrelationen mit dem klassendurchschnittlichen Lernzuwachs in Mathematik ( $r = 0,18$  bei  $p \leq 0,05$ ), die Korrelationskoeffizienten für die Fä-

cher Biologie und Physik weichen nicht signifikant von Null ab. Es zeigt sich also ein im Vergleich zu den Effekten der Binnendifferenzierung deutlich kleinerer Einfluss, der zudem nur im Fach Mathematik bedeutend von Null verschieden ist. Anzumerken sei an dieser Stelle jedoch, dass mit der Bezugsnormorientierung und diagnostischen Kompetenz von Lehrenden nur zwei Elemente individueller Förderung berücksichtigt werden. Wie in Kapitel 4 dargestellt, umfasst dieses Konzept noch weitere Merkmale, etwa die didaktischen Kompetenzen der Lehrenden, aber auch Formen des Feedbacks (im Gegensatz zu der Leistungsbeurteilung) und der Lernberatung. Aus der Untersuchung von Gruehn (2000) lassen sich drei Schlussfolgerungen ableiten:

- a) Dass die Autorin der individuellen Förderung keinen Einfluss auf den Schulerfolg von Kindern und Jugendlichen nachweisen konnte, muss hier zurückgewiesen werden. Zwar handelt es sich bei den untersuchten Effekten lediglich um eine schwache Korrelation mit der klassendurchschnittlichen Leistung in einem der drei untersuchten Schulfächer. Allerdings berücksichtigt der Indikator ‚Adaptivität des Lehrerverhaltens‘ auch nur Teildimensionen individueller Förderung.
- b) Eine Nutzung der hier dargestellten Skalen als Indikatoren für individuelle Förderung muss entsprechend mit Vorsicht erfolgen: Während der Faktor ‚Adaptivität des Lehrerverhaltens‘ zwei Teildimensionen individueller Förderung abdeckt, kann dies für die Skala der binnendifferenzierenden Maßnahmen nicht abschließend beurteilt werden.
- c) Die Maßnahmen der inneren Differenzierung scheinen, unabhängig davon, ob sie der individuellen Förderung zugeordnet werden können, vorrangig dann angewandt zu werden, wenn die durchschnittlichen Leistungen einer Klasse niedrig sind.

Aus den bisher dargestellten Untersuchungen vor allem aus dem deutschsprachigen Raum lässt sich bereits die hohe Varianz in den Befunden zu der Wirkung individueller Förderung ablesen: Während Helmke und Weinert (1997) ebenso wie Beck et al. (2008) und – für ein spezifisches Unterrichtskonzept, welches Teilkomponenten individueller Förderung beinhaltet – Reis et al. (2011) Effekte in einigen Dimensionen nachweisen können, zeigen sich keine bis negative Zusammenhänge bei Gruehn (2000) (welche ggf. in umgekehrter Kausalrichtung zu interpretieren sind). Solch einer Varianz in Forschungsergebnissen kann mit Metaanalysen begegnet werden, die zum Ziel haben, eine große Anzahl von Studienergebnissen zu einzel-

nen Indikatoren, üblicherweise Effektstärken wie das bereits vorgestellte Cohens  $d$ , zusammenzufassen. Dass diese gerade dann, wenn einer didaktischen Maßnahme kein einheitliches Verständnis und Konzept zugrunde liegt, mit Vorsicht herangezogen werden sollten, zeigt sich beispielsweise an der Metaanalyse von Hattie (2009). Untersucht wurde in dieser nur ein Aspekt der hier interessierenden Merkmale von Schulerfolg, die kognitiven Charakteristika von Schülerinnen und Schülern und ihre Beeinflussung durch unterschiedliche, vorrangig lehrerbezogene Merkmale von Unterricht. Herangezogen wurden hierfür die Daten aus über 800 Metaanalysen, in denen sich etwa 50.000 Studien mit insgesamt über 83 Millionen Untersuchungsteilnehmern finden lassen (vgl. Köller, 2012). Für die Bewertung der Effektstärken der jeweiligen Merkmalsträger orientiert sich Hattie an dem erwähnten Effektstärkemaß  $d$ , wobei Hattie Werte von  $d \geq 0,40$  als substantiell bezeichnet. Köller (2012) argumentiert, dass dieser Wert „eine gewisse Plausibilität [besitzt], da wir mittlerweile aus den großen nationalen und internationalen Studien Evidenz haben, dass in der Sekundarstufe I ein Jahr Unterricht in einem Fach einen Wissenszuwachs in der Größenordnung von ca.  $d = .40$  bis  $d = .50$  generiert. Für Hattie sind dementsprechend Maßnahmen von Bedeutung, die in einer Lerngruppe den Effekt normalen Unterrichts ‚schlagen‘“ (Köller, 2012, S. 74, Herv. i. O.). Dabei können pädagogische Maßnahmen mit Effektstärken von  $d < 0$  als schädend, Zusammenhänge von  $0 \leq d < 0,20$  als weder schädend noch helfend, Effekte von  $0,20 \leq d < 0,40$  als erfolgreich, zwischen  $0,40 \leq d < 0,60$  als sehr erfolgreich und Effektstärken von  $d \geq 0,60$  als äußerst erfolgreich bezeichnet werden, wobei der Wissenszuwachs in den letzten drei Kategorien dem eines halben, eines ganzen und dem Zuwachs entspricht, der über dem eines durchschnittlichen Schuljahres liegt (vgl. Köller, 2012). Für die in dieser Ausarbeitung interessierende individuelle Förderung (‚individualized instruction‘) ergibt sich eine durchschnittliche Effektstärke von  $d = 0,23$ , die also knapp in dem durch Köller (2012) als ‚erfolgreich‘ (weil der Lernzuwachs durch diese Maßnahme etwa dem Lernzuwachs eines halben Schuljahres entspricht) betitelten Bereich, jedoch unterhalb Hatties „Umschlagpunkt“ (Hattie, 2013, S. 21) substantieller Einflüsse von  $d \geq 0,40$  liegt. Dabei sei anzumerken, dass einer Meta-Metaanalyse solchen Umfangs naturgemäß eine starke Variabilität dessen, was unter dem ohnehin unterschiedlich verstandenen Konstrukt der individuellen Förderung zu verstehen ist, unterliegt. So betrachtet Horak (1981) als eine der von Hattie herangezogenen Metaanalysen etwa den „self-paced modularized approach“ (Horak, 1981, S. 249) und Programme wie die „Individually Prescribed Instruction (IPI)“ und das „Program for Learning in

Accordance with Needs (PLAN)“ (beides Horak, 1981, S. 250) im Fach Mathematik. Dabei differenziert sie die 129 gefundenen Effektstärken beispielsweise nach Outcome (Schulnotenäquivalent, Ergebnisse in standardisierten Tests), Dauer der individuellen Förderung (weniger oder mehr als acht Monate), nach Klassenstufen und ähnlichem und kommt entsprechend zu hochdifferenzierten Aussagen wie „individualized modules appear to work best at the junior high school level (7-9) whether the criterion measure is standardized achievement tests or teacher-developed measures“ (Horak, 1981, S. 252). Insgesamt findet Horak (1981) eine Effektstärke individueller Förderung auf die kognitiven Mathematikleistungen von Schülerinnen und Schülern in Höhe von Glass'  $\Delta^{17} = -0,07$ , wobei einzelne Effektstärken in differenzierteren Betrachtungen erheblich davon abweichen: So liegt beispielsweise der Einfluss individueller Förderung im ersten Schuljahr bei Glass'  $\Delta = 0,31$  und im achten Schuljahr bei Glass'  $\Delta = 0,53$ , in Klasse vier hingegen bei Glass'  $\Delta = -0,22$ . Aiello und Wolfle (1980) kommen bezüglich der heterogenen Befunde in ihrer Metaanalyse (115 Studien) sogar zu dem Schluss, dass ihre „initial attempts at explaining the variance observed in effect sizes using multiple regression analysis has [sic] failed to identify any instructional or design variables which account for a large proportion of the variance in effect size“ (Aiello & Wolfle, 1980, S. 8). Die durchschnittliche Effektstärke liegt für diese Metaanalyse bei Cohens  $d = 0,35$ . Untersucht wurden die „computer-assisted instruction (CAI)“, das „personalized system of instruction (PSI)“, „audio-tutorials (AT)“ und „programmed instruction (PI)“ (alle Aiello & Wolfle, 1980, S. 7), wobei die ersten beiden stärkere Effekte (beide  $d = 0,42$ ) aufwiesen als die letztgenannten ( $d = 0,21$  bzw.  $d = 0,27$ ). Einen anderen Ansatz wählten Waxman, Wang, Anderson und Walberg (1985), deren Ergebnisse ebenfalls von Hattie aufgenommen wurden. Als Auswahlkriterien für die insgesamt 35 Studien, die für diese Metaanalyse herangezogen wurden, diente nicht die Evaluation bestimmter Programme individueller Förderung, wie dies in den oben dargestellten, ebenfalls durch Hattie berücksichtigten Metaanalysen der Fall ist. Stattdessen führen die Autoren sieben Merkmale auf, von denen mindestens eines in den ausgewählten Studien berücksichtigt sein sollte. Zu diesen gehörte „1. Instruction is based on the assessed capabilities of each student. 2. Materials and procedures permit each student to progress in mastering the curriculum at a pace

---

<sup>17</sup> Im Gegensatz zu der Berechnung des bekannteren Effektstärkemaßes  $d$  nach Cohen schlägt Glass (1977) bei sehr unterschiedlichen Standardabweichungen in Kontroll- und Treatmentgruppe vor, statt der mittleren Standardabweichung die Standardabweichung der Kontrollgruppe zu verwenden, da diese nicht durch die Intervention beeinflusst wird (vgl. auch Psychometrica – Institut für psychologische Diagnostik, 2015).

suitable to his or her abilities and interests. 3. Periodic progress evaluations inform students of their mastery. 4. Students assume responsibility for diagnosing their current needs and abilities, planning and pursuing individual learning activities, and evaluating mastery. 5. Alternative learning activities and materials are available to help students acquire academic skills and content. 6. Students have a choice in selecting educational goals, outcomes, and activities. 7. Students help one another to pursue individual goals and cooperate in achieving group goals" (Waxman et al., 1985, S. 27–28). Hier zeigt sich ein sehr ausgedehntes Verständnis von individualisiertem Unterrichten, das sowohl Elemente der Passung der Lerninteressen und Lernangebote fokussiert, als auch entsprechende Materialien, Lernstanddiagnosen (sowohl durch Lehrende als auch Lernende), aber auch eine Freiheit der Schülerinnen und Schüler bei der Wahl ihrer Lernziele und Lernaktivitäten sowie ein Miteinander und gegenseitiges Unterstützen bei dem Erreichen der individuellen Ziele. Dabei ergibt sich im Mittel eine Effektstärke von 0,45, die bei einer Differenzierung nach Outcomes für die kognitive Facette mit 0,39 etwas schwächer und für die affektive Facette mit 0,60 deutlich stärker ausfällt<sup>18</sup>. Eine weitere Metaanalyse, die von Hattie herangezogen wurde, konzentriert sich auf die Effekte individueller Förderung in weiterführenden Schulen (Klasse 6 bis 12). Die 51 hierfür ausgewählten (quasi-)experimentellen Studien (publiziert zwischen 1960 und 1978) mussten die Kriterien „division of course content into chapter-length units, use of class time for individual work by students, and emphasis on formative testing" (Bangert, Kulik & Kulik, 1983, S. 145) erfüllen und sich auf Outcomes wie etwa der fachbezogenen Schulleistung oder dem Selbstkonzept beziehen. Für die kognitive Dimension ergibt sich dabei eine durchschnittliche Effektstärke von Glass'  $\Delta = 0,10$  (gegenüber den Kontrollgruppen), welche als trivial zu bezeichnen ist. Anzumerken sei jedoch, dass die Einflussstärken individueller Förderung auf die Leistung der Sekundarschülerinnen und -schüler in den einzelnen Studien zwischen Glass'  $\Delta = -0,84$  und Glass'  $\Delta = 1,24$  variieren. Für das Selbstkonzept wurden keine signifikanten Unterschiede zwischen Kontroll- und Treatmentgruppen festgestellt. Ähnlich wie in der Metaanalyse von Aiello und Wolfle (1980) konnten Bangert et al. (1983) darüber hinaus keine signifikanten Einflüsse unterschiedlicher Merkmale der Studiendesigns oder einzelner Aspekte individueller Förderung auf die Varianz in den Effektstärken finden. Wenngleich hier nicht vertiefend auf alle von Hattie (2009) genutzten Meta-

---

<sup>18</sup> Neben einer Beschreibung dessen, dass der Umfang der jeweiligen Studien in die Berechnung der Effektstärke als Gewichtsvariable aufgenommen wurde, findet sich keine Erläuterung, wie die Effektstärke berechnet wurde und um welchen Indikator (Cohens d, Glass'  $\Delta$  o.ä.) es sich handelt.

analysen eingegangen werden kann<sup>19</sup>, zeigt sich jedoch, dass hinter der ‚Hattie’schen‘ Effektstärke von  $d = 0,23$  Studien zu unterschiedlichen Aspekten individueller Förderung stehen, die in einigen Fällen sehr spezifische Unterrichtskonzepte evaluieren, in anderen Fällen abstraktere, dem spezifischen Fachunterricht übergeordnete Elemente der Individualisierung untersuchen, dabei unterschiedliche Schulformen und Schulfächer in den Blick nehmen und – selbst innerhalb der einzelnen Metaanalysen – zu hoch unterschiedlichen Effektstärken kommen. Wenngleich die von Hattie (2009) abgeleitete Effektstärke einen Eindruck über die allgemeine Tendenz der Wirkung individueller Förderung auf die (kognitive) Schulleistung von Kindern und Jugendlichen vermittelt, sollte ihrer Interpretation mit Vorsicht begegnet werden; insbesondere dann, wenn sich wie hier auf einen Teilaspekt von Schule, der vierten Klasse, und ein Land, für das in den herangezogenen Studien sicherlich nur wenige Untersuchungen vorlagen<sup>20</sup>, bezogen wird. Auch sei anzumerken, dass es sich in den Metaanalysen um Studien handelt, die bis zu 60 Jahre (mindestens jedoch 20 Jahre) alt sind – eine Zeitspanne, in der sich sicherlich unterrichtliche Rahmenbedingungen zu Gunsten oder Ungunsten individueller Förderung entwickelt haben können.

Eine aktuellere Metaanalyse von Van de Pol, Volman und Beishuizen (2010) bezieht sich auf eine Möglichkeit der Umsetzung individueller Förderung im Unterricht, dem Scaffolding. Scaffolding beruft sich auf einen zentralen Aspekt des Lernens: Dass es häufig durch andere angeleitet werde (vgl. Van de Pol et al., 2010; Stone, 1998). Das Konzept kann dabei anhand seiner metaphorischen Bedeutung erläutert werden: So sei der Begriff dem Baugewerbe entliehen, „where a scaffold is a temporary structure erected to help with the building or modification of another“ (Van de Pol et al., 2010, S. 271). In gleicher Weise könne die Rolle der Lehrenden bei problemlösenden Aktivitäten von Kindern beschrieben werden. Diese sei ein „temporary support provided for the completion of a task that learners otherwise might not be able to complete“ (Van de Pol et al., 2010, S. 272). Dabei werde dieses Konzept auf sehr unterschiedliche Weise im Schulunterricht umgesetzt. Lipowsky (2015) identifiziert in Anlehnung an Van de Pol et al. (2010) und Puntambekar und Hubscher (2005) die Hauptaspekte des Scaffolding, zu welchen

---

<sup>19</sup> Zwei weitere Studien konnten nicht eingesehen werden. Dabei handelt es sich um Hartley (1977) und Atash und Dawson (1986).

<sup>20</sup> Ein Auswahlkriterium für Hatties Meta-Metaanalyse war die Publikation der Studien in englischer Sprache, so dass diese vor allem U.S.-amerikanische Untersuchungen beinhaltet. Eine Analyse der in den jeweiligen Metaanalysen herangezogenen Studien könnte hier weitere Information darüber, ob sich darunter auch deutsche Untersuchungen befinden, bereitstellen.

- a) „eine fortlaufende prozessbegleitende Diagnostik der Lern- und Verstehensprozesse des einzelnen Schülers oder der Schülergruppe (ongoing diagnosis),
- b) eine am Lernstand und an den Lernvoraussetzungen des einzelnen Schülers oder der Schülergruppe ausgerichtete und kalibrierte Unterstützung der Lehrperson (adaptivity and calibrated support) und
- c) die schrittweise Ausblendung der Lehrerunterstützung in enger Verbindung mit einer zunehmenden Kontrolle des eigenen Lernprozesses durch den Lernenden (fading)“

gehören. Van de Pol et al. (2010) fassen in ihrer Metaanalyse vorrangig die Ergebnisse aus Studien mit quasi-experimentellen<sup>21</sup> und korrelativen Designs zusammen, von denen hier insbesondere die Einflüsse des Scaffolding auf die kognitiven Schülermerkmale von Bedeutung sind<sup>22</sup>. Dabei verweisen die Autoren explizit darauf, dass es bisher nur wenige empirische Arbeiten zu diesem Thema gebe, in denen Scaffolding zudem höchst unterschiedlich gemessen werde (sowohl durch Messinstrumente als auch in der Umsetzung in Treatments). Aufgeführt werden konnten entsprechend keine Effektstärken, wie dies häufig Ziel von Metaanalysen ist. Viel eher beschränken sich die Autoren auf die Darstellung der empirischen Herangehensweise und einer inhaltlichen Beschreibung der in den Studien festgestellten Effekte. So konnten für die Beurteilung des Einflusses des Scaffolding auf kognitive Schülercharakteristika eine Metaanalyse und vier Einzelstudien herangezogen werden (Veröffentlichungen zwischen 1998 und 2002). Dabei zeigte sich in allen vier Einzelstudien ein Zusammenhang zwischen höheren Schülerleistungen und Scaffolding. In der Metaanalyse (Swanson & Lussier, 2001) wurde lediglich das Strategietraining der Schülerinnen und Schüler durch die Lehrenden als wirksamer als das Scaffolding beschrieben. Van de Pol et al. (2010, S. 283) schließen hieraus, dass „[t]he results of studies on students' [...] cognitive activities indicate that scaffolding is effective.“ Zeitgleich sei betont, dass dies sicherlich eine generelle Tendenz darstellt, weitere Forschung in diesem spezifischen Anwendungsbereich individueller Förderung jedoch notwendig ist.

---

<sup>21</sup> Studien mit rein experimentellem Charakter konnten nicht identifiziert werden (vgl. Van de Pol, Volman & Beishuizen, 2010).

<sup>22</sup> Zu den weiteren untersuchten Merkmalen gehören die Effekte des Scaffolding auf metakognitive und affektiv-motivationale Merkmale. Während erste für die Analysen in dieser Ausarbeitung keine Relevanz haben, konnten Van de Pol et al. (2010) für die affektiv-motivationalen Schülercharakteristika lediglich zwei Studien identifizieren, die sich auf die Eingebundenheit der Schülerinnen und Schüler in den Unterricht und ihr dortiges ‚Vergnügen‘ („enjoyment“) fokussierten. Dies ist ebenfalls nicht Bestandteil der nachstehenden Analysen, so dass auf eine Darstellung der Ergebnisse verzichtet wird.

Deutlich ausdifferenzierter stellen Seidel und Shavelson (2007) die Ergebnisse ihrer Metaanalyse dar. In dieser Studie zu Lehrerwirksamkeit fassen die Autoren die in 112 korrelativen, quasi-experimentellen und experimentellen Forschungsarbeiten gefundenen Effekte unterschiedlicher Lehrervariablen auf die Lernprozesse und kognitiven wie motivational-affektiven Charakteristika des Schulerfolgs von Schülerinnen und Schülern zusammen. 32 dieser Studien wurden dabei der Komponente ‚Regulation and Monitoring‘ zugeordnet, zu denen die Lehrermerkmale ‚Support‘, ‚Feedback and Monitoring‘ sowie ‚Adaptivity‘ und ‚Self-Regulation‘ gehören. Da dies alles Elemente der individuellen Förderung (wie in Kapitel 4 dargestellt) sind, soll diese Komponente folgend als Indikator für den Einfluss der Individualisierung auf den Schulerfolg von Schülerinnen und Schülern herangezogen werden. Die nachstehende Tabelle 2 gibt dabei einen Überblick über die von Seidel und Shavelson (2007) identifizierten Effektstärken. Anzumerken sei vorab, dass die Autoren Fishers Z als Index der Effektstärke berichten<sup>23</sup>. Für die nachfolgende Übersicht wurden diese Kennwerte durch die Autorin (KSH) in das bekanntere Cohens d umgewandelt<sup>24</sup>, um eine leichtere Interpretation und eine Vergleichbarkeit mit anderen hier dargestellten Studien zu ermöglichen. In Tabelle 2 sind lediglich die Ergebnisse für die individuelle Förderung (‚Regulation and Monitoring‘) dargestellt<sup>25</sup>. Angegeben sind zunächst die Effekte über alle Studien hinweg (alle Outcomes/Lernerfolge und alle Dimensionen), in Folgeschritten werden diese differenziert nach Aspekten des Schulerfolgs (Einfluss der individuellen Förderung auf Lernprozesse, auf den motivational-affektiven sowie den kognitiven Schulerfolg) und spezifischen Unterrichtsdimensionen und Studiendesigns (Leseunterricht, Mathematik, Naturwissenschaften, Unterricht in Grundschulen und Sekundarschulen, Effekte in korrelativen und quasi-experimentellen Studiendesigns<sup>26</sup>).

---

<sup>23</sup> Alle in den Originalstudien publizierten Kennwerte wie Korrelations- und Regressionskoeffizienten, t-Werte etc. wurden in Fishers Z transformiert; die Effektstärken wurden zudem anhand des jeweiligen Stichprobenumfangs gewichtet.

<sup>24</sup>  $d = \frac{2r}{\sqrt{1-r^2}}$ , wobei Fishers Z wie r zu interpretieren ist.

<sup>25</sup> Zu den weiteren Komponenten des Lehrens und Lernens gehören „Time for Learning“, „Organization of Learning“, „Social Context for Learning“, „Goal-Setting/Orientation“, „Execution of Learning Activities“, „Basic Information Processing“, „Domain-specific Information Processing“ und „Evaluation of Learning“ (vgl. Seidel & Shavelson, 2007).

<sup>26</sup> Studien mit experimentellem Design konnten für die hier ausgewählten Aspekte nicht identifiziert werden.

**Tabelle 2:** Effektstärken (Cohens  $d$ ) des Einflusses der Lehrerfacette ‚Regulation and Monitoring‘ auf die verschiedenen Aspekte des Schulerfolgs in unterschiedlichen Dimensionen

	Alle Dimensionen	Lesen	Mathematik	Naturwissenschaften	Grundschule	Sekundarschule	Korrelative Studiendesigns	quasi-experimentelle Studiendesigns
Alle Outcomes	0,04	0,08	0,02	0,06	0,08	0,02	0,04	0,10
Lernprozesse	0,10	-	-	0,06	-	0,10	0,26	0,06
Motivational-affektiver Schulerfolg	0,16	0,14	0,20	-	0,12	0,18	0,16	0,22
Kognitiver Schulerfolg	0,02	0,08	-0,08	-	0,06	-0,02	0,02	0,06

Quelle: in Anlehnung an Seidel & Shavelson, 2007, adaptiert, eigene Darstellung

Bezüglich der Interpretation der Effektstärken verweisen Seidel und Shavelson (2007) auf Cohen (1969). Sie fügen dem jedoch an, dass „[t]he fact that effect sizes may appear small, however, does not necessarily mean that they are unimportant. This is especially the case in the study of teaching and learning, given the large number of factors affecting students' performance in school systems“ (Seidel & Shavelson, 2007, S. 471). Dies soll in der folgenden Zusammenfassung der oben dargestellten Effekte berücksichtigt werden.

Auffällig ist zunächst, dass die Einflüsse individueller Förderung auf affektive Merkmale des Schulerfolgs in allen Dimensionen deutlich stärker ausfallen als auf kognitive Merkmale des Schulerfolgs. Besonders deutlich wird dies für das Unterrichtsfach Mathematik: Hier scheint Individualisierung in einem negativen Zusammenhang mit der Schulleistung von Kindern und Jugendlichen zu stehen, während motivational-affektive Charakteristika mit  $d = 0,20$  vergleichsweise stark beeinflusst werden. Im Durchschnitt (d.h. über alle Dimensionen und Facetten des Outcomes von Schule hinweg) fällt die Wirkung von ‚Regulation and Monitoring‘ mit Effektstärken von maximal  $d = 0,10$  eher gering aus. Interessant ist zudem der Vergleich der Effekte in unterschiedlichen Studiendesigns: Während sich vergleichsweise starke Einflüsse individueller Förderung auf Merkmale des Lernprozesses in korrelativen Designs finden ( $d = 0,26$ ), können diese in quasi-experimentellen Studien kaum nachgewiesen werden. Anders hingegen die Wirkung auf den motivational-affektiven Schulerfolg: Dieser fällt mit  $d = 0,22$  (im Vergleich zu  $d = 0,16$ ) in Studien mit Versuchs- und Kontrollgruppen etwas stärker aus als in Studien, die Zusammenhänge zwischen verschiedenen Merkmalen bivariat testeten. Auch scheint sich für dieses Merkmal ein stärkerer Einfluss in Sekundarschulen als an

Grundschulen abzuzeichnen, während kognitive Merkmale an Grundschulen mit  $d = 0,06$  kaum und an Sekundarschulen minimal negativ ( $d = -0,02$ ) beeinflusst werden. Die hier besonders interessierende Lesedimension wird – vor allem verglichen mit Mathematik – positiv beeinflusst, wobei auch hier der Effekt individueller Förderung auf affektiv-motivationale Merkmale des Schulerfolgs der Schülerinnen und Schüler mit  $d = 0,14$  stärker ausfällt als auf seine kognitiven Merkmale ( $d = 0,08$ ). Zusammenfassend lässt sich (unter Berücksichtigung der für die weiteren Ausarbeitungen interessierenden Aspekte) feststellen, dass

- a) die hier betrachteten Komponenten individueller Förderung in allen Dimensionen einen stärkeren Einfluss auf motivational-affektive als auf kognitive Merkmale des Schulerfolgs aufweisen,
- b) Unterschiede in den Effektstärken auch auf die Studiendesigns zurückgeführt werden können, wobei die in a) dargestellte Verteilung konstant bleibt und
- c) individuelle Förderung durchaus in einem negativen Zusammenhang mit der kognitiven Leistung von Schülerinnen und Schülern stehen kann, wobei aus der obigen Darstellung nicht deutlich wird, ob dem beispielsweise korrelative Beziehungen zugrunde liegen, die auch durch eine umgekehrte Wirkungsrichtung (individuelle Förderung bei schlechten Leistungen) beeinflusst sein können – die Effektstärken der korrelativen vs. quasi-experimentellen Studiendesigns lassen darauf schließen.

Eine weitere, deutlich frühere Metaanalyse umfasst ebenfalls kognitive und affektive Schülermerkmale, stellt diese jedoch in Zusammenhang mit Aspekten *offenen Unterrichts*, für welchen erfolgversprechende Merkmale identifiziert werden sollen (vgl. Giaconia & Hedges, 1982). Dabei sei zunächst anzumerken, dass das Gesamtkonzept offenen Unterrichts hier nicht Bestandteil der Analysen sein wird, da dabei vor allem die Förderung von Selbstständigkeit und Mündigkeit der Schülerinnen und Schüler durch deren stärkere Einbindung in die Gestaltung des Unterrichts in den Blick genommen wird. Mehrfach nachgewiesen werden konnte, dass den Kindern und Jugendlichen – verglichen mit den Lernenden in Klassen traditionell lehrergelenkten Unterrichts – leichte bis erhebliche Nachteile in mehreren Leistungsbereichen (u.a. Lese- und Rechenfähigkeit, sprachliche Kompetenzen, vgl. etwa Bennett, 1976) ebenso wie im kreativen wie affektiv-motivationalen Bereich entstehen (Rauin, 1987, aktueller Gruehn, 2000 und Hartinger, 2005). Da Giaconia und Hedges (1982) jedoch einzelne Teilaspekte eines weitgefassten Konzepts offenen Unterrichts untersuchen, zu denen

auch Merkmale individueller Förderung gehören, sollen diese einzelnen Merkmale und ihre Wirkung auf kognitive wie affektiv-motivationale Charakteristika der Schülerinnen und Schüler im Folgenden dargestellt werden<sup>27</sup>. Die Grundlage hierfür bildeten 153 Studien, die ausgehend von den jeweiligen Effektstärken bezüglich einzelner Teilaspekte offenen Unterrichts in drei unterschiedliche Gruppen eingeteilt wurden. Dabei wurde das obere Drittel der Studien mit den größten Effektstärken der ‚large-effect‘-Gruppe zugeordnet, das untere Drittel der Studien mit den geringsten Effektstärken entsprechend der ‚smaller-effect‘-Gruppe. Studien mittlerer Effektstärke wurden nicht berücksichtigt, da die Autoren explizit den Unterschied zwischen den Studien mit den (in Relation betrachtet) stärksten bzw. schwächsten Effekten offenen Unterrichts in Augenschein nehmen wollten. Untersucht wurden die so ausgewählten Studien danach, welche Komponenten des Unterrichtskonzepts in die Analysen einbezogen wurden mit dem Ziel festzustellen, welche dieser Teilaspekte die erfolgreichen (= hohe Effektstärke) von den weniger erfolgreichen Programmen unterscheiden<sup>28</sup>. Für die Interpretation der in Tabelle 3 dargestellten (ausgewählten) Befunde benötigt es zunächst zweierlei Information, zu denen (1) die Anzahl der für die Messung des jeweiligen Outcomes berücksichtigten Studien gehört. So finden sich in beiden Effektstärkegruppen je 16 Studien, die Effekte einzelner Aspekte auf das Selbstkonzept der Schülerinnen und Schüler analysieren, 17 für den Einfluss auf die Leseleistung, 14 für die Leistungen in Mathematik und 8 für die Sprachkompetenz der Lernenden. Und (2) findet sich nachstehend eine kurze Erläuterung der hier interessierenden Unterrichts- bzw. Lehrermerkmale.

„Feature: Diagnostic Evaluation. Definition: Purpose of evaluation is to guide instruction; little or no use of conventional tests, but extensive use of work samples, observation, and written histories, of the student. Indicators and descriptive statements: charting of progress towards specific individual goals; evaluation used to facilitate and guide learning; child’s performance not compared to that of other children; teacher’s record-keeping combines constant jotting in class and thoughtful writing about each child; less standardized concept of

---

<sup>27</sup> Aufgrund der hohen Diversität des Verständnisses offenen Unterrichts, welches den für die oben dargestellte Metaanalyse herangezogenen Studien zugrunde liegt, empfehlen die Autoren die Verwendung eines Merkmalkatalogs einzelner Komponenten offenen Unterrichts. Dies ermögliche, etwaige Unterschiede zwischen Studien ggf. auf die unterschiedliche Betrachtung unterschiedlicher Komponenten des Unterrichtskonzepts zurückzuführen. Die oben dargestellten Befunde beziehen sich auf Aspekte dieses Katalogs, die mit individueller Förderung in Verbindung gebracht werden können.

<sup>28</sup> Giaconia und Hedges (1982) unterscheiden insgesamt die Teilbereiche „Role of Child in Learning“, „Diagnostic Evaluation“, „Materials to Manipulate“, „Individualized Instruction“, „Multiage Grouping of Students“, „Open Space“, „Team Teaching“ (vgl. Giaconia & Hedges, 1982).

student progress; nongraded approach to evaluate student's performance" (Giaconia & Hedges, 1982, S. 593)

„Feature: Individualized Instruction. Definition: Instruction based on the individual needs and abilities of each student; individualization of rate of, [sic] methods, and materials for learning; small group as opposed to large group instruction. Indicators and descriptive statements: individualized instruction; individualized approach; individualized work; environment responsive to individual learner needs; individualizing the curriculum; individualized goal setting; learning in accord with their own rate and style; small group or individual instruction" (Giaconia & Hedges, 1982, S. 593–594)

Es zeigt sich also, dass die zwei wie oben dargestellt definierten Teilaspekte offenen Unterrichts Merkmale individueller Förderung abdecken (rückbeziehend auf die Elemente individueller Förderung, wie sie in Kapitel 4 erläutert wurden, handelt es sich dabei z.B. um Lern-diagnosen, die es erlauben, die schülerspezifischen Stärken und Defizite festzustellen, Lern-aufgaben, die sich an den Lernvoraussetzungen und Zielen der einzelnen Lerner orientieren sowie adäquate Leistungsdokumentationen und -beurteilungen).

In Tabelle 3 finden sich die Häufigkeiten (Prozentwerte), mit denen die beiden hier interessierenden Unterrichtsaspekte in Studien mit starken Effekten gegenüber Studien mit schwachen Effekten vorkommen. Aufgeschlüsselt wird dies nach den (ausgewählten) Outcome-Variablen Selbstkonzept, Lese-, Mathematik- und Sprachleistung.

**Tabelle 3:** Prozentwerte der Studien mit großen (L) und kleinen (S) Effektgrößen für die Outcome-Variablen Selbstkonzept, Lese-, Mathematik- und Sprachleistung, die die Teilaspekte ‚Diagnostische Evaluation‘ und ‚Individualisierte Instruktion‘ aufweisen

		<b>Selbst- konzept</b>	<b>Lese- leistung</b>	<b>Mathematik- leistung</b>	<b>Sprach- leistung</b>
Diagnostische Evaluation	L	100	82	80	86
	S	46 **	77	83	83
Individualisierte Instruktion	L	100	85	83	86
	S	44 **	77	86	100

Legende: \*\* =  $p < 0,01$  für den  $\chi^2$ -Test (Fishers Exact Test)

Quelle: Giaconia & Hedges, 1982, S. 597, adaptiert

Zu interpretieren ist dies wie folgt: Alle (100%) Studien mit hoher Effektstärke (L), die sich auf den Outcome ‚Selbstkonzept‘ beziehen, beinhalten Aspekte der diagnostischen Evaluation und der individualisierenden Instruktion, während diese beiden Lehrermerkmale nur in

46 bzw. 44 Prozent der Studien mit geringer Effektstärke (S) Berücksichtigung finden. Dieser Unterschied ist nach Fishers Exact  $\chi^2$ -Test in beiden Fällen signifikant, entsprechend wahrscheinlicher sei es, dass Studien mit hoher Effektstärke bezüglich des Selbstkonzepts Elemente diagnostischer Evaluation und individualisierender Instruktion beinhalten. Giaconia und Hedges (1982) leiten daraus ab, dass beide Aspekte offenen Unterrichts hochwahrscheinlich einen Effekt auf das Selbstkonzept der Lernenden haben. Anders hingegen die Unterschiede zwischen den Studien hoher und geringer Effektstärken bezüglich der kognitiven Outcomes: Hier unterscheidet sich die Häufigkeit, mit der die beiden Lehrercharakteristika in Studien mit starken bzw. geringen Einflüssen auf die Leistungsdimensionen (Lesen, Mathematik, Sprache) Berücksichtigung finden, nicht signifikant. Ableiten lässt sich daraus, dass weder die diagnostische Evaluation noch die individualisierende Instruktion einen systematischen Einfluss auf die kognitiven Facetten des Schulerfolgs haben, da entsprechende Studien, in denen eine Leistungssteigerung nachgewiesen werden konnte, diese beiden Merkmale individueller Förderung nicht enthielten. Insgesamt bestätigt sich auch hier die Wirkung von Individualisierung im Unterricht, die vorrangig Einfluss auf affektiv-motivationale Charakteristika von Schulerfolg, und hier vor allem auf das Selbstkonzept der Schülerinnen und Schüler zu haben scheint, nicht so sehr jedoch in Verbindung gebracht wird mit (kognitiven) Leistungszuwächsen der Lernenden.

Eine Analyse der Wirkung individueller Förderung explizit auf die Selbstkonzeptentwicklung von Grundschülerinnen und -schülern führten Lipowsky, Kastens, Lotz und Faust (2011) durch. Untersucht wurde dabei die lese- und schreibbezogene Selbstkonzeptentwicklung von 735 Erst- und Zweitklässlern in 38 Klassen an privaten und staatlichen Grundschulen Deutschlands unter den Bedingungen leistungsbezogener Aufgabendifferenzierungen durch die Lehrperson. Erfasst wurde letztere anhand der Skala ‚Tendenz zur aufgabenbezogenen Differenzierung im Deutschunterricht‘, die – in Anlehnung an die Unterscheidung innerer Differenzierung und individueller Förderung (vgl. Kapitel 3) – explizit die Facetten der hier interessierenden Individualisierung enthält<sup>29</sup>. In Mehrebenenanalysen unter Kontrolle der Schulzugehörigkeit, des Geschlechts, der Lernvoraussetzungen und des Lernstands zu Beginn der ersten Klasse zeigte sich für die Selbstkonzeptentwicklung im Schreiben ein positiver Einfluss der individuellen Förderung ( $\beta = 0,14$  bei  $p < 0,05$ ), nicht jedoch für die lesebezogene Selbst-

---

<sup>29</sup> Beispielitem: „Bei der Unterrichtsvorbereitung überlege ich mir häufig, welche Aufgabenschwierigkeit für welche Schüler gerade richtig ist“ (Lipowsky, Kastens, Lotz & Faust, 2011, S. 876).

konzeptentwicklung; hier ließen sich keine systematischen Effekte ( $\beta = 0,05$  n.s.) nachweisen. Eine weitere Analyse galt der Überprüfung der Frage, ob Maßnahmen individueller Förderung zu einer Minderung des Big-Fish-Little-Pond-Effekts (BFLP-Effekt) führen. Der Effekt beruht darauf, dass soziale Vergleiche eine wichtige Quelle für die Ausbildung des eigenen Fähigkeitsselbstbildes sind – ein Vergleich der eigenen Leistungen (‘fish’) mit einer leistungsstarken Klasse (‘pond’) kann daher zu einer negativeren Einschätzung der eigenen Kompetenzen führen als ein Vergleich derselben Leistungen innerhalb einer leistungsschwächeren Klasse; bei letzterer kann es zu einer deutlichen Überschätzung der eigenen Fähigkeiten kommen (vgl. auch Marsh, 1984 und 1987, ausführlicher in Kapitel 8.2).

Lipowsky et al. (2011) stellen fest, dass der BFLP-Effekt bereits für das erste Schuljahr sowohl im Lesen als auch im Schreiben nachgewiesen werden kann. Analog zu den Ergebnissen des Einflusses individueller Förderung auf die Selbstkonzeptentwicklungen im Lesen und Schreiben lässt sich lediglich ein positiver Effekt der unterrichtlichen Maßnahme auf den BFLP-Effekt im Schreiben nachweisen ( $\beta = 0,14$  bei  $p < 0,01$ ), für die Beeinflussung des BFLP-Effekts für das Lesen finden sich keine systematischen Effekte. Damit reihen sich diese Ergebnisse in die Befunde einiger der oben dargestellten Studien ein: Individuelle Förderung im Deutsch- (bzw. landessprachlichen) Unterricht scheint insbesondere mit schreibbezogenen Merkmalen in Zusammenhang zu stehen, wenig hingegen mit lesebezogenen Charakteristika der Schülerinnen und Schüler.

In eine ähnliche Richtung deuten auch die Befunde einer Studie von Krätzschar (2010), die in mehrebenenanalytischen Auswertungen den Zusammenhang zwischen den Ausprägungen des schulischen und einiger fachbezogener Selbstkonzepte und dreier Unterrichtsarrangements (herkömmlich lehrergelenkter, altersgemischter und individualisierter Unterricht), in welchen die 620 Hamburger Gesamtschülerinnen und -schüler zwei Jahre lang unterrichtet wurden, untersuchte. Dabei zeigte sich „für *individualisierten Unterricht* ein Einfluss nur für das *schulische Selbstkonzept* und das *Selbstkonzept im Lesen*: Schüler im individualisierten Unterricht verfügen über höhere Selbstkonzepte als Schüler im unveränderten Unterricht. [...] Für die *Selbstkonzepte in Mathematik und Englisch* findet sich kein Unterschied zwischen Schülern im unveränderten und individualisierten Unterricht“ (Krätzschar, 2010, S. 177, Herv. i. O.). Allerdings zeigte sich ebenfalls, dass die gefundenen Unterschiede durch die Hinzunahme von Personenmerkmalen (Geschlecht, SES und Leistung) aufgeklärt

werden konnten und sich entsprechend nicht signifikant zwischen den Gruppen unterscheiden. Die Autorin begründet dies (neben möglichen stichprobenimmanenten Ursachen) damit, dass Selbstkonzepte bei Sekundarschülerinnen und -schülern bereits relativ stabil sind und ihre Beeinflussung unter Umständen einen längeren Zeitraum als den hier untersuchten beansprucht.

Einem anderen Diskurs folgten Helmke und Schrader (1990) mit ihrer Untersuchung bezüglich Unterrichtsmerkmalen, die sowohl die kognitive als auch die affektiv-motivationale Entwicklung der Lernenden überdurchschnittlich günstig beeinflussen. Gegenstand der Untersuchung, die auf den Daten der Studie ‚Unterrichtsqualität und Leistungszuwachs‘ basierte, waren dabei unter anderem die Leistungsentwicklung im Fach Mathematik und die Entwicklung des Fähigkeitsselbstbilds in Mathematik in den fünften Klassen von 39 bayrischen Hauptschullehrenden. Auf Grundlage von niedrig-inferenten Unterrichtsbeobachtungen (lehrerzentriertes Beobachtungssystem, Bewertung in neun Unterrichtsstunden durch trainierte Unterrichtsbeobachter), hoch-inferenten Unterrichtsbeurteilungen (Beurteilung auf Ratingskalen durch externe Beobachter) sowie anhand von Fragebögen erfasste Schüler- und Lehrerangaben wurden sechs Konstrukte zu Merkmalen des Unterrichts gebildet. Zu diesen gehören neben Management, Aufgabenorientierung, Klarheit, Langsamkeitstoleranz und affektivem Klima auch die hier interessierende Adaptivität, verstanden als „Anteil von Kleingruppenunterricht sowie Anteil individueller Hilfestellungen an allen registrierten Interaktionen (niedrig-inferente Beobachtungen); angemessene Variation des Lehrerverhaltens aus Schülersicht (Schülerfragebogen, 5 Items); Präferenz eines kleingruppenorientierten Unterrichts sowie Bereitschaft zur Variation des Unterrichtsangebots (Lehrerfragebogen, 2 Items)“ (Helmke & Schrader, 1990, S. 184–185). Dabei zeigt sich zunächst nur ein leicht signifikanter Zusammenhang der Adaptivität mit der Mathematikleistung ( $r = 0,25$ ,  $p < 0,10$ ), die Korrelation mit dem Selbstkonzept weicht nicht bedeutend von Null ab. In vertiefenden Analysen bilden die Autoren klassenbezogene Cluster, die sich hinsichtlich der Unterrichtsmerkmale unterscheiden. Dabei bildete sich ein ‚Positivcluster‘ heraus, welches hohe Werte in der Aufgabenorientierung, der Adaptivität und der Langsamkeitstoleranz aufwies. Die fünf diesem Cluster zugewiesenen Klassen sind zahlenmäßig klein und haben einen hohen Migrationsanteil; bezüglich der Intelligenz, des Vorkenntnisniveaus und der Vorkenntnisstreuung lagen diese Klassen im Mittelfeld. Für sie zeigte sich eine positive Entwicklung sowohl in der kognitiven wie affektiv-motivationalen Domäne. In einem Vergleich mit einem weiteren Posi-

tivcluster, in welchem jedoch lediglich eine positive Entwicklung im kognitiven, nicht jedoch im affektiv-motivationalen Bereich nachgewiesen werden konnte, zeigte sich darüber hinaus, dass alle Unterrichtsmerkmale bis auf die Klassenführung stärker ausgeprägt waren; der deutlichste Unterschied zeigt sich in der Langsamkeitstoleranz: Diese ist deutlich ausgeprägter in dem effektstärkeren Cluster. Die Autoren ziehen drei Schlüsse für einen Unterricht, der sowohl kognitiv als auch affektiv-motivational anregend sei: So müsse solch ein Unterricht (1) aufgaben- und lehrstofforientiert sein, wobei die Unterrichtszeit vor allem für die Durchnahme des Lerngegenstands genutzt werde. Er müsse zudem (2) den Schülerinnen und Schülern ausreichend Zeit für die Beantwortung von Fragen geben und sie keinem Tempodruck aussetzen. Und (3) müsse sich der Lehrer „um eine Anpassung der Unterrichtsinhalte und der Schwierigkeit seiner Anforderungen an unterschiedliche Schülervoraussetzungen [bemühen], [...] öfter die Unterrichtsform [variieren] und [...] aktive individuelle fachliche Hilfestellungen [leisten]“ (Helmke & Schrader, 1990, S. 194). Die Autoren verweisen zudem darauf, dass eine möglichst gleichmäßige, ausgewogene Ausprägung dieser Unterrichtsmerkmale zu einer überdurchschnittlichen Entwicklung bezüglich kognitiver wie affektiv-motivationaler Zielkriterien führe.

In ähnlicher Weise analysiert Gruehn (1995) anhand eines Teildatensatzes der bereits vorgestellten BIJU-Studie und dort anhand von 73 siebten Klassen in drei Bundesländern, welche Unterrichtscharakteristika zu einer positiven kognitiven wie nicht-kognitiven Entwicklung der Schülerinnen und Schüler beiträgt. Diese Entwicklung wird dabei anhand von mehreren Messzeitpunkten innerhalb eines Schuljahres nachgezeichnet. Als kognitive Facette dient ein Mathematiktest, als nicht-kognitive Merkmale werden das Fähigkeitsselbstkonzept und die Schulfreude der Jugendlichen herangezogen. Die Unterrichtsmerkmale wurden alleinig aus Perspektive der *Lernenden* erfasst. Auch hier werden Cluster anhand der Schülermerkmale gebildet mit dem Ziel zu prüfen, wie die Unterrichtsmerkmale in diesen Clustern verteilt sind. Herangezogen werden vier Gruppen, die sich (1) durch hohe Merkmale in der kognitiven wie nichtkognitiven Domäne auszeichnen, entweder (2) hohe kognitive, aber geringe nichtkognitive Merkmale oder (3) geringe kognitive, dafür aber hohe nichtkognitive Ausprägungen aufweisen oder (4) niedrige Werte in beiden Facetten zeigen. Dabei wird eine Individualisierung des Unterrichts (verstanden als leistungsbezogene Binnendifferenzierung) signifikant weniger von Schülerinnen und Schülern wahrgenommen, deren Klassen dem ‚Positivcluster‘ (hohe kognitive wie nichtkognitive Entwicklung) zugeordnet sind als im ‚Nega-

tivcluster' (negative Entwicklung in beiden Kriterien). Der Vergleich zwischen Positivcluster und den jeweiligen ‚Teilclustern' (entweder positive kognitive oder nichtkognitive Entwicklung) zeigte keine signifikanten Unterschiede. Gruehn schlussfolgert, dass ein individualisierender Unterricht möglicherweise nachteilig auf die kognitive Entwicklung der Lernenden wirke, da „eine solche Unterrichtsgestaltung wertvolle Unterrichtszeit bindet, die zusammen mit anderen zeitintensiven Unterrichtssequenzen wie ausdauerndem Wiederholen und Üben zu einem Zeitverzug führt, der sich ab einem bestimmten Zeitpunkt nicht mehr kompensieren läßt“ (Gruehn, 1995, S. 546).

Aus den hier dargestellten Forschungsergebnissen zu der Wirkung individueller Förderung auf kognitive wie affektive Schülermerkmale und ihre Entwicklung wird vor allem eines deutlich: Einheitliche Wirkzusammenhänge lassen sich nicht ableiten. Dies hat, wie oben bereits dargestellt, eine mögliche Ursache darin, dass individuelle Förderung unterschiedlich verstanden und folglich unterschiedlich operationalisiert wird. So finden sich neben wenig spezifischen Konstruktbeschreibungen wie ‚individualized instruction', ‚individualisierter Unterricht' und ‚Adaptivität' (Hattie, 2009; Giaconia & Hedges, 1982; Krätzschar, 2010) vor allem Teilaspekte individueller Förderung wie die Handlungs- und Planungskompetenz von Lehrenden (Beck et al., 2008), (individuelle fachliche) Unterstützung (Helmke & Weinert, 1997a; Seidel & Shavelson, 2007; Helmke & Schrader, 1990; Reis et al., 2011), Feedback (Seidel & Shavelson, 2007), Diagnostik (Giaconia & Hedges, 1982; Helmke & Weinert, 1997a) sowie eine individuelle Bezugsnormorientierung (Gruehn, 2000). Die leistungsbezogene Aufgabendifferenzierung bzw. Binnendifferenzierung (Lipowsky et al., 2011; Gruehn, 1995 und 2000; Reis et al., 2011) sowie differenzierende Settings wie Kleingruppenunterricht (Helmke & Schrader, 1990) müssen nicht – wie dies in Kapitel 3 beschrieben wurde – zwangsweise Elemente individueller Förderung beinhalten, reicht es hier doch beispielsweise, wenn unterschiedlichen Schülergruppen unterschiedlich schwere Aufgaben zur Verfügung gestellt werden, die sich nicht an den Interessen, Lernvoraussetzungen und Neigungen der Lernenden orientieren müssen (keine Planung der Passung der Unterrichtsangebote). So ist es nicht verwunderlich, dass sich die *Ergebnisse* dieser Studien ebenfalls in hohem Maß unterscheiden. Ein genereller Trend lässt sich dennoch ableiten: Individuelle Förderung scheint affektiv-motivationale Schülercharakteristika wie Motivation, Selbstkonzept und Lernfreude stärker zu beeinflussen als kognitive Merkmale (meist gemessen über fachspezifische Leistungstests).

Bezüglich der *Wirkung auf kognitive Schülermerkmale* finden sich domänenspezifische Effekttstärken und Korrelationen von  $-0,30$  (zum Beispiel bei Gruehn, 2000 für ‚Verwendung binnendifferenzierender Maßnahmen‘) und bis zu  $+0,33$  (etwa bei Reis et al., 2011 für ‚differentiated instruction, including specific reading strategy instruction‘, ähnlich auch bei Helmke & Weinert, 1997a für ‚individuelle fachliche Unterstützung‘ und Hattie, 2009 für ‚individualized instruction‘). Statistisch kaum bedeutsame Effekte nahe oder nicht signifikant von Null abweichend finden sich zudem ebenfalls, etwa in der Metaanalyse von Seidel und Shavelson (2007) für ‚Support, Feedback and Monitoring, Adaptivity, Self-Regulation‘ und bei Giaconia und Hedges (1982), welche die ‚Diagnostic Evaluation‘ und ‚Individualized Instruction‘ als Facetten offenen Unterrichts hinsichtlich ihrer Wirkung auf kognitive Merkmale untersuchten. Und auch *fachspezifische* Wirkzusammenhänge differieren in ähnlicher Spannweite<sup>30</sup>, so dass auch diese nicht zu einer einheitlichen Wirkrichtung zusammengefasst werden können. Auf zwei Besonderheiten sei in diesem Kontext jedoch hingewiesen: Die zum Teil negativen Zusammenhänge (etwa bei Seidel & Shavelson, 2007 sowie Gruehn, 1995 und 2000) ergeben sich, wie Gruehn (1995) vermutet, zum einen möglicherweise daraus, dass individuelle Förderung unter den gegebenen schulischen Bedingungen (eine Lehrkraft pro Klasse u.ä.) sehr (zeit-)aufwändig ist, so dass der damit einhergehende Zeitverzug im Vergleich zu herkömmlich unterrichteten Klassen irgendwann nicht mehr kompensiert werden kann, die Schülerinnen und Schüler in gleicher Zeit weniger lernen, was schließlich mit vergleichsweise geringeren Testleistungen einhergeht. Andererseits können diese Befunde dahingehend interpretiert werden, dass individuelle Förderung gerade dann zum Einsatz kommt, wenn das Leistungsniveau einer Klasse unterdurchschnittlich ist, was ebenfalls zu negativen Korrelationsparametern führen würde. Deutlich wird, dass gerade hier (quasi-)experimentelle Studien mit im Längsschnittdesign von Nöten sind, um solche angenommenen Wirkrichtungen empirisch nachweisen zu können.

Bezüglich der *Wirkung individueller Förderung auf affektiv-motivationale Schülermerkmale* zeigt sich ein etwas abweichendes Bild: Hier bewegen sich alle Befunde im positiven Bereich zwischen  $+0,14$  und  $+0,20$ , etwa in der Metaanalyse von Seidel und Shavelson (2007) und Giaconia und Hedges (1982). Bei Lipowsky et al. (2011, ‚leistungsbezogene Aufgabendifferenzierung‘) findet sich zwar ein positiver Einfluss auf das Selbstkonzept im Schreiben

---

<sup>30</sup> Für Mathematik:  $-0,20$  bis  $+0,32$ , für den muttersprachlichen Unterricht (Lese- und Schreibkompetenzen): nicht signifikant bis  $+0,33$ , für die naturwissenschaftlichen Fächer:  $-0,30$  bis  $+0,10$ .

( $\beta = 0,14$ ), nicht jedoch auf das Selbstkonzept im Lesen. Reis et al. (2011) untersuchen zwar nicht explizit die affektiv-motivationale Wirkung des SEM-R-Ansatzes (Leseunterricht), interpretieren ihre Befunde aber dahingehend, dass das Konzept, welches Aspekte individueller Förderung beinhaltet, sich positiv auf das Leseengagement und die Lesefreude von Kindern und Jugendlichen auswirke. Und auch Giaconia und Hedges (1982) stellen in ihrer Metaanalyse fest, dass 100% der Studien, denen ein hoher Einfluss auf bzw. Zusammenhang mit dem Selbstkonzept nachgewiesen wurde, Aspekte der individuellen Förderung (diagnostische Evaluation und individualisierte Instruktion) untersuchten. Keine signifikanten Zusammenhänge finden sich lediglich bei Krätzschar (2010), bei der alle Effekte durch Hintergrundmerkmale wie Geschlecht, sozioökonomischer Status und Leistung aufgeklärt werden konnten sowie bei Helmke und Schrader (1990), bei denen die Korrelation zwischen dem Selbstkonzept in Mathematik und der Adaptivität der Lehrperson (Kleingruppenunterricht, individuelle Hilfestellungen, Variation des Unterrichtsangebots) nicht bedeutend von Null abwich. Es lässt sich also für diese Domäne vorläufig festhalten, dass individuelle Förderung einen schwach positiven Einfluss auf die affektiv-motivationale Entwicklung von Schülerinnen und Schülern hat. Doch auch hier sei darauf verwiesen, dass solche Zusammenhänge nur bedingt anhand von sechs Studien, die zudem unterschiedliche Aspekte nicht-kognitiver Schülermerkmale untersuchten, verallgemeinert werden können – weitere Forschungsarbeiten sind auch in diesem Teilgebiet von Nöten.

Dennoch wird aus den hier dargestellten Befunden bisheriger Forschungsarbeiten zweierlei abgeleitet: Es wird angenommen, dass individuelle Förderung in einem (schwach) positiven Zusammenhang mit der kognitiven (hier: Leseleistung) und (etwas stärker) mit der affektiven Domäne (hier: Selbstkonzept im Lesen) steht.

Dass individuelle Förderung auch eine Möglichkeit darstellen könnte, der in Kapitel 2 beschriebenen Heterogenität im Klassenzimmer zu begegnen, wurde nur in einer der vorgestellten Studien am Rande aufgegriffen: So konnten Beck et al. (2008) zeigen, dass adaptiv unterrichtende Lehrende größere Lernfortschritte in leistungsheterogenen Klassen erzielten als Lehrende, die kaum adaptiv unterrichteten und nicht entsprechend qualifiziert wurden.

Bisher noch gar nicht ersichtlich wird aus den bisherigen Forschungsarbeiten, ob dieses didaktische Unterrichtssetting es darüber hinaus ermöglicht, dem Herkunftseffekt ausgleichend zu begegnen: Wie in dem folgenden Kapitel gezeigt wird, beeinflusst die sozioökono-

mische Lage einer Familie den Schulerfolg von Kindern in beträchtlichem Ausmaß, was sich bereits in höchst unterschiedlichen Lern- und Leistungsständen bei Schuleintritt zeigt und sich im Verlauf der Schullaufbahn manifestiert. Ein Unterricht, der sich auf diese unterschiedlichen Voraussetzungen der Lernenden einstellt und darüber hinaus ihre individuellen Neigungen und Interessen berücksichtigt, könnte sich positiv vor allem auf die affektiv-motivationale, aber auch auf die kognitionsbezogene Entwicklung von Schülerinnen und Schülern auswirken und die Relevanz herkunftsbedingter Voraussetzungen für den Schulerfolg der Lernenden abschwächen.

## **7. Herkunftseffekt**

Aus den vorangegangenen Kapiteln lassen sich folgende Kernelemente des schulischen Lernens ableiten: (1) Lernen ist ein hochindividueller Prozess, der gerade dann gelingt, wenn der Lerngegenstand für den Lernenden bedeutungsvoll ist. Er ist damit in hohem Maße abhängig von dem Vorwissen, den Lernvoraussetzungen, Interessen und Begabungen der Schülerinnen und Schüler. (2) Maßnahmen der individuellen Förderung, die sich vor allem dadurch auszeichnen, dass sie sich an diesen einstellungs- und voraussetzungsbezogenen Charakteristika der Schülerinnen und Schüler orientieren, beeinflussen kognitive und affektive Aspekte des Schulerfolgs von Kindern und Jugendlichen positiv.

Dabei ist bekannt, dass einigen Lernenden ein erheblicher Vorteil in der Schule zugutekommt: Bedingt durch familiäre Herkunftsmerkmale fällt es diesen Schülerinnen und Schülern deutlich leichter, sich dem in der Schule vorherrschenden Lernhabitus anzupassen – ein Umstand, der sich in systematisch größeren Schulerfolgen ausdrückt. Auf diesen sogenannten Herkunftseffekt soll im Folgenden vertiefend eingegangen werden, bevor daraus die hier zentrale Forschungsfrage abgeleitet werden soll.

Dass der Familie als primärer Sozialisationsinstanz eine entscheidende Rolle für den schulischen Lernerfolg von Kindern und Jugendlichen zufällt, konnte in einer Vielzahl von Studien nachgewiesen werden (vgl. etwa Becker, 2010; Wild & Lorenz, 2010; Pekrun, 2001). Dabei sind es vorrangig sozioökonomisch bedingte Merkmale wie der Sozialstatus und die familiären Ressourcen, aber auch der häusliche Anregeungsgehalt, denen eine prognostische Bedeutsamkeit im Hinblick auf die Kompetenzentwicklung, Bildungsaspiration sowie den erreichten Bildungsabschluss zukommt (Boudon, 1974; Wild & Lorenz, 2010; Becker, 2010).

Auch wirkt sich der kulturelle Sozialisationseffekt, welcher sich durch Unterschiede etwa in der Vermittlung von Sprachkultur, Lern- und Bildungsmotivation sowie habitualisierten Lerngewohnheiten zwischen den Sozialschichten zeigt, auf die schulische Leistung und die Bildungsverläufe von Kindern und Jugendlichen aus (Becker, 2010). Dies zeigt sich ebenfalls in der Ausgestaltung und Kompatibilität informeller und formaler Bildungskontexte: Während Kinder aus bildungsferneren Milieus ihre Freizeit häufig in unorganisierten Kontexten verbringen, orientieren sich die alltäglichen außerschulischen Erfahrungswelten von Kindern bildungsnäherer Milieus stärker an schulischen Anforderungen und Inhalten. Hierzu gehören insbesondere die Einbindung in Vereine, deren Aktivitäten organisiert, zeitlich fixiert und oftmals durch Erwachsene angeleitet sind ebenso wie außerschulische Unterrichtsstunden im Musizieren, der Hausaufgabenbetreuung oder Nachhilfe, welche für die Kompensation schulischer Rückstände genutzt werden können. Während sich also die schulischen und außerschulischen Bildungsbereiche von Kindern bildungsferner Familien deutlich unterscheiden, zeichnet sich der Habitus von Kindern bildungsnaher Familien durch eine höhere Konformität mit dem Schulhabitus aus. Dies hat zur Folge, dass Kinder bildungsferner Milieus häufiger das Gefühl äußern, Probleme zu haben, dem Unterricht zu folgen und insgesamt weit größere Angst davor haben, in der Welt formaler Bildung Fehler zu machen (vgl. Betz, 2006b). Diese Kinder bewegen sich also im Schulkontext unsicherer und ängstlicher als ihre Mitschülerinnen und Mitschüler, was auch zur Folge hat, dass sie weniger zuversichtlich sind, bei Anstrengung gute Noten zu bekommen. (vgl. Betz, 2006a)

Um diese und ähnliche herkunftsspezifischen Unterschiede in der Bildungsbeteiligung erklären zu können, wird häufig auf Boudons (1974) Unterscheidung primärer und sekundärer Effekte als prominenteste Theorie zur Wahl von Bildungswegen zurückgegriffen<sup>31</sup>. Zu den primären Sozialschichteffekten werden insbesondere die Folgen unterschiedlicher familiärer Ressourcen ökonomischer, sozialer und kultureller Natur gezählt. Die Unterscheidung dieser drei Kapitalsorten erläutert Bourdieu (1983) ausführlich. Er erweitert den aus der Wirtschaftstheorie stammenden Begriff des Kapitals, um „jene Unterscheidungsprinzipien auszumachen, mit denen so umfassend wie möglich die größtmögliche Zahl an festgestellten Unterschieden zu erklären [ist]“ (Bourdieu, 1989, S. 407) mit dem Ziel, unterschiedliche gesellschaftliche Klassen zu differenzieren und in einem sozialen Raum zu verorten. Die relative

---

<sup>31</sup> Diese bildet entsprechend häufig die Grundlage von Studien zu Bildungsungleichheiten und konnte in verschiedenen empirischen Arbeiten bestätigt werden (vgl. etwa Maaz & Nagy, 2010; Betz, 2006a; Lörz & Schindler, 2011; Siewert, 2013; Kristen, 1999).

Position einer Person in diesem sozialen Raum wird dabei durch bestimmte individuelle Merkmale wie etwa das Einkommen (ökonomisches Kapital) oder den Bildungsabschluss (institutionalisiertes kulturelles Kapital) bestimmt. Dabei wird das *ökonomische Kapital* als unmittelbar und direkt in Geld konvertierbar verstanden (vgl. Bourdieu, 1983, 1992). Unter kulturellem Kapital versteht Bourdieu (1983) zum einen *objektiviertes kulturelles Kapital* in Form von kulturellen Gütern wie Bücher oder Gemälde, wobei deren eigentliche Aneignung (im Gegensatz zum juristischen Eigentum) *inkorporiertes Kulturkapital* voraussetzt, also „die Verfügung über kulturelle Fähigkeiten, die den Genuß eines Gemäldes oder den Gebrauch einer Maschine erst ermöglichen“ (Bourdieu, 1983, S. 189). *Institutionalisiertes kulturelles Kapital* als drittes Merkmal wird durch schulische bzw. Bildungstitel erlangt, die als „Zeugnis[se] für kulturelle Kompetenz [...] seinem Inhaber einen dauerhaften und rechtlich garantierten konventionellen Wert über[tragen]“ (Bourdieu, 1983, S. 190) und den Zugang zu Berufen und damit einhergehend zu ökonomischem Kapital legitimieren und erleichtern können (vgl. Van Essen, 2013). Bei dem *sozialen Kapital* handelt es sich um „Ressourcen, die auf der Zugehörigkeit zu einer Gruppe beruhen“ (Bourdieu, 1983, S. 191, Herv. i. O.). Die Ausprägung des Sozialkapitals einer Person hängt dabei davon ab, wie groß das Netzwerk an tatsächlich mobilisierbaren Beziehungen ist und über wie viel ökonomisches und kulturelles Kapital die Personen innerhalb des Netzwerkes verfügen (vgl. Bourdieu, 1992).<sup>32</sup> Alle drei Kapitalsorten sind mit Einschränkungen ineinander konvertierbar. So kann etwa ein hoher Bildungsabschluss als institutionalisiertes kulturelles Kapital zu einem gut bezahlten Arbeitsverhältnis (ökonomisches Kapital) führen. Die Aufnahme von Personen mit spezifischem Wissen in das eigene Netzwerk (soziales Kapital) kann zu der (teilweisen) Aneignung dieses Wissens (inkorporiertes kulturelles Kapital) beitragen, während der Verkauf etwa eines Kunstgegenstandes (objektiviertes kulturelles Kapital) einen Zugewinn an ökonomischem Kapital bewirkt (vgl. Van Essen, 2013). Bei der Umwandlung von Kapitalsorten entstehen Transformationskosten, die sich je nach Zweck unterscheiden: So ist als Umwandlung objektivierten kulturellen Kapitals in ökonomisches Kapital der Verkauf eines Musikinstruments beispielsweise deutlich weniger mit Arbeit (im Sinne von zeitlicher Verzögerung oder sekun-

---

<sup>32</sup> Bourdieu führt zudem das *symbolische Kapital* an, bei welchem es sich um eine übergeordnete Kapitalform handelt, die unter Nutzung von ökonomischem, kulturellem und sozialem Kapital zum Gewinn und Erhalt von Anerkennung beiträgt. Personen, die über ein hohes symbolisches Kapital verfügen, „haben weitreichende Gestaltungsfreiheiten“ (Van Essen, 2013, S. 20), da sie auf der Basis von kollektiver Anerkennung „über legitime (also allgemein anerkannte) Deutungshoheit [verfügen]“ (ebd.). Diese Kapitalform ist für die folgende Ausarbeitung und insbesondere die Betrachtung primärer und sekundärer Herkunftseffekte nach Boudon nicht von Bedeutung und kann daher im Folgenden vernachlässigt werden.

dären Kosten) verbunden als etwa die Nutzung des eigenen sozialen Netzwerks bei der Jobsuche. Dieses „[kann] nur dann kurzfristig, zum richtigen Zeitpunkt, eingesetzt werden, wenn [es] bereits seit langem etabliert und lebendig gehalten worden [ist], als sei[...] [es] ein Selbstzweck“ (Bourdieu, 1983, S. 196). Die Nutzung des sozialen Netzwerks ist also mit deutlich mehr Transformationsarbeit verbunden, was verdeutlicht, dass die Kapitalsorten nicht uneingeschränkt und beliebig ineinander transformierbar sind.

Die Höhe und Ausprägungen der unterschiedlichen Kapitalsorten in Familien und die damit verbundene Zugehörigkeit zu einer gesellschaftlichen Schicht werden genutzt, um den Zusammenhang zwischen familiärer Herkunft und Bildungsbeteiligung zu erklären. Sie zählen nach Boudon (1974) zu den *primären Herkunftseffekten (primary effects of stratification)* und beeinflussen maßgeblich den kindlichen Sozialisationsprozess innerhalb der Familie. Nachgewiesen werden konnte dabei, dass sich die Sozialschichten „in der Vermittlung von Sprachkultur, in der Lern- und Bildungsmotivation hin zum selbstregulierten Handeln und Lernen sowie in den habitualisierten Lerngewohnheiten“ (Becker, 2010, S. 169) unterscheiden. Eine besondere Rolle spielt dabei – und im Kontext der nachfolgenden Analysen auf Basis der Daten der Internationalen Grundschul-Lese-Untersuchung – die häusliche Lesesozialisation, also Prozesse, die die Entwicklung von Lesemotivation und Lesekompetenz unterstützen (Rosebrock, 2006; Stubbe, Buddeberg, Hornberg & McElvany, 2007). Dabei sind es vor allem die unterschiedlichen familiären Sozialisationsumwelten, Kommunikationsstrukturen und die individuellen Leseerfahrungen (Bonfadelli & Fritz, 1993; McElvany, 2008), die sich je nach sozioökonomischer wie kultureller häuslicher Ausstattung unterscheiden und maßgeblich die Lesekompetenz und mit ihr den Schulerfolg von Kindern und Jugendlichen beeinflussen (Senechal & LeFevre, 2002; Carter, Chard & Pool, 2009; Stephenson, Parrila, Georgiou & Kirby, 2008). Am deutlichsten wird dies an dem Beispiel des Vorlesens von Büchern (eine weitergehende Übersicht kindlicher Lesesozialisationsprozesse im familiären Umfeld findet sich bei Stubbe et al., 2007), welches nicht nur das Interesse des Kindes am Lesen fördert, sondern auch ein Bewusstsein für Buchstaben-Laut-Verbindungen schafft, den Wortschatz erweitert und Kenntnisse über Strukturen, Schemata und Schriftkonventionen vermittelt, die grundlegende Voraussetzung für das Leseverständnis sind (Bus, Van Ijzendoorn & Pellegrini, 1995; Cochran-Smith, 1984; Stubbe et al., 2007). Zudem kann durch das Vorlesen von Büchern, die sich an dem Interesse des Kindes orientieren, die Lesemotivation gefördert werden. Durch die Interaktion mit dem bzw. der Vorlesenden „erfährt und übt

[das Kind] die Kommunikation über das Gelesene, die Verknüpfung mit vorhandenem Wissen und das Reflektieren und Bewerten von Inhalten und Sprache“ (Stubbe et al., 2007, S. 303). Dabei (und auch bezüglich ihres eigenen Leseverhaltens) agieren die Eltern als Vorbild. Dass auch hier substanzielle Unterschiede in dem lesebezogenen *Home Literacy Environment* (HLE, zu Deutsch: häusliche Lernumgebung) zwischen Familien mit hohem und niedrigen sozioökonomischen Status bestehen, konnte in einer Vielzahl von Studien festgestellt werden (vgl. etwa Rashid, Morris & Sevcik, 2005; Roberts, Jurgens & Burchinal, 2005; Stephenson et al., 2008; Burgess, Hecht & Lonigan, 2002; Van Steensel, 2006; Stubbe, 2010, eine Übersicht findet sich auch bei McElvany, 2008). Als prominentes Beispiel lässt sich die IGLU-basierte Untersuchung von Stubbe et al. (2007) aufführen, die einen Index der Lesesozialisation bildeten, welcher angibt, wie intensiv die Lesesozialisation in den Elternhäusern der befragten Viertklässlerinnen und Viertklässler war und ist (Mittelwert: 300, Standardabweichung: 60). Die Intensität der Lesesozialisation in Familien der obersten Sozialschicht (Akademiker, Techniker und Führungskräfte) liegt bei 321 Punkten, die Intensität in Familien von manuell Tätigen liegt bei 272 Punkten, die Differenz beträgt 48 Punkte. Zudem zeigt sich eine Varianzaufklärung des Leseverständnisses durch den Lesesozialisationsindex in Höhe von 6,3%. Diese Befunde verdeutlichen einerseits, dass die Intensität der häuslichen Lesesozialisation von Bedeutung ist für die Entwicklung des Leseverständnisses als zentraler Kulturtechnik, andererseits aber auch, dass diese zwischen Familien unterschiedlicher sozioökonomischer Ressourcen zu Ungunsten der bildungsferneren Haushalte stark variiert. Die sich daraus ergebenden Unterschiede führen zwangsläufig zu unterschiedlichen Eingangsvoraussetzungen für die Schulbildung (vgl. etwa Becker, 2010; Büchner, 2003; Bradley & Corwyn, 2002; Lee & Burkam, 2002; Baumert & Schümer, 2001). Boudon reduziert dieses Wirkungsgefüge: „The lower the social status, the poorer the cultural background – hence the lower the school achievement, and so on“ (Boudon, 1974, S. 29). Kinder insbesondere der untersten Sozialschicht wachsen in einem vergleichsweise ungünstigen häuslichen Sozialisationskontext auf und zeigen aufgrund der daraus resultierenden kognitiven Nachteile schlechtere Schulleistungen als Kinder aus höheren Sozialschichten – und scheitern eher an den dem Bildungssystem immanenten Selektionshürden, wie etwa dem Übergang von der Grundschule auf das Gymnasium (vgl. auch Becker, 2010).

Demgegenüber wirkt sich der *sekundäre Herkunftseffekt* kurzfristig und unmittelbar auf die Bildungschancen aus. Dabei handelt es sich um die sozialstatusabhängige Entscheidung ins-

besondere für den weiterführenden Bildungsweg, die – bei gegebenen Schulleistungen – abhängig ist von der individuellen Bewertung der Kosten und des Nutzens der zu wählenden Schullaufbahn (Boudon, 1974). Boudon orientiert sich dabei an den Annahmen der Werterwartungstheorie, nach welcher eine Option gewählt wird, die unter Abwägung der Kosten den höchsten Nutzen verspricht (vgl. Boudon, 1974, auch Kristen, 1999 und Maaz & Nagy, 2010). Neben den Kosten finanzieller Natur weist Boudon auf die „social cost[s]“ (Boudon, 1974, S. 30, zu Deutsch häufig zusammenfassend *soziale Nähe bzw. Distanz*, vgl. etwa Becker, 2010) hin. So sei „the choice of a curriculum [...] in most cases related to anticipated mobility, [...] [thus representing] a cost for the individual [...]“ (Boudon, 1974, S. 38). Dabei ist die zu überwindende soziale Distanz zu höheren Bildungsformen in höheren Sozialschichten aufgrund familiärer Bildungsvorstellungen, -erwartungen und verinnerlichter Sozialschichtzugehörigkeit deutlich geringer ausgeprägt. Zudem sind in dieser Gruppe die subjektiv erwarteten Kosten höherer Bildung niedriger und ihr erwarteter Nutzen deutlich größer als in den unteren Sozialschichten. Letztere bedürfen daher ein höheres Aspirationsniveau sowie mehr Ressourcen und Anstrengungen, um sich – bei gleichen Leistungen – ebenfalls für eine höhere Bildungslaufbahn zu entscheiden. (vgl. Becker, 2010) Dies zeigt sich deutlich in den aktuellen Analysen der Internationalen Grundschul-Lese-Untersuchung (IGLU 2011) für Deutschland. So präferieren Eltern der oberen Dienstklasse<sup>33</sup> 3,76 mal eher den Übergang ihres Kindes auf ein Gymnasium als Eltern, die als (Fach-)Arbeiter tätig sind – und dies bei Annahme gleicher kognitiver Fähigkeiten sowie Lese-, Mathematik- und Naturwissenschaftskompetenzen (vgl. Stubbe, Bos & Euen, 2012)<sup>34</sup>. Wie viel sekundäre Herkunftseffekte bei der tatsächlichen Übergangsrate auf das Gymnasium erklären können, zeigt Neugebauer (2010) mittels kontrafaktischer Dekompositionsanalysen anhand der Daten der bundesweiten Schulkinderkohorte des DJI-Kinderpanels (n = 538). In dieser Stichprobe wechseln 73% der Kinder der oberen Bildungsschicht auf das Gymnasium. Unter Annahme gleicher Leistungen würden lediglich 48% dieser Kinder auf die gleiche Schulform wechseln, würden sie der unteren Bildungsschicht angehören. Dies entspricht 59% der herkunftsbedingten Bildungsun-

---

<sup>33</sup> Diese Unterteilung orientiert sich an der Unterscheidung unterschiedlicher sozioökonomischer Stellungen nach Erikson, Goldthorpe und Portocarero (1979), wie sie auch in Kapitel 12.3.1 beschrieben werden.

<sup>34</sup> Anzumerken sei an dieser Stelle, dass Bildungsübergangentscheidungen in Deutschland nicht alleine durch Eltern gefällt werden sondern dass in der Mehrzahl der Bundesländer ein Konsens zwischen den Vorstellungen der Eltern und den Grundschulempfehlungen der Lehrerinnen und Lehrer angestrebt wird. Die Einschätzungen der Lehrpersonen sind hingegen in Baden-Württemberg, Bayern, Brandenburg, Sachsen, Saarland und Thüringen verbindlich (vgl. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2010).

gleichheiten am Übergang auf das Gymnasium, die auf sekundäre Effekte zurückgeführt werden können (primäre Effekte: 41%). Darüber hinaus zeigt Neugebauer (2010), dass die Übergangsentscheidung vorrangig im mittleren Notenbereich besonders ins Gewicht fällt. Hier kommt es besonders häufig zu unterschiedlichen Entscheidungen in bildungsnahen und bildungsfernen Elternhäusern – zu Ungunsten der Kinder unterer Bildungsschichten. Dahingegen differiert das Entscheidungsverhalten an den Rändern der Notenverteilung kaum – der Übergang sehr guter und sehr schlechter Schülerinnen und Schülern wird nahezu unabhängig von ihrer sozialen Herkunft getroffen. Dies erscheint, rückbeziehend auf die Werterwartungstheorie, schlüssig: Ein Kind mit einer durchschnittlichen Schulleistung von 1,0 wird, unabhängig von dem sozialen Status seiner Eltern, kaum auf eine Real- oder Hauptschule geschickt, während ein Kind, welches knapp einer Nichtversetzung entgangen ist, nicht auf das Gymnasium wechseln wird, unabhängig der ggf. hohen Bildungsaspirationen seiner Eltern (vgl. Neugebauer, 2010).

Dass neben dem hohen sekundären Einfluss auch der primäre Effekt der sozialen Herkunft auf den Schulerfolg weitreichende Folgen aufweist, konnten etwa Bos und Kollegen (2003, aktueller auch Wendt et al., 2012) zeigen: So liegen Viertklässlerinnen und Viertklässler aus Elternhäusern der untersten Sozialschicht im Leseverständnis etwa eine halbe Standardabweichung (2001: 56 Punkte, 2011: 53 Punkte) hinter ihren Mitschülerinnen und Mitschülern der obersten Sozialschicht, was in etwa dem Leistungszuwachs eines ganzen Schuljahres entspricht.

Abgeleitet werden kann hieraus eine dreifache Benachteiligung von Kindern aus sozial schwachen Familien innerhalb des deutschen Bildungssystems: Erstens verfügen diese über geringere schulische Kompetenzen, zweitens haben sie – verglichen mit Schülerinnen und Schülern, die ähnliche Leistungen zeigen – eine geringere Chance, eine Gymnasialempfehlung zu erhalten und drittens haben sie selbst bei einer Gymnasialempfehlung eine geringere Chance, von ihren Eltern tatsächlich auf ein Gymnasium geschickt zu werden (vgl. Stubbe, 2010; Pietsch & Stubbe, 2007).

Diese enge Kopplung von sozialer Herkunft und Schulerfolg in Deutschland soll zum Anlass genommen werden zu analysieren, ob die individuelle Förderung eine Möglichkeit darstellt, diesem Herkunftseffekt ausgleichend zu begegnen. Dabei wird (aus den obigen Darstellungen ableitend) angenommen, dass Kinder und Jugendliche aus bildungsnahen Elternhäusern

zum einen über mehr Vorwissen, zum anderen über lerngünstigere Einstellungen verfügen als ihre Mitschülerinnen und Mitschüler aus bildungsfernen Familien. Ein Unterricht, der keine Maßnahmen ergreift, den jeweiligen Lerngegenstand für die Schülerinnen und Schüler einer Klasse bedeutsam zu machen (ihn also an ihre Interessen, Neigungen, Begabungen, Lernvoraussetzungen und an ihr Vorwissen anzupassen), könnte hier zu einer systematischen Benachteiligung der Kinder und Jugendlichen führen, die aufgrund familiärer Merkmale schwächere Lern- und Leistungsvoraussetzungen aufweisen. Letzteren könnte es unter Umständen also schwerer fallen, dem Unterricht zu folgen, wenn sie über keine Strategie verfügen, sich einen Lerngegenstand anzueignen, der nicht mit bestehendem Wissen verknüpft werden kann und dem möglicherweise wenig überschulische Relevanz beigemessen wird. Eine Folge dessen wäre, dass die bestehenden Leistungsunterschiede zwischen beiden Schülergruppen ausgeweitet würden. Explizit anzumerken sei hierbei, dass individuelle Förderung nicht zum Ziel haben muss, die leistungsbezogene Heterogenität einer Klasse zu verringern<sup>35</sup>. Dies würde (im Extremfall) bedeuten, dass die leistungsschwächsten Schülerinnen und Schüler derlei zu fördern seien, dass sie das klassendurchschnittliche Leistungsniveau erreichen, während die übrigen Kinder und Jugendlichen keine leistungsförderliche Unterstützung erhalten würden. Ziel kann es stattdessen sein, *alle* Schülerinnen und Schüler einer Klasse in dem Maße zu fördern, dass sie den für sich je maximal möglichen Schulerfolg erreichen. Diesem letztgenannten Verständnis wird sich im Folgenden gewidmet; betrachtet werden sollen explizit nicht die Auswirkungen individueller Förderung auf die Schülerinnen und Schüler bildungsferner Familien, sondern der allgemeine Effekt der didaktischen Maßnahme auf den Schulerfolg und ihr Einfluss auf den damit verknüpften Herkunftseffekt. Dabei wird im Folgenden zunächst auf das den Analysen zugrunde liegende Verständnis von Schulerfolg eingegangen, bevor das Modell um eine soziale Facette erweitert und schließlich einer Prüfung unterzogen wird.

## 8. Schulerfolg

Der Schulerfolg von Schülerinnen und Schülern gilt als zentrales Kriterium für die Bewertung unterrichtlicher Wirkungen. Dabei können schul-, unterrichts- und lehrerbezogene Merkmale kognitive Lern- und Leistungszuwächse bei Kindern und Jugendlichen hervorrufen; genau-

---

<sup>35</sup> Diese Wirkung von individueller Förderung konnte bisher auch nicht nachgewiesen werden (vgl. etwa Warwas, Hertel & Labuhn, 2011; Weinert & Helmke, 1996; Lipowsky, 2015).

so aber auch ihre motivational-affektive Entwicklung und andere Facetten von Lernprozessen beeinflussen. Das Ziel guten Unterrichts besteht demnach nicht nur in einem guten Abschneiden bzw. dem Zugewinn in Testleistungen, sondern beispielsweise auch in der Förderung von Selbstvertrauen, Motivation, Interesse und dem Abbau von Leistungsangst. Gerade die Ausbildung eines positiven Fähigkeitsselbstkonzepts erscheint – neben fachlichen Leistungen – einen wichtigen ‚Ertrag‘ von Schule darzustellen, berücksichtigt man, dass dieses nicht nur einen wesentlichen Beitrag für das Erreichen eines (kognitiven) Schulerfolgs leistet, sondern auch schulübergeordnet von Bedeutung ist: So verfestigen sich Selbsteinschätzungen in einer hierarchischen Struktur im Laufe der Zeit zu einer allgemeinen Einschätzung der eigenen Fähigkeiten, die wiederum in einem engen Zusammenhang mit der allgemeinen Lern- und Leistungsbereitschaft und der damit verbundenen Motivation steht. Es ist daher erstrebenswert, Kinder und Jugendliche nicht nur hinsichtlich ihrer schulischen Leistungen zu fördern, sondern auch eine positive Entwicklung ihres Selbstvertrauens anzustreben. Im Folgenden sollen entsprechend beide Dimensionen – die kognitive wie die affektive Komponente dessen, was Schulerfolg ausmacht, berücksichtigt werden. Dabei wird sich auf die Leseleistung und das Leseselbstkonzept der Schülerinnen und Schüler bezogen, da diese Kernelemente der hier genutzten Internationalen Grundschul-Lese-Untersuchung (IGLU) sind.

### **8.1 Leseleistung und ihre Konzeption in IGLU**

Die Betrachtungen von Schulleistungen orientieren sich auch heute noch an einer Definition von Krapp aus dem Jahr 1976, in welchem diese allgemein als das Ergebnis von Lernprozessen dargestellt werden, die durch Unterrichtsmaßnahmen initiiert oder gesteuert werden. Unterschieden wird in diesem Kontext zudem ein dynamisches Leistungsverständnis, welches sich auf den Lernprozess als solchen fokussiert und welchem ein statisches bzw. (lern-)produktorientiertes Verständnis gegenüber steht (vgl. etwa Beckmann, 1978; Heller, 1984). Bezüglich der Bedeutung von Schulleistungen für empirische Untersuchungen schlägt Krapp (1984) eine weitere Differenzierung in ein hypothetisches und ein empirisches Verständnis von Schulleistungen vor: So stelle das hypothetische Konstrukt die überdauernden, situations- und zeitübergeordneten Fähigkeiten von Schülerinnen und Schülern dar. Krapp (1984) bezeichnet diese als allgemeine Schulleistungsfähigkeit. Bei dem empirisch beobachtbaren Konstrukt handelt es sich hingegen um die konkret und situationsbezogen beobachtbaren Einzelleistungen von Lernenden, die Krapp (1984) als Schulleistungsverhalten bezeichnet – und die allenfalls eine Momentaufnahme der allgemeinen Schulleistungsfähigkeit darstellen

können. Empirische Untersuchungen sind – diesem Verständnis folgend – auf die Messung des Outputs von Lernprozessen beschränkt, vernachlässigen also die dynamische Komponente der Schulleistung (Nikolova, 2011). Internationale Studien wie PISA, TIMSS und IGLU, die im Rahmen des Konstanzer Beschlusses (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 1997) auch in Deutschland etabliert wurden, orientieren sich daher am Kompetenz- statt an dem Schulleistungsbegriff. Kompetenzen seien dabei zu verstehen als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2001, S. 27–28). Im Zuge dieser Umorientierung werden in Deutschland seit Beginn der 2000er Jahre bundesweit einheitliche Bildungsstandards definiert und implementiert, aus welchen sich für die unterschiedlichen Schulfächer Kompetenzstufenmodelle<sup>36</sup> ableiten lassen und welche „die fachbezogenen Kompetenzen einschließlich zugrunde liegender Wissensbestände [beschreiben], die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt ihres Bildungsganges erreicht haben sollen“ (Köller, 2004b, S. 12). Sie ermöglichen es, individuelle Leistungsergebnisse innerhalb eines systematischen und fachlich verankerten Konzepts von Kompetenzstufen zu verorten und zu vergleichen (Köller, 2004b).

Dabei lässt sich dieses Verständnis von Schulleistung als Fähigkeit der Schülerinnen und Schüler, ihre Kenntnisse und Fertigkeiten zur Bewältigung realitätsnaher Herausforderungen einzusetzen (Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung, 2001), auch auf das angloamerikanische *literacy*-Konzept (in der Regel als ‚Grundbildung‘ übersetzt, vgl. Baumert, Stanat & Demmrich, 2001) zurückführen, das von der Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung (OECD) adaptiert wurde. Definiert werden kann es als „Verwendung von gedruckten und geschriebenen Informationen, um in der Gesellschaft zurechtzukommen, eigene Ziele zu erreichen und eigenes Wissen sowie die individuellen Möglichkeiten zu entwickeln“ (Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung, 1998, S. 16). Damit geht der *literacy*-Begriff über die Fähigkeit des Lesens und Schreibens hinaus; es finden sich Aussagen wie „being technologically literate, scientifically literate, even politically and socially literate“ (Harlen, 2001, S. 83).

---

<sup>36</sup> Dies geschah bisher vor allem für die Hauptfächer Deutsch, Mathematik und erste Fremdsprache für die unterschiedlichen Schulabschlüsse (vgl. Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2003a, 2003b, 2003c, 2004a, 2004b, 2004c, 2004d, 2004e, 2012a, 2012b, 2012c).

Für PISA wurden drei literacy-Domänen identifiziert, die für das Zurechtkommen in der Wissensgesellschaft bedeutsam sind. Hierzu gehören die Lese-, die mathematische sowie die naturwissenschaftliche Grundbildung, die im dreijährigen Turnus überprüft werden. Ergänzt wird die Befragung Fünfzehnjähriger in PISA beispielsweise durch die Internationale Grundschul-Lese-Untersuchung (IGLU, englisches Akronym: PIRLS), welche die reading literacy bzw. Lesekompetenz bei Viertklässlerinnen und Viertklässlern in 56 Ländern bzw. Bildungssystemen (Stand: Erhebung 2011, Tarelli, Valtin, Bos, Bremerich-Vos & Schwippert, 2012) erfasst. Definiert wird die Lesekompetenz dabei in Anlehnung an die IEA Reading Literacy Study von 1991 als „the ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment“ (Mullis, Martin, Kennedy, Trong & Sainsbury, 2009, S. 11). Dabei wird die Lesekompetenz als zentrale Kulturtechnik und Schlüsselkompetenz für die Teilnahme an nahezu allen gesellschaftlichen Lebensbereichen verstanden als konstruktiver, interaktiver Prozess: Der bzw. die Lesende verfügt über effektive Lesestrategien und die Fähigkeit, das Gelesene zu reflektieren und zu interpretieren. Aus der Interaktion zwischen dem Text und der Person mit ihrer individuellen Leseerfahrung, ihrem Hintergrundwissen und Fertigkeiten konstruiert sie aktiv die Bedeutung des Gelesenen. Relevant ist außerdem die Beschaffenheit des Textes: Er beinhaltet spezifische Sprach- und Strukturelemente, bezieht sich auf ein bestimmtes Thema und entstammt einer bestimmten Quelle, beispielsweise traditionellen Medien wie etwa Bücher, Zeitschriften, Dokumente und Zeitungen oder Informations- und Kommunikationstechnologien wie dem Internet, E-Mails, SMS-Nachrichten und video- oder fernsehbezogenen Medien. Zudem kann der Kontext der Lesesituation das Engagement und die Motivation zu lesen fördern oder hemmen (Bos, Lankes, Prenzel, Valtin & Walther, 2002; Mullis et al., 2009). In IGLU werden drei Bereiche der Lesekompetenz erfasst:

1. Die Leseintention, mit der ein Text gelesen wird; hier wird unterschieden zwischen literarischen und informierenden (Sach-)Texten,
2. die Verstehensprozesse, wobei zwischen textimmanentem und wissensbasiertem Verstehen unterschieden wird und
3. das Leseverhalten und die Einstellungen zum Lesen (Bremerich-Vos, Tarelli & Valtin, 2012).

Erhoben werden die erstgenannten Bereiche in IGLU anhand eines Lesetests, der wiederum beide Leseintentionen abdeckt: Eingesetzt werden zum einen narrativ-fiktionale Texte bzw. Kurzgeschichten sowie Testfragen, die die literarischen Erfahrungen mit fiktionalen Geschehnissen, die Kinder beim Lesen dramatischer Texte machen, abbilden. Hierzu gehören etwa Nachfragen zu den Gefühlen und Gedanken der Protagonisten sowie zu Aspekten der Handlung, der räumlichen und zeitlichen Gestaltung. Zudem lesen Kinder, um Informationen zu gewinnen. Hierfür werden faktuale bzw. Sachtexte aufgenommen, die chronologisch bzw. logisch (z.B. Berichte, Instruktionen) oder diskontinuierlich (z.B. Listen, Tabellen, Grafiken) aufgebaut sein können (Bremerich-Vos et al., 2012). Insgesamt beziehen sich die Testaufgaben in IGLU auf vier Verstehensprozesse:

- explizit angegebene Informationen lokalisieren
- einfache Schlussfolgerungen ziehen
- komplexe Schlussfolgerungen ziehen, interpretieren und kombinieren
- Inhalt und Sprachgebrauch prüfen und bewerten

Dabei werden die ersten zwei Verstehensleistungen als textimmanent, die letzten beiden als wissensbasiert zusammengefasst (Bremerich-Vos et al., 2012).

Das Leseverhalten und die Einstellungen zum Lesen als dritter der oben genannten Bereiche der Lesekompetenz werden mit Hilfe von Rahmenfragebögen erhoben. Hierüber wird auch das lesebezogene Selbstkonzept der Schülerinnen und Schüler (in der vorliegenden Ausarbeitung verstanden als Merkmal affektiven Schulerfolgs) erfasst, welches im Folgenden näher erläutert wird.

## **8.2 Leseselbstkonzept und seine Konzeption in IGLU**

Wie dies eingangs bereits erläutert wurde, stellt nicht nur die schulische Leistung von Schülerinnen und Schülern einen Erfolgsindikator guten Unterrichts dar. In einer Analyse von Schulgesetzen, Richtlinien und Lehrplänen deutscher Bundesländer zeigt Gruehn (1995), dass auch affektive Wirkungen von Unterricht und Lehrerhandeln auf Seite der Schülerinnen und Schüler in der Schulpraxis als wichtige Lernergebnisse bewertet werden. So werden als Ziele des allgemeinen schulischen Bildungsauftrags nicht nur die Vermittlung grundlegender Kenntnisse, Fähig- und Fertigkeiten, sondern auch Hilfen zur Entwicklung einer mündigen und sozial verantwortlichen Persönlichkeit, die in der Lage ist, selbstbestimmt und selbstverantwortlich zu lernen, aufgeführt (vgl. exemplarisch Kultusministerium des Landes Nord-

rhein-Westfalen, 1989 und 1993 sowie Gruehn, 1995). Hieraus werden spezifischere Ziele wie „die Förderung von Selbstbewusstsein und Selbstvertrauen, die Entwicklung von intrinsischer Handlungsorientierung sowie interessierter und lustvoller Auseinandersetzung mit fachlichen Inhalten [abgeleitet], die neben der konkreten, fachleistungsbezogenen Wissensaneignung nichtkognitive Aspekte des Lernprozesses“ (Gruehn, 1995, S. 532) abbilden. Und auch der aktuelle Bildungsplan für Hamburger Grundschulen legt beispielsweise einen Schwerpunkt auf überfachliche Kompetenzen „als gemeinsame Aufgabe und Ziel aller Unterrichtsfächer sowie des gesamten Schullebens“ (Freie und Hansestadt Hamburg & Behörde für Schule und Berufsbildung, 2011, S. 10). Dabei werden die Bereiche Selbstkonzept und Motivation, soziale Kompetenzen und lernmethodische Kompetenzen unterschieden. So sollen die Schülerinnen und Schüler Vertrauen in ihre Fähigkeiten entwickeln, lernen selbstkritisch zu sein, ihre Meinung zu vertreten sowie sich Ziele zu setzen und sie zu verfolgen. Ebenso steht der angemessene Umgang mit anderen und hier insbesondere die Fähigkeit zu kommunizieren, zu kooperieren, Rücksicht zu nehmen sowie Hilfe zu leisten und sich in Konflikten angemessen zu verhalten im Fokus. Zu den lernmethodischen Kompetenzen gehören die Förderung eines systematischen, zielgerichteten Lernens sowie die Fähigkeit, Strategien und Medien zur Beschaffung und Darstellung von Informationen zu nutzen (vgl. Freie und Hansestadt Hamburg & Behörde für Schule und Berufsbildung, 2011). Dass affektive und motivationale Persönlichkeitsmerkmale für schulische Lernprozesse von Bedeutung sind, spiegeln auch die Erkenntnisse der empirischen Schul- und Unterrichtsforschung wider. So konnten Sauer und Gattringer (1986) mit Hilfe von multiplen Regressionsanalysen zeigen, dass durch motivationale Aspekte, hier insbesondere das Ausmaß, Erfolge der eigenen Fähigkeit und Misserfolge nicht mangelnder eigener Fähigkeit zuzuschreiben (Attribuierungsvoreingenommenheiten), etwa 9% der Varianz eines Schulleistungstests bei Grundschulern erklärt werden – gegenüber 35% aufgeklärter Varianz durch einen mehrfaktoriellen Intelligenztest und 10% durch die Messung der kognitiven Leistungs- und Verarbeitungsgeschwindigkeit. Hingewiesen wird jedoch auf eine mögliche Konfundierung kognitiver und motivationaler Merkmale insbesondere bei der Intelligenzmessung, auf die auch Quack (1979) verweist: So läge der Richtwert des durch Intelligenz aufgeklärten ‚reinen‘ Anteils an Schulleistungsvarianz bei etwa 25 bis 30% und bei etwa 50% für den mit nicht-kognitiven Variablen unkontrolliert konfundierten Anteil. Anhand multipler Regressionsanalysen konnte nachgewiesen werden, dass die Schulleistung im Fach Deutsch der fünften Klasse mit ähnlicher Ge-

naugigkeit durch kognitive Variablen (49%, konfundiert) und nicht-kognitive Prädiktoren (46%), zu denen etwa das Selbstkonzept der eigenen Leistung gehören, vorhergesagt werden kann (Quack, 1979). Das leistungsbezogene Selbstkonzept spielt auch in weiteren Studien zu schulischen Lern- und Leistungsprozessen eine zentrale Rolle. Sein Einfluss auf den Schulerfolg liegt dabei üblicherweise zwischen  $r = 0,40$  und  $r = 0,60$  und kann daher als moderat bezeichnet werden (Guay, Marsh & Boivin, 2003; Skaalvik & Valas, 1999; Marsh, Parker & Smith, 1983). Höhere Korrelationen zeigen sich insbesondere dann, wenn fachspezifische Fähigkeitsüberzeugungen und Noten statt Leistungstests für die Überprüfung des Zusammenhangs herangezogen werden (Schicke & Fagan, 1994; Valentine, DuBois & Cooper, 2004; Weber, 2012). Relevant ist die Ausprägung des Selbstkonzepts für die schulischen Leistungen, da sie leistungsbezogenes Verhalten initiieren und aufrechterhalten kann. Niedrige Fähigkeitsüberzeugungen führen nach Helmke und Weinert (1997b) zu der Entwicklung von Vermeidungsstrategien, mit Hilfe derer Schülerinnen und Schüler versuchen, sich vor Misserfolgen zu schützen um auf diese Weise den Selbstwert zu stabilisieren. Es erscheint daher sinnvoll, der Leseleistung als kognitiven Indikator für den Schulerfolg das (lesebezogene) Selbstkonzept der Schülerinnen und Schüler heranzuziehen, um hiermit die affektive Komponente des Konstrukts abdecken zu können.

Definiert wird das Selbstkonzept als „geordnete Menge aller im Gedächtnis gespeicherten selbstbezogenen Informationen“ oder als „organisiertes Wissen über die eigene Person“ (Filipp, 1979; Markus & Wurf, 1987, zitiert nach Jerusalem & Schwarzer, 1991, S. 115). Diese Forschungstradition fand ihren Ursprung in der 1890 von James vorgestellten Theorie des Selbst (vgl. James, 1890). Dieser unterscheidet zwei Bestandteile des Selbst: zum einen das *Selbst als Subjekt des Erkennens (Self as knower)* und zum anderen das *Selbst als Objekt der Erkenntnis (Self as known)*, ausgehend von der Tatsache, dass der Mensch zur Selbstreflexion fähig ist. Bis heute schließen sich hieran zahlreiche Untersuchungen und theoretische Vertiefungen an, die sich vorrangig mit dem Self as known und dort insbesondere mit dem Wissen über die eigenen Fähigkeiten beschäftigen. „Vorstellungen über die Höhe eigener Fähigkeiten (Fähigkeitsselbstkonzepte) sind ein zentraler Inhalt selbstbezogenen Wissens, weil Personen Fähigkeiten häufig als wichtige Eigenschaften des Selbst empfinden: Man ist in der Regel lieber fähig als unfähig“ (Dickhäuser, 2006, S. 5). Darüber hinaus werden Fähigkeitssselbstkonzepte nicht nur für sich, sondern auch als über unterschiedliche vermittelnde Fak-

toren auf Lernprozesse begünstigend oder hemmend wirkend betrachtet (vgl. Dickhäuser, 2006).

Unter den verschiedenen Modellen zur Erklärung und Darstellung des Selbstkonzepts hat sich das hierarchische Selbstkonzeptmodell von Shavelson, Hubner und Stanton (1976) bewährt. Diese definieren das Selbstkonzept zunächst sehr allgemein als „a person’s perception of himself. These perceptions are formed through his experience with his environment, [...] and are influenced especially by environmental reinforcements and significant others“ (Shavelson et al., 1976, S. 411). Dabei wird von einem reziproken Zusammenhang zwischen Selbstkonzept und allgemeinen Verhaltensweisen ausgegangen: Die Selbstwahrnehmung wirkt sich auf die Verhaltensweisen aus, während diese wiederum die Selbstwahrnehmung beeinflussen. Shavelson, Hubner und Stanton (1976, S. 411) beschreiben das Selbstkonzept als „organized, multifaceted, hierarchical, stable, developmental, evaluative, differentiable“, worauf im Folgenden näher eingegangen wird<sup>37</sup>.

Als strukturiert (organized) beschreiben Shavelson et al. (1976) das Selbstkonzept, da unterschiedliche individuelle Erfahrungen, welche zur Selbstwahrnehmung beitragen, in weniger komplexe Kategorien reduziert bzw. rekodiert werden, die in ein facettenreiches Kategoriensystem (multifaceted) aufgenommen werden. Diese Facetten können dabei auf Individual- als auch auf Gruppenebene geteilt werden. Zudem sei das Selbstkonzept hierarchisch (hierarchical) aufgebaut mit den Facetten individueller Erfahrungen auf der untersten bis hin zu dem allgemeinen Selbstkonzept (general self-concept) auf der höchsten Ebene. Dieses sei zudem stabil (stable), wobei mit abnehmender Hierarchie eine Abnahme der Stabilität einhergeht, da in den niederen Ebenen das Selbstkonzept von Situation zu Situation variieren kann. Dennoch werden diese unteren Ebenen vom allgemeinen Selbstkonzept beeinflusst, so dass eine gewisse Resistenz gegenüber Veränderungen anzunehmen ist: „To change general self-concept, many situation-specific instances, inconsistent with general self-concept, would be required“ (Shavelson et al., 1976, S. 414). Bei dem Selbstkonzept handelt es sich also nicht um ein unveränderliches Persönlichkeitskonstrukt; viel eher unterliegt es einer konstanten (wenn auch nicht beliebigen) Entwicklung (developmental), die sich insbesondere im Kindesalter vollzieht. „With increasing age and experience [...] self-concept becomes increasingly differentiated“ (Shavelson et al., 1976, S. 414). Kleinkinder können beispielsweise noch

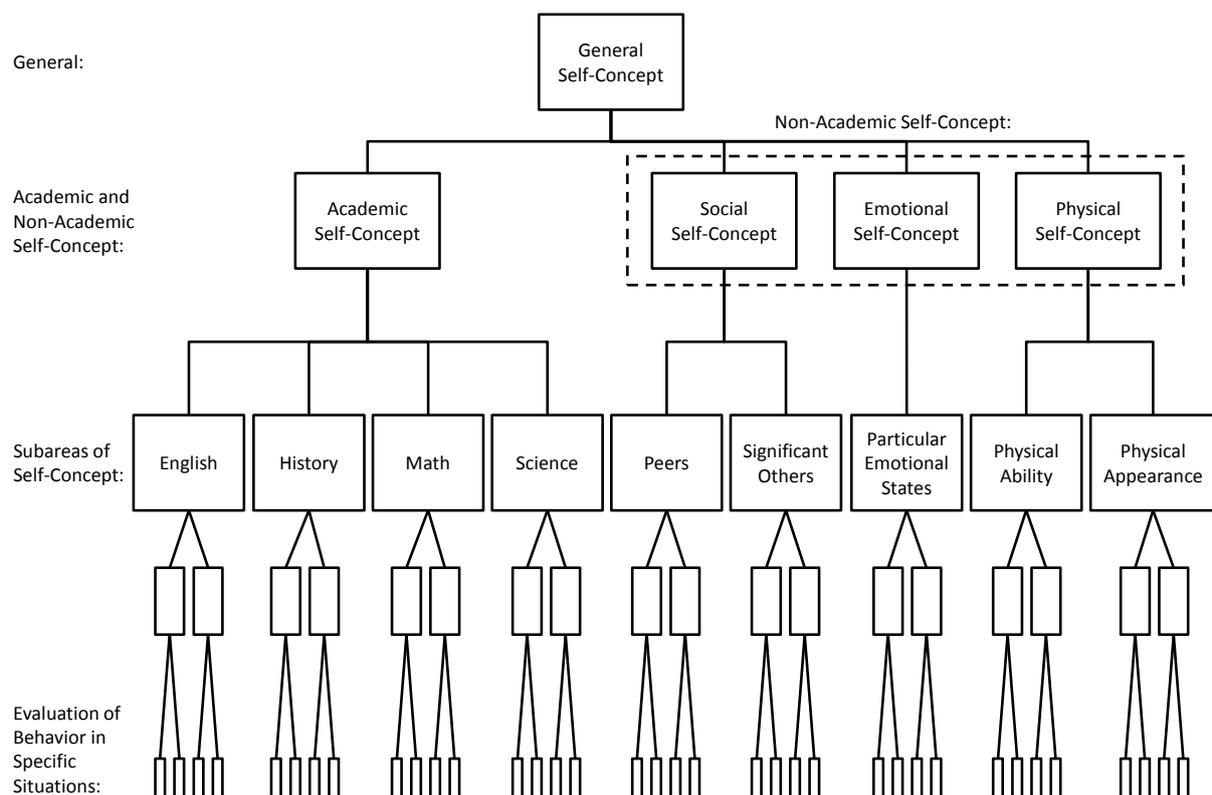
---

<sup>37</sup> Dieser Abschnitt orientiert sich an den Darstellungen von Schulz-Heidorf (2012).

nicht zwischen sich selbst und ihrer Umwelt unterscheiden, entwickeln später zunächst ein globales, undifferenziertes und situationsspezifisches Selbstkonzept und schließlich das hierarchische Selbstkonzept wie es hier vorgestellt wird. Letzteres hat nach Shavelson et al. (1976) auch bewertenden (evaluative) Charakter. So gelangt das Individuum in einer spezifischen Situation nicht nur zu einer Selbstbeschreibung (description of himself), sondern auch zu einer Selbstbewertung. Dabei kann diese Bewertung „against absolute standards, such as the ‘ideal,’ [sic!] [...] against relative standards such as ‘peers’ or perceived evaluations of ‘significant others’” (Shavelson et al., 1976, S. 414) erfolgen. Zuletzt beschreiben die Autoren das Selbstkonzept als differenzierbar (differentiable) gegenüber ähnlichen Konstrukten, mit welchem das Modell theoretisch verwandt ist.

Abbildung 4 zeigt das hierarchische Modell von Shavelson und Kollegen (1976).

**Abbildung 4:** Hierarchisches Modell des Selbstkonzepts nach Shavelson et al. (1976)

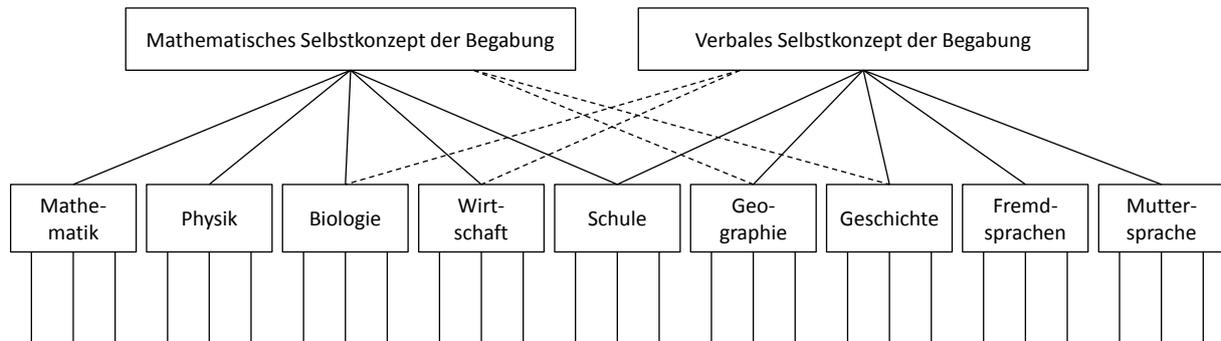


Quelle: Shavelson et al., 1976, S. 413, adaptiert

Abbildung 4 verdeutlicht, dass das ursprüngliche Modell von Shavelson und Kollegen (1976) in Bezug auf die schulische Komponente lediglich ein allgemeines akademisches Selbstkonzept vorsah, in welches auf untergeordneten Ebenen unterrichtsfachspezifische Selbstkonzepte integriert waren. Anhand späterer Analysen konnten Marsh et al. (1988) hingegen zeigen, dass das verbale und das mathematische Selbstkonzept kaum miteinander korrelier-

ten, so dass diese auf der Ebene globaler schulischer Faktoren als getrennte schulische Selbstkonzepte eingehen (vgl. Abbildung 5).

**Abbildung 5:** Struktur des schulischen Selbstkonzepts im revidierten Modell



Quelle: Möller & Trautwein, 2009, S. 188, adaptiert

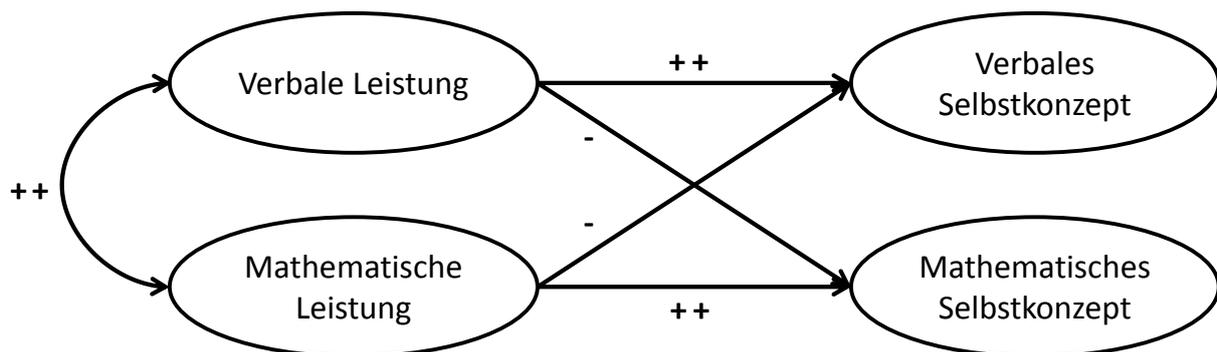
Der Befund, nach welchem das verbale und mathematische Selbstkonzept nicht miteinander zusammenhängen, erscheint zunächst nicht trivial, konnte doch gezeigt werden, dass die beiden Leistungen in diesen Domänen (Muttersprache/Fremdsprachen und Mathematik/Physik) deutlich positive Korrelationen zwischen  $r = 0,50$  und  $r = 0,80$  aufweisen (vgl. Köller & Möller, 2001; Köller, Klemmert, Möller & Baumert, 1999; Marsh, 1986). Darüber hinaus bestehen *innerhalb* einer Domäne üblicherweise ebenfalls positive Korrelationen zwischen  $r = 0,40$  und  $r = 0,60$  (vgl. Köller & Möller, 2001; Köller & Möller, 2000; Hansford & Hattie, 1982). Dass sich dennoch kein Zusammenhang zwischen dem mathematischen und verbalen Selbstkonzept zeigt, veranlasste Marsh, Byrne und Shavelson (1988) zur Entwicklung des Internalen/Externalen Bezugsrahmen-Modells (*Internal/External Frame of Reference-Model*, kurz I/E-Modell). Demnach sind für die Bildung fachspezifischer Selbstkonzepte zwei Informationsquellen von Bedeutung: Bei dem *external* frame of reference handelt es sich um interindividuelle bzw. soziale Vergleiche, bei denen die eigene Leistung in einem Schulfach mit den Leistungen der Mitschülerinnen und Mitschüler verglichen wird. Insbesondere Marsh (1984, 1987) prägte dies unter dem Begriff ‚Big-fish-little-pond-effect‘: Erzielt ein Schüler bzw. eine Schülerin („fish“) bessere Leistungen als die Mitschülerinnen und Mitschüler („pond“), wirkt sich dies positiv auf das Selbstkonzept in diesem Fach aus. Bei Misserfolgen bzw. vergleichsweise geringeren Leistungen ergeben sich entsprechend negative Effekte auf das fachspezifische Selbstkonzept. Eine Folge dessen wäre, dass ein Schüler mit einer durchschnittlichen Leistung in einer leistungsschwachen Klasse ein hohes Selbstkonzept entwickelt, während derselbe Schüler in einer leistungsstarken Klasse ein niedriges Selbstkonzept

entwickeln würde. Dem sozialen Auf- bzw. Abwärtsvergleich kommt daher eine tragende Rolle in der Ausbildung von Fähigkeitsselbstbildern zu.

Der *internal* frame of reference als intraindividuellem bzw. dimensionalem Vergleich wird hingegen für Vergleiche der Leistungsergebnisse eines Faches mit den *eigenen* Leistungen eines anderen Faches herangezogen (vgl. Köller & Möller, 2001; Marsh et al., 1988). Wenn ein Kind also bessere Leistungen in Deutsch als in Mathematik zeigt, wird es nach dieser Überlegung ein besseres verbales als ein mathematisches Selbstkonzept aufweisen, selbst dann, wenn es gemessen an seinen Mitschülerinnen und Mitschülern (oder auch an objektiven Standards) nicht besonders erfolgreich in Deutsch ist.

Dieses Muster der Beziehungen zwischen verbalem und mathematischem Selbstkonzept ist in Abbildung 6 dargestellt.

**Abbildung 6:** Typisches Ergebnismuster in Pfadanalysen zu der Untersuchung des internalen frame-of-reference-Modells



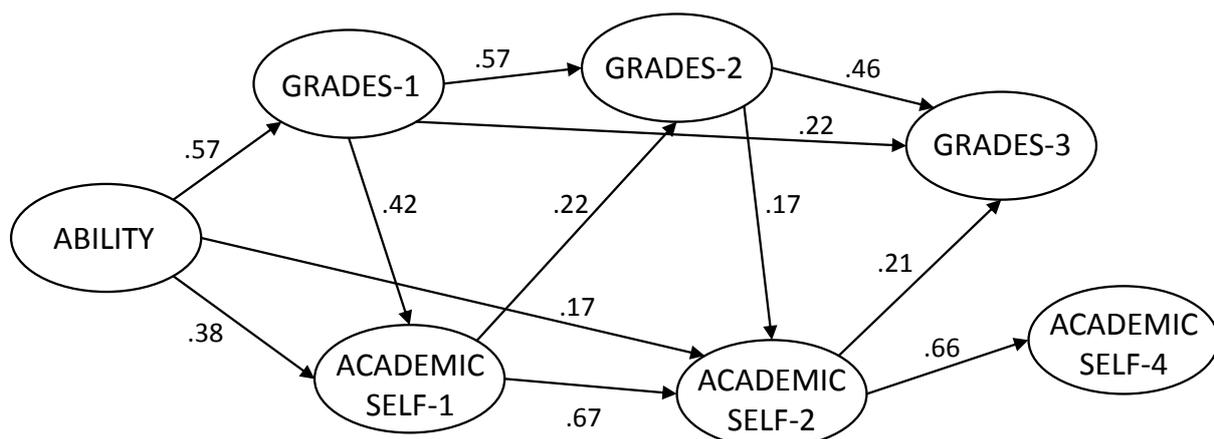
Quelle: Köller & Möller, 2001, S. 382, adaptiert

Nachgewiesen werden konnte die Struktur des in Abbildung 6 dargestellten Modells in einer Vielzahl von Untersuchungen (vgl. etwa Möller & Trautwein, 2015; Marsh & Köller, 2004; Dickhäuser, 2003; Möller & Köller, 2001; Köller & Möller, 2001; Köller, 2000; Köller et al., 1999), wobei die positiven Effekte der Leistung auf das Selbstkonzept *innerhalb* einer Domäne deutlich stärker ausfallen als die negativen Effekte *zwischen* den Domänen (in Abbildung 6 durch ‚+‘ und ‚-‘ dargestellt).

Dies unterstützen auch die Befunde zahlreicher weiterer Untersuchungen zu dem Zusammenhang zwischen Selbstkonzept und Leistung, wobei über die Richtung dieses Zusammenhangs nach wie vor Uneinigkeit besteht (Green, Nelson, Martin & Marsh, 2006; Marsh & Yeung, 1998; Abt Gürber, 2011). Der Frage nachgehend, ob sich fähigkeitsbezogene Selbsteinschätzungen auf künftige Schulleistungen auswirken oder ob es diese Leistungen sind, die

die Selbstkonzepte beeinflussen, werden grundsätzlich drei Ansätze unterschieden. Der *Skill-Development-Ansatz* geht davon aus, dass „fachbezogene Selbstkonzepte von schulischen und außerschulischen Rückmeldungen beeinflusst werden, dass also Leistungen ursächlich für Selbstkonzepte sind“ (Möller & Trautwein, 2009, S. 197). Der *Self-Enhancement-Ansatz* hingegen nimmt an, dass das Selbstkonzept die Leistungen beeinflusst. Marsh (2003) kritisiert diese „overly simplistic [...] models“ (Marsh, 2003, S. 13), die lediglich ‚entweder-oder‘-Schlüsse zuließen und favorisiert einen *reziproken Ansatz*, der von einer wechselseitigen Beziehung und gegenseitigen Beeinflussung der Konstrukte ausgeht: „Prior academic self-concept has a positive effect on subsequent achievement beyond what can be explained in terms of prior academic achievement. Subsequent academic self-concept is also affected by priori achievement beyond what can be explained in terms of prior academic self-concept“ (Marsh, 2003, S. 2). Nachweisen konnte Marsh dies in einer Längsschnittstudie mit 1.456 U.S.-amerikanischen Schülerinnen und Schülern, die sich zu den vier Messzeitpunkten in Klasse 10, 11 und 12 befanden und ebenfalls ein Jahr nach ihrem Schulabschluss befragt wurden. Als Leistungsindikatoren wurden dabei sowohl die Ergebnisse standardisierter Tests als auch die Schulnoten der Jugendlichen herangezogen (vgl. Marsh, 1990). Abbildung 7 zeigt das berechnete Modell und seine standardisierten Regressionsparameter.

**Abbildung 7:** Strukturgleichungsmodell der Classic Causal Ordering Study von Marsh (1990) zu vier Messzeitpunkten



Anmerkungen: Standardisierte Parameter; nicht signifikante Pfade nicht abgebildet. Zu T3 wurde das akademische Selbstkonzept nicht erhoben.

Quelle: Marsh, 1990, S. 121, adaptiert

Von besonderer Bedeutung für den oben dargestellten Diskurs sind hier die Effekte zwischen den Leistungsindikatoren und dem akademischen Selbstkonzept zwischen den Erhebungszeitpunkten: So wird vor allem deutlich, dass das akademische Selbstkonzept die (über Schulnoten erfasste) Leistung des jeweils nächsten Messzeitpunkts beeinflusst. Und auch der

gegenläufige Effekt lässt sich beobachten: Das akademische Selbstkonzept kann durch die Leistung (erfasst über standardisierte Tests) des vorangegangenen Messzeitpunkts vorhergesagt werden (wobei anzumerken ist, dass sich keine solche Pfade für den Einfluss der Schulnoten auf das Selbstkonzept zu späteren Zeitpunkten finden lassen). Diesen Befunden schließen sich auch Skaalvik und Hagtvet (1990) an, die in einer längsschnittlichen Untersuchung von 635 Schülerinnen und Schülern in zwei Kohorten über zwei Messzeitpunkte (Klasse 3, später 4 und Klasse 6, später 7) beide gegenläufige Effekte nachweisen konnten:

So zeigte sich ein starker Einfluss der zum ersten Messzeitpunkt erhobenen Schulleistung<sup>38</sup> auf das akademische Selbstkonzept zum zweiten Messzeitpunkt für die jüngere Kohorte (Klasse 3, später 4), während sich der umgekehrte Effekt (Selbstkonzept auf Leistung) für diese Gruppe nicht nachweisen ließ. Für die zweite Kohorte konnte ein gegenläufiger Effekt festgestellt werden: „[...] the obtained relationship between self-concept of ability at T1 and academic achievement at T2 in Cohort 72 [Klasse 6, später 7, Anm. KSH] was rather strong, and self-concept of ability was a stronger predictor of academic achievement than vice versa“ (Skaalvik & Hagtvet, 1990, S. 304). Dabei bewegten sich die standardisierten Regressionsparameter zwischen  $\beta = 0,24$  für die Leistung zu T1 als Prädiktor des Selbstkonzepts zu T2 in der jüngeren Kohorte und  $\beta = 0,27$  für die ältere Kohorte, während der Einfluss des Selbstkonzepts zu T1 dort mit  $\beta = 0,49$  deutlich die Leistung zu T2 beeinflusst. Wie beschrieben, konnte dieser Zusammenhang für die jüngere Kohorte nicht gefunden werden. Skaalvik und Hagtvet (1990) interpretieren dies damit, dass das Selbstkonzept gerade in jüngeren Jahren noch variabel ist und sich eher mit Leistungsrückmeldungen verändert als dies für ältere Schülerinnen und Schüler, deren Selbstkonzept bereits größere Stabilität aufweist, der Fall zu sein scheint. Ein gefestigtes Selbstkonzept stünde dann zudem in einem Zusammenhang mit leistungsbezogenen Erwartungen und dem Leistungsverhalten, welche die tatsächliche Leistung beeinflussen – dies würde sich in entsprechenden Analysen in einem stärkeren Einfluss des Selbstkonzepts auf die Leistung manifestieren. Die Interpretation der Zusammenhänge für die Sechst- bzw. späteren Siebtklässler führen die Autoren insbesondere auf die

---

<sup>38</sup> Als Schulleistung wird eine durch die Lehrperson vorgenommene Kategorisierung der Schülerinnen und Schüler herangezogen: Hierfür wurde diese gebeten, die Kinder und Jugendlichen ihren Leistungen entsprechend in eine von fünf Kategorien (hohe bis niedrige Leistungen) einzuordnen. Dies liegt vor allem darin begründet, dass es in Norwegen, dem Erhebungsland, in den entsprechenden Jahrgängen keine Noten und keine standardisierten Tests gab und diese aufgrund schulindividueller Curricula schwierig durchzuführen gewesen wären.

Einführung von Benotungen in Klasse 7 zurück<sup>39</sup>. Mit ihr einher gehen vermehrte Hausaufgaben und Tests, die eine Verfestigung des akademischen Selbstkonzepts nach sich ziehen könnten. Auch könnten diese Änderungen in den schulischen Anforderungen zu einer stärkeren schulbezogenen Angst führen, die sich insbesondere für Kinder mit niedrigem Selbstkonzept und niedrigen Leistungserwartungen negativ auf das Leistungsverhalten und entsprechend die tatsächlichen Schulleistungen auswirken könnte. Einschränkend verweisen die Autoren jedoch auch darauf, dass die Noteneinführung in Klasse 7 die Vergleichbarkeit der beiden Kohorten mindert. Diese Befunde, nach welchen für jüngere Jahrgänge aufgrund des weniger gefestigten Selbstkonzepts eher der Skill-Development-Ansatz, für ältere Schülerinnen und Schüler hingegen reziproke Zusammenhänge nachgewiesen werden konnten, konnten dennoch repliziert werden (vgl. etwa Muijs, 1997). Wigfield und Karpathian (1991) kommen in einem Review bestehender Forschungsergebnisse bezüglich des Zusammenhangs schulischer Leistung und akademischem Selbstkonzept zu dem Schluss, dass „[o]nce ability perceptions are more firmly established the relation likely becomes reciprocal: Students with high perceptions of ability would approach new tasks with confidence, and success on those tasks is likely to bolster their confidence in their ability“ (Wigfield & Karpathian, 1991, S. 255).

Wenngleich weitere klärende Untersuchungen in diesem Bereich notwendig sind, kann zusammenfassend festgestellt werden, dass sowohl das akademische Selbstkonzept als auch die schulische Leistung als erklärungs mächtige Variablen für die Ausprägung des jeweiligen Gegenkonstrukts gelten. Zudem scheint es vermehrt Hinweise darauf zu geben, dass ihnen ein reziproker Zusammenhang zugrunde liegt (der ggf. in jüngeren Jahren zugunsten eines stärkeren Einflusses der schulischen Leistung auf die Ausbildung und Verfestigung des akademischen Selbstkonzepts ausfällt). Da für die folgenden Analysen keine Längsschnittdaten vorliegen, die Aussagen über gerichtete kausale Beziehungen zulassen, soll der Zusammenhang zwischen Leseselbstkonzept und Leseleistung korrelativ analysiert und entsprechend als reziprok interpretiert werden. Eine Beschreibung dessen, wie das Leseselbstkonzept anhand der hier verwendeten IGLU-Daten modelliert werden kann, findet sich in Kapitel 12.3.4.

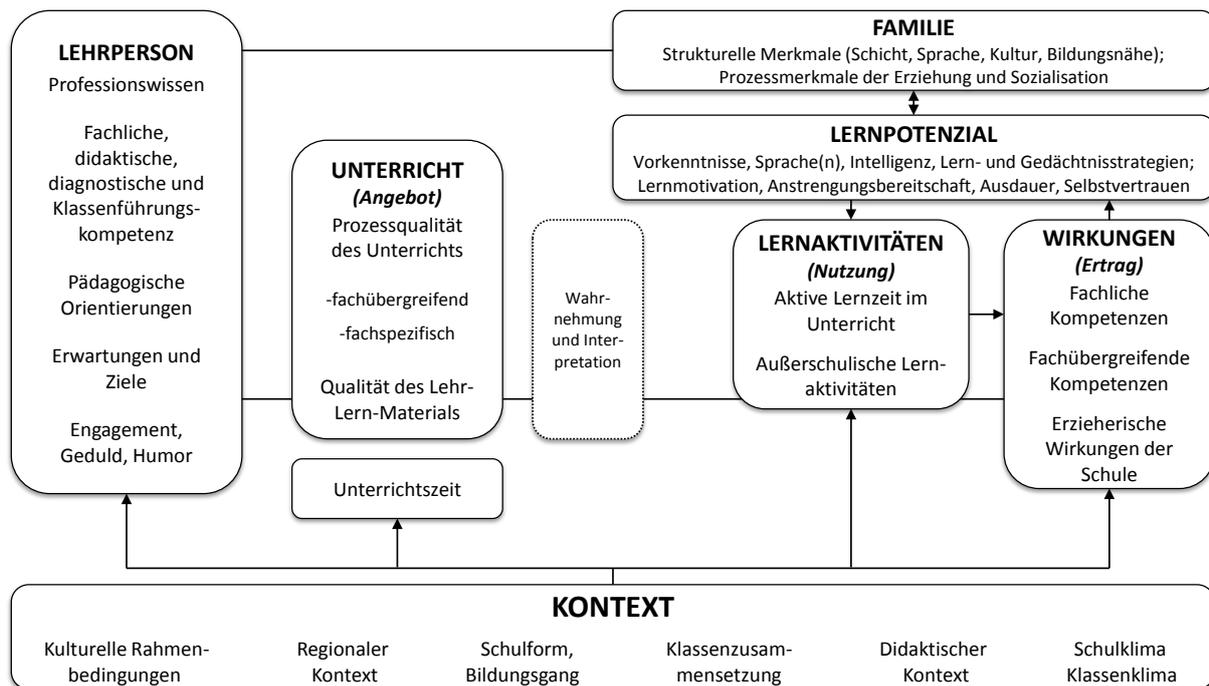
---

<sup>39</sup> Die Studie wurde in Norwegen durchgeführt. Mittlerweile werden dort Noten in Klasse 8 eingeführt.

### **8.3 Rahmenmodelle des Schulerfolgs**

Oben wurde dargestellt, dass sich Schulerfolg nicht nur in Leistungen ausdrückt, sondern unter anderem auch in der Entwicklung eines gestärkten Selbstkonzepts. Dass diese allerdings nicht linear und monokausal auf den durch die Lehrkraft angebotenen Unterricht zurückgeführt werden können, beschreibt beispielsweise Helmke (2006) in seinem Angebot-Nutzungs-Modell des Unterrichts (vgl. Abbildung 8, auch Kapitel 2). Hier kommt den individuellen Voraussetzungen („Lernpotenzial“) der Schülerinnen und Schüler eine besondere Bedeutung zu: „Ob, wie lange und wie erfolgreich jemand lernt und was er leistet, hängt ganz wesentlich von den kognitiven, motivationalen und volitionalen Lernvoraussetzungen ab: Intelligenz, Vorkenntnisniveau, Lernstrategien [...], Leistungsangst, Lernmotivation und Lernemotion“ (Helmke, 2009, S. 82). Dass diese in besonderem Maße durch das elterliche Erziehungs- und Unterstützungsverhalten sowie den sozioökonomischen Hintergrund der Familie beeinflusst werden, konnte bereits in Kapitel 7 gezeigt werden. Für das Unterrichtsangebot und seine Nutzung bedeutet dies vor allem eine Abhängigkeit von individuellen Wahrnehmungs- und Interpretationsprozessen auf Seiten der Schülerinnen und Schüler. Darüber hinaus sind nach Helmke alle schulischen Settings (vor-)strukturiert durch Kontextmerkmale, zu denen etwa regionale und kulturelle Gegebenheiten, die Klassenzusammensetzung sowie die Wahrnehmung des Schul- und Klassenklimas gehören.

**Abbildung 8:** Erweitertes Angebot-Nutzungs-Modell des Unterrichts



Quelle: Helmke, 2006, S. 262

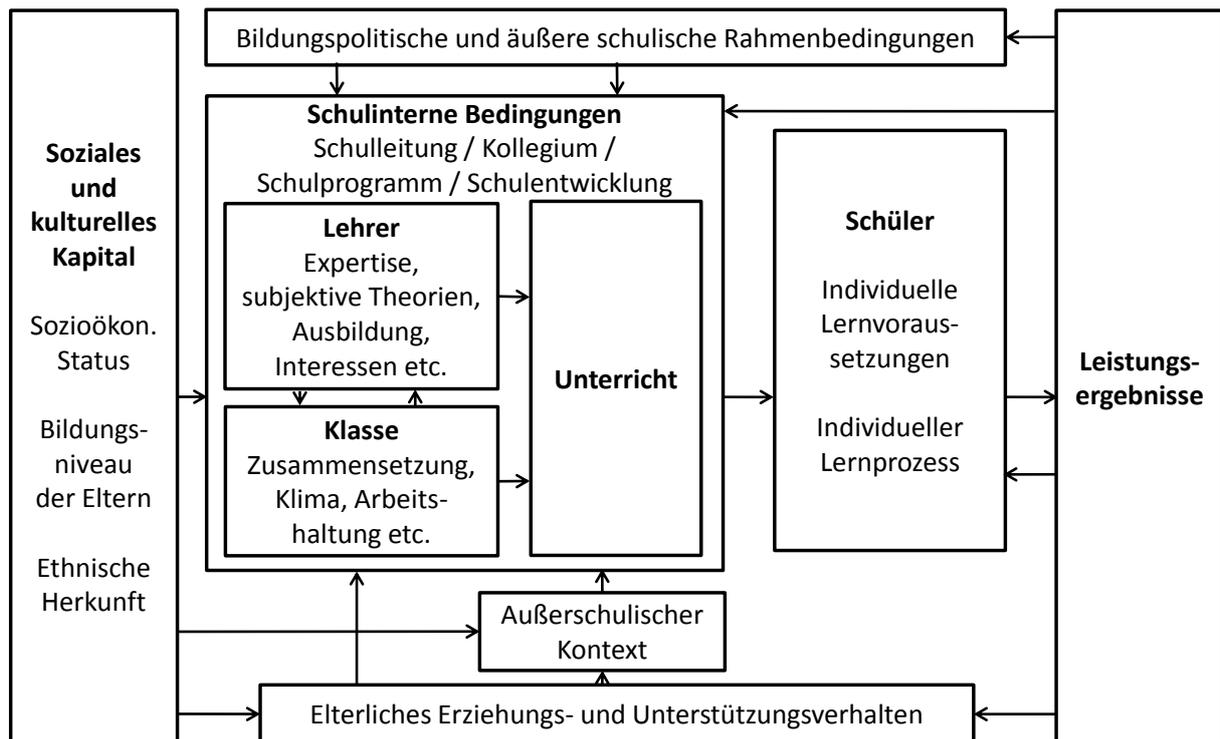
Auch der Vergleich mit weiteren Modellen schulischer Wirkungsweisen (etwa Carroll, 1963; Harnishfeger & Wiley, 1975; Bloom, 1976; Atkinson, 1974; Helmke & Weinert, 1997b; Ditton, 2000; Scheerens, 1990; Gerecht, Steinert, Klieme & Döbrich, 2007; Slavin, 1996; Stringfield, 1994) verdeutlicht, dass Schulerfolg nicht nur als Lehr-Lern-Prozess zwischen Lehrkräften und Schülerinnen bzw. Schülern verstanden wird, sondern – je nach Dimension<sup>40</sup> – als Resultat eines komplexen Wirkungsgefüges aus Lernvoraussetzungen, sozialer Herkunft, familiärem Anregungsgehalt, schulischen Rahmenbedingungen, Merkmalen der Klassenzusammensetzung, der Wahrnehmung der sozialen Lernumwelt und weiterer schulischer wie außerschulischer Faktoren.

Dies spiegelt sich ebenfalls in dem Rahmenmodell der Internationalen Grundschul-Lese-Untersuchung (IGLU) wider, deren Daten für die folgend dargestellten Analysen herangezogen werden (vgl. Abbildung 9). Ähnlich wie bei Helmke stellen die individuellen Lernvoraus-

<sup>40</sup> Ditton (2000) unterscheidet zwei Dimensionen, anhand welcher sich die Modelle der Schuleffektivität differenzieren lassen. In der strukturellen Dimension wird das schulische Wirkungsgefüge als Mehrebenensystem wahrgenommen, in welchem individuelle, unterrichtliche und schulische Merkmale in einen sozial-regionalen Kontext eingebettet sind, wobei diese höheren Ebenen den Rahmen und die Bedingungen für Wirkmechanismen auf den unteren Ebenen stellen (vgl. auch Creemers, Scheerens & Reynolds, 2000). Die dynamische Dimension fokussiert die Wirkungsweisen des Schulwesens unter Einbeziehung der individuellen Eingangsvoraussetzungen der Schülerinnen und Schüler. Die meisten aktuellen Modelle, darunter auch das IGLU-Rahmenmodell sowie Helmkes Angebot-Nutzungs-Modell verknüpfen beide Dimensionen, meist jedoch mit unterschiedlichen Schwerpunkten (vgl. auch Thillmann, 2012).

setzungen und Lernprozesse des Schülers bzw. der Schülerin den zentralen Bestandteil des Modells dar. Sie stehen in einem direkten Zusammenhang mit den Leistungsergebnissen und werden durch schulische Bedingungen beeinflusst. Zu diesen zählt die Lehrperson mit ihren didaktischen Kompetenzen, ihrer Ausbildung, Interessen sowie ihrer Expertise ebenso wie die Klasse, deren Zusammensetzung, Klima, Arbeitshaltung und ähnliches einen Einfluss auf den Lehrer bzw. die Lehrerin haben, in gleichem Maße aber auch durch sie bzw. ihn beeinflusst werden kann. Beides bedingt den Fachunterricht und seine Qualität. Zudem zählen das Schulprogramm und die Schulentwicklung sowie die Zusammensetzung des Lehrerkollegiums und der Schulleitung zu den Facetten schulinterner Bedingungen. Familiäre Merkmale wie etwa das elterliche Erziehungs- und Unterstützungsverhalten sowie sozioökonomische und kulturelle Aspekte werden ebenso wie institutionelle Merkmale im Sinne von bildungspolitischen und äußeren schulischen Rahmenbedingungen in das Modell einbezogen.

**Abbildung 9:** IGLU-Rahmenmodell für den Zusammenhang zwischen Schülerleistungen und deren Bedingungen



Quellen: Bos & Lankes et al., 2003, S. 16; Hornberg, Bos, Buddeberg, Potthoff & Stubbe, 2007, S. 22; Tarelli, Wendt, Bos & Zylowski, 2012, S. 35

Das in Abbildung 9 visualisierte komplexe Modell erlaubt bereits einen Einblick in die Vielschichtigkeit der Studie, deren Kernelement – die Feststellung des Leseverständnisses von Viertklässlerinnen und Viertklässlern – ergänzt wird durch Rahmenfragebögen, die eine Befragung der unterschiedlichen in Schule involvierten Parteien und entsprechend die Analyse

komplexer schulbezogener Zusammenhänge, wie sie in dem oben abgebildeten Rahmenmodell dargestellt werden, ermöglichen. So können durch den *Schülerfragebogen* vor allem Lernaktivitäten, fachspezifische Einstellungen, Selbstkonzepte und leistungsbezogene Verhaltensweisen erfasst werden ebenso wie der familiäre Hintergrund und soziodemografische Merkmale der Kinder. Ergänzt wird dieser durch einen *Elternfragebogen* zu familiären Ressourcen, dem Bildungshintergrund der Eltern, Lernaktivitäten mit dem Kind inklusive Angaben zu der familiären Lesesozialisation und vertiefenden soziodemografischen Merkmalen. Dem *Fragebogen für die Deutschlehrkräfte* können neben Personenmerkmalen u.a. die Einstellungen zu unterrichtsrelevanten Aspekten, ihre Unterrichtspraxis und Lehrmethoden sowie lernrelevante Voraussetzungen der Schülerinnen und Schüler entnommen werden. Im *Schulfragebogen* stellen die Schulleitungen Informationen etwa über allgemeine Schulcharakteristika, den Kontext und die Ausstattung der Schule, die Kooperation im Lehrerkollegium und die Einbindung der Eltern bereit. Der *Curriculumfragebogen* wird durch Fachexpertinnen und Fachexperten ausgefüllt und enthält Fragen zu den Rahmenbedingungen der Umsetzung des Curriculums, also etwa über Entscheidungsinstanzen, mögliche Evaluationen der Implementation und Unterstützungsmöglichkeiten für die Umsetzung der Curricula durch die Lehrerinnen und Lehrer (Tarelli & Wendt et al., 2012). Somit besteht mit IGLU ein umfassender internationaler Datensatz, der es ermöglicht, die Lesekompetenz und das Leseselbstkonzept als kognitive wie affektive Facette des Schulerfolgs von Viertklässlerinnen und Viertklässlern in Verbindung zu bringen mit schulischen und familiären Kontextmerkmalen und Lernbedingungen, die zu der Erklärung von Leistungsunterschieden beitragen können. Auf diese Möglichkeit wird im folgenden Kapitel vertiefend eingegangen.

## **9. Wahrnehmung der schulischen Lernumwelt**

Neben der beschriebenen Verankerung der Förderung nicht-kognitiver Merkmale wie etwa das Selbstkonzept und soziale Kompetenzen der Schülerinnen und Schüler in Schulgesetzen und Schulprofilen (vgl. Kapitel 8.2), nimmt das individuelle Wohlbefinden sowie die Ermöglichung nicht-kognitiver persönlicher Entwicklung von Kindern und Jugendlichen in der Diskussion um Schule als Lern- und Lebensort einen immer höheren Stellenwert ein. Als Zielgröße konnte sich dabei vor allem das Konzept der Wahrnehmung der Lernumwelt, welches der Klimaforschung entstammt, etablieren (vgl. etwa Eder, 2010), wenngleich sich für dieses Konstrukt eine Menge heterogener Erklärungsansätze finden lassen. Eder (1996, 2010) un-

terscheidet beispielsweise drei Verwendungsweisen des Begriffs: So charakterisiere dieser zum einen die emotionale Grundtönung einer pädagogischen Gesamtatmosphäre, also die „emotionale Qualität der sozialen Beziehungen zwischen Lehrern und Schülern sowie die Befindlichkeit der ganzen Klasse“ (Eder, 1996, S. 424). Zudem wird der Klimabegriff genutzt für die Beschreibung der in einer erzieherischen Umwelt herrschenden Grundorientierungen und Werthaltungen, zu denen etwa die Schülerkultur, der Schulethos und andere gehören. In einer dritten Verwendungsweise wird unter Klima die subjektiv wahrgenommene (Lern-) Umwelt subsumiert. Dreesmann et al. (1992) hingegen fokussieren bei der Definition individuell oder kollektiv geteilte kognitive Verarbeitungen der Umweltwahrnehmungen der Mitglieder einer Gruppe, während Schreiner (1973) sowie Lange, Kuffner und Schwarzer (1983) die Gesamtheit der sozialen Erfahrungen und schulischen Merkmale in der Wahrnehmung der Schülerinnen und Schüler in den Mittelpunkt stellen. Pekrun (1983) ergänzt diese Modelle um die Voraussetzung, das Klima müsse relativ zeitüberdauernd und somit charakteristisch sein. Saldern (1987, S. 17) fasst diese Ansätze zusammen. So sei „[...] das Sozialklima von Schulklassen als hypothetisches Konstrukt [...] ein auf die Schulklasse bezogenes, relativ überdauerndes, molares und mehrdimensionales Aggregat situativer Wahrnehmung und kognitiver Verarbeitung von situationalen Reizen, das sich in der Beschreibung von Umwelten, Strukturen und Verhalten in der Schulklasse bzw. einem ihrer Subsysteme (z.B. Cliques) durch das Individuum widerspiegelt und die Bildung von Einstellungen zur Lernsituation sowie individuelles Verhalten beeinflusst“. Bei den wesentlichen Merkmalen des Klimabegriffs handelt es sich demnach um die subjektive Wahrnehmung, die Kontinuität, die auf dem relativ stabilen Charakter des Konstrukts beruht, die Molarität, da sich das Klima auf alle Mitglieder einer Gruppe bezieht sowie die Mehrdimensionalität, weil Klima als Oberbegriff für unterschiedliche Subdimensionen dient, etwa Gemeinschaft oder Sozial- und Leistungsdruck (vgl. Janke, 2006).

Festzustellen ist in diesem Zusammenhang, dass die Modelle, Definitionen und Instrumente zur Erfassung des Klimakonstrukts im Schulkontext insbesondere in den 1970er und 1980er Jahren erstellt und weiterentwickelt wurden. Neuere Studien beziehen sich oft auf diese älteren Ansätze; aktuelle Modelle finden sich kaum.

Im Folgenden wird die Entwicklung der Klimaforschung grob umrissen, bevor im Anschluss daran zentrale Kritikpunkte an dem bisherigen Klimakonstrukt ebenso erläutert werden wie

die Konsequenzen, die sich daraus ergeben. Schließlich wird auf einen Bestandteil des Klimabegriffs – der Wahrnehmung der sozialen Lernumwelt – vertiefend eingegangen. Dabei wird sich an den Darstellungen von Schulz-Heidorf (2012) sowie Schulz-Heidorf und Schwip-pert (2014) orientiert.

### **9.1 Die Entwicklung der Klimaforschung**

Ihren Anfang fand die Forschungstradition in Murrays need-press-Modell (1938, nach Eder, 2010). Dieses lässt sich dem interaktionspsychologischen Ansatz zuschreiben und stellt ein sehr basales Modell zur Erklärung individuellen Verhaltens dar: Dieses entstehe durch das Zusammenwirken von individuellen Bedürfnissen (*needs*) und Umwelteinflüssen (*press*), wobei unter *needs* weitestgehend die Ziele und Absichten einer Person verstanden werden, während unter *press* Umwelteinflüsse subsumiert werden, die die Befriedigung individueller Bedürfnisse einer Person behindern oder fördern können (vgl. Janke, 2006). Verhalten ent-stehe dabei, „wenn eine durch *needs* aktivierte Person bestimmte Arten von *press* sucht o-der vermeidet“ (Janke, 2006, S. 25, Herv. i. O.). Unklar bleibt in Murrays Modell, wie das In-dividuum und seine wahrgenommene Umwelt zusammenwirken (vgl. Saldern, 1987).

Weitergeführt wurde dieser in den 1930er Jahren entwickelte Murray'sche Ansatz von Stern (1970), der den College Characteristics Index (CCI) sowie den High School Characteristics Index (HSCI) entwickelte, mithilfe welcher Personen in ihren schulischen Umwelten anhand von 30 need- sowie 30 press-Dimensionen charakterisiert werden können. Dabei unter-scheidet sich dieser Ansatz von Murrays Ausgangsmodell dahingehend, dass die press-Dimension unterteilt wurde in von einem Kollektiv geteilte und eine individuelle Wahrneh-mung der Umwelt. Ziel ist es dabei, das Ausmaß der *needs* und *press* als Beschreibung des Klimas zu erfassen und zu analysieren, was sich praktisch als wenig schulbezogen sowie als schwer handhabbar und interpretierbar erwies (vgl. Eder, 2010; Janke, 2006).

Ebenfalls reiht sich das Modell der Klasse als soziales System von Getzels und Thelen (1960, nach Janke, 2006) in die Weiterentwicklungen des Murray'schen need-press-Modells ein, erweitert es in mehrere Richtungen und stellt die Klasse als soziales System in den Mittel-punkt. Dabei lassen sich die Modellmodifikationen vor allem in der Ergänzung zweier Facet-ten identifizieren: Bei der nomothetischen Dimension handelt es sich um die soziologischen Bedingungen des Verhaltens (Rollen, Erwartungen, geknüpft an eine Institution, hier: Schu-le). Anhand der idiographischen Dimension werden die individuellen Bedürfnisse sowie die

Persönlichkeit des Individuums erfasst. Dabei übernimmt die Gruppe, hier also die Schulklasse, die vermittelnde Komponente zwischen den beiden Dimensionen. Klima entstehe dabei durch das Zusammenwirken personaler Gegebenheiten und äußerer Faktoren (vgl. Janke, 2006).

Auf diesen Ansatz baut der Learning Environment Inventory (LEI) von Anderson (1973, nach Eder, 2010) auf. Dieser ist insbesondere dazu geeignet, Struktur- sowie Prozessmerkmale von Schulklassen abzubilden (vgl. Eder, 2010). Im deutschsprachigen Raum lassen sich hierunter die Arbeiten von Kahl, Buchmann und Witte (Lernsituationstest, 1977, nach Eder, 2010) und als elaboriertestes Verfahren die Landauer Skalen zum Sozialklima (LASSO, Saldern & Littig, 1985) subsumieren, die auf dem LEI von Anderson (1973) aufbauen.

Ein weiteres Modell konnte sich insbesondere in der U.S.-amerikanischen Klimaforschung etablieren. Dabei handelt es sich um das Lebensraum-Konzept, in welchem Lewin (1963) als erster die Beschreibung der Umwelt einer Person (psychologische Umwelt) mit ihrem individuellen Verhalten (psychologische Person: Bedürfnisse, Ziele, Wünsche) und dem Wissen über ihre Umwelt verknüpft. Dies ermöglichte erstmalig Interpretationen dahingehend, wie gleiche Umwelten für unterschiedliche Personen unterschiedliche Bedeutungen aufweisen können; ein zentraler Aspekt für die Erklärung der Wirkung von Klima auf das mitunter divergente Verhalten verschiedener Personen (vgl. Gruehn, 2000; Janke, 2006).

Die Frage nach der Entstehung kollektiven Klimas aus der individuellen Wahrnehmung einer (Lern-)Umwelt kann anhand Dreesmanns sozial-kognitivem Modell (1982) analysiert werden. Dabei bildet die Annahme, dass die Schülerinnen und Schüler einer Klasse im Zeitverlauf durch geteilte Erlebnisse, Interaktionen und Abgrenzungen gegenüber anderen Klassen zu einer Gruppe zusammenwachsen und so gemeinsame Erlebens- und Verarbeitungsstrukturen etablieren, die Grundlage für sein Modell (vgl. Dreesmann, 1982). Das Klima entstehe dabei aus den individuellen Wahrnehmungen der objektiven Lernumwelt als „Resultat der teilweisen Verschmelzung von individuellen Prozessen der Wahrnehmung, Informationsverarbeitung, Urteilsbildung etc., die durch den sozial-dynamischen Prozeß zwischen den Schülern einer Klasse, der sie zu einer Gruppe macht, initiiert und getragen wird“ (Dreesmann, 1982, S. 52).

Das transaktionalistische Modell schulischer Klimawahrnehmung nach Pekrun (1985) als weiterer Ansatz der Klimaforschung stellt einen komplexen Rahmen verschiedener Umweltas-

pekte verknüpft mit personeninneren Eigenschaften dar, in welchem „gesellschaftliche Bedingungen, schulische Umwelten und Merkmale des Individuums“ (Janke, 2006, S. 32–33) verortet werden können. Die Wahrnehmung klimatischer Umweltbedingungen nimmt dabei eine vermittelnde Rolle zwischen Person und Umwelt ein: Individuelles (schulisches) Klima entstehe dann, wenn Umwelteindrücke als häufig auftretend und als charakteristisch im Gegensatz zu ausnahmsweise auftretenden Ereignissen interpretiert werden. Kollektives Klima entstehe hieraus, wenn diese Wahrnehmung mit weiteren Personen des Systems geteilt wird. Pekrun (1985) unterscheidet dabei zwischen *Klimawahrnehmungen* als aktuelle Eindrücke der Umwelt und *Klimakognitionen* als über einen längeren Zeitraum stabile, abrufbare Klimavorstellungen, wobei letztere oftmals mit Hilfe von Fragebögen erfasst würden.

Die Einbeziehung und Erfassung des Verhaltens von Lehrerinnen und Lehrern im schulischen Klimakontext gelang Moos und Trickett (1974, nach Eder, 2010) mit der Classroom Environment Scale (CES). Bessoth (1989, Unterrichtsklima-Instrument) adaptierte dieses Instrument für den deutschsprachigen Raum. Mit den oben bereits erwähnten Landauer Skalen zum Sozialklima (Saldern & Littig, 1985) stagnierte die Klimaforschung im nationalen sowie internationalen Kontext zum Ende der 1980er Jahre vorerst. Vereinzelt lassen sich aktuellere Instrumente finden, etwa bei Eder und seinem Linzer Fragebogen zum Schul- und Klassenklima (LFSK, Eder, 1998), der sozialisations- und schultheoretischen Ansätzen zugeordnet werden kann und erstmals berücksichtigt, dass Schülerinnen und Schüler oftmals von mehreren Lehrerinnen und Lehrern unterrichtet werden. Dieser ermöglicht darüber hinaus, das Klima auf Schul- und Klassenebene zu erfassen. Ein Verfahren zur Beschreibung des Klimas ab Beginn der Grundschule entwickelten Rauer und Schuck (2003) mit dem Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschul- und mittlerweile auch Sekundarschülern (FEES).

Für einen Überblick über die in den oben aufgeführten Instrumenten genutzten Facetten des Klimas an Schulen wird im Folgenden auf ihre Bestandteile eingegangen. Verzichtet wird dabei explizit auf eine nähere Erläuterung der Dimensionen, insbesondere auch, weil im Rahmen der vorgesehenen Analysen lediglich auf einige wenige Klimafacetten, die in IGLU-E 2011 erhoben wurden, zurückgegriffen werden kann.

## 9.2 Dimensionen des schulischen Klimas

Nach Eder (1996, S. 51) werden „empirisch aufgefundene latente Strukturen einer Menge von Klima-Elementen [als Dimensionen des Klimas] verstanden, wie sie z.B. mit der Faktorenanalyse nachgewiesen werden können. Bei entsprechenden empirischen Belegen können auch die Skalen einzelner Erhebungsinstrumente als Dimensionen bezeichnet werden“. Die oben vorgestellten größtenteils U.S.-amerikanischen und deutschen Instrumente der Klimaforschung lassen sich den vier Bereichen Schüler-Lehrer-Beziehung, Schüler-Schüler-Beziehung, Unterrichtsmerkmale sowie Lernhaltungen der Schülerinnen und Schüler zuordnen (vgl. Tabelle 4). Deutlich wird hieraus, dass insbesondere die Schüler-Lehrer-Beziehung sowie Merkmale des Unterrichts und weiterhin auch die Schüler-Schüler-Beziehung in fast allen Studien erfasst werden, wenngleich die Zuordnung der Untersuchungen zu diesen vier Bereichen nicht immer eindeutig durchgeführt werden kann (vgl. Eder, 1996). Dass die Lernhaltungen der Schülerinnen und Schüler als Bestandteil des Klimas in die Übersicht aufgenommen wurden, kritisiert Gruehn (2000) an dieser Zuordnung nach Eder (1996), „stellen sie doch m. E. eher Folgen als Konstituenten des Klimas dar“ (Gruehn, 2000, S. 72). Dies erklärt möglicherweise, warum nur wenige Dimensionen diesem Bereich zugeordnet werden können (vgl. Tabelle 4).

**Tabelle 4:** Komponenten wichtiger Klima-Instrumente

Verfahren	Schüler-Lehrer-Beziehung	Schüler-Schüler-Beziehung	Merkmale des Unterrichts	Lernhaltungen der Schüler
<b>LEI</b> Learning Environment Inventory (Anderson, 1973)	Formalität; Demokratisches Verhalten; Bevorzugung	Vielfalt; Gemeinschaft; Cliquesbildung; Spannungen; Konkurrenz; Apathie	Lerntempo; Ziel-/Aufgabenorientierung; Desorganisation; Schwierigkeit; Zufriedenheit; Ausstattung	
<b>CES</b> Classroom Environment Scale (Moos & Trickett, 1974)	Lehrerunterstützung	Klassengemeinschaft; Wettbewerbsverhalten	Aufgabenorientierung; Ordnung und Organisation; Regelmäßigkeit; Innovation; Lehrerkontrolle	Beteiligung
<b>FUK</b> Fragebogen zum Unterrichts-klima (Dreesmann, 1979)	Kooperation zwischen Lehrer und Schüler	Kameradschaft; Konkurrenz	Schwierigkeit; Verständlichkeit; Erfahrungsnähe; Individualisierungsmangel; Erfolgsaussicht von Anstrengungen; Disziplin	
<b>FEND</b> Klima-Skalen (Fend, 1977)	Restriktive Kontrolle; Mitbestimmung; Lehrerengagement; Anonymität		Leistungsdruck; Disziplin- druck; Unterrichtsverlauf; Anpassungsdruck; Sozialbeziehungen	
<b>LASSO</b> Landaauer Skalen zum Sozialklima (Saldern & Littig, 1987)	Fürsorglichkeit des Lehrers; Aggression gegen den Lehrer; Zufriedenheit mit dem Lehrer; Autoritärer Führungsstil des Lehrers; Bevorzugung und Benachteiligung durch den Lehrer	Ausmaß der Cliquesbildung; Hilfsbereitschaft der Mitschüler; Aggression gegen Mitschüler; Diskriminierung von Mitschülern; Zufriedenheit von Mitschülern; Konkurrenzverhalten von Mitschülern	Leistungsdruck; Zufriedenheit mit dem Unterricht; Disziplin und Ordnung; Fähigkeit des Lehrers zur Vermittlung von Lehrinhalten	Resignation; Reduzierte Unterrichtsteilnahme

Quelle: Eder, 1996, S. 52–53, in Auszügen

Tabelle 4 ist zu entnehmen, dass sich die Skalen in den einzelnen Bereichen inhaltlich sehr unterscheiden, sofern dies ohne nähere Beschreibung der konkreten Skaleninhalte nachvollziehbar ist<sup>41</sup>. Lediglich der Bereich der Schüler-Schüler-Beziehungen erscheint vergleichsweise homogen. Gruehn (2000, S. 72) führt dies auf eine „offensichtlich [weit auseinandergehende] Auffassung[...] über einzelne Elemente und konstituierende Bestandteile des Klimas“ zurück und vermutet hierin einen Hinweis darauf, dass in der Klimaforschung keine einheitliche Konzeption des Konstrukts vorliegt. Im Folgenden sollen weitere zentrale Kritikpunkte zusammengefasst werden.

### 9.3 Kritik an der Klimaforschung

Dass die Aktivitäten innerhalb der Klimaforschung seit den 1980er Jahren zusehends abnehmen, führt Gruehn (2000) auf mehrere Aspekte zurück. Zum einen wurde eine Vielzahl von Instrumenten entwickelt, die teilweise unter ähnlichen Namen unterschiedliche Skalen enthielten (vgl. etwa die Skalen *Task Orientation* des CES nach Moos & Trickett, 1974 und *Goal Direction* des LEI nach Anderson, 1973). Ihre Einordnung in inhaltsbezogene Klassifikations-schemata (vgl. Kapitel 9.2) ist entsprechend schwierig und gelingt nur dort vollständig, wo eine Einteilung anhand ihrer Adressaten unternommen wird (siehe etwa Saldern, Littig & Ingenkamp, 1986). Dabei wird unterschieden nach Skalen zu der Schüler-Schüler-Beziehung, der Schüler-Lehrer-Beziehung und Merkmalen des Unterrichts. Nach Gruehn fokussieren die unterschiedlichen Skalen innerhalb dieser Dimensionen jedoch auch auf unterschiedliche Aspekte des Klimas, was eher zu einer Kumulation unterschiedlich bedeutsamer Klimaaspekte statt einer Replikation und damit Verdichtung bereits gefundener Zusammenhänge führe (vgl. Gruehn, 2000). Zudem wird angeführt, dass Befunde aus der Unterrichts- bzw. Lehr-Lern-Forschung beispielsweise zu nicht-klimatischen Kognitionen von Schülerinnen und Schülern sowie Lehrerinnen und Lehrern oder Effekten direkter und indirekter Instruktion im Unterricht in bisherige Klimamodelle kaum integriert wurden. Für eine Zusammenführung der bestehenden klimabezogenen Einzelbefunde und bestehenden Theorien etwa der Motivationsforschung, der Sozial- oder der Kognitionspsychologie wäre daher die Formulierung explikativer Theorien zu *ausgewählten Aspekten der Lernumwelt* – in Abhängigkeit des zu untersuchenden Effektmerkmals – von zentraler Bedeutung. Ebenfalls wirft Gruehn (2000) die Frage auf, ob der Begriff des Klimas dieses Konstrukt analytisch sinnvoll beschreibt, wird doch – wie oben dargestellt – davon ausgegangen, dass eine Vielzahl unterschiedlicher As-

---

<sup>41</sup> Eine ausführliche Beschreibung der Skalen findet sich bei Lange, Kuffner und Schwarzer (1983).

pekte dem Klimabegriff untergeordnet werden, deren Erhebung zudem unterschiedliche und unterscheidbare Skalen und Instrumente erfordert. Darüber hinaus bliebe bei der Beschreibung vorherrschenden Klimas, insbesondere auch im Zusammenhang mit Leistungsmerkmalen häufig die mehrebenenspezifische Struktur der Daten unberücksichtigt, was zu einer Verzerrung der gefundenen Effekte führen kann (vgl. Gruehn, 2000).

Aufgrund der Heterogenität der unter Klima subsumierten Theorien, Messinstrumente und Wahrnehmungsperspektiven legt Gruehn nahe, „auf den Begriff des Klimas zugunsten der neutraleren Bezeichnung Wahrnehmung [zu verzichten]“. Dies „hätte [...] den Vorteil, den Unterricht und die dort stattfindenden Interaktionsprozesse unter dem Gesichtspunkt unterschiedlicher Wahrnehmungsperspektiven der daran Beteiligten untersuchen zu können“ (beides Gruehn, 2000, S. 90). Zudem würde dies die Anschlussfähigkeit dieser Forschungstradition an die auf Beobachtungsdaten beruhende Unterrichtsforschung wahren.

Für die vorliegende Arbeit hat dies vor allem eine Implikation: Zugunsten einer stärkeren Anschlussfähigkeit insbesondere auch an die in Kapitel 8.3 dargestellten Rahmenmodelle des Schulerfolgs und schulischen Lernens soll Gruehns Vorschlag gefolgt werden. Statt die hier interessierenden Klimadimensionen der Schüler-Schüler- und Schüler-Lehrer-Beziehung einem allgemeinen Klimakonstrukt (und damit einer der u. U. konkurrierenden Klimatheorien) zuzuordnen, sollen diese als Facetten der Wahrnehmung der schulischen Lernumwelt durch die Schülerinnen und Schüler einzeln aufgenommen werden. Dies hat zudem den Vorteil, dass dem Klima nicht zugeordnete Theorien, die möglicherweise jedoch Erklärungsansätze für die Wahrnehmung der beiden sozialen Facetten bereitstellen, für ihre Beschreibung herangezogen werden können<sup>42</sup>. Auf sie wird im Folgenden näher eingegangen.

#### **9.4 Wahrnehmung der sozialen Lernumwelt durch Schülerinnen und Schüler**

Wie in Kapitel 9.2 beschrieben, lassen sich die Dimensionen der Klimaforschung in drei adressatenbezogene Dimensionen unterteilen<sup>43</sup> (vgl. Eder, 1996; Saldern et al., 1986; Gruehn,

---

<sup>42</sup> Die Mehrebenenstruktur der Daten wird ebenfalls bei der Auswertung berücksichtigt – nähere Beschreibungen hierzu finden sich in Kapitel 12.

<sup>43</sup> Eine ähnliche Dimensionierung findet sich bei Fend (1977). Dieser unterscheidet zwischen dem Inhaltsaspekt (normative Erwartungen und Werte wie Leistungs- und Disziplindruck), dem Interaktionsaspekt (Diskussionen über Formen von Autorität und Mitbestimmung in der Interaktion zwischen Lehrenden und Schülerinnen bzw. Schülern) sowie dem Beziehungsaspekt (soziale Beziehungen zwischen Lehrenden und Schülerinnen bzw. Schülern) (vgl. auch König, 2006). Da in den folgenden Analysen die Schüler-Schüler- wie auch die Schüler-Lehrer-Beziehung als getrennte Konstrukte eingehen, wird sich an der oben aufgeführten adressatenbezogenen Unterteilung orientiert.

2000). Zu diesen gehören die Schüler-Schüler-Beziehung, die Schüler-Lehrer-Beziehung sowie Merkmale des Unterrichts wie etwa Leistungsdruck, Disziplin und Unterrichtszufriedenheit (vgl. Saldern & Littig, 1987). Im Rahmen der vorliegenden Arbeit wird – wie oben erwähnt – die *soziale* Facette der Lernumwelt fokussiert. Hierunter werden die Schüler-Schüler- sowie die Schüler-Lehrer-Beziehung verstanden als zentrale Komponenten der Wahrnehmung der sozialen, schulbezogenen Lernumwelt durch Schülerinnen und Schüler. Eine Erfassung aller drei Dimensionen ist zunächst aus ökonomischen Gründen nicht zweckmäßig – die vielen unterschiedlichen Merkmale des Unterrichts, die es aufzunehmen gälte, würden den Rahmen des Mehrebenenstrukturgleichungsmodells sowie die hier vorliegende Arbeit überschreiten. Zum anderen erscheint es sinnvoll, sich in Anlehnung an Gruehns Kritik (vgl. Kapitel 9.3) auf einige zentrale Merkmale der Lernumwelt zu fokussieren und theoriegeleitet ihren Einfluss auf spezifische Effektmerkmale – hier auf den Schulerfolg der Schülerinnen und Schüler unter Berücksichtigung des individualisierenden Lehrerhandelns – zu modellieren. Da angenommen werden kann, dass sich individualisierte Unterrichtssettings durch weniger Konkurrenzsituationen unter den Schülerinnen und Schülern sowie direktere Schüler-Lehrer-Interaktionen auszeichnen, stellen die Schüler-Schüler- sowie die Schüler-Lehrer-Beziehung die lernumweltbezogenen Kernkonstrukte der folgenden Analysen dar. Auf ihre Bedeutung und ihren Zusammenhang mit den Dimensionen des Schulerfolgs wird im Folgenden gesondert eingegangen.

### **9.5 Schüler-Schüler-Beziehung<sup>44</sup>**

Eine gelungene Integration in soziale Gruppen zählt zu den wichtigsten Voraussetzungen für das Wohlbefinden und ein motiviertes, selbstverwirklichendes Verhalten (Gerecht, 2010). Dabei stellt die Klassengemeinschaft eine der bedeutendsten Sozialräume für Schülerinnen und Schüler dar, in welchem Rückhalt, emotionale Unterstützung, Sicherheit und Bestätigung (Siegler, DeLoache & Eisenberg, 2011) geboten und schwierige Übergangsphasen, wie etwa der Schuleinstieg oder -wechsel erleichtert werden können, zum Beispiel wenn sich unter den neuen Mitschülerinnen und Mitschülern bereits Freunde befinden (Ladd & Coleman, 1997). Schulkinder mit mindestens einem Freund bzw. einer Freundin zeigen ein Jahr nach Mobbing-Erfahrungen deutlich weniger Anpassungsprobleme wie Aggression und Ängstlichkeit als Kinder die angeben, keine Freunde zu haben (Hodges, Boivin, Vitaro & Bu-

---

<sup>44</sup> Die Darstellungen in diesem Kapitel sind teilweise den Erläuterungen bei Schulz-Heidorf und Schwippert (2014) entlehnt.

kowski, 1999). Zudem konnte gezeigt werden, dass das Fehlen von freundschaftlichen Beziehungen im Kindes- und Jugendalter zu Depressionen und sozialem Rückzug führen kann (Ladd & Troop-Gordon, 2003; Pedersen, Vitaro, Barker & Sorge, 2007). Der Aufbau sozialer Beziehungen zu Gleichaltrigen, so genannten *Peers*, stellt darüber hinaus einen wichtigen Grundbaustein für die soziale Entwicklung dar. Hier finden sich Kinder erstmals in einem sozialen Gefüge wider, in welchem ihre Rolle nicht vorab festgelegt ist: Sie sind in Kindergarten und Schule – anders als im familiären Umfeld – nicht *Kind* sondern gleichberechtigtes Mitglied einer Gruppe und müssen ihre Position (Anführer, Sprecher, Freund u. ä.) zunächst finden. Dabei erfahren Kinder, dass die Anerkennung innerhalb dieser Gruppe abhängig von dem eigenen sozialen Verhalten ist. Auch lernen sie soziale Umgangsformen wie Streiten, Rücksicht nehmen, Vertrauen aufbauen, sich einbringen (Rohlf's, 2009; Zinnecker, Behnken, Maschke & Stecher, 2003; Krappmann, 2002). Diese Lerngelegenheiten, die sich aus dem Eingehen von freundschaftlichen Beziehungen ergeben, bezeichnet Du Bois-Reymond (2000) als „Peerkapital“, welches auch im späteren Leben für den Aufbau sozialer Netzwerke von großer Bedeutung ist.

Im Kontext Schule ist der Rückhalt in der Klassengemeinschaft insbesondere notwendig, um Kapazitäten für Lerninhalte zu schaffen, die sich andernfalls auf die Erfüllung des Bedürfnisses nach sozialer Integration konzentrieren würden (Pekrun, 1998). Bedingt wird dies durch das grundlegende psychische Bedürfnis nach sozialer Eingebundenheit, welches neben dem Erleben der eigenen Autonomie und Kompetenz erfüllt sein muss, damit neue Kenntnisse und Fähigkeiten angeeignet werden können (Selbstbestimmungstheorie nach Deci & Ryan, 1993). Zu den schulbezogenen Merkmalen, die ein positives Erleben der eigenen sozialen Eingebundenheit begünstigen, zählt unter anderem eine wertschätzende Beziehung der Schülerinnen und Schüler untereinander (Kunter, 2005; Rakoczy, 2006), auch, weil diese die sozioemotionale Verarbeitung schulischer Leistungsanforderungen unterstützen können (Krüger & Deppe, 2009).

Diese Befunde unterstreichen die Bedeutung einer gelungenen sozialen Integration innerhalb der Klasse (Schüler-Schüler-Beziehung) nicht nur für die soziale Entwicklung und das allgemeine Wohlbefinden von Kindern und Jugendlichen, sondern auch für ihre Lernbereitschaft und damit den Schulerfolg.

In Verbindung gebracht werden kann dies zudem mit dem akademischen Selbstkonzept, also (schul-)fachspezifischen Fähigkeitseinschätzungen der Schülerinnen und Schüler. Wie in Kapitel 8.2 (Internal/External Frame of Reference-Model nach Marsh, 1986) beschrieben, sind für deren Ausbildung unter anderem soziale Vergleiche von Bedeutung: So werden die eigenen Leistungen denen der Mitschülerinnen und Mitschüler gegenüber gestellt. Neben dem Vergleich der eigenen Leistung über unterschiedliche Schulfächer hinweg dient dieser externe Bezugsrahmen als Grundlage für die Bildung und Festigung akademischer Selbstkonzepte. Doch welche Rolle spielt hierbei die soziale Integration innerhalb der Klasse? Angenommen werden kann ein reziproker Zusammenhang mit dem Selbstkonzept der Schülerinnen und Schüler: So lässt sich einerseits vermuten, dass freundschaftliche Beziehungen innerhalb der Klasse als eine Art Puffer dienen und ein Leistungsvergleich zu Ungunsten der eigenen Person weniger ins Gewicht fällt als dies bei Kindern der Fall sein könnte, die über einen geringen Rückhalt innerhalb der Klassengemeinschaft verfügen und zudem unterhalb des (wahrgenommenen) Leistungsdurchschnitts liegen. Andererseits könnten ein anhaltend ungünstig ausfallender Leistungsvergleich und die damit einhergehende Wahrnehmung, dauerhaft schlechter zu sein als die Mitschülerinnen und Mitschüler, das Gefühl der sozialen Integration innerhalb der Klasse mindern. Da ein positives akademisches Selbstkonzept wie in Kapitel 8.2 dargestellt in einem positiven Zusammenhang mit der Schulleistung steht, unterstreicht dies noch einmal die Wichtigkeit einer gelungenen Integration im Klassenkontext.

## **9.6 Schüler-Lehrer-Beziehung**

In der Forschung zu der Schüler-Lehrer-Beziehung zeigt sich am deutlichsten der zentrale Kritikpunkt an der Klimaforschung, nach welchem diese uneinheitliche Konzepte und Teilfacetten aufweist, die nur schwer verdichtet werden können (siehe Darstellungen in Kapitel 9.3). So wird unter der Schüler-Lehrer-Beziehung beispielsweise die emotionale Grundtönung des Unterrichtsklimas verstanden ebenso wie Lehrerfürsorglichkeit oder Schülerzentriertheit, Bevorzugung bzw. Benachteiligung, emotionale Wärme, das Interesse der Lehrperson an den Belangen der Schülerinnen und Schüler, der geduldige und konstruktive Umgang mit Fehlern und mitunter sogar eine individuelle Bezugsnormorientierung (vgl. bspw. Saldern & Littig, 1987; Satow, 2000; Lipowsky, 2015; Jerusalem & Klein-Heßling, 2002). Krammer (2009) differenziert zwei Betrachtungsweisen der Interaktion: Demnach lässt sich die Schüler-Lehrer-Beziehung einerseits als Beziehungsgestaltung interpretieren, deren Wirkung vorrangig auf das Wohlbefinden und die Motivation der Lernenden im Klassenzimmer ausge-

richtet ist. Zum anderen lässt sich die Schüler-Lehrer-Interaktion im Sinne eines Unterrichtsgesprächs als Lernunterstützung verstehen, welche „die zentrale didaktische Funktion der Anregung und Begleitung des Lernens [...]“ (Krammer, 2009, S. 36) innehat. Dieser ersten pädagogischen Perspektive soll hier gefolgt werden, da sie eher als die didaktische Perspektive dem Klimagedanken folgt. In Anlehnung an Oser und Baeriswyl (2001) und Seidel (2003) folgen Kunter und Voss (2011) einer ähnlichen (wenn auch nicht deckungsgleichen) Einteilung in Sicht- und Tiefenstrukturen der Unterrichtsqualität, die auf verschiedenen Wegen die Initiierung und Aufrechterhaltung von Lernprozessen bei Schülerinnen und Schülern ermöglichen sollen. Dabei bezeichnen Sichtstrukturen dem Unterricht übergeordnete Organisationsmerkmale wie Rahmenbedingungen, Unterrichtsmethoden und Sozialformen, Tiefenstrukturen hingegen Merkmale direkter Lehr-Lern-Prozesse, zu welchen nicht nur der Umgang mit dem Lernstoff und der Grad der kognitiven Anregung gehören, sondern auch Aspekte der Schüler-Schüler- wie Schüler-Lehrer-Beziehung. Letztere beschreiben Kunter und Voss (2011) als „emotionale und motivationale Unterstützung der Lernenden durch die Lehrkraft. Hierzu gehören zum Beispiel der respektvolle und geduldige Umgang mit den Lernenden (insbesondere bei Fehlern oder Verständnisschwierigkeiten) oder eine hohe Sozialorientierung der Lehrkräfte, das heißt Ansprechbarkeit bei persönlichen oder sozialen Schwierigkeiten“ (Kunter & Voss, 2011, S. 90). Dies deckt sich größtenteils mit den von Helmke (2014) zusammengefassten Aspekten, die sich in bisherigen Untersuchungen als wichtig für ein lernförderliches Klima erwiesen haben. Zu ihnen gehören der angemessene Umgang mit Fehlern, der Abbau von Angst, Geduld, Humor und Toleranz für langsamere Schülerinnen und Schüler auf Seiten der Lehrkraft. Mit den hier zur Verfügung stehenden Items aus IGLU-E wird die Qualität der Schüler-Lehrer-Beziehung ausschließlich aus der Perspektive der Lernenden erfasst. Dabei ist anzumerken, dass diese nicht explizit die einzelnen oben genannten Facetten der Schüler-Lehrer-Beziehung abdecken, sondern allgemeine Einstellungen der Viertklässlerinnen und Viertklässler gegenüber ihren Lehrenden („Mit meinem Klassenlehrer komme ich gut aus“) erfragen. Wenngleich also keine differenzierteren Aussagen zu den Teildimensionen getroffen werden, bilden die Schülereinschätzungen doch ein Abbild ihres Eindrucks der Lehrperson, ihrer Fürsorge, Unterstützung bei Verständnisfragen und Fehlern und ihrer Ansprechbarkeit bei persönlichen oder sozialen Schwierigkeiten. Dies soll (und einzig: kann) entsprechend für die empirische Überprüfung der Schüler-Lehrer-Beziehung herangezogen werden.

Dabei zeigt sich – analog zu einer gelungenen Integration in die Klassengemeinschaft (vgl. Kapitel 9.5) – die Bedeutung einer positiven Schüler-Lehrer-Beziehung vor allem für die Lernmotivation und die allgemeine Einstellung gegenüber der Schule, die, wie dies in einer Reihe von Untersuchungen gezeigt werden konnte (etwa Den Brok, Brekelmans & Wubbels, 2004; Furrer & Skinner, 2003; Kunter & Voss, 2011; Opdenakker, Maulana & Den Brok, 2012, vgl. auch ausführlichere Darstellung in Kapitel 9.7), den Schulerfolg von Kindern und Jugendlichen beeinflussen kann. Dies stellt sich ebenso in einem systematischen Literaturreview von Osterman (2000) dar, die zusammenfasst, dass „students' experience of acceptance is associated with a positive orientation towards school, class work, and teachers. Students who experienced a greater sense of acceptance by peers and teachers were more likely to be interested in and enjoy school and their classes. These perceptions of school were also reflected in their commitment to their work, higher expectations of success, and lower levels of anxiety“ (Osterman, 2000, S. 331). Auch Lipowsky (2015) kommt zu diesem Fazit und stellt diesem an, dass sich das durch eine positive Schüler-Lehrer-Beziehung gestärkte schulische Engagement und die höhere Motivation entsprechend positiv auf den Lernerfolg der Schülerinnen und Schüler auswirkten. Eine rein motivationale Dimension soll in dem hier vorgestellten Modell zwar nicht berücksichtigt werden, da dies den Umfang der Arbeit übersteigen würde – eine entsprechende Erweiterung des Modells stellt eine Perspektive für mögliche Anschlussarbeiten dar. Aufgenommen wird jedoch der Einfluss der Schülerwahrnehmung der Schüler-Schüler- und Schüler-Lehrer-Beziehung auf das Selbstkonzept und die Leistung der Kinder. Studien, die sich mit einer oder beiden Klimafacetten und ihrem Einfluss auf den Schulerfolg auseinandersetzen, sollen im Folgenden eingehender erläutert werden.

### **9.7 Soziale Lernumwelt und Schulerfolg**

Wenngleich Gruehn (2000, vgl. Kapitel 9.3) aus unterschiedlichen Gründen die geringe Anschlussfähigkeit der Klimaforschung an bestehende Theorien der Unterrichts- sowie Lehr-Lernforschung kritisiert und eine alternative Verwendungsweise vorschlägt, wird sich für die folgende Darstellung der Zusammenhänge zwischen Facetten der schulischen Lernumwelt und des Schulerfolgs zunächst weiterhin am Klimabegriff orientiert, da dieser nach wie vor in empirischen Untersuchungen herangezogen wird. Fokussiert werden dabei vor allem Studien, die sich explizit auf die in den Kapiteln 9.5 und 9.6 vorgestellten sozialen Komponenten und ihre Einflüsse auf den kognitiven wie affektiven Schulerfolg beziehen. Eine Trennung dieser Ergebnisdarstellung nach den beiden Dimensionen des Schulerfolgs – Leistung und

Selbstkonzept – gelingt nur bedingt, da einige Studien beide Effekte untersuchen – so soll der aktuelle Forschungsstand entsprechend in einem gemeinsamen Kapitel dargelegt werden. Dabei muss einer Verallgemeinerung der Befunde aus Studien, die sich dem Zusammenhang zwischen sozialen Merkmalen der Wahrnehmung der Lernumwelt und den *Schulleistungen* widmen, aufgrund methodischer Unterschiede mit Vorsicht begegnet werden: In U.S.-amerikanischen Untersuchungen werden beispielsweise häufig die Ergebnisse standardisierter Leistungstests, die jährlich schulübergreifend durchgeführt werden, als Leistungskriterium herangezogen, während Untersuchungen im deutschsprachigen Raum oftmals Schulnoten als Referenz nutzen. Letztere stellen jedoch einen weniger gut geeigneten Indikator dar, da sie klassenreferenziell, also auf den Vergleich innerhalb einer Schulklasse ausgelegt und zudem von der unterrichtenden und bewertenden Lehrkraft sowie der Schulform abhängig sind (vgl. Gruehn, 2000, auch Trapmann, Hell, Weigand & Schuler, 2007). Dementsprechend fallen die Korrelationen zwischen den einzelnen Klimadimensionen und Schulnoten höher aus als würden standardisierte Leistungstests als Referenz für die Schulleistung genutzt werden (vgl. etwa Helmke, 1992). Dies muss bei dem Versuch, einen Überblick über die Ergebnisse der verschiedenen Studien zu gewinnen, entsprechend berücksichtigt werden.

Saldern (1987) fasst die Klimabereiche, denen ein positiver Zusammenhang mit der Schulleistung nachgewiesen wurde, zusammen: Zu diesen gehören – neben der Präferenz für das jeweilige Fach, Zufriedenheit, Kohäsion, Interesse und Engagement, Möglichkeiten zur Selbstbestimmung, Frömmigkeit, Umgebung, einem demokratischen Führungsstil und Zielgerichtetheit – auch ein positives Lehrer-Schüler-Verhältnis sowie sozialer Anschluss. Ergänzt wird diese Übersicht durch Eder (1996), der die Ergebnisse der Untersuchung von Moos (1979, nach Eder, 1996) in die Reihe der klimabezogenen Facetten aufnimmt, die einen positiven Einfluss auf die Leistung haben. Neben einem hohen Maß an Aufgabenorientierung, Wettbewerb, Organisation, Klarheit, der Betonung schulischer Aufgaben und Leistungen, einem mittleren Ausmaß an Klarheit, Ordnung und Strukturierung zählt hierzu auch ein freundliches und unterstützendes Lehrerverhalten. Für diese Facetten zeigen sich – neben einer positiven Wirkung auf die Schulleistung – unter anderem auch positive Einflüsse auf die Kreativität, Selbstachtung und Unabhängigkeit der Schülerinnen und Schüler (vgl. Saldern, 1987). Dabei kritisiert Gruehn (2000) Befunde, nach welchen sich eine positiv wahrgenommene schulische Lernumwelt positiv auf die *Leistungen* der Schülerinnen und Schüler

auswirkt. Die in diesem Zusammenhang oftmals zitierten bis zu 35% Varianzaufklärung der Schulleistung durch Klimamerkmale (Stand: 2000) können auf methodische Besonderheiten zweier Studien zurückgeführt werden. Demnach lassen sich die hohen Korrelationen in der Untersuchung von Walberg, Singh und Rasher (1977, nach Gruehn, 2000) durch die Stichprobenauswahl begründen: Die Analyseeinheiten bildeten hier keine ganzen Klassen, sondern jeweils zehn besonders fleißige und besonders faule Schülerinnen und Schüler. „Da aus anderen Studien bekannt ist, daß Leistungsmotivation und Klimawahrnehmung zusammenhängen [...], überraschen diese überdurchschnittlich hohen Korrelationen auf der Basis solcher Extremgruppen nicht“ (Gruehn, 2000, S. 81). In einer weiteren Untersuchung (Walberg & Anderson, 1972, nach Gruehn, 2000) wurden als Indikatoren der Leistung die Abschlussprüfungen kanadischer Sekundarstufen verwendet, welche nach Gruehns Überlegungen vermutlich näher am schulischen Curriculum lägen als die Leistungstests, welche in anderen Studien eingesetzt würden. Neben diesen beiden und den oben dargestellten Untersuchungen finden sich wenige statistisch signifikante Zusammenhänge zwischen Leistung und Klimawahrnehmung, etwa bei Fraser und Fisher (1982, nach Gruehn, 2000) mit einer signifikant positiven Korrelation zwischen der Leistung und der Aufgabenorientierung auf Individualebene, bei Power und Tisher (1979, nach Gruehn, 2000) eine signifikant negative Korrelation auf Klassenebene mit der wahrgenommenen Apathie und in der Untersuchung von Dreesmann (1982) zwei statistisch signifikante Korrelationen auf der Individualebene mit den Skalen Förderung des Selbstvertrauens und wahrgenommene Schwierigkeit. Trotz der „überzogenen Behauptungen zur Bedeutung des Klassenklimas für die Schulleistung“ geht Gruehn (2000, S. 85) von einem „zwar schwachen, trotzdem nicht bedeutungslosen Klimaeinfluß auf schulisches Lernen aus“. Dieser Einschätzung, die zunächst wenig nach den unterschiedlichen Klimafacetten, wie sie oben beschrieben wurden, differenziert, sollen im Folgenden Ergebnisse aus Studien nebengestellt werden, die sich explizit mit den Auswirkungen einer gelungenen *sozialen* Integration (Schüler-Schüler- wie Schüler-Lehrer-Beziehung) als ein Teilaspekt des Klimabegriffs beschäftigen und nicht nur als Prädiktoren der Schulleistung, sondern auch affektiv-motivationaler Merkmale der Lernenden herangezogen werden.

So konnte Moos (1979) anhand der Classroom Environment Scale (CES) feststellen, dass „ein Klima mit starker Betonung sozialer Beziehungen („Relationship“) [...] zu größerer Zufriedenheit bei den Schülern, zu gesteigertem Interesse am Fach, zu positiverem sozialen Verhalten

(Hilfsbereitschaft und Freundlichkeit) und zu größerer Unabhängigkeit, Selbstachtung und Kreativität bei den Schülern führt als ein Klima mit anderer Ausprägung“ (Eder, 1996, S. 72). Fend (1977) bestätigt, dass die soziale Beziehung zwischen Lehrerinnen bzw. Lehrern und ihren Schülerinnen und Schülern von Bedeutung ist. Demnach mindere sich abweichendes und störendes Verhalten unter den Lernenden, je stärker die Zuwendung seitens des Lehrers bzw. der Lehrerin ausfällt. Außerdem konnte anhand dieser Untersuchung in 135 Klassen verschiedener Schulformen unter Konstanzhaltung der Klassengröße, sozialen Herkunft, dem Geschlechterverhältnis innerhalb der Klassen, des Alters sowie der Schulleistungen eine positive Korrelation des Selbstbewusstseins mit der Qualität der sozialen Beziehungen innerhalb der Klasse in Höhe von  $r = 0,26$  festgestellt werden. Und auch Saldern (1987) bestätigt diese Befunde zu den Beziehungen der Klassenkameradinnen und -kameraden untereinander in ähnlicher Stärke und Richtung.

Bezüglich der Wirkung einer positiven Schüler-Lehrer- und Schüler-Schüler-Beziehung vor allem auf Merkmale der *Schulleistung* kann zunächst Hatties Metaanalyse (Hattie, 2009, vgl. auch Kapitel 6) herangezogen werden. Hier findet sich eine positive Schüler-Lehrer-Beziehung auf Rang 11 von 138 Aspekten, die auf ihre lernförderliche Wirkung hin überprüft und ihrer Effektgröße nach in eine Reihenfolge gebracht wurden – der Einfluss von Gleichaltrigen rangiert auf Platz 41. Mit Effektstärken von  $d = 0,72$  (Schüler-Lehrer-Beziehung) und  $d = 0,53$  (Schüler-Schüler-Beziehung) können diese den sehr erfolgreichen Maßnahmen zugeordnet werden, durch welche der Lernzuwachs in etwa dem eines ganzen Schuljahres entspricht. In Anlehnung an Cornelius-White (2007) differenziert Hattie zudem sieben<sup>45</sup> Teilaspekte, die eine wirksame Schüler-Lehrer-Beziehung ausmachen. Zu diesen gehören (mit abnehmender Bedeutung; Effektstärken zwischen  $d = 0,75$  und  $d = 0,28$ ) ein non-direktiver Lehrstil, Empathie, Wärme, kognitive Aktivierung („encouragement of higher order thinking“), adaptive Anpassung an Unterschiede, Glaubwürdigkeit. Hattie fasst die Wirkung einer positiven Schüler-Lehrer-Beziehung zusammen. So zeichnen sich entsprechende Klassen aus durch „more engagement, more respect of self and others, there are fewer resistant behaviors, there is greater non-directivity (student-initiated and student-regulated activities), and there are higher achievement outcomes“ (Hattie, 2009, S. 119). Demnach würde sich eine positive Schüler-Lehrer-Beziehung vor allem durch zwischenmenschliche Empathie und eine

---

<sup>45</sup> Ein achttes Merkmal, „learner-centered beliefs“, weist eine Effektstärke von  $d = 0,10$  auf, muss daher den weder schädlichen noch helfenden Maßnahmen zugeordnet werden und wird folglich oben nicht aufgeführt.

Sensibilität und Adaptivität an die (Lern-)Bedingungen der Kinder und Jugendlichen auszeichnen, die dann nicht nur den (bei Hattie v.a. kognitiven) Schulerfolg, sondern auch das Verhältnis der Schülerinnen und Schüler untereinander günstig beeinflusst. Auch bezüglich der Schüler-Schüler-Beziehung identifiziert Hattie (2009) in Anlehnung an Wilkinson und Fung (2002) eine Reihe von Merkmalen, die das schulische Lernen von Kindern und Jugendlichen beeinflussen können. Zu diesen gehören eine lernbezogene Unterstützung („helping“), das Beibringen von Lernstoff („tutoring“), der Aufbau von Freundschaften, das Geben von Feedback und die allgemeine Färbung der Lernumwelt, die die Schulfreude bei den Schülerinnen und Schülern stärken kann. Darüber hinaus konnte in einigen der 13 durch Hattie (2009) analysierten Studien nachgewiesen werden, dass eine geringe Akzeptanz bei den Klassenkameraden mit einem geringeren lernbezogenen Verhalten und schwächeren Schulleistungen einherginge (Ladd, 1990; Ladd, Kochenderfer & Coleman, 1997; Buhs & Ladd, 2001). Hier zeigt sich (wenn dies auch nicht explizites Ziel der Meta-Metaanalyse von Hattie war), dass sich eine gelungene soziale Integration in den Klassenkontext nicht nur auf kognitive, sondern auch auf affektiv-motivationale Merkmale des Schulerfolgs auswirken kann.

Eingehender untersuchen dies Roorda et al. (2011) ebenfalls in einer Metaanalyse. Gegenstand war dabei der Einfluss positiver und negativer Schüler-Lehrer-Beziehungen auf Schulleistungen und die schulischen Einstellungen der Lernenden vom Kindergarten bis zum zwölften Schuljahr. Die schulischen Einstellungen („engagement“) werden dabei definiert als „the quality of a student’s connection or involvement with the endeavor of schooling and hence with the people, activities, goals, values, and place that compose it“ (Skinner, Kindermann & Furrer, 2009, S. 494, nach Roorda et al., 2011, S. 495) und umfassen verhaltensbezogene („participation in academics [...], social and extracurricular activities“), emotionale („positive and negative feelings and reactions to academics, teachers, classmates, and school“) und kognitive („thoughtfulness and willingness to invest in mastering of difficult skills and comprehension of complex ideas“, alle Roorda et al., 2011, S. 495) Aspekte. Für die Analyse wurden die Ergebnisse aus 99 Studien, die zwischen 1990 und 2011 auf Englisch publiziert wurden, berücksichtigt. Insgesamt wurden so Daten von 129.423 Lernenden erfasst; für das Effektstärkemaß wurde der Korrelationskoeffizient Pearsons  $r$  herangezogen. Berichtet werden im Folgenden (vgl. Tabelle 5) die Effektstärken für die Gesamtstichprobe und für

die Untergruppe der Studien, die sich auf die Grundschule (Kindergarten bis Ende der Grundschule) beziehen<sup>46</sup>.

**Tabelle 5:** Effektstärken der Fixed Effects Modelle der Zusammenhänge zwischen einer positiven und negativen Ausprägung der Schüler-Lehrer-Beziehung, der Leistung und den schulbezogenen Einstellungen in der Gesamtstichprobe und grundschulbezogenen Untersuchungen

	Positive SLB und Einstellung			Negative SLB und Einstellung			Positive SLB und Leistung			Negative SLB und Leistung		
	r	k	N	r	k	N	r	k	n	r	k	n
Gesamtstich- probe	0,39	61	88.417	-0,32	18	5.847	0,16	61	52.718	-0,15	28	18.944
Grundschule	0,26	35	13.750	-0,34	15	4.823	0,14	42	14.127	-0,19	24	6.726

Legende: SLB = Schüler-Lehrer-Beziehung, k = Anzahl der Studien, alle Effektstärken auf  $p < 0,01$  signifikant

Quelle: Roorda et al., 2011, adaptiert

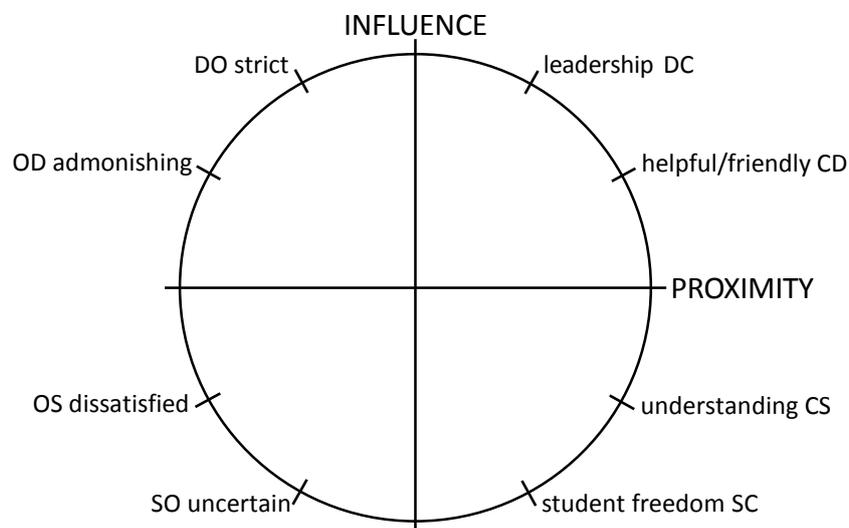
Zunächst zeigt sich theoriekonform, dass eine positive Schüler-Lehrer-Beziehung mit einer positiven und eine negative Schüler-Lehrer-Beziehung mit einer negativen Ausprägung sowohl der Leistung als auch den Einstellungen gegenüber der Schule einhergeht. Dabei fallen die Effekte in beiden Stichproben für den Zusammenhang mit der Leistung deutlich geringer aus. Bedeutsame Unterschiede zwischen der Gesamtstichprobe und der Teilstichprobe der grundschulbezogenen Studien zeigen sich nur für den Zusammenhang einer positiven Schüler-Lehrer-Beziehung mit den schulbezogenen Einstellungen der Kinder: Dieser fällt in jüngeren Jahren geringer aus (für Sekundarschulen:  $r = 0,30$ ). Die Autoren nehmen in einer weiteren Analyse Moderatorvariablen in das Modell auf und untersuchen dabei unter anderem den Einfluss der als Leistungsindikatoren genutzten Schulnoten und standardisierten Leistungstests. Dabei bestätigen sich die eingangs berichteten Effekte teilweise: So fällt der Zusammenhang zwischen einer positiven Schüler-Lehrer-Beziehung und den Schulnoten mit  $r = 0,24$  deutlich stärker aus als der Einfluss auf die anhand standardisierter Tests erfassten Leistung ( $r = 0,07$ ). Allerdings zeigte sich dieser Effekt nicht für das Zusammenhangsmaß zwischen negativer Schüler-Lehrer-Beziehung und Leistung ( $\beta = -0,006$  bei  $p = 0,970$ ). Insgesamt

<sup>46</sup> Die Autoren geben darüber hinaus für jeden Pfad zwei Effektstärken an: Dabei ermöglichen es die Koeffizienten aus den Random Effects Models, die Zusammenhangsstärken über die Metaanalyse hinaus zu verallgemeinern, allerdings zu Lasten der statistischen Power, was dazu führen kann, dass signifikante Effekte nicht identifiziert werden können. Die Koeffizienten aus Fixed Effects Models hingegen beziehen sich nur auf die in der Metaanalyse verwendeten Studien; ihre Verwendung schränkt die Verallgemeinerbarkeit der Ergebnisse ein, allerdings weisen sie eine größere Teststärke auf. Da Metaanalysen in der Regel stark verallgemeinerte Konstrukte verwenden, um eine möglichst breite Stichprobe zu generieren, können auf ihrer Grundlage ohnehin nur sehr allgemeine Aussagen getroffen werden, so dass hier auf die Angaben aus den Fixed Effects Models zurückgegriffen wird.

zeigt sich also auch in dieser Metaanalyse, dass die Qualität der Beziehung der Schülerinnen und Schüler zu ihren Lehrenden insbesondere von Bedeutung ist für affektiv-motivationale Merkmale, jedoch in einem weniger starken Zusammenhang mit der schulischen Leistung steht, was Roorda et al. (2011) in Anlehnung an Hamre und Pianta (2001) damit begründen, dass „TSRs [Teacher-Student-Relationship, Anm. KSH] are partly a measure of social adjustment and therefore more proximal to behavioral than academic outcomes. Furthermore, engagement has been found to act as a mediator between TSRs and achievement“ (Roorda et al., 2011, S. 515).

Belegen können dies ebenfalls die Befunde von Den Brok et al. (2004), welche sich weniger an der Klimaforschung denn an dem systemischen Ansatz der Kommunikation orientieren, den insbesondere Watzlawick, Beavin und Jackson (1967) prägten, um ein Modell der Lehrer-Schüler-Interaktion zu prüfen. Abbildung 10 stellt dieses in Anlehnung an Leary (1957) und Wubbels, Créton und Hoymayers (1985) genutzte Modell sozialen Lehrerverhaltens (im Original: Model for Interpersonal Teacher Behaviour, vgl. Den Brok et al., 2004, S. 413) dar.

**Abbildung 10:** Modell des sozialen Lehrerverhaltens



Quelle: Den Brok et al., 2004, S. 413

Diese circumplexe Darstellung ermöglicht es, acht Merkmale des Lehrerverhaltens auf zwei Achsen (Influence – Proximity) abzubilden und so ihre jeweilige Nähe zu diesen Ausprägungen zu kennzeichnen. Dabei bewegt sich die Achse „Influence“ (Einflussnahme) zwischen Kontrolle und Unterordnung (Dominance/Submission), die Achse „Proximity“ (Verbundenheit) hingegen zwischen Kooperation und Opposition (Cooperation/Opposition). Diesen beiden Aspekten des Lehrerverhaltens folgt der bei Den Brok et al. (2004) eingesetzte Questi-

onnaire on Teacher Interaction (QTI). Überprüft wird anhand zweier Stichproben (Physikunterricht,  $n = 826$ , 43 Lehrende in 43 Klassen an 43 Schulen und Fremdsprachenunterricht Englisch,  $n = 941$ , 32 Lehrende in 52 Klassen an 21 Schulen, drittes Sekundarschuljahr), inwieweit die beiden Dimensionen Einflussnahme und Verbundenheit kognitive wie affektive Merkmale der Schülerinnen und Schüler beeinflussen können. Für die Erfassung der kognitiven Leistung wurden ein Problemlösetest (Physik) und ein Leseverständnistest (Englisch als Fremdsprache), für die affektiven Charakteristika der Schülerinnen und Schüler Fragebögen zu den fachspezifischen motivationalen Einstellungen herangezogen. Die entsprechenden Ergebnisse, welche durch Mehrebenenanalysen ermittelt wurden, geben die Autoren in dem Effektstärkenmaß Cohens  $d$  an (vgl. hierfür Ausführungen in Kapitel 6). Aufgrund der starken Ausdifferenziertheit werden diese in gekürzter Form in der folgenden Tabelle 6 wiedergegeben.

**Tabelle 6:** Effektstärken des Einflusses der Charakteristika des Lehrerverhaltens auf kognitive und affektive Merkmale der Schülerinnen und Schüler

	Kognitive Tests		Schulfreude		Relevanz des Faches		Selbstvertrauen		Leistungsbereitschaft	
	Phy	Eng	Phy	Eng	Phy	Eng	Phy	Eng	Phy	Eng
Einflussnahme	0,103	-	0,132	-	0,332	-	-0,539	-0,261	0,275	0,217
Verbundenheit	-	-	0,671	0,577	0,284	-	0,257	0,199	0,280	0,510

Quelle: Den Brok et al., 2004, S. 442, adaptiert und gekürzt

Deutlich wird hier, dass die kognitive Leistung der Schülerinnen und Schüler nur in der Physik-Stichprobe und dort auch nur durch die Lehrerdimension der Einflussnahme beeinflusst wird, nicht jedoch durch das Ausmaß der Verbundenheit – diesem konnte gar kein Effekt auf die kognitive Facette nachgewiesen werden. Anders hingegen die affektiven Merkmale der Jugendlichen: Hier zeigen sich für fast alle Subdimensionen z.T. beachtliche Zusammenhänge, insbesondere für den Einfluss der Verbundenheit mit der Lehrperson auf die Schulfreude ( $d_{\text{Physik}} = 0,671$  und  $d_{\text{Englisch}} = 0,577$ ) und auf die Leistungsbereitschaft im Fach Englisch als Fremdsprache ( $d = 0,510$ ). Besonders hervor sticht ebenfalls der negative Effekt der Einflussnahme auf das Selbstvertrauen in beiden Stichproben. Dies ist damit zu interpretieren, dass ein durch Kontrolle gekennzeichnete Unterricht (maximale Ausprägung auf der Skala ‚Einflussnahme‘) scheinbar zu einer Schwächung des Selbstvertrauens sowohl im Fach Physik als auch im Fremdsprachenunterricht Englisch führt. Allgemein bestätigt sich also auch hier das Bild, dass eine gestärkte Bindung zwischen Lehrenden und Lernenden vor allem einen Ein-

fluss auf affektive Merkmale von Schülerinnen und Schülern hat, nicht so sehr jedoch auf kognitive Merkmale – wobei sich die Höhe der Effekte in verschiedenen Unterrichtsfächern unterscheiden können.

So finden Kunter und Voss (2011) in einer auf den Daten der COACTIV-Studie basierenden Analyse beispielsweise keine Effekte der als konstruktiven Unterstützung bezeichneten Facette der Schüler-Lehrer-Beziehung auf die Schulleistung im Fach Mathematik. Entsprechend zeigte sich in der quasiexperimentellen Längsschnittuntersuchung in 194 Sekundarschulklassen und unter Anwendung eines Mehrebenenstrukturgleichungsansatzes ein Einfluss der konstruktiven Unterstützung (modelliert durch Schülerangaben zu dem Umgang mit Fehlern, Kränkungen, adaptiven Erleichterungen, Geduld und der Sozialorientierung ihrer Lehrerinnen und Lehrer) auf einen curriculumorientierten IRT-skalierten Mathematiktest in Höhe von  $r = 0,11$ , welcher sich nicht signifikant von Null unterscheidet. Dagegen ergeben sich für die motivational-affektiven Zielvariablen Leistungsängstlichkeit und Freude an Mathematik signifikante Effekte in Höhe von  $r = -0,42$  (Angst) und  $r = 0,46$  (Freude).

Ähnliche Zusammenhänge stellt auch Satow (2000) unter anderem anhand von Mehrebenenanalysen fest. Überprüft wurde hier, wie sich die Schülerwahrnehmung sozialer Klimamerkmale, zu denen die Lehrerfürsorglichkeit, die (individuelle) Bezugsnormorientierung und das Schüler-Sozialklima (Unterstützung, Vertrauen und Rücksichtnahme unter den Schülern) gehörten, auf die Selbstwirksamkeitserwartung (Zuversicht in die eigenen Fähig- und Fertigkeiten) und die Leistung (Schülerangaben zu Noten in Mathematik, Deutsch und Englisch; Zusammenfassung in umkodierten Leistungsindex, hohe Werte entsprechen guten Schulnoten) auswirken. Herangezogen wurden hierfür die Daten der Begleitforschung des Modellversuchs ‚Verbund Selbstwirksamer Schulen‘<sup>47</sup>, wobei die Schulen nicht zufällig ausgewählt wurden, es sich bei den neun Haupt-, Real-, Gesamtschulen und Gymnasien also nicht um eine repräsentative Stichprobe handelt. Aus diesen Schulen wurden Jugendliche, die sich zu allen drei Messzeitpunkten (über drei Jahre) in der Sekundarstufe I (Klassen 7 bis 10) befanden, berücksichtigt. Dies entsprach 921 Schülerinnen und Schülern. Für diese ging eine sich zwischen den drei Messzeitpunkten positiv entwickelte individuelle Klimawahrnehmung mit einer günstigen individuellen Entwicklung der Selbstwirksamkeitserwartungen

---

<sup>47</sup> Der Autor verweist explizit darauf, dass – wenngleich der Name des Projekts dies nahelegt – keine explizite Förderung mit dem Ziel der Veränderung von Kontroll- und Selbstwirksamkeitsüberzeugungen an den Schulen durchgeführt wurde.

einher ( $\gamma = 0,31$ ). Ein ähnlicher Einfluss auf die Leistung konnte nicht nachgewiesen werden ( $\gamma = -0,01$ ). Da zudem keine ähnliche Entwicklung zwischen Klimawahrnehmung und Leistung zu beobachten war, schlussfolgert Satow (2000), dass die Befunde „zudem nahe[legen], daß der Effekt auf die Selbstwirksamkeitserwartungen nicht über bessere Schulnoten erzielt wurde“ (Satow, 2000, S. 175). Ein Einfluss der klassendurchschnittlichen Klimawahrnehmung auf die individuelle Selbstwirksamkeitserwartung (Cross-Level-Interaktionseffekt) konnte zwar nachgewiesen werden, muss jedoch – gerade im Hinblick auf den oben dargestellten Haupteffekt – als unbedeutend eingeschätzt werden. Weiterhin konnte Satow (2000) zeigen, dass eine individualisierte Schüler-Lehrer-Beziehung und eine als unterstützend wahrgenommene Schüler-Schüler-Beziehung einen positiven Effekt auf die psychische und physische Gesundheit der Lernenden und ihren Optimismus hat. Zudem zeigte sich für das kollektive Klima auf Klassenebene ein indirekter Effekt auf die Leistung: So beeinflussen die auf Klassenebene positiv wahrgenommene Schüler-Schüler- wie Schüler-Lehrer-Beziehung die Selbstwirksamkeitserwartungen positiv, was wiederum mit besseren Schulleistungen, weniger Prüfungsangst und einem gesteigerten psychischen wie physischen Befinden der Lernenden einherging. Darüber hinaus wirkte sich – im Gegensatz zu der kollektiven Klimawahrnehmung – das kollektiv wahrgenommene Anforderungsniveau (Eindruck, für gute Noten viel leisten zu müssen. Beispielitem: „Der Unterricht geht so schnell, dass wir oft nicht mitkommen“, Satow, 2000, S. 115) negativ auf die Selbstwirksamkeitserwartungen aus. Satow konnte den beiden Merkmalen der Schüler-Lehrer- wie Schüler-Schüler-Beziehung also eine Reihe von direkten wie indirekten Effekten insbesondere auf die Selbstwirksamkeitserwartungen, weniger auf die Leistung nachweisen.

Dem entgegen findet sich ein positiver Zusammenhang zwischen Merkmalen der Schüler-Lehrer-Beziehung und Schulnoten bei Wentzel (1997). Diese analysierte in einer längsschnittlichen Untersuchung unter anderem das psychologische Belastungsempfinden, die Kontrollüberzeugung, die schulische Anstrengungsbereitschaft sowie die Schulleistung (gemessen am Grade Point Average (GPA), also Schulnoten in den Hauptfächern) von 248 Sechstklässlern und ihre Entwicklung in diesen Dimensionen bis zur achten Klasse unter Einfluss der wahrgenommenen Fürsorge der Lehrkraft. Die wahrgenommene Fürsorge wurde dabei anhand des Teacher Social and Academic Support Skala der Classroom Life Measure (Johnson, Johnson, Buckman & Richards, 1985) mit den Beispielitems „My teacher really cares about me“ und „My teacher cares about how much I learn“ (vgl. Wentzel, 1997) erfasst. Sie steht in

einem hoch signifikanten, negativen Zusammenhang mit dem Belastungserleben der Jugendlichen ( $r = -0,23$ ,  $p < 0,001$ ); eine Interpretation erfolgt in dem Sinne, dass ein stärkeres Fürsorgeerleben der Jugendlichen durch die Lehrkraft mit einem geringeren psychologischen Belastungsempfinden (Angst, Depression, Selbstwertgefühl, Wohlbefinden) einhergeht. Ähnliche Effekte zeigen sich für die Kontrollüberzeugung, die die Autorin in die drei theoriegeleiteten Attribuierungsformen *internal* („If I want to do well in school, it’s up to me to do it“, ein Ereignis ist Konsequenz des eigenen Verhaltens), *sozial-external* („When I do well in school, it’s because the teacher likes me“, ein Ereignis wird auf andere Personen zurückgeführt) und *fatalistisch-external* („When I get a good grade in school I usually don’t know why I did so well“, ein Ereignis wird auf einen Zufall, das Schicksal oder andere externe Ursachen zurückgeführt) unterteilt und getrennt analysiert. Dabei zeigen sich negative Zusammenhänge für beide Formen der externalen Kontrolle ( $r = -0,25$  mit  $p < 0,001$  für sozial-external und  $r = -0,22$  mit  $p < 0,001$  für fatalistisch-external) und ein positiver Zusammenhang mit der internalen Kontrolle ( $r = 0,27$  mit  $p < 0,001$ ). Und auch die Leistung, gemessen am Notendurchschnitt, korreliert positiv mit der Lehrerfürsorge ( $r = 0,18$  mit  $p < 0,001$ ). Der theoriegeleiteten Argumentation der Autorin folgend kann demnach eine positive Einschätzung des Fürsorgeverhaltens der Lehrkraft nicht nur das Belastungsempfinden verringern, sondern auch die schulische Leistung verbessern, internale Kontrollüberzeugungen stärken und externe verringern. Zu kritisieren ist allerdings, dass diese Ergebnisse lediglich einer Korrelationsmatrix entstammen; für Hintergrundmerkmale wurde hierbei (anders als für andere, hier nicht relevante Analysen) nicht kontrolliert, ebenso wenig wurde die geclusterte Zweiebenenstruktur (Lehrer- und Schülermerkmale) berücksichtigt. Die angegebenen Parameter sollen daher lediglich als Hinweise für das Vorliegen von Zusammenhängen der Lehrerfürsorge mit dem affektiven wie kognitiven Schulerfolg interpretiert werden, über deren Richtung hier zudem keine Aussagen getroffen werden können<sup>48</sup>.

Eine weitere Untersuchung in diesem Zusammenhang widmet sich ausschließlich motivational-affektiven Merkmalen der Lernenden unter Einfluss der Schüler-Schüler- und Schüler-Lehrer-Beziehung. Ryan, Stiller und Lynch (1994) zogen hierfür die Daten von 606 Jugendlichen der siebten und achten Klassenstufe in einem querschnittlichen Design heran. Als Maße für die Beziehungsqualität wurden anhand unterschiedlicher Instrumente unter anderem

---

<sup>48</sup> Zwar entstammen die von Wentzel (1997) genutzten Daten einer längsschnittlichen Untersuchung, die Aussagen über die Richtung von Zusammenhängen prinzipiell zulassen würde – allerdings finden sich alle oben beschriebenen Merkmale in einem einzigen Messzeitpunkt.

die Dimensionen *gefühlte Geborgenheit* („My friend/teacher understands me“), *emotionale Verbundenheit* („When I am feeling happy [...], I go to my friend/teacher“) und *schulbezogene Verbundenheit* („If I had a problem with my school work I would share it with my friend/teacher“) erhoben. In Zusammenhang gebracht wurden diese mit mehreren affektiven Schülermerkmalen, zu denen u.a. positive Bewältigungsstrategien, die schulische Motivation, Kontrollüberzeugungen, schulbezogene Einstellungen und Engagement im Unterricht sowie das Selbstwertgefühl der Schülerinnen und Schüler gehören. In einer Regressionsanalyse, bei der für Geschlecht und Schulnoten kontrolliert wurde, wurden die Ausprägungen der jeweiligen affektiv-motivationalen Schülercharakteristika in Beziehung gesetzt zu den drei Dimensionen der Beziehungsqualität (Geborgenheit, emotionale und schulbezogene Verbundenheit), die wiederum für die Beziehung zu den Klassenkameraden und den Lehrenden aufgeschlüsselt wurden. Dabei zeigte sich ein deutliches Bild: Während ein Effekt aller drei Facetten der Schüler-Schüler-Beziehungsqualität (Gefühl der Geborgen- und Verbundenheit) lediglich auf das Selbstwertgefühl nachgewiesen werden konnte ( $r = 0,13$  bis  $0,24$  bei  $p < 0,001$ ), ergaben sich für die Einschätzung der Beziehungsqualität bezüglich der Lehrperson für alle<sup>49</sup> anderen affektiven Facetten, nicht jedoch für das Selbstwertgefühl signifikante Zusammenhänge. Demnach beeinflusst das Gefühl der Geborgenheit und emotionalen wie schulbezogenen Verbundenheit mit der Lehrkraft die Ausbildung bzw. Stärkung von Bewältigungsstrategien, die schulische Motivation, Kontrollüberzeugungen sowie die schulbezogenen Einstellungen und das Engagement im Unterricht, nicht jedoch das Selbstwertgefühl der Jugendlichen, welches hingegen stark in einem Zusammenhang mit der Qualität der Schüler-Schüler-Beziehungen steht. Anzumerken ist, dass es sich auch hier um eine Querschnittuntersuchung handelt, in der Aussagen über Wirkzusammenhänge nicht getroffen werden können.

Zwei weiteren motivational-affektiven Merkmalen und deren Beeinflussung durch soziale Merkmale der Lernumwelt widmet sich eine Untersuchung von Furrer und Skinner (2003). In einem längsschnittlichen Design wurden hier 641 Schülerinnen und Schüler von der dritten bis zur sechsten Klasse dahingehend untersucht, wie sich ihr emotionales<sup>50</sup> und lernbezoge-

<sup>49</sup> Einzige Ausnahme: Der Einfluss einer positiven emotionalen Verbundenheit mit der Lehrperson auf die Kontrollüberzeugungen der Schülerinnen und Schüler unterschied sich nicht signifikant von Null.

<sup>50</sup> Beispielitems: „When we start something new in school, I feel interested“ und „When working on classwork, I feel mad“ (Furrer & Skinner, 2003, S. 153).

nes<sup>51</sup> Verhalten im Unterricht unter Aspekten der Schüler-Schüler-, Schüler-Lehrer-Beziehung und weiteren, hier nicht relevanten Merkmalen verändert. Dabei wurden das emotionale sowie lernbezogene Verhalten der Kinder jeweils durch die Lehrperson sowie durch die Schülerinnen und Schüler selbst eingeschätzt. Für den Einfluss einer gelungenen sozialen Integration in den Verbund der Mitschülerinnen und Mitschüler auf das emotionale Erleben des Unterrichts ergeben sich hoch signifikante Regressionskoeffizienten von  $\beta = 0,21$  (Lehrereinschätzung) bzw.  $\beta = 0,50$  (Schülereinschätzung) und auf das lernbezogene Verhalten in Höhe von  $\beta = 0,24$  (Lehrereinschätzung) bzw.  $\beta = 0,42$  (Schülereinschätzung). Und auch der Schüler-Lehrer-Beziehung wird ein hoch signifikanter Einfluss auf das emotionale Unterrichtsfinden in Höhe von  $\beta = 0,24$  (Lehrereinschätzung) bzw.  $\beta = 0,55$  (Schülereinschätzung) nachgewiesen, während sich für das lernbezogene Verhalten lediglich ein hoch signifikanter Zusammenhang für die Selbsteinschätzung der Schülerinnen und Schüler in Höhe von  $\beta = 0,43$  zeigte. Furrer und Skinner (2003) schließen hieraus vor allem, dass „[c]hildren who felt appreciated by teachers were more likely to report that involvement in academic activities was interesting and fun and that they felt happy and comfortable in the classroom. In contrast, children who felt unimportant or ignored by teachers reported more boredom, unhappiness, and anger while participating in learning activities“ (Furrer & Skinner, 2003, S. 159). Zusammen mit der Feststellung, dass sich der Einfluss der Schüler-Schüler-Beziehung, wie sie in anderen Studien bereits nachgewiesen wurde, auch bei Furrer und Skinner (2003) replizieren ließ, zeigt sich auch hier die Bedeutung einer gelungenen Integration im Klassenkontext sowie einer wertschätzenden Beziehung zu der Lehrperson für die affektiv-motivationale Komponente des schulischen Lernens.

Explizit nur eines dieser beiden Merkmale der sozialen Lernumwelt, der Bedeutung einer gelungenen sozialen Integration innerhalb der Klassengemeinschaft (Schüler-Schüler-Beziehung), wird in einer rezenten Analyse aufgegriffen. Untersucht wurde anhand der IGLUE-Daten aus 2011 und mittels eines Strukturgeleichungsmodells<sup>52</sup> der Zusammenhang zwischen der Schüler-Schüler-Beziehung, dem akademischen Selbstkonzept und der Leseleistung. In Anlehnung an die Selbstbestimmungstheorie der Motivation (Deci & Ryan, 1993) wurde zunächst angenommen, dass der kognitive Schulerfolg (unter anderem) durch die

---

<sup>51</sup> Beispielitems: „I participate when we discuss new material“ und „In class, I just act like I’m working“ (Furrer & Skinner, 2003, S. 153).

<sup>52</sup> Der Mehrebenenstruktur der Daten wurde dabei mit der Berücksichtigung der Clusterstruktur (Schüler in Klassen) begegnet; alle Variablen fanden sich auf Individualebene.

soziale Integration der Kinder in den Kreis ihrer Mitschülerinnen und Mitschüler bedingt ist, da sie „ein grundlegendes psychologisches Bedürfnis darstellt, ohne dessen Erfüllung Maßnahmen eher für das Erlangen von Anerkennung, Wertschätzung und Freundschaft statt für die Wahrnehmung von Lerngelegenheiten ergriffen werden“ (Schulz-Heidorf & Schwippert, 2014, S. 598). Erwartungsgetreu zeigte sich hierfür ein positiver, wenngleich nur schwacher Effekt in Höhe von  $\beta = 0,16$  (bei  $p < 0,001$ ). Für den Zusammenhang zwischen dem akademischen Selbstkonzept und der sozialen Integration wurden zwei gegenläufige Effekte abgeleitet. Zum einen wurde das Selbstkonzept etwa zur Hälfte durch Items abgebildet, die einen fähigkeitsbezogenen Vergleich mit den Mitschülerinnen und Mitschülern einer Klasse fokussieren. Demnach könnte ein positiv ausfallender Leistungsvergleich die Kinder in der Bewertung ihrer Fähigkeiten stärken, was sich in der Ausbildung eines höheren Selbstkonzepts niederschlagen würde. Andererseits wurde auch angenommen, dass ein positives akademisches Selbstkonzept die Einbindung in den Klassenkontext stärken könnte, etwa wenn das Vertrauen in die eigenen Fähigkeiten die Schülerinnen und Schüler ermutigt, ihren Mitschülern schwierige Sachverhalte zu erklären und bei der Lösung von Aufgaben zu helfen, was dazu führen könnte, dass diese Kinder vermehrt durch ihre Klassenkameradinnen und -kameraden angesprochen und um Hilfe gebeten werden, welches eine Integration in den Klassenkontext zusätzlich verstärkt. Das (dauerhafte) Gefühl, schlechter zu sein als die Mitschülerinnen und Mitschüler könnte hingegen einen sozialen Rückzug nach sich ziehen (vgl. Schulz-Heidorf & Schwippert, 2014). Die Wechselwirkung der beiden Facetten Selbstkonzept und Schüler-Schüler-Beziehung unterstreichen zudem die Annahme, dass sich Freunde gegenseitig um Unterstützung in schul- und leistungsbezogenen Situationen bitten werden, was dem Gegenüber das Zutrauen entsprechender Fähigkeiten suggeriert, ebenso wie die Vermutung, dass klassenbezogene Freundschaften die Möglichkeit bieten, Schulisches zu besprechen und daraus Rückschlüsse auf die eigene Person erlauben – etwa, wenn eine als besonders schwierig empfundene Prüfung durch Freunde ebenfalls als schwierig eingeschätzt wird und daraus resultierend sich die Attribuierung weg von einer internalen (eigene Fähigkeiten) hin zu einer externalen (Schwierigkeit des Tests) Bewertung verschiebt. Überprüft wurde dies anhand einer Korrelation<sup>53</sup>, die mit  $r = 0,51$  (bei  $p < 0,001$ ) überraschend hoch ausfiel – wobei nochmals auf die Nähe einiger Selbstkonzept-Items zu der sozia-

---

<sup>53</sup> Ursache-Wirkungs-Beziehungen können in querschnittlich angelegten Studiendesigns nicht überprüft werden, so dass die hier möglicherweise zugrunde liegenden gegenläufigen Effekte nicht einzeln geprüft werden konnten.

len Dimension (sozialer Vergleich) verwiesen sei. Angenommen werden muss, dass sich dieser Zusammenhang schwächer darstellt, da wo das akademische Selbstkonzept keine soziale Vergleichskomponente beinhaltet. Dennoch untermauern die dargestellten Befunde die hohe Bedeutsamkeit des Leistungsvergleichs innerhalb einer Klasse vor allem für die Selbstkonzeptentwicklung. Da bereits mehrfach nachgewiesen werden konnte, dass gerade für leistungsschwache Kinder und Jugendliche solche Vergleiche dauerhaft ein geringeres Selbstbild, vermehrte Schul- und Leistungsangst sowie geringere Motivation und Desinteresse an schulischen Belangen bedingen (vgl. etwa Möller & Trautwein, 2015; Streblov, 2004; Jerusalem, 1985), erfahren Unterrichtsmethoden und Lehrstile, die mit einer Reduzierung solcher Vergleichsmomente einhergehen, besondere Bedeutung.

Der soziale Vergleich unter den Schülerinnen und Schülern einer Klasse ist unter anderem auch Gegenstand der Analysen von König (2006). Dieser untersuchte die Wirkungsrichtungen zwischen der durch die Lernenden wahrgenommenen Konkurrenzorientierung unter den Schülerinnen und Schülern (als ein Indikator der Schüler-Schüler-Beziehung), dem Engagement der Lehrpersonen (als eine Facette der Schüler-Lehrer-Beziehung), dem Selbstkonzept der Begabung und der Misserfolgsattribuierung (Zuschreibung auf eigene Fähigkeiten oder Zufall), auch unter Kontrolle der Schulformzugehörigkeit (Gymnasium, Realschule, Gesamtschule) und des elterlichen Bildungsstands. Herangezogen wurden hierfür die Daten von 816 Berliner Schülerinnen und Schülern der siebten (später achten) Klasse zu zwei Messzeitpunkten; diese entstammen der Studie AIDA (Adaption in der Adoleszenz). Durch das längsschnittliche Design und die statistische Berücksichtigung der Mehrebenenstruktur der Daten konnten die Wirkungsrichtungen und -ebenen zwischen den oben genannten Konstrukten erörtert werden; eine Herangehensweise, die bis zu diesem Zeitpunkt in der schulischen Klimaforschung weitestgehend vernachlässigt wurde. Im Folgenden wird nur auf die Ergebnisse des Lehrerengagements, der Konkurrenzorientierung und des Selbstkonzepts der Begabung eingegangen, da diese für die nachstehenden Analysen (auf Grundlage von Grundschuldaten) von besonderer Bedeutung sind. In einem ersten Schritt untersuchte König (2006), welche Wirkungen in welche Richtung zwischen den unterschiedlichen Konstrukten über die beiden Messzeitpunkte hinweg bestehen. Dabei zeigte sich – teilweise entgegen der Vorannahmen –, dass Schülerinnen und Schüler mit einem hohen Selbstkonzept mehr Konkurrenz zwischen ihren Mitschülern wahrnehmen als Jugendliche mit einem geringeren Selbstkonzept der Begabung ( $\beta = 0,10$  auf  $p < 0,05$  auf Individualebene). Umgekehrt stellt

sich heraus, dass das mittlere Selbstkonzept einer Klasse ansteigt, je mehr Konkurrenz durch die Klasse wahrgenommen wird ( $\beta = 0,29$  mit  $p < 0,05$  auf Klassenebene). Gerade dieser zweite Befund auf Kontextebene erscheint (rückbeziehend auf das durch König, 2006 modifizierte handlungstheoretische Modell von Fend, 1990) erwartungswidrig, sollte sich das Selbstkonzept doch positiv entwickeln je *günstiger* sich die sozialen Ressourcen darstellten. König (2006) führt dies auf Merkmale der Items zurück<sup>54</sup>, aber auch auf den möglichen Umstand, dass soziale Vergleiche in Klassen, die sich durch eine starke Konkurrenz auszeichnen, häufiger vorgenommen werden und ihnen eine größere Bedeutung zukommt als in Klassen mit geringerer Konkurrenzorientierung. „Dass verstärkte Vergleiche dann in einem etwas höheren Klassen-Selbstkonzept resultieren, könnte damit zusammenhängen, dass Schüler/innen bei sozialen Vergleichen tendenziell positive Selbstbeurteilungen vorziehen [...], um ihrem Bedürfnis nach Selbstwertschutz oder Selbstwerterhöhung nachzukommen“ (König, 2006, S. 192). Dass Schülerinnen und Schüler darüber hinaus bestrebt sind, ein positives Selbstbild beizubehalten, könne dagegen erklären, warum die klassenbezogene Lernumwelt eher als kompetitiv wahrgenommen wird, je höher das Selbstkonzept ist: So könnten sich Jugendliche mit positiver Selbsteinschätzung eher herausgefordert sehen, ihre Leistungen zu halten (und die Lernumwelt entsprechend als kompetitiver wahrnehmen) als Schülerinnen und Schüler, deren Selbstkonzept weniger von der schulischen Leistungseinschätzung abhängt (und die entsprechend die Lernumwelt als durch weniger Konkurrenz geprägt wahrnehmen, vgl. König, 2006). Überraschenderweise zeigte sich kein direkter Einfluss des wahrgenommenen Engagements der Lehrkraft auf die Ausprägung des Selbstkonzepts. Der Autor nimmt an, dass sich dieses möglicherweise indirekt über das erlebte Konkurrenzverhalten auf die Ausbildung des Selbstkonzepts auswirken könnte. Neben diesen analysiert König zudem die Beziehung zwischen beiden Facetten der Lernumwelt. Hier zeigte sich ein signifikant negativer Einfluss des kollektiv eingeschätzten Lehrerengagements auf die kollektiv wahrgenommene Konkurrenzorientierung unter den Mitschülerinnen und Mitschülern ( $\beta = -0,29$  bei  $p < 0,05$ ) auf Klassenebene, nicht jedoch auf Individualebene – spätere Mehrebenenanalysen konnten diesen Effekt zudem auf Schul- und nicht auf Klassenebene verorten. Demnach habe das kollektiv wahrgenommene Engagement eines Lehrerkollegiums einer Schule Einfluss

---

<sup>54</sup> So beinhalten die verwendeten Items der Selbstkonzept-Skala ausschließlich soziale Vergleichsperspektiven (etwa „Ich habe den Eindruck, dass ich klüger bin als andere in meinem Alter“, vgl. König, 2006, S. 128); hohe Selbstkonzept-Werte gehen hier also aus dem Vergleich mit der Leistung der Klassenkameraden hervor. Zudem seien alle Selbstkonzept-Items positiv formuliert, was generell eher zu einer stärkeren Zustimmung führe.

auf das Klima (nämlich die kollektiv wahrgenommene Konkurrenzorientierung) einer Klasse. Wenngleich dies ein Hinweis darauf ist, dass die Wirkungen einzelner Klimadimensionen generell differenziert betrachtet werden sollten, müssen diese letztgenannten Befunde aufgrund einer nicht hinreichenden Güte des Modells mit Vorsicht interpretiert werden.

Zusammenfassend lässt sich Folgendes festhalten:

- (1) Wie dies in den Kapiteln 9.5 und insbesondere 9.6 bereits dargestellt wurde, handelt es sich bei der Schüler-Schüler- und Schüler-Lehrer-Beziehung nicht um über verschiedene Studien hinweg eingrenzbare Konstrukte. So werden dem lern- und leistungsförderlichen Lehrerverhalten in den oben dargestellten Studien Merkmale wie der Umgang mit Fehlern, Kränkungen (negative Polung), Geduld, Engagement, Fürsorge, Kontrolle und Opposition (beides negative Polung), Geborgenheit, Verbundenheit, Freundlichkeit, Empathie, Wärme, kognitive Aktivierung, adaptive Anpassung an die Unterschiede der Lernenden, Glaubwürdigkeit, Zuwendung und eine individuelle Bezugsnormorientierung zugeordnet. Bei der Schüler-Schüler-Beziehung geschieht dies weniger ausdifferenziert, hier finden sich Aspekte wie die Konkurrenzorientierung (negative Polung), Unterstützung, Vertrauen, Rücksichtnahme, Geborgenheit und Verbundenheit, Freundschaft und das Geben von Feedback, denen eine lern- und leistungsbezogen positive Wirkung nachgewiesen werden sollte und z.T. nachgewiesen werden konnte.
- (2) Insbesondere bezüglich der Wirkung der sozialen Merkmale auf die Schulleistung finden sich heterogene Ergebnisse, die sich von stark positiven bis hin zu nicht signifikanten Zusammenhängen erstrecken, wobei die Mehrzahl der oben genannten Studien leicht positive Effekte postulieren. Sicherlich stehen diese auch mit der methodischen Herangehensweise (Berücksichtigung der Mehrebenenstruktur der Daten, Repräsentativität der Stichprobe, Kontrolle von Hintergrundvariablen und Eingangsniveaus etc.) in Verbindung. Zu entnehmen ist diesem die Tendenz, dass beide sozialen Merkmale einen schwach positiven Einfluss auf den kognitiven Schulerfolg von Kindern und Jugendlichen haben, der insbesondere dann stärker ausfällt, wenn er anhand von Noten (statt mithilfe standardisierter Leistungstests) gemessen wird. Anzunehmen ist zudem, dass sich beide sozialen Merkmale indirekt, etwa über motivationale Charakteristika der Schülerinnen und Schüler auf die Leistung auswirken.

- (3) Deutlich stärkere Beziehungen finden sich zwischen der Schüler-Schüler- bzw. Schüler-Lehrer-Beziehung und affektiv-motivationalen Merkmalen der Lernenden, wenngleich letztere ebenfalls weit gefasst sind. So zählen zu diesen die Kreativität, Selbstachtung und Unabhängigkeit der Schülerinnen und Schüler ebenso wie ein störendes Verhalten (negative Polung), Leistungs- bzw. Prüfungsangst, Lernfreude und das lernbezogene Verhalten, das Selbstkonzept, Selbstbewusstsein, Selbstwirksamkeitserwartungen, Optimismus, Kontrollüberzeugungen, Schulfreude, Zufriedenheit, allgemeine schulische Einstellungen, Motivation und ähnliches. Auch die Wirkungen auf die psychische und physische Gesundheit der Kinder und Jugendlichen wurde untersucht. Dabei finden sich lediglich in einem Fall keine signifikant positiven Effekte (vgl. Ausführungen zu Ryan et al., 1994). Wenngleich diese nachgewiesenen Zusammenhänge kaum zu einem einheitlichen Wirkschema zusammengefasst werden können, lässt sich daraus jedoch zumindest ableiten, dass das (individuell wahrgenommene) Lehrerverhalten sowie eine (wahrgenommene) gelungene Integration in den Klassenkontext stärker mit Effekten auf affektiv-motivationale Schülercharakteristika einhergehen als dass diese die kognitive Leistung der Lernenden beeinflussen.
- (4) Bezüglich des Wirkzusammenhangs zwischen Schüler-Schüler- und Schüler-Lehrer-Beziehung kann auf Grundlage der oben dargestellten Studien ebenfalls eine Tendenz abgeleitet werden: Wie dies die Untersuchungen von Moos (1979), Hattie (2009) und König (2006) nahelegen, beeinflusst ein soziales, sich u.a. durch Empathie, Wärme, adaptive Anpassung an die Unterschiede der Lernenden und Engagement auszeichnendes Lehrerverhalten die Hilfsbereitschaft, Freundlichkeit und Konkurrenzorientierung (negative Polung) der Schülerinnen und Schüler untereinander. Anzumerken ist hier, dass es sich lediglich bei König (2006) um eine längsschnittliche Untersuchung, die solche Kausalaussagen zulässt, handelt. Die Annahmen der beiden übrigen Studien erfolgen theoretisch und auf Grundlage von Korrelationen. Für die folgenden Analysen wird dennoch ein regressiver Zusammenhang (Beeinflussung der Schüler-Schüler-Beziehung durch die Schüler-Lehrer-Beziehung) angenommen, wenngleich aufgrund der querschnittlichen Datengrundlage auch hier keine Aussagen über Wirkrichtungen getroffen werden können.
- (5) Die Berücksichtigung der Mehrebenenstruktur der Daten und ihre zum Teil differierenden Wirkzusammenhänge auf und zwischen den Ebenen findet nur in wenigen

der oben genannten Untersuchungen Anwendung, was zum Teil sicherlich darauf zurückzuführen ist, dass entsprechende Analysemethoden erst seit etwa 15 Jahren weitestgehend etabliert, einige Studien jedoch älter sind. Wie dies Satow (2000) und König (2006) zeigen, weist die *kollektive* Wahrnehmung von Merkmalen der Lernumwelt zum Teil andere (meist klassenbezogene) Effekte auf als diese auf Individual-ebene zu finden sind – wenngleich diese schwächer ausfallen als auf L1. Wie König (2006) anmerkt, sollte die Wirkung von Facetten der Lernumwelt (in Anlehnung an entsprechende theoretische Annahmen) auf und zwischen den unterschiedlichen Ebenen daher differenziert betrachtet werden.

Aus dem in diesem Kapitel dargestellten Forschungsstand kann also abgeleitet werden, dass beide soziale Facetten der Schüler-Lehrer- und Schüler-Schüler-Beziehung in einem (wenn auch schwachen) positiven Zusammenhang mit der kognitiven Leistung und (vergleichsweise stärker) mit affektiv-motivationalen Schülercharakteristika stehen und die soziale Integration in den Klassenkontext darüber hinaus durch die Ausprägung der Schüler-Lehrer-Beziehung beeinflusst wird. Das Folgekapitel widmet sich dem Einfluss individueller Förderung auf diese beiden Lernumweltmerkmale.

### **9.8 Individuelle Förderung und soziale Merkmale der Lernumwelt**

Aus den Darstellungen des vorangegangenen Kapitels wird deutlich, welche Merkmale Schüler-Schüler- und Schüler-Lehrer-Beziehungen aufweisen, denen ein bedeutsamer Effekt auf den Schulerfolg von Kindern und Jugendlichen nachgewiesen werden konnte. Dabei zeichnet sich eine positive Schüler-Lehrer-Beziehung vor allem durch Fürsorge aus: Der Lehrer bzw. die Lehrerin widmet sich den einzelnen Lernenden und ihren spezifischen Voraussetzungen und Bedürfnissen, bietet einen geborgenen Raum, in dem Fehler erlaubt sind und zeigt Geduld und Unterstützung für individuelle Lernprozesse und -erfolge, die idealerweise anhand vorheriger Leistungen des Einzelschülers bewertet und rückgemeldet werden (individuelle Bezugsnormorientierung). Ein solches Lehrerverhalten mindert zudem soziale (leistungsbezogene) Vergleiche, sodass Wettbewerbs- und Konkurrenzmomente unter den Lernenden reduziert werden können, was gleichzeitig der Ausbildung von positiven Schüler-Schüler-Beziehungen zugutekommt (vgl. Wagner, 1982). Ähnlich wird dies durch Rosenholtz und Simpson (1984) beschrieben, nach welchen soziale Vergleiche gerade dann verstärkt werden, wenn sich der Unterricht unter anderem durch gleiche Anforderungen an alle Lernen-

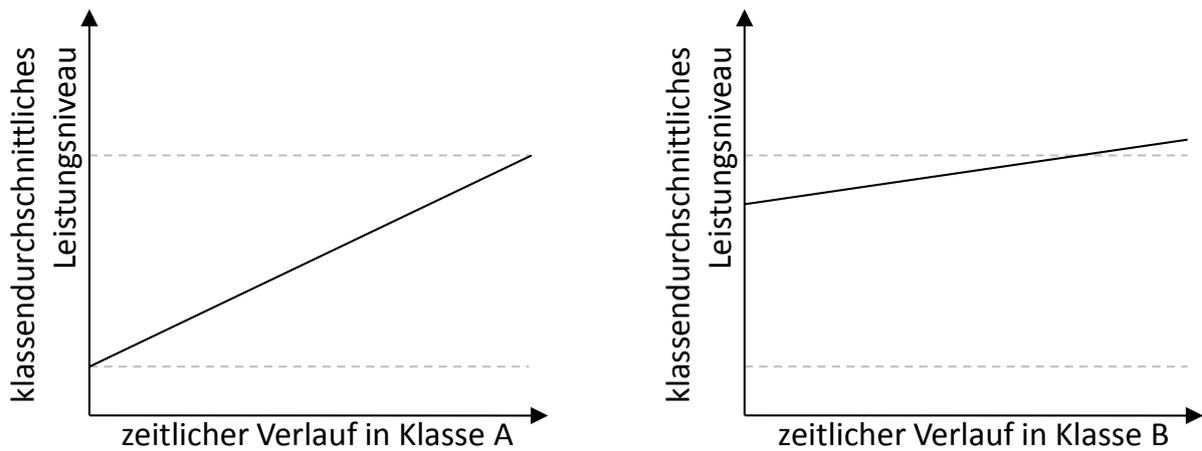
den, öffentliche Leistungsbewertungen und die Bildung von leistungshomogenen Lerngruppen auszeichnet – also Elementen, wie sie in individualisierten Settings explizit nicht eingesetzt werden. Die lern- und leistungsförderliche Wirkung einer gelungenen sozialen Integration in den Klassenkontext zeigt sich vor allem dann, wenn die Schülerinnen und Schüler das Gefühl haben, ihren Klassenkameraden vertrauen zu können, Hilfe und Unterstützung leisten und annehmen zu können, Rücksicht zu erfahren und mit ihren Mitschülerinnen und Mitschülern nicht in Konkurrenz zu stehen. Anzunehmen ist folglich, dass ein Unterricht, der sich vorrangig durch individuelle Förderung auszeichnet, einer positiven Ausgestaltung der Schüler-Lehrer- wie Schüler-Schüler-Beziehung zuträglich ist. Herleiten lässt sich dies vor allem daraus, dass individualisierte Lernsettings, in denen im Extremfall alle Schülerinnen und Schüler an unterschiedlichen, an ihre Voraussetzungen und Interessen angepassten Aufgaben arbeiten und ggf. entsprechend individualisierte Leistungsrückmeldungen erhalten, weniger Möglichkeiten der sozialen Vergleiche und damit der Ausbildung von als konkurrenzorientiert wahrgenommenen Schüler-Schüler-Beziehungen bieten. Dazu könnte Individualisierung die Kinder und Jugendlichen dafür sensibilisieren, dass Leistungsergebnisse nicht Hauptaugenmerk von Unterricht sein müssen, sondern beispielsweise auch das Lernen an sich und das Interesse hieran. Dies könnte es den Kindern und Jugendlichen erleichtern, Hilfe und Unterstützung ihrer Klassenkameraden anzunehmen (und anzubieten): Dieselbe Handlung eines Kindes – nämlich die Bitte um Unterstützung – könnte dann also nicht nur dahingehend interpretiert werden, dass es nicht in der Lage sei, alleine eine Leistung zu erzielen, sondern auch oder vermehrt in dem Sinne, dass es Hilfe dabei benötige, sich einen Lerngegenstand anzueignen. Solch eine alternative Interpretation würde dann als weniger selbstwertbedrohlich eingestuft werden und zudem die Einschätzung der Schüler-Schüler-Beziehung als unterstützend fördern.

Darüber hinaus ist denkbar, dass in 1:1-Interaktionen zwischen Lehrenden und Lernenden, die den individualisierten Unterricht mitunter charakterisieren, lernbezogene wie außerschulische Probleme und Voraussetzungen, aber auch besondere Neigungen der Schülerinnen und Schüler vermehrt thematisiert und berücksichtigt werden können, was im Umkehrschluss zu einer gestärkten Beziehung zwischen Lernenden und Lehrenden führen kann. Überprüft wird im Folgenden also, ob sich eine individuelle Förderung auf die durch die Schülerinnen und Schüler wahrgenommene Schüler-Schüler- und Schüler-Lehrer-Beziehung positiv auswirken kann. In Anlehnung an die in Kapitel 8.3 dargestellten Rahmenmodelle des

Schulerfolgs, die u.a. konstatieren, dass sich das Lehrverhalten nicht linear und direkt auf den Schulerfolg von Kindern und Jugendlichen auswirken muss, und unter Berücksichtigung der Tatsache, dass sowohl bezüglich einer positiven Schüler-Schüler- als auch Schüler-Lehrer-Beziehung ein positiver Effekt auf die Leseleistung und das Leseselbstkonzept angenommen wird, wird zudem überprüft, ob sich die individuelle Förderung möglicherweise vermittelt über die beiden hier berücksichtigten sozialen Facetten der schulischen Lernumwelt *indirekt* auf den Schulerfolg auswirkt.

## **10. Zusammensetzung der Klasse**

Wie dies in Kapitel 8.1 dargestellt wurde, stellt die Schulleistung das zentrale Kriterium für die Bewertung von unterrichtlicher Wirkung dar. Berücksichtigt werden muss in diesem Kontext jedoch, dass Unterricht immer in einem bestimmten Setting stattfindet, welches eine maßgebliche Bedingung für sein Gelingen darstellt. So sei nach Weinert (1997, S. 51) die Qualität des Unterrichts „nicht nur von der Persönlichkeit und Kompetenz des einzelnen Lehrers abhängig, sondern wird z.B. auch vom durchschnittlichen Niveau und der Variationsbreite kognitiver Leistungsvoraussetzungen der Schüler in einer Klasse beeinflusst“. Dass die Klassenzusammensetzung eine wichtige Rahmenbedingung darstellt, die den Unterrichtserfolg beeinflusst, argumentiert auch Helmke (2009), der neben Merkmalen des sozialen Hintergrunds (Bildungsferne, kognitiver Anregeungsgehalt, Ausbildungsniveau der Eltern) auch auf den Sprachhintergrund der Schülerinnen und Schüler als Klassenmerkmal verweist. Erst die Berücksichtigung solcher Aspekte der Klassenzusammensetzung führt zu einem fairen Urteil dessen, was Unterricht leistet. Abbildung 11 verdeutlicht dies an einem verwandten Beispiel.

**Abbildung 11:** Leistungsentwicklung in Klassen unterschiedlicher Eingangsvoraussetzungen

Dargestellt sind zwei Klassen, die sich durch unterschiedliche leistungsbezogene Eingangsvoraussetzungen und eine unterschiedlich starke Leistungsentwicklung auszeichnen. Die Schülerinnen und Schüler der Klasse A weisen zu Beginn eines Beobachtungszeitraums ein geringeres durchschnittliches Leistungsniveau auf als die Schülerinnen und Schüler der Klasse B. Würde nun anhand des Leistungsniveaus *am Ende* des Messzeitraums ein didaktisches Unterrichtsetting auf seine Qualität hin beurteilt werden, würde Klasse B besser abschneiden. Würde allerdings berücksichtigt, unter welchen Bedingungen (hier: klassendurchschnittliches Leistungsniveau *zu Beginn* des Beobachtungszeitraums) unterrichtet wurde, würde das Urteil zugunsten der Klasse A ausfallen, da hier die Leistungsentwicklung über den Erhebungszeitraum deutlich positiver ist; die didaktische Maßnahme (bzw. die Lehrperson) deutlich mehr leistet als in der Vergleichsgruppe B<sup>55</sup>. Dies kann übertragen werden auf den oben vorgestellten Diskurs: Eine Lehrperson, die in einer Klasse unterrichtet, in der beispielsweise ein geringes Vorkenntnisniveau, geringe familiäre Anregungsgehalte und ungünstige sozioökonomische Herkunftsmerkmale vorherrschen, müsste deutlich mehr leisten, um dasselbe Leistungsniveau zu erreichen, das eine Lehrkraft in einer herkunftsbezogen privilegierten Klasse womöglich mit vergleichsweise geringer Anstrengung erreicht.

In der vorliegenden Arbeit sollen ebenfalls Merkmale der Klasse berücksichtigt werden, unter denen individuelle Förderung stattfindet. In Anlehnung an die oben vorgestellten Aspekte werden die klassendurchschnittlichen familiären Ressourcen (Economic, Social and Cultural Status, ESCS, s. Kapitel 12.3.1) und der klassendurchschnittliche Sprachgebrauch („Wie oft

<sup>55</sup> Anzumerken sei, dass es sich hierbei um eine stark verkürzte, modellhafte Darstellung handelt. Unberücksichtigt bleibt dabei natürlich, dass sich eine Lehrkraft in leistungsstarken Klassen beispielsweise weiteren Lernzielen widmen kann, die dann nicht in die Messung eines vorab definierten Outputs einfließen – was im Umkehrschluss also nicht bedeuten muss, dass eine Lehrperson didaktisch weniger geleistet hat.

sprichst du zu Hause Deutsch?“) aufgenommen. Nicht berücksichtigt wird das Leistungsniveau der Klassen, da hier nicht – wie beispielsweise in Studiendesigns mit mehreren Messzeitpunkten – für Eingangsniveaus und eine entsprechende Leistungsentwicklung kontrolliert werden kann. Die klassendurchschnittliche Lesekompetenz wird hier als abhängige, durch individuelle Förderung zu erklärende Variable und nicht als Merkmal der Klassenzusammensetzung, welche die Häufigkeit individueller Förderung beeinflusst, herangezogen. Eine Kontrolle von Unterrichtseffekten bleibt Studien vorbehalten, die Verlaufs- bzw. Prozessdaten erheben.

## **11. Fragestellung und Ableitung der Hypothesen**

In den vorangegangenen Kapiteln wurde zunächst dargestellt, dass die sich auf Schulklassen bezogene Heterogenität, hier verstanden als die kulturellen, ökonomischen und sprachlichen Unterschiede sowie die daraus und aus den unterschiedlichen häuslichen Anregungsgehalten resultierenden leistungs- und selbstbezogenen Ausgangsbedingungen von Schülerinnen und Schülern trotz Homogenisierungsbestrebungen nach wie vor ein charakterisierendes Merkmal des Schulalltags darstellt. Gerade im Zuge der Inklusions- und aktueller der Flüchtlingsdebatte ist nicht davon auszugehen, dass sich Klassenzusammensetzungen zukünftig vereinheitlichen werden. Doch selbst wenn dies gelinge, würde es das Lehren im Unterricht nicht erleichtern: Wie die dargestellten lernpsychologischen, konstruktivistischen, reformpädagogischen und neurowissenschaftlichen Perspektiven konstatieren, zeigen Schülerinnen und Schüler unter denselben (äußerlichen) Rahmenbedingungen unterschiedliche Lernerfolge, was auf ihre unterschiedlichen Lernvoraussetzungen, Interessen, Neigungen, ihr Vorwissen und ihre Lernmotivation zurückgeführt werden kann. Aus diesen Umständen heraus werden hohe Erwartungen an das didaktische Konzept der individuellen Förderung gestellt, welches als „Schlüssel“ (Klieme & Warwas, 2011, S. 805) für den Umgang mit den unterschiedlichen lern- und leistungsbezogenen Voraussetzungen der Schülerinnen und Schüler bezeichnet wird und jüngst aufgenommen wurde in nahezu alle bundeslandspezifischen Schulgesetze (vgl. Fischer, 2014). Diesen vor allem bildungstheoretischen und politischen Forderungen gegenübergestellt werden muss, dass individuelle Förderung bisher keinem einheitlichen, praxiserprobten Konzept folgt und unterschiedlichen Definitionen unterliegt. Gerade in Abgrenzung zu Formen der inneren Differenzierung (auch: Binnendifferenzierung) wurde hier die *Planung der Passung* der Unterrichtsangebote an die individuellen lern- und

leistungsbezogenen Voraussetzungen der Schülerinnen und Schüler in das Zentrum der theoretischen Betrachtung gerückt; dies zeigt jedoch bei Weitem keine Allgemeingültigkeit. Dass sich die Definitionen und das Verständnis individueller Förderung so sehr unterscheiden, ist möglicherweise auch ein Grund dafür, dass die wenigen bisher hierzu vorliegenden Forschungsergebnisse kaum verallgemeinert werden können. Lediglich festhalten lässt sich, dass dieses didaktische Unterrichtssetting scheinbar einen größeren Einfluss auf nicht-kognitive Merkmale der Lernenden (Motivation, Lernfreude, Selbstkonzept und ähnliche) zu haben scheint als auf kognitive Dimensionen (fachspezifische Testleistungen) – für letztere fallen Korrelationsparameter in einigen wenigen Untersuchungen sogar negativ aus, was in umgekehrter Kausalrichtung dahingehend interpretiert wird, dass individuelle Förderung gerade dann angewendet wird, wenn das Leistungsniveau einer Klasse gering ist. Längsschnittliche Untersuchungen, in welchen solche Zusammenhänge *nachgewiesen* werden könnten, finden sich bislang kaum. Ebenfalls unberücksichtigt bleibt bisher, ob individuelle Förderung möglicherweise eine ausgleichende Wirkung auf den Herkunftseffekt – also den Einfluss familiärer Merkmale auf den Schulerfolg von Kindern und Jugendlichen – hat. Wie vielfach gezeigt werden konnte, verfügen Schülerinnen und Schüler aus bildungsnahen Elternhäusern in der Regel über mehr Vorwissen und lerngünstigere Einstellungen, bewegen sich im Schulkontext sicherer, da sie mit den dort vorherrschenden Regeln vertrauter sind und zeigen als ein Resultat im Durchschnitt größere Schulerfolge als Schülerinnen und Schüler bildungsferner Familien. Daher wird hier der übergeordneten Frage nachgegangen, ob ein Unterricht, der den Lerngegenstand für Kinder und Jugendliche bedeutsam macht indem er an ihre individuellen Interessen, Neigungen, Begabungen, Lernvoraussetzungen und an ihr Vorwissen angepasst wird (individuelle Förderung), die systematische Abhängigkeit des Schulerfolgs von der sozioökonomischen Herkunft der Schülerinnen und Schüler abschwächen kann. Ziel ist es dabei nicht, die Leistungsstreuung innerhalb einer Klasse zu verringern (also möglicherweise nur die leistungsschwächeren Schülerinnen und Schüler zu fördern), sondern den individuell größtmöglichen Schulerfolg für jedes Kind zu begünstigen. Dabei wird in der hier vorgestellten Untersuchung ein Fokus nicht nur auf kognitive Merkmale des Schulerfolgs gelegt (hier: Leseleistung), sondern auch auf die Entwicklung eines gestärkten Selbstkonzepts (hier: im Lesen). Dass diese beiden Konstrukte miteinander in hierarchischem Verhältnis zueinander stehen können, soll aufgrund der hier nur querschnittlichen Betrachtung vernachlässigt und stattdessen von einem reziproken Zusammen-

hang ausgegangen werden, wie ihn Marsh (1990, 2003), Muijs (1997), Wigfield und Karpathian (1991) und Skaalvik und Hagtvet (1990) (zumindest für einige Klassenstufen) konstatieren. Daraus ergeben sich folgende Annahmen:

### **11.1 Herkunftseffekt**

Die soziale Herkunft, verstanden als die unterschiedlichen Ausprägungen des sozialen, kulturellen und ökonomischen Kapitals einer Familie, beeinflusst sowohl die Leseleistung (Hypothese 1) als auch das Selbstkonzept im Lesen (Hypothese 2) positiv, die beide als Indikatoren des Schulerfolgs herangezogen werden. Ebenso beeinflusst das lesebezogene Anregungspotenzial der häuslichen Lernumgebung die Leseleistung (Hypothese 3) sowie das Selbstkonzept im Lesen (Hypothese 4) positiv. Da gezeigt werden konnte, dass die familiäre Lesesozialisation durch das sozioökonomische Kapital einer Familie beeinflusst wird, wird angenommen, dass die für Hypothese 1 und Hypothese 2 beschriebene Ausprägung der Kapitalsorten einen positiven Einfluss auf die häusliche Lernumgebung zeigt (Hypothese 5).

### **11.2 Effekte der Wahrnehmung der schulischen Lernumwelt**

Ziel dieser Ausarbeitung ist es, die Abhängigkeit des Schulerfolgs von dem didaktischen Unterrichtssetting individueller Förderung zu analysieren. In Anlehnung an die komplexen Modelle schulischer Wirkungsweisen kann jedoch nicht davon ausgegangen werden, dass das Lehrverhalten direkte und lineare Lerneffekte bei den Schülerinnen und Schülern hervorruft. Neben individuellen Interpretationsprozessen des Unterrichtsangebots ist es auch die Wahrnehmung der schulischen Lernumwelt, die den Schulerfolg von Kindern und Jugendlichen beeinflusst. Der Forschung zum Schul- und Klassenklima entlehnt, wird der Fokus auf die sozialen Facetten der Schüler-Schüler- sowie Schüler-Lehrer-Beziehung gelegt. Dabei wird angenommen, dass eine gelungene soziale Integration in den Kreis der Klassenkameraden das Bedürfnis nach sozialer Eingebundenheit (vgl. Deci & Ryan, 1993) erfüllt, was Kapazitäten für Lerninhalte als Grundlage der Lernbereitschaft schafft und die sozioemotionale Verarbeitung der schulischen Leistungsanforderungen erleichtert. Es ist entsprechend davon auszugehen, dass eine als positiv eingeschätzte Schüler-Schüler-Beziehung die Leistung der Kinder und Jugendlichen positiv beeinflusst (Hypothese 6). Für ihren Zusammenhang mit dem Selbstkonzept wird von einem reziproken Effekt ausgegangen: So ist es denkbar, dass negativ ausfallende Vergleiche mit den Leistungen der Mitschülerinnen und -schüler weniger stark das Selbstkonzept beeinflussen, wenn ein Kind über eine gestärkte soziale Anbindung

im Klassenkontext verfügt als wenn ein Lerner keinen sozialen Rückhalt erfährt. Andererseits könnten anhaltend ungünstig ausfallende Vergleiche und die damit verbundene Wahrnehmung, dauerhaft schlechter zu sein als die Klassenkameraden, die soziale Integration im Klassenkontext mindern. Da solche Kausalrichtungen im vorliegenden Querschnitt nicht festgestellt werden können, soll hier von einem reziproken, positiven Zusammenhang zwischen der Schüler-Schüler-Beziehung und der Ausprägung des Selbstkonzepts (im Lesen) ausgegangen werden (Hypothese 7).

Die Schüler-Lehrer-Beziehung wirkt sich dagegen vor allem auf motivationale Aspekte aus, die hier nicht expliziter Bestandteil der Analyse sein sollen. Dennoch ist davon auszugehen, dass sich eine (durch die Schülerinnen und Schüler wahrgenommene) Akzeptanz durch die Lehrperson vermittelt durch eine gesteigerte Motivation und ein stärkeres schulisches Engagement positiv auf die Leistung von Kindern und Jugendlichen auswirkt (Hypothese 8). Für den Einfluss auf das Selbstkonzept sollten sich gemäß bisherigen Forschungsergebnissen stärkere positive Zusammenhänge finden lassen (Hypothese 9). Ebenfalls aus diesen abgeleitet ist, dass sich ein soziales, empathisches, adaptives Lehrerverhalten positiv auf die Hilfsbereitschaft, Freundlichkeit und negativ auf die Konkurrenzorientierung der Schülerinnen und Schüler auswirkt, sodass im Folgenden von einer positiven Beeinflussung der als durch die Lernenden positiv eingeschätzten Schüler-Lehrer-Beziehung auf die Schüler-Schüler-Beziehung ausgegangen wird (Hypothese 10).

### **11.3 Schulerfolg**

Wie eingangs beschrieben, wird der Schulerfolg von Viertklässlerinnen und Viertklässlern hier nicht nur anhand kognitiver Merkmale (Testergebnisse, hier: Leseleistung), sondern auch anhand der Entwicklung des Selbstkonzepts (hier: im Lesen) bewertet, da neben der Leistung beispielsweise auch die Stärkung der Selbstwahrnehmung als wichtiges Ziel von Schule verstanden wird. Da aufgrund des querschnittlich angelegten Studiendesigns zeitlich hierarchische Wirkrichtungen nicht überprüft werden können, soll folgend von einem reziproken Zusammenhang zwischen Selbstkonzept und Leistung ausgegangen werden (Hypothese 11).

### **11.4 Effekte der Klassenzusammensetzung**

Ein wichtiges Kriterium für die Bewertung der Qualität von Unterricht ist die Klassenzusammensetzung. Ihre Berücksichtigung ermöglicht eine faire Beurteilung dessen, was Unter-

richtsmethoden wie die individuelle Förderung leisten können, da schwierige ebenso wie ideale Eingangsvoraussetzungen der Schülerinnen und Schüler in die Beurteilung didaktischer Settings einfließen und idealerweise um sie korrigiert werden können. Wenngleich ein solch kontrolliertes Design hier nicht vorliegt (benötigt würde dafür ein Prä-Post-Design, welches die Kontrolle von Eingangswerten erlaubt), soll doch zumindest der Einfluss einiger Klassenmerkmale auf die Häufigkeit individueller Förderung betrachtet werden. Dabei handelt es sich um das Klassenaggregat der sozioökonomischen Familienmerkmale der Schülerinnen und Schüler (Hypothese 12) und das Klassenaggregat der Häufigkeit, mit der zu Hause Deutsch gesprochen wird (Hypothese 13). Da aus rezenten Arbeiten (vgl. Schulz-Heidorf & Solheim, 2016) bekannt ist, dass Lehrende bisher zögerlich Formen der individuellen Förderung im Unterricht anwenden, wird davon ausgegangen, dass eine geringe Ausprägung dieser Merkmale (eine im Klassenmittel geringe soziale Herkunft der Schülerinnen und Schüler, geringer Gebrauch der deutschen Sprache im häuslichen Umfeld) eine häufigere individuelle Förderung nach sich zieht, da sie dann möglicherweise den Lehrenden als besonders angebracht erscheint. Wäre individuelle Förderung bereits ein etabliertes und weitläufig implementiertes Unterrichtssetting, könnte hier einer anderen Argumentation gefolgt werden: Dann dürften solche oben beschriebenen Merkmale der Klassenzusammensetzung keinen Einfluss auf die Häufigkeit individueller Förderung zeigen.

### **11.5 Effekte der individuellen Förderung**

Der theoretischen Annahmen folgend wird eine positive Beeinflussung der Leistung (Hypothese 14) und des Selbstkonzepts (Hypothese 15) der Schülerinnen und Schüler durch die individuelle Förderung angenommen, da Kinder und Jugendliche entsprechend ihren eigenen Lernständen, Voraussetzungen, Neigungen und Interessen lernen können, was selbiges bedeutsamer und entsprechend erfolgreicher machen sollte. Zugleich wird angenommen, dass durch individualisierte Unterrichtssettings soziale Vergleiche mit den Leistungen der Mitschülerinnen und Mitschüler gemindert werden, was gerade einer positiven Selbstkonzeptentwicklung zugutekommen könnte. Aus bisherigen Forschungsergebnissen ableitend sollte dieser Effekt für das affektive Merkmal des Schulerfolgs (Leseselbstkonzept) stärker ausfallen als für die kognitive Dimension (Leseleistung).

Hauptforschungsinteresse im Rahmen der vorliegenden Arbeit ist jedoch nicht der direkte Einfluss der individuellen Förderung auf den Schulerfolg, sondern ihre potenziell moderie-

rende Funktion bezüglich des Herkunftseffekts: Wird das Lernangebot so aufbereitet, dass es den (durch die soziale Herkunft bedingten) Lernständen, -voraussetzungen, Interessen und Neigungen der einzelnen Schülerinnen und Schüler in einem individuell(er)en Interaktionsrahmen entgegenkommt, könnte die systematische Benachteiligung von Kindern und Jugendlichen aus bildungsfernen Familien, die mit dem schulimmanenten Habitus, seinen Regeln und Strukturen nicht vertraut sind und darüber hinaus im familiären Kontext weniger Unterstützung und Anregung erfahren, gemindert werden. Dies würde sich in einem negativen Einfluss der individuellen Förderung auf den Herkunftseffekt zeigen. Berechnet wird dies für alle vier hier differenzierten Herkunftseffekte: für den Effekt des sozioökonomischen Kapitals einer Familie auf die Leistung (negative Beeinflussung durch individuelle Förderung: Hypothese 16) und das Selbstkonzept (Hypothese 17) ebenso wie für den Effekt der häuslichen Lesesozialisation auf die Leistung (Hypothese 18) und das Selbstkonzept (Hypothese 19).

Eine weitere hypothetische Annahme ist die Beeinflussung der schülerbezogenen Wahrnehmung der sozialen Umwelt durch die individuelle Förderung: Durch die individualisierten Interaktionssettings zwischen Lehrenden und Lernenden könnte die Schüler-Lehrer-Beziehung gestärkt werden (Hypothese 20). Mit ihr einher geht eine Minderung der sozialen Vergleiche innerhalb der Klasse, so dass die soziale Integration (Schüler-Schüler-Beziehung) möglicherweise weniger durch diese belastet wird (Hypothese 21). Vermittelt über diese beiden Facetten könnte sich die individuelle Förderung zudem indirekt auf die Leseleistung (vermittelt über die Schüler-Lehrer-Beziehung: Hypothese 22, über die Schüler-Schüler-Beziehung: Hypothese 23) und das Leseselbstkonzept (Hypothese 24 für die Schüler-Lehrer-Beziehung, Hypothese 25 für die Schüler-Schüler-Beziehung) auswirken.

Diese in den vorausgegangenen Kapiteln theoretisch und aus bisherigen Forschungsarbeiten abgeleiteten Zusammenhänge sollen im Folgenden einer empirischen Überprüfung unterzogen werden. Anzumerken sei an dieser Stelle nochmals explizit, dass es sich dabei um eine querschnittliche Betrachtung handelt, die Aussagen über Kausalrichtungen nicht erlaubt. Mit Vorsicht müssen daher die theoretisch begründeten gerichteten Effekte der konfirmatorischen Strukturgleichungsmodellierung interpretiert werden.

## **12. Empirische Analysen**

Das folgende Kapitel widmet sich zunächst der Datengrundlage, die für die nachfolgenden Analysen genutzt werden soll. Dem schließt sich die Darstellung der Analysemethode und die mit ihr einhergehenden Kennwerte und Spezifika an, bevor das zu analysierende Modell in Einzelschritten aufgebaut und geprüft wird.

### **12.1 Datengrundlage: IGLU-E 2011**

Während in Kapitel 8.1 bereits die Struktur, das Rahmenmodell und das Verständnis von reading literacy, das der Internationalen Grundschul-Lese-Untersuchung (IGLU, im internationalen Kontext: PIRLS) zugrunde liegt, erläutert wurden, werden im Folgenden die zur Verfügung stehenden Daten und ihre Charakteristika beschrieben.

Für die nachstehenden Analysen werden Daten aus IGLU-E aus der aktuellsten Erhebungswelle 2011 herangezogen. Die begriffliche Differenzierung der Studie erfolgt dabei nach der Verfügbarkeit der Daten: Während in IGLU bzw. PIRLS (Progress in International Reading Literacy Study) kaum länderspezifische Items eingesetzt wurden, die (frei) verfügbaren Daten also eine hohe internationale Vergleichbarkeit gewährleisten, handelt es sich bei IGLU-E um eine nationale, deutschlandspezifische Erweiterung der eingesetzten Items und der Stichprobe. Letzteres ermöglicht innerdeutsche Vergleiche, etwa zwischen den verschiedenen Bundesländern. IGLU-E verfügt im internationalen Vergleich zudem über mehr Fragen z.B. im Rahmenfragebogen der Schülerinnen und Schüler, etwa zu individuellen Lernbedingungen wie Leistungsangst und Misserfolgsattributionen, zu den Wahrnehmungen der Schule, beispielsweise die Einschätzung der eigenen sozialen Integration und zu den Lebens- und Lernbedingungen in der Familie, zu denen etwa das Vorhandensein bestimmter Wohlstandsgüter und das Interesse der Eltern an den Aktivitäten der Kinder gehört. Auch der Lehrerfragebogen ist umfangreicher, hier finden sich beispielsweise mehrere Items zur individuellen Förderung, während international lediglich ein Item verfügbar ist.

Deutschland nimmt seit 2001 an PIRLS, einer Studie der International Association for the Evaluation of Educational Achievement (IEA) teil, deren Erhebungen alle fünf Jahre stattfinden. Die Teilnahme wird durch das Bundesministerium für Bildung und Forschung sowie durch die Kultusministerkonferenz gefördert.

Um einen Überblick über die verfügbaren Daten zu gewinnen, werden im Folgenden kurze deskriptive Statistiken angeführt. Dem schließt sich der Aufbau der Einzelmodelle und schließlich der Aufbau des gesamten Mehrebenenstrukturgleichungsmodells an.

Die gesamtdeutsche Stichprobe besteht aus 4.000 Schülerinnen und Schülern, die im Mittel 10,4 Jahre alt sind (Standardabweichung: 0,5). Etwa die Hälfte (49,5%) der Viertklässlerinnen und Viertklässler sind weiblich. Die Stichprobe ist in 205 Klassen an 197 Schulen gegliedert, 222 Lehrerinnen und Lehrer wurden befragt. 85% der Schülerinnen und Schüler werden von Lehrerinnen unterrichtet, 10% von Lehrern, für die übrigen Lehrkräfte liegen keine Angaben zu dem Geschlecht vor. In 14 Klassen erhalten die Lernenden im sogenannten Teamteaching Unterricht durch je zwei Lehrpersonen, in einer Klasse durch drei Lehrende. Im Durchschnitt werden die Viertklässlerinnen und Viertklässler beschult von Lehrerinnen und Lehrern<sup>56</sup>, die über 18 Jahre Unterrichtserfahrung (Standardabweichung: 12,2) verfügen und im Mittel zwischen 40 und 49 Jahre<sup>57</sup> alt sind.

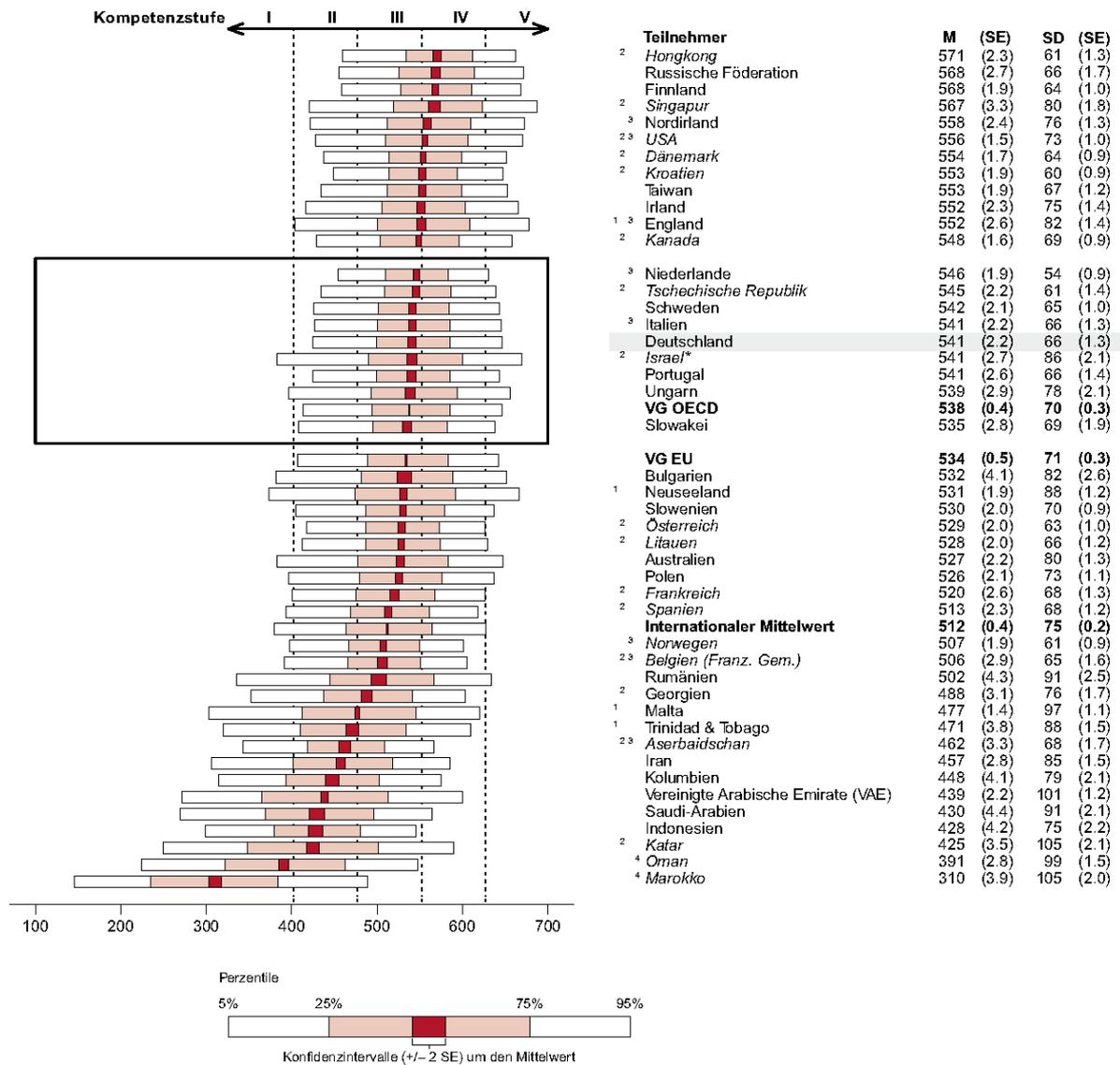
Die Leseleistung der deutschen Schülerinnen und Schüler liegt im internationalen Vergleich im oberen Drittel: Mit durchschnittlich 541 Punkten auf der Kompetenzskala (Mittelwert 500, Standardabweichung 100) schneiden die deutschen Viertklässlerinnen und Viertklässler dabei signifikant besser ab als die Vergleichsgruppe der Europäische Union (VG<sub>EU</sub>, 534 Punkte) und liegen ebenfalls oberhalb des internationalen Mittelwerts (512 Punkte). Von dem Gesamtmittelwert der OECD unterscheidet sich Deutschland nicht signifikant und liegt auf einer Kompetenzstufe mit bspw. den Niederlanden, Schweden, Ungarn und der Slowakei. Die Spitzengruppe, welche mit durchschnittlich 558 Punkten deutlich bessere Leistungen zeigt (Maximum: Hongkong mit 571 Punkten), erreicht Deutschland jedoch nicht (vgl. Abbildung 12).

---

<sup>56</sup> Zu beachten ist bei allen Aussagen, die über Lehrpersonen im Rahmen von IGLU getroffen werden, dass diese Stichprobe nicht zufällig gezogen wurde, also keine repräsentative Stichprobe der *Lehrpersonen* (wohl aber der Viertklässlerinnen und Viertklässler) in Deutschland bildet. Die Angaben der Lehrpersonen wurden immer in Referenz zu der jeweiligen Klasse getroffen, sind daher ausschließlich in Zusammenhang mit der Schülerstichprobe zu interpretieren (als Lehrpersonen, die eine repräsentative *Schülerstichprobe* unterrichten) und können nicht verallgemeinert werden (weitere Informationen zur Stichprobenziehung in Kapitel 12.2).

<sup>57</sup> Das Alter der Lehrpersonen wurde in Kategorien erfasst: unter 25, 25-29, 39-39, 40-49, 50-59 und 60 oder älter. Der Mittelwert liegt bei 4,02, was der vierten Kategorie (40-49) entspricht (Standardabweichung: 1,182).

**Abbildung 12:** PIRLS-2011-Testleistungen der Schülerinnen und Schüler im internationalen Vergleich – Gesamtskala Lesen



- Nicht statistisch signifikant vom deutschen Mittelwert abweichende Staaten ( $p > .05$ ).  
 Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.  
 1= Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.  
 2= Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.  
 Teilnehmer mit sehr hohen Ausschlussquoten ( $> 20\%$ ) sind mit einem \* gekennzeichnet.  
 3= Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.  
 4= Sehr hoher Anteil an Schülerinnen und Schülern mit nicht skalierbaren Leistungswerten.

Quelle: Bos, Bremerich-Vos, Tarelli & Valtin, 2012, S. 97, gekürzt

Im Vergleich zu den früheren Untersuchungen aus 2001 und 2006 musste eine rückläufige Entwicklung festgestellt werden: So konnte das von 2001 auf 2006 verbesserte Leseverständnis der Viertklässlerinnen und Viertklässlern nicht gehalten werden; die Leistungen der aktuellsten Erhebungswelle entsprechen in etwa denen aus 2001 (Tarelli & Valtin et al., 2012). Auch besteht ein Unterschied im Leseverständnis zwischen Jungen und Mädchen. Letztere erreichen durchschnittlich 545 Punkte und damit 8 Punkte mehr als ihre Mitschüler. Im Vergleich zu 2001 (13 Punkte) hat sich dieser Unterschied jedoch signifikant verringert.

Statt des Migrationshintergrunds (Geburtsland des Kindes, der Eltern und Großeltern oder ähnlich) wird in IGLU der Sprachgebrauch als Indikator herangezogen, da gezeigt werden konnte, dass dieser einen höheren Einfluss auf die Testleistungen hat als der Migrationshintergrund nach Geburtsland (vgl. Schwippert, Wendt & Tarelli, 2012). Im internationalen Vergleich zeigt sich kein markantes Ergebnis für Deutschland: Gemessen über das Item „Gebrauch der Testsprache in der Familie“ mit den Ausprägungen „nie“, „manchmal“ und „immer oder fast immer“ ergibt sich ein äußerst geringer Prozentsatz derer, die zu Hause nie Deutsch sprechen (0,8%), 18,9% der befragten Kinder sprechen zu Hause manchmal, 80,4% immer oder fast immer Deutsch. Dabei unterscheiden sich die Anteile der Kinder, die zu Hause nie oder manchmal die Testsprache sprechen (19,6%) nicht signifikant von Ländern wie Frankreich (22,4%), England (20,8%), Italien (20,6%), der Slowakei (20,4%), Schweden (19,9%), Norwegen (18,5%) und Dänemark (18,1%). Schwippert et al. (2012) kommen zu dem Schluss, dass „[h]insichtlich der allgemeinen Anteile an Schülerinnen und Schülern, die nur manchmal oder nie die Testsprache zu Hause sprechen, [...] somit für Deutschland von keiner besonders auffälligen Situation gesprochen werden [kann]. Vielmehr zeigen die Ergebnisse, dass Mehrsprachigkeit von Schülerinnen und Schülern mittlerweile in vielen europäischen Ländern für ein Viertel bis ein Sechstel aller Grundschulkindern zur Alltagsrealität gehört. In Deutschland lebt jedes fünfte Grundschulkind in mehrsprachigen Familien“ (Schwippert et al., 2012, S. 193).

Dieser allgemeinen Beschreibung der hier vorliegenden Daten wird im folgenden Kapitel 12.2 eine Erläuterung der genutzten Analysemethoden nebengestellt. Kapitel 12.3 beschäftigt sich dann mit der Prüfung der Einzelmodelle, bevor in Kapitel 12.4 die Einzelbeziehungen und in Kapitel 12.5 schließlich das Gesamtmodell analysiert werden.

## **12.2 Analysemethoden und Kennwerte**

Die in den vorherigen Abschnitten aus Theorie und Forschungsstand abgeleiteten Zusammenhänge werden im Folgenden einer empirischen Prüfung unterzogen. Berücksichtigt werden müssen dabei eine Reihe von Spezifika, die mit der (z.T. PIRLS-immanenten) Datenstruktur zusammenhängen.

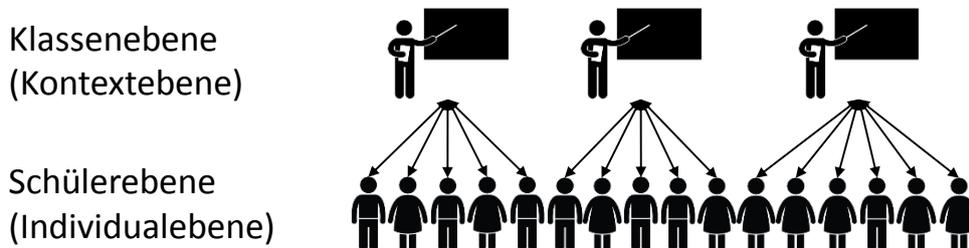
Eine dieser Besonderheiten ergibt sich aus der geclusterten, hierarchischen Schachtelung der Daten: So wurden bei der Stichprobengenerierung nicht *einzelne Schülerinnen und Schüler* sondern *ganze Klassen* aus der Grundgesamtheit (Viertklässlerinnen und Viertklässler an

staatlichen Grundschulen in Deutschland) gezogen. Dies hat vorrangig ökonomische Gründe, da auf diese Weise statt etwa 400 einzelner Schülerinnen und Schüler an 400 unterschiedlichen Schulen 205 Klassen (mit insgesamt 4.000 Kindern) an 197 Schulen befragt wurden – ein erheblich geringerer Aufwand – wenngleich dadurch insgesamt deutlich mehr Viertklässlerinnen und Viertklässler befragt werden mussten. Die Ursache hierfür liegt darin, dass sich Kinder und Jugendliche *innerhalb* von Klassen ähnlicher sind als *zwischen* Klassen und Schulen. Begründet werden kann dies damit, dass die Schülerinnen und Schüler einer Klasse als soziales System ähnliche Erfahrungen machen, die sich von anderen Klassen (auch der gleichen Schule) unterscheiden können: So weist jede Klasse unterschiedliche Schülermerkmale auf (unterschiedlich aktive Schülerinnen und Schüler, ‚Störenfriede‘, sprachliche Niveaus etc.), wird durch ein klassenindividuell zusammengesetztes Lehrerkollegium unterrichtet und macht gemeinsame Erfahrungen z.B. auf Wandertagen und Klassenfahrten, die ebenfalls stark zwischen den Klassen variieren. Zudem verfügen die Kinder und Jugendlichen desselben Wohneinzugsgebiets einer Schule mitunter über ähnliche sozioökonomische Hintergründe, die sich von denen der Schülerinnen und Schüler anderer Schulen deutlich unterscheiden können. Dieser Ähnlichkeit der Personen in Klassen muss bei der (möglichst ökonomischen) Stichprobenziehung Rechnung getragen werden, da davon ausgegangen werden muss, dass die zufällige Varianz in den Personenmerkmalen, wie wir sie finden würden, würden wir *einzelne* Kinder an *einzelnen* Schulen befragen, durch die Befragung *ganzer Klassen* aus oben genannten Gründen stark eingeschränkt wird. In Analysen, die auf solch einem fehlerhaften Datensatz aufgebaut würden, würden Standardfehler und Varianz entsprechend systematisch unterschätzt werden, was zu signifikanten Ergebnissen führen kann, da wo eigentlich die  $H_0$  angenommen werden müsste. Erst durch die Ziehung zusätzlicher Klassen kann dieser verminderten Varianz begegnet werden. Zudem muss für die Ähnlichkeit der Schülerinnen und Schüler innerhalb von Klassen in Analysen, die auf Daten mit solchen Clusterstrukturen basieren, kontrolliert werden.

Eine weitere Besonderheit findet sich in den unterschiedlichen Ebenen, auf denen die folgenden Analysen durchgeführt werden: So werden zum einen Angaben der Schülerinnen und Schüler und ihrer Eltern genutzt (sozioökonomischer Status, familiäre Unterstützung, Selbstkonzept und Leseleistung), zum anderen die Angaben der Lehrpersonen zur individuellen Förderung sowie Merkmale der Klasse. Für die Zuordnung der Eltern-, Lehrer- und Klasseninformationen zu den Schülerdaten bedeutet dies Unterschiedliches: Während ein El-

ternfragebogen immer einem einzelnen Schüler bzw. einer Schülerin zugeordnet werden kann, gelten die Angaben der Lehrperson (etwa darüber, wie häufig im Unterricht individuell gefördert wird) immer für *ganze Klassen* – müssen also immer mehreren Schülerinnen und Schülern zugeordnet werden. Diese beiden Fälle sind auch in den Analysen zu unterscheiden: Die Angaben der Schülerinnen und Schüler sowie ihrer Eltern findet sich auf der Individualebene (auch L1 oder within-Ebene); Angaben, die für ganze Klassen gelten (zum Beispiel – aber nicht ausschließlich – die Angaben der Lehrpersonen) werden der übergeordneten Kontextebene (auch L2 oder between-Ebene) zugeordnet. Abbildung 13 verdeutlicht diese hierarchische Schachtelung, die – theoretisch – beliebig erweiterbar ist<sup>58</sup>.

**Abbildung 13:** Hierarchische Schachtelung von Schülerdaten auf Individualebene in Klassen auf der Kontextebene



Quelle: eigene Darstellung, Piktogramme: Khoon Lay Gan/123rf.de

Generell können Analysen dabei auf beiden Ebenen durchgeführt werden, wobei mögliche Zusammenhänge auf den unterschiedlichen Levels mathematisch unabhängig sein können. Solche Mehrebenenanalysen (MEA) bieten gegenüber herkömmlichen Regressionsanalysen zudem den Vorteil, dass Interaktionen *zwischen* den Ebenen (sogenannte Cross-Level-Interaktionseffekte) untersucht werden können. Der oben angesprochenen Klumpung der Stichprobe wird, wie dies auch in Abbildung 13 erkennbar ist, in Mehrebenenanalysen ebenfalls Rechnung getragen. Für diese Zuordnung von Personen zu Gruppen (hier also Schülerinnen und Schüler zu Klassen) wird eine Cluster-Variable benötigt, die Auskunft über die Gruppenzugehörigkeit gibt (z.B. eine Klassen-ID). Auf dies wird später nochmals eingegangen.

Bezüglich der Interpretation von Korrelations- und Regressionsparametern in Mehrebenenmodellen ist es hilfreich, einen inhaltlich plausiblen Nullpunkt der unabhängigen Variablen

<sup>58</sup> Eine Erweiterung nach oben macht zum Beispiel da Sinn, wo unterschiedliche Schulen oder Länder (auch Bundesländer, regionale Strukturen etc.) auf L3 untersucht werden sollen. Das Modell ist auch nach unten erweiterbar: Möglich ist etwa eine Schachtelung mehrerer Messzeitpunkte in (übergeordneten) Personen, wie sie Längsschnittanalysen erfordern. Ein Beispiel findet sich in Abbildung 24 im Anhang 1.

festzulegen („Zentrierung“). Dabei werden zwei<sup>59</sup> Verfahren unterschieden: Bei der grand-mean-Zentrierung wird der Gesamtmittelwert (der ganzen Stichprobe) einer unabhängigen Variable von den individuellen Beobachtungswerten abgezogen. Dabei erhält eine Beobachtungseinheit (z.B. eine Schülerin<sup>60</sup>) den Wert 0, wenn ihr Wert genau dem Mittel der gesamten Stichprobe entspricht; höhere Werte indizieren stichprobenbezogen überdurchschnittliche, geringere Werte stichprobenbezogen unterdurchschnittliche Werte. Die groupmean-Zentrierung folgt derselben Logik, allerdings werden hier die Durchschnittswerte einer unabhängigen Variablen *eines Clusters* (z.B. einer Klasse) von den individuellen Beobachtungswerten (z.B. einer Schülerin) abgezogen. Der Wert 0 entspricht dann dem Mittelwert des jeweiligen Clusters. Während die grandmean-Zentrierung auf beiden Ebenen durchgeführt werden kann, kann die groupmean-Zentrierung nur auf Individualebene vorgenommen werden, da auch nur dort Beziehungen *innerhalb von Clustern* analysiert werden können. Je nach gewählter Zentrierung kann eine Regressionskonstante interpretiert werden als Erwartungswert der abhängigen Variablen, wenn die unabhängige Variable dem Stichprobenmittelwert (grandmean-Zentrierung) oder dem kontext- bzw. klassenspezifischen Mittelwert (groupmean-Zentrierung) entspricht. Werden Einflüsse von Variablen auf Individual- und Kontextebene betrachtet, können prinzipiell beide Zentrierungsformen, auch in Kombination, herangezogen werden (wobei die groupmean-Zentrierung, wie dargestellt, nur auf Individualebene Anwendung finden kann). Enders und Tofighi (2007) empfehlen aber gerade bei der Analyse von Cross-Level-Interaktionseffekten (Beeinflussung eines L1-Zusammenhangs durch eine oder mehrere L2-Variablen), wie sie hier Gegenstand des Vorhabens sind, die Zentrierung am groupmean auf L1 und am grandmean auf L2, da hierbei Mittelwerte der unabhängigen Variablen um Kontextunterschiede adjustiert sind, so dass der L2-Prädiktor nicht durch Interaktionen mit diesen konfundiert ist. Dieser Empfehlung soll hier gefolgt werden; alle unabhängigen Variablen auf Individualebene werden also derart zentriert, dass ihr Nullpunkt dem jeweiligen Klassenmittel entspricht; auf Kontextebene entspricht der Nullpunkt der unabhängigen Variablen dem Stichprobenmittel.

---

<sup>59</sup> Ein drittes Verfahren benötigt keine der oben dargestellten nachträglichen Zentrierungen – etwa dann, wenn Variablen über einen natürlichen, direkt interpretierbaren Nullpunkt verfügen (z.B. Einkommen oder Angaben in Jahren) oder aufgrund von Rekodierungen bereits vor den Analysen einen interpretierbaren Nullpunkt zugewiesen bekommen haben (z.B. Geschlecht: 0 = Mädchen, 1 = Junge oder die Bücheranzahl, beginnend bei 0 = 0 bis 10 Bücher, 1 = 11 bis 20 Bücher etc.).

<sup>60</sup> Diese Darstellungen beziehen sich auf die den hier vorgestellten Analysen zugrunde liegende Struktur, d.h. Zweiebenenmodelle, in denen Individuen in Klassen geclustert sind.

Für die folgenden Mehrebenenanalysen wird das Programm Mplus (Muthén & Muthén, 2012) genutzt. Dies bietet vor allem den Vorteil, Mehrebenenanalysen mit dem Strukturgleichungsansatz zu kombinieren: So können Konstrukte wie das Selbstkonzept, der sozioökonomische Status und die familiäre Herkunft latent modelliert und simultan getestet werden. Mplus erlaubt die Analyse solcher Mehrebenenstrukturgleichungsmodelle (MSEM<sup>61</sup>) auf bis zu drei Ebenen; für die folgenden Analysen werden, wie in Abbildung 13 erkennbar, lediglich zwei Ebenen unterschieden.

Eine Voraussetzung für Mehrebenenanalysen ist die eindeutige Verknüpfung von Daten auf L1 und L2 (vgl. Abbildung 13). Demnach ist es nicht möglich, dass ein Schüler zwei Klassen zugeordnet bzw. von zwei Lehrpersonen unterrichtet wird. In dem für die Analyse vorliegenden IGLU-E-2011-Datensatz finden sich jedoch 215 Schülerinnen und Schüler in 15 Klassen, die im Leseunterricht von jeweils mindestens zwei Lehrpersonen im sogenannten Teamteaching unterrichtet werden. Dies wird im Rahmen von PIRLS wie folgt beschrieben: „Students who lag behind in reading first receive special attention from the classroom teacher. Many schools also have a teacher or a team of teachers engaged in special education who can assist the classroom teacher when necessary“ (Mullis, Martin, Minnich, Drucker & Ragan, 2012, S. 483). In der deutschen Stichprobe werden 5,4% der Kinder von mehr als einer Lehrperson im Lesen unterrichtet, davon die meisten (94%, 14 Klassen) von zwei und eine Klasse von drei Lehrpersonen. Im Datensatz spiegelt sich dies wie folgt wider: Die Fälle der Schülerinnen und Schüler, die von zwei oder drei Lehrenden im Lesen unterrichtet werden, werden verdoppelt (bzw. verdreifacht); ihnen werden jeweils die Daten der unterschiedlichen Lehrer ‚angespielt‘. Eine GewichtungsvARIABLE sorgt später dafür, dass die verdoppelten (verdreifachten) Schülerangaben nicht doppelt (dreifach) sondern lediglich mit dem Ursprungsgewicht in die Analysen einfließen. Die Nutzung dieses entsprechend adjustierten Gewichts ist für die Analyse auf Individualebene unproblematisch; erst die Verknüpfung verschiedener Ebenen verletzt die Voraussetzung für Mehrebenenanalysen, nach welcher alle Fälle auf L1 eindeutig den Fällen auf L2 (und umgekehrt) zugeordnet werden müssen – also Lehrerangaben beispielsweise nicht innerhalb von Klassen variieren dürfen. Für die Behebung dieses Umstands finden sich in entsprechenden Untersuchungen mit PIRLS- und TIMSS-Daten unterschiedliche Herangehensweisen: Erstens könnten diese 215

---

<sup>61</sup> Für Mehrebenenstrukturgleichungsmodelle hat sich die Abkürzung der englischen Bezeichnung (MSEM: Multilevel-Structural-Equation-Model) auch im Deutschen etabliert.

Fälle aus den Analysen ausgeschlossen werden. Eine Prüfung dieser Option ergab jedoch, dass sich die Teamteaching-Stichprobe in einigen der zu verwendenden Konstrukte signifikant von der Gesamtstichprobe unterscheidet, so dass auf diese Option verzichtet wurde. Auch der Ansatz, die Lehrperson zu ermitteln, die mehr Unterrichtszeit in einer Klasse verbringt und diese dann für die Analysen zu nutzen, wurde verworfen: Zwar ist solch eine Analyse prinzipiell möglich, allerdings liegt ihr die Annahme zugrunde, dass Lehrzeit mit Einfluss gleichgesetzt werden kann. Lehrpersonen, die unter Umständen weniger Zeit in der Klasse verbringen, dafür jedoch einen höheren Einfluss auf die Leseleistung haben (z.B. weil sie expliziten Leseunterricht geben, während die andere Lehrkraft eher den Schriftspracherwerb fokussiert), würden hier aus der Analyse herausgenommen. Drittens könnten die für einige Klassen mehrfach vorliegenden Lehrerangaben gemittelt werden. Da dies zu einer Verminderung der Varianz in den Lehrerangaben führen würde, wurde diesem Ansatz ebenfalls nicht nachgegangen. Behoben werden kann dieser Umstand hingegen mit der Verwendung einer anderen Cluster-Variablen, welche Informationen über die Gruppenzugehörigkeit beinhaltet. Gängig ist die Verwendung der IDCLASS, einer Variablen, die für alle Schülerinnen und Schüler einer Klasse denselben Wert annimmt (vgl. Tabelle 7). Diese Variable nimmt aber auch dann identische Werte an, wenn eine Klasse von zwei oder mehr Lehrpersonen unterrichtet wird. Sie kann also nicht für die eindeutige Verknüpfung von L1- und L2-Informationen herangezogen werden. Anders die Gruppierungsvariable IDTEALIN: Diese wird aus der IDTEACH, über welche die Lehrpersonen im Gesamtdatensatz identifizierbar sind, und der IDLINK, einer klasseninternen Nummerierung der Lehrpersonen zusammengesetzt. Tabelle 7 gibt einen exemplarischen Überblick über die unterschiedlichen Kodierungen für Klassen, die von einer Lehrperson unterrichtet werden (Fall 1) und Klassen, die von zwei Lehrpersonen im Teamteaching unterrichtet werden (Fall 2).

**Tabelle 7:** Beispielhafte Kodierung der Klassenzugehörigkeit für zwei Klassen mit jeweils drei Schülern, ohne und mit Teamteaching

Situation	IDCLASS	IDSTUD	IDTEACH	IDLINK	IDTEALIN
<b>Fall 1: Ein Lehrer unterrichtet alleine eine Klasse im Lesen</b>					
Schüler 1 in Klasse 1 mit Lehrer 1	503	50301	501	1	50101
Schüler 2 in Klasse 1 mit Lehrer 1	503	50302	501	1	50101
Schüler 3 in Klasse 1 mit Lehrer 1	503	50303	501	1	50101
<b>Fall 2: Zwei Lehrer unterrichten im Teamteaching eine Klasse im Lesen</b>					
Schüler 1 in Klasse 1 mit Lehrer 1	603	60301	601	1	60101
Schüler 1 in Klasse 1 mit Lehrer 2	603	60301	602	2	60202
Schüler 2 in Klasse 1 mit Lehrer 1	603	60302	601	1	60101
Schüler 2 in Klasse 1 mit Lehrer 2	603	60302	602	2	60202
Schüler 3 in Klasse 1 mit Lehrer 1	603	60303	601	1	60101
Schüler 3 in Klasse 1 mit Lehrer 2	603	60303	602	2	60202

Deutlich wird, dass – obwohl in beiden Klassen exemplarisch drei Schüler zu finden sind, im zweiten Fall (Teamteaching) insgesamt sechs Fälle identifizierbar sind. Dies liegt, wie oben erläutert, an der Dopplung der Schülerfälle, denen die unterschiedlichen Lehrerinformationen angefügt werden. Erkennbar ist zudem, dass die IDCLASS für jeden Schüler innerhalb der beiden Klassen den gleichen Wert annimmt. Erst die aus IDTEACH und IDLINK wie oben beschrieben zusammengesetzte IDTEALIN-Variable enthält für *jeden Lehrer einer Klasse* eine eindeutige Kennung. Diese soll für die folgenden Analysen als Gruppierungsvariable herangezogen werden. Dies bedeutet zunächst, dass für den Fall 2 (Teamteaching) zwei statt einer ‚Klasse‘ in die Berechnungen einfließen (differenziert über die IDTEALIN: 60101 und 60102). Über die zu verwendende Gewichtungvariable TCHWGT (Teacherweight; Summe der Gewichte der unterrichteten Schülerinnen und Schüler) werden diese doppelten Klassen (und mit ihnen die doppelten Schülerangaben) heruntergewichtet, so dass sie nicht mehrfach in die Analysen eingehen. Auf diese Weise bleiben die Angaben der Teamteaching-Lehrenden ohne Varianzverlust erhalten.

Eine weitere Überlegung vor den Analysen gilt dem Umgang mit fehlenden Werten. Nach Lüdtke, Robitzsch, Trautwein und Köller (2007), die sich in ihrer Systematisierung auf Rubin (1976) beziehen, können drei Typen fehlender Werte unterschieden werden. Bei den vollständig zufällig fehlenden Werten (Missing Completely At Random, MCAR) hängt der Ausfallprozess nicht systematisch mit der Ausprägung auf den fehlenden Werten zusammen. Zufällig fehlende Beobachtungen (Missing At Random, MAR) hängen mit der Ausprägung einer anderen Variablen zusammen, werden jedoch nicht von der Ausprägung dieser Variab-

len beeinflusst (Lüdtke et al., 2007). Ist letzteres hingegen der Fall, wird von nicht zufällig fehlenden Werten (Missing Not At Random, MNAR) gesprochen. Daraus leiten sich nach Lüdtke et al. (2007) unterschiedliche Verfahren für den Umgang mit fehlenden Werten ab: Bei den klassischen Verfahren werden Fälle, die fehlende Werte aufweisen, ausgeschlossen. Bei dem imputationsbasierten Vorgehen werden die fehlenden Werte vor den tatsächlichen Analysen ersetzt während bei den modellbasierten Verfahren Modellschätzungen und die Behandlung der fehlenden Werte gleichzeitig durchgeführt werden. Für die folgenden Analysen soll auf das modellbasierte Verfahren der Full Information Maximum Likelihood-Methode (FIML) zurückgegriffen werden. Hierbei wird das zu analysierende Modell für alle Fälle, also auch solche mit fehlenden Werten spezifiziert. Dabei wird auf Basis der individuellen (vorhandenen) Werte *für jeden Fall* ein wahrscheinlicher Modellparameter (Likelihood) geschätzt, wobei die fehlenden Werte – unter MAR-Annahme – explizit nicht berücksichtigt („ignoriert“) werden. Die Schätzung der *Gesamtmodellparameter* erfolgt anschließend anhand des Produkts der einzelnen Likelihoods. Bei der FIML-Methode werden also keine fehlenden Werte imputiert; die Modellparameter werden anhand der beobachteten Daten geschätzt.

Eine Besonderheit in IEA-Studien, zu denen auch PIRLS gehört, wurde oben bereits angesprochen: die Verwendung von GewichtungsvARIABLEN, die aufgrund der Besonderheiten der Stichprobe (disproportionale Stichprobenziehung bezüglich Untergruppen, Umgang mit fehlenden Werten etc.) angewendet werden müssen. Durch das Boston College (vgl. Martin, Mullis & Kennedy, 2007) werden mehrere Gewichte zur Verfügung gestellt, deren Auswahl sich an den verwendeten Daten und dem Ziel der Auswertung orientiert. Für die Nutzung von Lehrerdaten ist dies das oben bereits erwähnte Teacherweight (TCHWGT, overall teacher weight). Dabei entspricht die gewichtete Stichprobengröße der Schülerzahl in der Population (Viertklässlerinnen und Viertklässler in Deutschland), wobei die aufgrund des Teamteachings mehrfach vorkommenden Schülerangaben entsprechend adjustiert werden. Dieses Gewicht wird für alle der folgenden Analysen herangezogen (und ist der Grund dafür, warum die Stichprobengröße in deskriptiven Analysen deutlich größer ist als die 4.000 Schülerdaten, die sich tatsächlich im Datensatz finden).

Für diese benötigt es der Erläuterung einiger Kennwerte, welche für die Beurteilung der Modellgüte herangezogen werden sollen. Dabei handelt es sich zunächst um den  $\chi^2$ -Kennwert

(Chi Square Test of Goodness of Fit). Dieser stellt ein gebräuchliches Maß zur Überprüfung der Gesamtmodellpassung dar und „assesses the magnitude of discrepancy between the sample and fitted covariance matrices“ (Hu & Bentler, 1999, S. 2). Durch einen signifikanten  $\chi^2$ -Wert kann die Nullhypothese, nach welcher das Modell in der Population exakt passt, abgelehnt werden. Zu beachten ist hierbei, dass der  $\chi^2$ -Wert bei großen Stichproben schnell signifikant wird. Bei der Beurteilung der Modellgüte sollen daher weitere Indizes herangezogen werden, etwa der Root Mean Square Error of Approximation (RMSEA). Mit diesem kann überprüft werden, wie gut das Modell bei unbekanntem aber optimal gewählten Parameterschätzungen auf die Kovarianzmatrix der Population passt. Gewählt wird dabei das Modell, welches die geringste Parameteranzahl aufweist (Hooper, Coughlan & Mullen, 2008). Nach Hu und Bentler (1999) sollte der RMSEA-Wert ( $0,00 \leq \text{RMSEA} \leq 1,00$ ) unter 0,06, bei Geiser (2011) unter 0,05 liegen. Für die folgende Analyse soll daher von einem guten Modellfit bei einem Wert  $\leq 0,05$  und von einem akzeptablen Fit bei einem Wert  $\leq 0,06$  ausgegangen werden.

Für die Berechnung eines weiteren Indizes, dem Comparative Fit Index (CFI), wird zunächst angenommen, dass alle latenten Variablen eines Modells nicht miteinander korrelieren. Der Fit dieses Unabhängigkeitsmodells wird dann mit dem Fit des Zielmodells verglichen. Dabei passt das Zielmodell im Vergleich zum Unabhängigkeitsmodell umso besser auf die Daten, je höher der CFI-Wert ( $0,00 \leq \text{CFI} \leq 1,00$ ) ist. Bei CFI-Werten  $\geq 0,90$  wird von einem akzeptablen (Backhaus, 2006; Hooper et al., 2008), bei CFI-Werten  $\geq 0,95$  von einem guten (Hu & Bentler, 1999; Hooper et al., 2008) Modellfit ausgegangen.

Diese Cut-off-Werte gelten ebenfalls für den Tucker Lewis Index (TLI), der von der gleichen Vorannahme wie der CFI ausgeht und ebenfalls Auskunft darüber gibt, wie gut ein theoretisches Modell mit den empirischen Daten vereinbart werden kann (vgl. Hu & Bentler, 1999).

Darüber hinaus hat sich für die Bewertung eines Modells auf unterschiedlichen Ebenen der SRMR (Standardised Root Mean Square Residual) etabliert, der die (standardisierte) Differenz zwischen den beobachteten und den im Modell geschätzten Kovarianzen misst. Er wird getrennt für die unterschiedlichen Ebenen angegeben und nimmt ebenfalls Werte zwischen 0,00 und 1,00 an. Nach Byrne (1998) sowie Diamantopoulos und Siguaw (2000) ist ein Modell mit einem SRMR  $< 0,05$  als gut zu bewerten, Hu und Bentler (1999) geben darüber hinaus einen SRMR  $< 0,08$  als akzeptablen Cut-off-Wert an (vgl. auch Hooper et al., 2008).

Für die Bewertung der Signifikanzen wurden a priori konservative Niveaus festgesetzt, da diese durch die Stichprobengröße beeinflusst werden (je größer, desto eher signifikant). Den hier folgenden Analysen liegt mit 4.000 Schülerinnen und Schülern eine umfangreiche Stichprobe zugrunde, so dass die Signifikanzniveaus auf  $\alpha \leq 0,001$  (hoch signifikant),  $0,001 < \alpha \leq 0,010$  (sehr signifikant) und  $0,010 < \alpha \leq 0,050$  (signifikant) festgelegt wurden (vgl. hierfür auch Bortz & Döring, 2009). Hingegen soll für die Skalenbildung ein liberaler  $\alpha$ -Wert  $\geq 0,650$  als akzeptabel bewertet werden, da es sich inhaltlich vorrangig um persönliche Einstellungen, also ‚weiche‘ Persönlichkeitsmerkmale handelt, die mitunter die Bewertung der Qualität einer Skala erschweren (vgl. etwa DeVellis, 2012).

Im folgenden Kapitel werden die jeweiligen für das Gesamtmodell heranzuziehenden Konstrukte zunächst einzeln betrachtet. Im anschließenden Kapitel 12.4 werden diese dann – den oben beschriebenen theoretischen Annahmen folgend – untereinander in Beziehung gesetzt.

### **12.3 Prüfung der Einzelmodelle**

Ziel der vorliegenden Analyse ist die Bewertung eines Strukturgleichungsmodells auf mehreren Ebenen. Solche Modelle setzen sich aus zwei Teilen zusammen: aus Messmodellen, die die einzelnen latenten Konstrukte (hier also bspw. ‚Selbstkonzept im Lesen‘) darstellen und aus einem Strukturmodell, in welchem die (theoretisch begründeten) Beziehungen zwischen den Messmodellen abgebildet werden. Dieses Kapitel widmet sich dem Aufbau der einzelnen Messmodelle, die für die Überprüfung der theoretischen Annahmen im Rahmen von Strukturanalysen herangezogen werden müssen. Dabei handelt es sich um einen Index der sozioökonomischen Stellung einer Familie (ESCS, vgl. Kapitel 12.3.1), um die häusliche Lese-sozialisation (Home Literacy Environment, HLE, vgl. Kapitel 12.3.2), um die Leseleistung und das Leseselbstkonzept als abhängige Variablen (vgl. Kapitel 12.3.3 und 12.3.4), um die sozialen Merkmale der schulischen Lernumwelt (Schüler-Schüler-Beziehung und Schüler-Lehrer-Beziehung, vgl. Kapitel 12.3.5 und 12.3.6), die Merkmale der Klassenzusammensetzung (vgl. Kapitel 12.3.7) und schließlich um den Prädiktor Individuelle Förderung (vgl. Kapitel 12.3.8). Sie alle sollen anschließend theoriegeleitet miteinander in Beziehung gesetzt werden. Dabei wird das Kernelement des Modells – der Herkunftseffekt – derlei konstruiert, dass es für spätere Arbeiten international anschlussfähig ist. Solch ein universelles Modell hat dann den Vorteil, dass es zum Beispiel für den Vergleich mit Bildungssystemen, in denen in anderem

Umfang individuell gefördert wird oder die einen anderen Umgang mit Heterogenität zeigen, direkt in Verbindung gebracht werden kann. Entsprechend werden für die Modellierung des ESCS, des HLE und des Leseselbstkonzepts nur international verfügbare Daten herangezogen.

### **12.3.1 ESCS: Die sozioökonomische Stellung der Familie**

In einem Großteil der empirischen Bildungsstudien werden eine Reihe unterschiedlicher Variablen für die Messung der familiären Herkunft von Schülerinnen und Schülern eingesetzt, die Aussagen z.B. über das Ausmaß sozialer Disparitäten ermöglichen. Ihre Konstruktion orientiert sich dabei an Bourdieus Kapitaltheorie (vgl. Kapitel 7) und fokussiert insbesondere Aspekte des ökonomischen wie des kulturellen Kapitals. Variablen zum *sozialen Kapital* der Familie sind in internationalen Studien selten und beschränken sich in Deutschland meist auf im Rahmen nationaler Erweiterungen zusätzlich eingesetzte Skalen. Eine Ursache hierfür liegt in der Schwierigkeit, das soziale Netzwerk einer Person bzw. einer Familie in Large-Scale-Untersuchungen möglichst ökonomisch zu erfassen und zu analysieren. Frühe Ansätze beschäftigen sich hierbei insbesondere mit egozentrierten Netzwerken wie dem *name generator/interpreter* von McCallister und Fischer (1978), mithilfe welchem erfasst wird, wer beispielsweise auf die Wohnung des Befragten aufpassen würde, wenn dieser verreist oder mit wem die befragte Person über persönliche Probleme sprechen würde (vgl. auch Bos, Stubbe & Buddeberg, 2010). Zu jeder Person, die der Befragte angibt, werden Zusatzinformationen (Geschlecht, Entfernung vom Wohnort des Befragten u.ä.) erfasst, was verdeutlicht, dass die Erhebung der Daten relativ umfangreich und aufwändig ist. Unklar bleibt darüber hinaus, wie die gewonnenen Informationen in Verbindung gebracht werden können mit der theoretischen Struktur des Sozialkapitals (vgl. Van Der Gaag & Snijders, 2005; Bos et al., 2010). Der deutlich ökonomischere Ansatz des *Positionsgenerators* von Lin und Dumin (1986) erfasst anhand einer Liste mit 20 vorgegebenen Berufen, welche dieser Berufe von Personen des Bekanntenkreises eines Befragten ausgeübt werden. Anhand ihres Prestiges werden dann Rückschlüsse auf die sozialen Ressourcen des Befragten gezogen. Van Der Gaag und Snijders (2005) kritisieren den unspezifischen Charakter dieses Verfahrens, welches mit den Ressourcen, die mit einer beruflichen Tätigkeit einhergehen, nur einen Ausschnitt des gesamten sozialen Kapitals erfasst. Snijders (1999) entwickelt ein Instrument, das beide Ansätze verknüpft: Der *Resource Generator* erfasst anhand der Frage „Kennen Sie jemanden, der ...“ die im Bekanntenkreis des Befragten verfügbaren sozialen Ressourcen und deckt dabei Aspekte wie Unterstützung bei rechtlichen Problemen, bei der Wohnungs- und Jobsuche, Hilfe in

finanziellen Notlagen, Reparaturen im Haushalt und Unterstützung der Kinder, falls diese Probleme in der Schule haben, ab. Teile dieses Instruments wurden auch in der nationalen Erweiterung von IGLU 2011 eingesetzt (Wendt, Bos, Tarelli, Vaskova & Walzebug, 2016). Da das folgend abgeleitete Mehrebenenstrukturgleichungsmodell den Anspruch erfüllen soll, für internationale Vergleiche genutzt werden zu können, wurde die soziale Facette der familiären Ressourcen für die Modellierung der sozioökonomischen Herkunft eines Kindes nicht herangezogen.

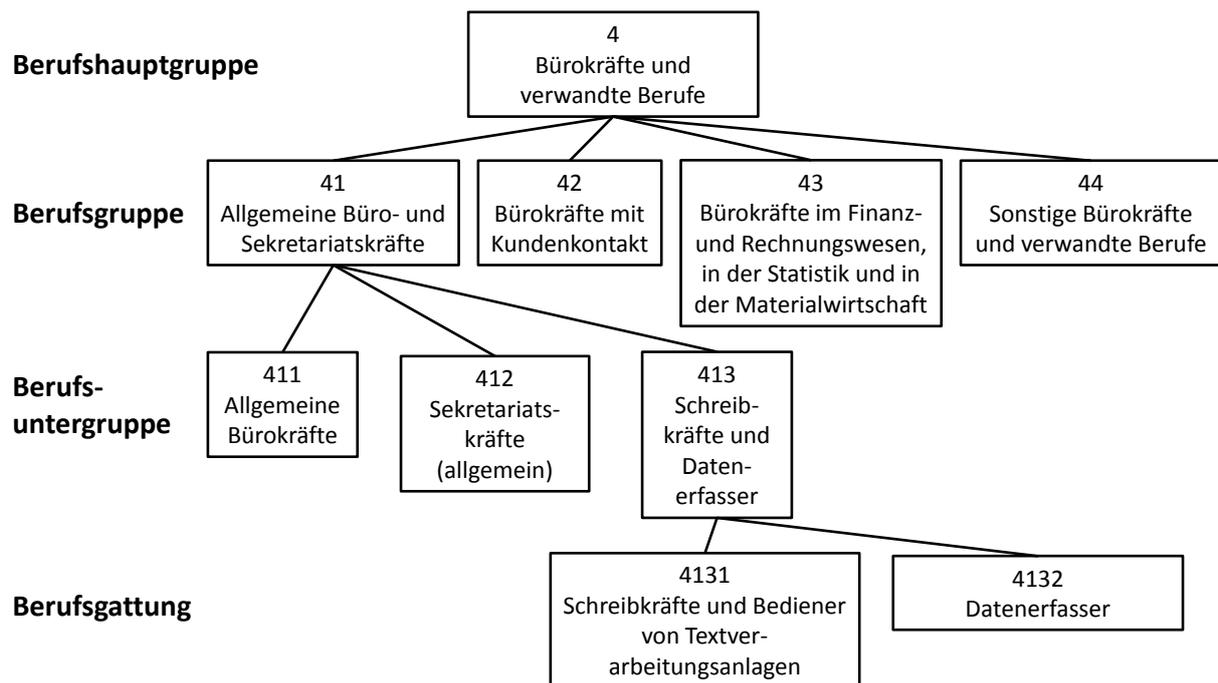
Für die Erfassung des ökonomischen und kulturellen Kapitals einer Familie steht hingegen eine Vielzahl von Indikatoren in internationalen Schulleistungsvergleichsstudien zur Verfügung. Diese werden im Folgenden kurz erläutert und auf ihre Verwendbarkeit für die nachstehenden Analysen geprüft.

Die *ökonomische Position* einer Familie kann beispielsweise durch die finanziellen Mittel, die einem Haushalt zur Verfügung stehen und dem relativen Wohlstand einer Familie im Vergleich zu anderen erfasst werden (Selbstauskunft). Darüber hinaus haben sich eine Reihe von Items etabliert, mit deren Hilfe die Ausstattung eines Haushalts erfragt und damit indirekt Auskunft über monetäre Ressourcen eingeholt werden kann. Zu diesen zählt etwa ein Computer, ein Handy, aber auch ein Rasenmäher als Indikator für ein eigenes Grundstück sowie ein zweites Auto. Und auch die berufliche Tätigkeit der Eltern gibt Auskunft über die ökonomische Situation (Einkommen) einer Familie – wobei hier die Trennung von ökonomischem und kulturellem Kapital nicht eindeutig ist: Da für bestimmte Berufe ein bestimmtes Bildungsniveau Voraussetzung ist, könnte dieser Indikator auch dem (institutionalisierten) kulturellen Kapital zugeordnet werden.

Die Erfassung – und insbesondere der Vergleich – von Berufen im nationalen wie internationalen Kontext birgt einige Hürden. So erschweren nicht nur unterschiedliche Bezeichnungen und die hohe Anzahl unterschiedlicher Berufe, sondern vor allem auch unterschiedliche ökonomische Verhältnisse zwischen Ländern einen systematischen Vergleich. Um dem entgegen zu wirken, wurde 1968 die internationale Standardklassifikation von Berufen (International Standard Classification of Occupations, ISCO-68) durch das Internationale Arbeitsamt veröffentlicht (International Labour Organization, 1969). Diese Klassifikation ordnet Berufe hierarchisch an, wobei auf der untersten Ebene die ausgeübten Tätigkeiten mit den wahrzunehmenden Aufgaben und Pflichten einer Person verzeichnet sind. Ähnliche Tätigkeiten werden

auf höheren Ebenen zu insgesamt 436 Berufsgattungen, 130 Berufsuntergruppen, 43 Berufsgruppen und schließlich zehn Berufshauptgruppen zusammengefasst, deren Kodierung numerisch ist (International Labour Office, 2012). Abbildung 14 gibt anhand der Berufshauptgruppe „Bürokräfte und verwandte Berufe“ einen Einblick in die hierarchische Struktur der ISCO. Eine Übersicht der Aufgaben der in diesem Beispiel aufgeführten Berufsgattung „Datenerfasser“ findet sich im Anhang (vgl. Anhang 2).

**Abbildung 14:** Beispiel der ISCO-08-Kodierung für Bürokräfte und verwandte Berufe, Auszüge

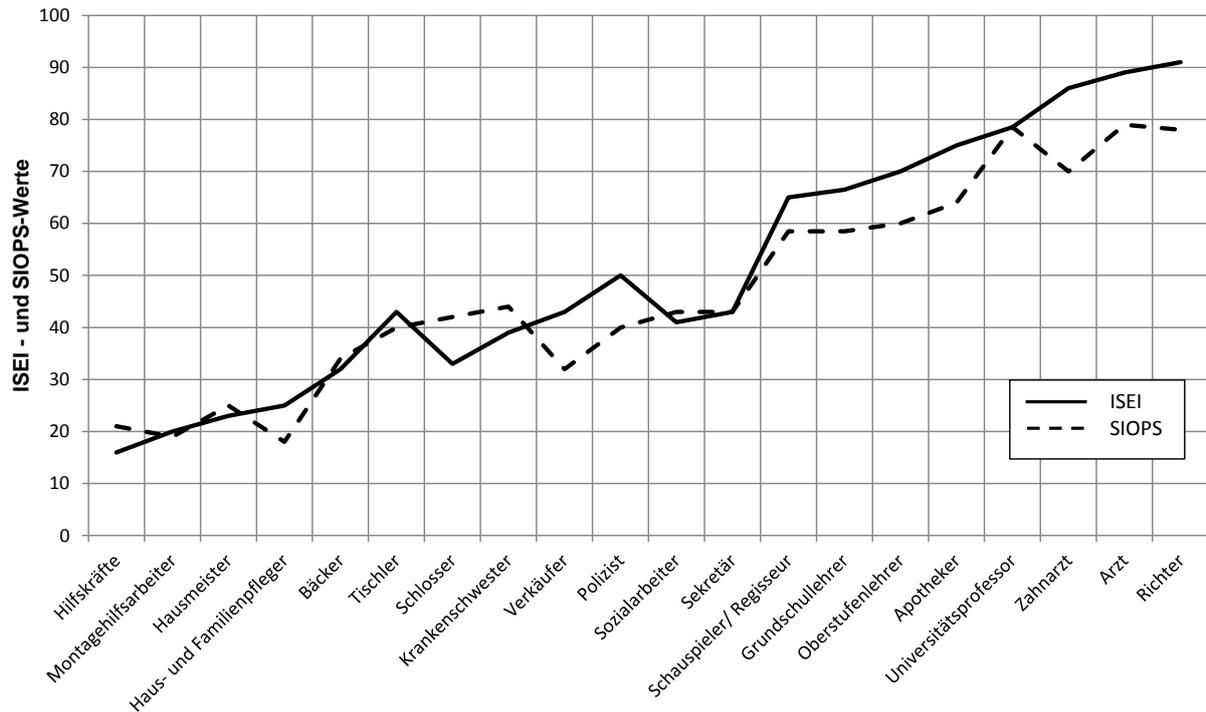


Quelle: Statistik Austria, o. J.a, eigene Abbildung

Die Klassifikation nach dem ISCO wird alle 20 Jahre an die Veränderungen der Arbeitswelt in den Industrienationen angepasst, die aktuellste Überarbeitung liegt aus dem Jahr 2008 (für den deutschsprachigen Raum vgl. Statistik Austria, o. J.a) vor. Die Erfassung der Berufe erfolgt zunächst in einem offenen Format und wird anschließend anhand der vierstelligen Kennziffern (vgl. Abbildung 14, bspw. „4132“ für Datenerfasser) kodiert. Aussagen über den Status eines Berufs erlauben diese Codes nicht; hierfür gibt es unterschiedliche Indizes, in welche die ISCO-Angaben überführt werden können. Zu den in der empirischen Bildungsforschung am häufigsten verwendeten Indizes gehören der International Socio-Economic Index of Occupational Status (ISEI), der Standard Index of Occupational Prestige Scale (SIOPS) sowie das Erikson-Goldthorpe-Portocarero-Modell (EGP-Klassifikation), die im Folgenden erläutert und für die Anwendung in einem international vergleichbaren Modell geprüft werden sollen.

Theoretische Grundlage für die Konstruktion des SIOPS bildet die Einschätzung der gesellschaftlichen Anerkennung unterschiedlicher Berufe, die in 85 Untersuchungen in etwa 60 Ländern erfasst wurde. Aus der Einstufung beruflicher Tätigkeiten hinsichtlich ihres sozialen Ansehens wird ein Berufsprestigemaß konstruiert, welches Werte zwischen 0 und 100 annehmen kann und, wie Treiman (1977) zeigen konnte, weitestgehend unabhängig vom untersuchten Land ist. Kritisiert werden die Indizes zum Berufsprestige insbesondere dafür, dass oftmals unklar bleibt, welche sozialstrukturell relevanten Merkmale in ihre Konstruktion eingehen (Baumert & Maaz, 2006).

Ein weiterer Index zur Messung des sozioökonomischen Status der beruflichen Tätigkeit stellt der International Socio-Economic Index of Occupational Status dar. Der ISEI, erstellt von Ganzeboom und Kollegen (1992), gibt an, „wie gut ein Beruf geeignet ist, um die Ausbildung einer Person in Einkommen umzuwandeln“ (Wendt et al., 2012, S. 177). Seine Entwicklung beruht auf Daten zu Beruf, Bildung und Einkommen von etwa 74.000 Beschäftigten in 16 Ländern und trennt die ökonomische Stellung vom Berufsprestige. Anhand des ISEIs ist es möglich, die aus den ISCO-Angaben entnommenen Berufe auf einer linearen Skala anzuordnen, die Werte zwischen 16 und 90 annehmen kann (Baumert & Maaz, 2006; Ganzeboom et al., 1992). Abbildung 15 verdeutlicht den Unterschied zwischen SIOPS und ISEI: Berufsprestige und sozioökonomischer Status verlaufen nicht parallel, kovariieren jedoch erheblich. Unterschiede zwischen beiden Indizes lassen sich vor allem bei Krankenschwestern bzw. -pflegern und Verkäuferinnen bzw. Verkäufern feststellen: Während die erstgenannte Tätigkeit mit höheren Prestigewerten einhergeht, erreichen Verkäuferinnen und Verkäufer höhere ISEI-Werte (Baumert & Maaz, 2006).

**Abbildung 15:** Sozioökonomischer Status (ISEI) und Berufsprestige (SIOPS) für ausgewählte Berufe

Quelle: Baumert & Maaz, 2006, S. 16, adaptiert

Im Gegensatz zu dem SIOPS hat sich der ISEI in internationalen Vergleichsuntersuchungen bewährt und wird daher häufig als Standardindikator für den Vergleich beruflicher Tätigkeiten über mehrere Länder hinweg genutzt (vgl. etwa Wendt et al., 2012; Ditton, Krüsken & Schauenberg, 2005; Autorengruppe Bildungsberichterstattung, 2010; Marks, Cresswell & Ainley, 2006; Ehmke & Siegle, 2005). Da insbesondere in Schulleistungsvergleichsstudien meist der ISEI von Mutter *und* Vater erfragt wird, werden beide Angaben oftmals zu einem Index zusammengefasst, dem Highest International Socio-Economic Index of Occupational Status (HISEI). Dieser gibt Auskunft über die *höchste* berufliche Stellung, die in einer Familie vorliegt, reduziert also die Informationsmenge auf einen einzigen, für den gesamten Haushalt geltenden Indikator.

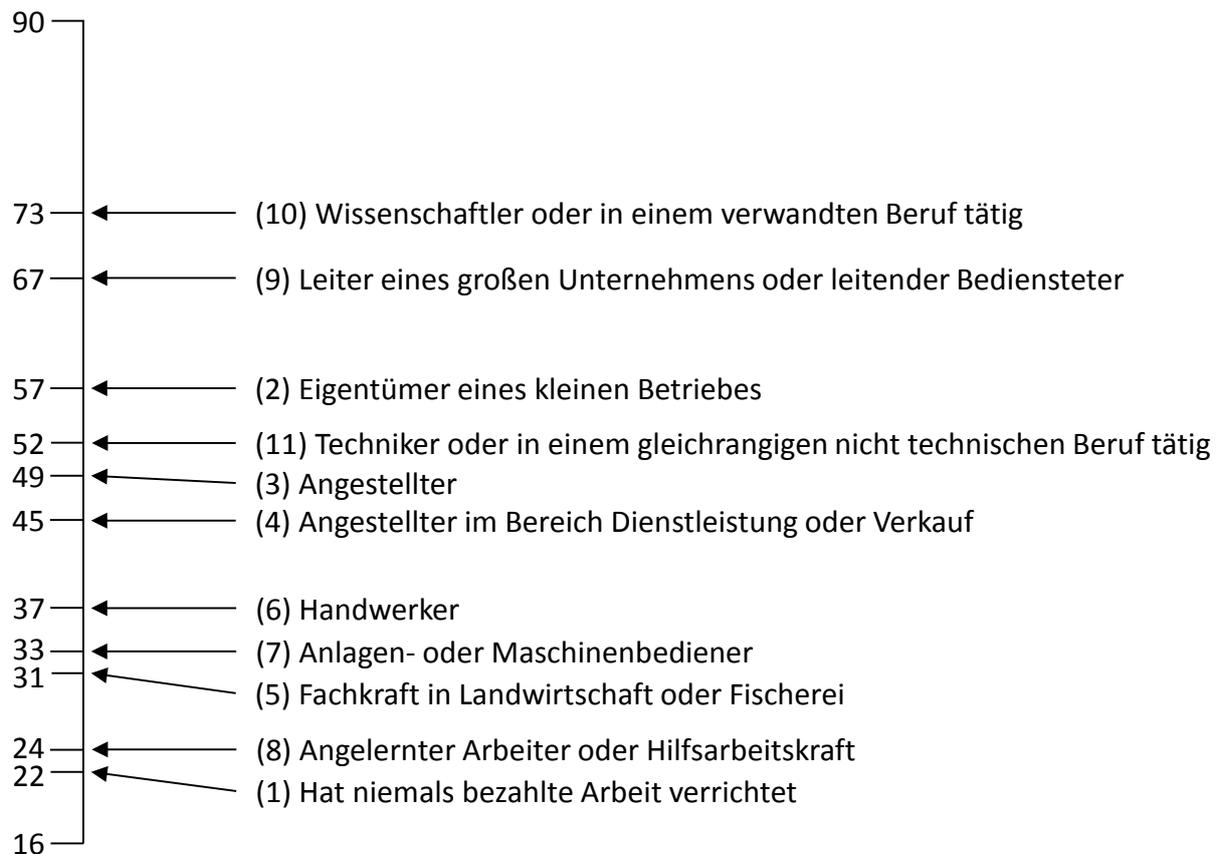
Einen dritten Ansatz zur Klassifikation unterschiedlicher sozioökonomischer Stellungen stellen die EGP-Klassen nach Erikson, Goldthorpe und Portocarero (1979) dar. Ihnen liegt der theoretische Ansatz zu Grunde, nach welchem die unterschiedlichen gesellschaftlichen Gruppierungen unterschiedliche Zugänge zu Einkommen, Macht, Bildung und gesellschaftlicher Anerkennung haben, die sich innerhalb der sozialen Schichten wenig, zwischen ihnen jedoch deutlich unterscheiden. Diese Differenzen zeichnen sich z.B. in der Verfügung über Kapital und Arbeitskraft sowie akademischer oder praktischer Bildung ab. Das EGP-Modell

verortet die unterschiedlichen Berufe in einem Kategoriensystem, das sich an der Art der Tätigkeit (manuell, nichtmanuell, landwirtschaftlich), der Stellung im Beruf (selbstständig, abhängig beschäftigt), den Weisungsbefugnissen (keine, geringe, große) und den für die Berufsausübung erforderlichen Qualifikationen (keine, niedrige, hohe) orientiert. Für den deutschsprachigen Raum lassen sich daraus sechs Klassen ableiten (für eine ausführliche Beschreibung s. Baumert & Schümer, 2001, S. 339): Die obere (I) und die untere (II) Dienstklasse, Routinedienstleistungen in Handel und Verwaltung (IIIa und IIIb), Selbstständige („Kleinbürgertum“) und selbstständige Landwirte (IVa-IVc), Facharbeiter und Arbeiter mit Leitungsfunktionen sowie Angestellte in manuellen Berufen (V-VI) und schließlich un- und angelernte Arbeiter sowie Landarbeiter (VIIa und VIIb).

Diese Klassifikation nach Erikson, Goldthorpe und Portocarero (1979) hat den Vorteil, dass sie im Unterschied zu ISEI und SIOPS anschauliche Beschreibungen konkreter Personengruppen erlaubt und theoretisch besser fundiert ist als erstgenannte. Zudem ist sie unabhängig von Bildungs- und Einkommensmaßen, bei deren Erhebung mit erheblichen Messfehlern zu rechnen ist (Baumert & Maaz, 2006). Allerdings müssen für die Einordnung von Berufen in das EGP-Modell neben den ISCO-Angaben die berufliche Stellung und Weisungsbefugnisse der Befragten mit erhoben werden, was insbesondere im internationalen Kontext häufig nicht der Fall ist – so auch bei IGLU. Für die nachstehenden Analysen wird daher der HISEI als Indikator für die ökonomische Stellung einer Familie herangezogen. Zu beachten ist hier, dass für den HISEI ebenfalls keine internationalen Daten vorliegen. Allerdings lassen sich Angaben zu den von Vater und Mutter ausgeübten Berufen (Item ASBH20A bzw. ASBH20B), die international verfügbar sind, in Anlehnung an Caro und Cortés (2012) in den HISEI überführen: Diese nutzen die in IGLU vorhandenen Berufskategorien und ordneten diese inhaltlich den auf einer kontinuierlichen Skala verorteten Berufskategorien von Ganzeboom et al. (1992) zu. Die folgende Abbildung 16 gibt einen Überblick über die Kodierung der ursprünglichen Variablen ASBH20A/B (Angaben in Klammern auf der rechten Seite), ihrer Bezeichnung und der Zuordnung auf der ISEI-Skala, die minimal 16 und maximal 90 Punkte annehmen kann. Dabei entsprechen hohe ISEI-Werte einer guten<sup>62</sup> beruflichen Stellung. Beispiele zu den Berufsklassifizierungen der IGLU-Variablen ASBH20A bzw. ASBH20B finden sich im Anhang 3.

---

<sup>62</sup> Diese Bewertung bezieht sich dabei auf die Eignung des Berufes, die Ausbildung einer Person in Einkommen umzuwandeln.

**Abbildung 16:** Anordnung der ursprünglichen Kodierung des von Mutter bzw. Vater ausgeübten Berufs auf der ISEI-Skala

Legende: Variablen ASBH20A bzw. ASBH20B im IGLU-Datensatz, Angaben in Klammern entsprechen der Originalkodierung

Quelle: eigene Abbildung, in Anlehnung an Caro und Cortés, 2012

Durch diese metrische Verortung der ursprünglich nominalen Skalierung der von Vater und Mutter ausgeübten Berufe auf der ISEI-Skala kann letztere als Indikator für die berufliche Stellung der Eltern in die Analysen aufgenommen werden. Anzumerken sei hier, dass die ursprüngliche Kodierung „(12) Nichts trifft zu“ als missing value kodiert wurde, da hier keine inhaltliche Einordnung auf der ISEI-Skala vorgenommen werden konnte. Dies entspricht 14,1% der Fälle, die zusätzlich zu den tatsächlichen missings (23,3%<sup>63</sup>) den fehlenden Werten zugeordnet wurden. Aus den beiden Angaben zu dem Beruf der Mutter und des Vaters wurde der Maximalwert des Elternpaares (*Highest International Socio-Economic Index of Occupational Status*) errechnet. Daraus ergeben sich für den HISEI folgende (gewichtete) Verteilungen (vgl. Tabelle 8):

<sup>63</sup> Dieser Wert der ‚tatsächlich‘ fehlenden Angaben erscheint hoch, ist aber für eine Skala, die nach der beruflichen Stellung der Eltern fragt, durchaus normal.

**Tabelle 8:** Häufigkeitsverteilung des HISEI in Deutschland

Ausprägung	Häufigkeit	Prozent
Hat niemals bezahlte Arbeit verrichtet	3.679	0,5
Angelernter Arbeiter oder Hilfsarbeitskraft	11.504	1,5
Fachkraft in Landwirtschaft oder Fischerei	4.746	0,6
Anlagen- oder Maschinenbediener	19.677	2,6
Handwerker	39.191	5,3
Angestellter im Bereich Dienstleistung oder Verkauf	89.461	12,0
Angestellter	130.355	17,5
Techniker od. in einem gleichrangigen nicht technischen Beruf tätig	44.485	6,0
Eigentümer eines kleinen Betriebes	71.779	9,7
Leiter eines großen Unternehmens oder leitender Bediensteter	41.365	5,6
Wissenschaftler oder in einem verwandten Beruf tätig	84.903	11,4
Fehlend	202.342	27,2
Gesamtsumme	743.487	100,0

Quelle: Eigene Berechnung, nach TCHWGT gewichtet, daraus ergeben sich die dem Populationsumfang entsprechenden Stichprobenzahlen

Deutlich wird hier, dass sich die Stichprobe vorrangig aus Eltern zusammensetzt, die im Angestelltenverhältnis arbeiten, zu der Gruppe der Lehrer, Juristen, Ärzte und Wissenschaftler (s. Erläuterungen der Kodierungen im Anhang 3) gehören oder leitende Funktionen im eigenen oder fremden Betrieb innehaben. Die Gruppe der angelernten bzw. Hilfsarbeiter, Maschinenbediener und Fachkräften in Landwirtschaft und Fischerei ist, ebenso wie die Personen, die bisher niemals bezahlte Arbeit verrichtet haben, für Deutschland erwartungsgetreu niedrig. Der HISEI soll also als Indikator für die ökonomische Stellung einer Familie herangezogen werden.

Wie eingangs jedoch erwähnt, ist diese Zuordnung der beruflichen Tätigkeit zu dem ökonomischen Kapital nicht eindeutig, da für bestimmte Berufe ein bestimmtes Bildungsniveau, also institutionalisiertes kulturelles Kapital von Nöten ist. Der Bildungsabschluss der erwachsenen Bezugspersonen eines Haushalts wird in IGLU und vielen weiteren Large-Scale-Untersuchungen jedoch auch gesondert erfasst. Konstruiert wird dieser Index anhand der Angaben zu Schul- und Berufsbildung der Eltern, welche auf Grundlage der *International Standard Classification of Education* (ISCED; UNESCO Institute for Statistics, 2012) kodiert werden. Tabelle 9 gibt einen Überblick über die unterschiedlichen ISCED-Stufen und zeigt einige Beispielabschlüsse für Deutschland auf. Angegeben ist zudem ihre Umrechnung in Schul- bzw. Bildungsjahre, die – anders als die reinen Berufsabschlüsse – mit anderen Ländern verglichen werden können (Ehmke & Siegle, 2005).

**Tabelle 9:** Internationale Klassifizierung von Berufsabschlüssen nach ISCED

ISCED-Stufe	Beschreibung	Beispielabschluss	Umrechnung in Schuljahre
ISCED 0	Elementarbereich		0
ISCED 1	Primarbereich	Abschluss an einer Grundschule, Sonder-/Förderschule	4
ISCED 2	Sekundarbereich I	Abschluss an einer Polytechnischen Oberschule, Hauptschul-/Volksschulabschluss, Realschulabschluss	10
ISCED 3B, 3C	Sekundarbereich II	Berufsgrundbildungsjahr, Berufsschule, Berufsfachschule, abgeschlossene Lehre, Abschluss an einer Handelsschule	11
ISCED 3A, 4	Postsekundärer Bereich	Fachhochschulreife, Hochschulreife, Abitur, Abschluss an einer Fachoberschule	12
ISCED 5B	Nicht-universitärer Tertiärbereich	Abschluss an einer Fachschule, Meister- oder Technikerschule, einer Schule des Gesundheitswesens, Abschluss an einer Berufsakademie, Fachakademie	15
ISCED 5A, 6	Universitärer Tertiärbereich	Fachhochschulabschluss, Diplom (FH), Hochschulabschluss, Promotion	17

Quelle: Eigene Darstellung, in Anlehnung an UNESCO Institute for Statistics, 2012

Tatsächlich erhoben wurde der ISCED in davon leicht abweichenden Stufen, wie dies in Tabelle 10 dargestellt ist. Dabei ist zu berücksichtigen, dass die tatsächlich im Fragebogen zu findende Skala für internationale Vergleiche nach Vorgaben der IEA umkodiert wurde. Dabei konnte Option 5 nicht auf Deutschland übertragen werden, so dass auf dieser Stufe für Deutschland entsprechend keine Daten zu finden sind. Tabelle 10 zeigt die neue Kodierung, ihre Zuordnung zu den ISCED-Stufen und die Häufigkeitsverteilungen für Deutschland.

**Tabelle 10:** IGLU-Kodierung und Häufigkeitsverteilung des ISCED für Deutschland

Code	Bezeichnung	ISCED-Stufe	Häufigkeit	Prozent
1	Ist nicht zur Schule gegangen	NO SCHOOL	1.923	0,3
2	Hat die Schule ohne Abschluss verlassen	ISCED 1 OR 2	6.414	0,9
3	Volksschulabschluss/Hauptschulabschluss/Ab- schluss der POS nach Klasse 8/Realschulab- schluss/Mittlere Reife/Fachschulreife/Ab- schluss der POS nach Klasse 10/Abschluss- zeugnis Berufsvorbereitungsjahr	ISCED 2	186.133	25,0
4	Abitur/Allgemeine Hochschulreife	ISCED 3	37.386	5,0
5	Option ist nicht erhoben oder Daten sind nicht verfügbar	ISCED 4	0	0,0
6	Fachhochschulreife/Berufsschulabschluss (Leh- re, Ausbildung)/Berufsfachschule/Abschluss an einer Fachschule, Meister- oder Techniker- schule oder an einer Schule des Gesundheits- wesens	ISCED 5B	153.657	20,7
7	Hochschulabschluss (Bachelor)/Abschluss an einer Berufsakademie/Diplom (FH)/Fachhoch- schulabschluss	ISCED 5A, 1ST DEGREE	49.076	6,6
8	Hochschulabschluss (Master, Magister, Dip- lom, Staatsexamen)/höher als Hochschulab- schluss, z.B. Promotion	BEYOND ISCED 5A, 1ST DEGREE	102.204	13,7
	Fehlend		206.694	27,8
	Gesamtsumme		743.487	100,0

Quelle: Eigene Berechnung, nach TCHWGT gewichtet, daraus ergeben sich die dem Populationsumfang entsprechenden Stichprobenzahlen

Es zeigt sich eine für Deutschland erwartungsgerechte Verteilung: Die meisten Eltern weisen einen Abschluss auf ISCED-Level 2 (v.a. Abschlüsse nach der 10. Klasse, z.B. Realschulabschluss) und ISCED 5B (v.a. Lehre, Ausbildung) auf, gefolgt von Eltern mit universitärem Abschluss (über ISCED 5A). Die Gruppe derer, die die Schule nicht besucht oder ohne Abschluss verlassen haben, ist mit insgesamt 1,2% ebenfalls erwartungsgerecht klein. Die ISCED-Angaben sollen entsprechend ebenfalls zur Modellierung des sozioökonomischen Hintergrunds der Schülerinnen und Schüler herangezogen werden.

Neben den Bildungsabschlüssen und der Klassifizierung der beruflichen Stellung hat sich die Frage nach der Anzahl der Bücher, die eine Familie besitzt, in nationalen wie internationalen Schulleistungsstudien als Proxy für das *kulturelle Kapital* der Eltern bewährt. Dabei „lässt sich annehmen, dass die Eltern eher komplexe Sachkenntnisse auf ihre Kinder übertragen kön-

nen, wenn sie über ausreichendes kulturelles Kapital verfügen. Eine große Anzahl an Büchern kann als ein Indikator für eine fördernde familiäre Situation gelten, in der die Bildung hoch geschätzt wird“ (Weißeno & Eck, 2013, S. 38). Nach Bourdieu (1983) können kulturelle Besitztümer wie Bücher zudem nur durch Personen adäquat genutzt werden, die über das notwendige inkorporierte Kulturkapital verfügen (vgl. auch Wendt et al., 2012). Angaben zu der Teilnahme an kulturellen Veranstaltungen ergänzen die Erfassung inkorporierten kulturellen Kapitals. Erfragt wird hier, wie häufig die Eltern im letzten Jahr mit ihrem Kind im Theater, im Museum/einer Kunstaussstellung, in der Oper/im Ballett/klassischen Konzert, auf einem Rock-/Pop-/Jazz-Konzert, einer Sportveranstaltung, im Kino, einem Stadtteil-/Volksfest/Jahrmarkt und bei einem Vortrag/einer Literaturvorlesung waren (Wendt et al., 2016). Diese Frage wird jedoch im internationalen Kontext in IGLU nicht eingesetzt, soll also für die Modellierung des (zunächst für einen internationalen Vergleich konzipierten) Mehrebenenstrukturgleichungsmodells nicht herangezogen werden. Im Gegensatz dazu findet sich die Angabe über die Anzahl der Bücher in den international verfügbaren Daten und zwar sowohl im Eltern- als auch im Schülerfragebogen. In der Regel wird davon ausgegangen, dass die Antworten der Eltern valider sind als die von Viertklässlerinnen und Viertklässlern, so dass das Item ASBH14 aus dem Elternfragebogen herangezogen werden soll<sup>64</sup>. Die folgende Tabelle 11 gibt einen Überblick über die Skalierung und die Häufigkeitsverteilung des Items. Der Fragestamm lautet „Wie viele Bücher gibt es in Ihrem Haushalt ungefähr (ohne Zeitschriften, Zeitungen oder Kinderbücher)?“ (Wendt et al., 2016).

---

<sup>64</sup> Zudem wurde geprüft, wie sehr sich die Eltern- und Schülerangaben unterscheiden. Dies ist lediglich in sehr geringem Maße der Fall, so dass die Elternangaben für die nachstehenden Analysen herangezogen werden können.

**Tabelle 11:** Häufigkeitsverteilung der Büchervariablen (Elternangaben) für Deutschland

Ausprägung	Häufigkeit	Prozent
0-10 Bücher	38.627	5,2
11-25 Bücher	67.941	9,1
26-100 Bücher	193.022	26,0
101-200 Bücher	107.412	14,4
Über 200 Bücher	181.768	24,4
Fehlend	154.717	20,9
Gesamtsumme	743.487	100,0

Quelle: Eigene Berechnung, nach TCHWGT gewichtet, daraus ergeben sich die dem Populationsumfang entsprechenden Stichprobenzahlen

Es zeigt sich, dass über ein Drittel der Eltern von Viertklässlerinnen und Viertklässlern in Deutschland über mehr als 100 Bücher verfügt (38,8%), es hingegen nur wenige Haushalte gibt, die weniger als 25 Bücher besitzen (14,3%). Hieraus kann – im Anbetracht der obigen Systematisierung der Kapitalsorten – davon ausgegangen werden, dass die Eltern der hier untersuchten Schülerinnen und Schüler über ein leicht überdurchschnittliches kulturelles Kapital verfügen.

Für die Messung des sozioökonomischen Hintergrunds in einem international anwendbaren Modell sollen also drei Indikatoren herangezogen werden: Der HISEI, der ISCED und die Angaben der Eltern zum Bücherbesitz. Dessen Modellierung orientiert sich dabei an den Überlegungen von Ehmke und Siegle (2005), die aus den Indikatoren der unterschiedlichen Kapitalsorten einen gemeinsamen Index, den sogenannten *Economic, Social and Cultural Status* (ESCS) bildeten und prüften, wie viel Varianz in der Schülerleistung sich durch diesen Index aufklären lässt – gegenüber der Modellierung auf Einzelitemebene. Herangezogen wurden hierfür Herkunftsvariablen zum sozioökonomischen Hintergrund, gemessen über die höchste berufliche Stellung von Mutter oder Vater (HISEI), dem höchsten elterlichen Bildungsabschluss, umgerechnet in Bildungsjahre (s. Ausführungen zum ISCED) sowie einen Index für häusliche Besitztümer. Letzterer setzt sich aus den klassisch-kulturellen (Bücher, klassische Literatur, Kunstwerke etc.), lernbezogenen (Taschenrechner, Schreibtisch, Wörterbuch etc.) und computerbezogenen (Lernsoftware, Internetanschluss etc.) Besitztümern einer Familie zusammen und enthält damit auch Angaben zum relativen Wohlstand eines Haushalts (Ehmke & Siegle, 2005). In der regressionsanalytischen Untersuchung verglichen Ehmke und Siegle (2005) diese verschiedenen Herkunftsindikatoren entsprechend dahingehend, wie viel Varianz der mathematischen Kompetenz in den unterschiedlichen Ländern der PISA-

Erhebung aus 2003 durch jeden einzelnen Indikator, alle drei Indikatoren gleichzeitig sowie durch den konstruierten ESCS aufgeklärt werden können. Tabelle 12 gibt hierfür einen Überblick für Deutschland und – als Vergleichsmaß – den OECD-Durchschnitt.

**Tabelle 12:** Varianzaufklärung in der mathematischen Kompetenz in PISA durch verschiedene Prädiktoren, differenziert für Deutschland und den OECD-Durchschnitt

Staaten	Modell 1 HISEI	Modell 2 PARED	Modell 3 HOMEPOS	Modell 4 HI + PA + HO	Modell 5 ESCS
Deutschland	15,0%	10,3%	7,7%	20,3%	22,6%
OECD	11,5%	12,3%	13,0%	20,3%	22,1%

Legende: PARED = Parental Education, höchster Bildungsabschluss von Vater oder Mutter, umgerechnet in Bildungsjahre; HOMEPOS = Home Possessions, häusliche Besitztümer

Quelle: nach Ehmke & Siegle, 2005, gekürzt

Deutlich wird, dass sowohl in Deutschland als auch im OECD-Durchschnitt die Varianzaufklärung aller drei Einzelindikatoren deutlich geringer ausfällt als bei dem ESCS. So liegt diese im Durchschnitt bei etwa 12 Prozent, während der ESCS etwa 22 Prozent zu der Aufklärung mathematischer Kompetenz beiträgt. Zudem liegt der ESCS in allen an PISA teilnehmenden OECD-Ländern bei bis zu 9 Prozent über dem jeweils besten Einzelindikator. Dieser Befund bestätigt sich auch in der Analyse aller deutschen Bundesländer anhand der PISA-E-Stichprobe (vgl. Ehmke & Siegle, 2005). Es zeigt sich also, dass der ESCS soziale Disparitäten klarer aufzudecken vermag als die drei Einzelindizes (HISEI, Schuljahre des höchsten Bildungsabschlusses von Vater oder Mutter, häusliche Besitztümer), wenngleich nur anhand dieser der Einfluss *einzelner* Bereiche der sozialen Herkunft deutlich wird (Ehmke & Siegle, 2005).

Für die folgenden Analysen soll sich unter Verwendung der international verfügbaren Daten an diesem Vorgehen orientiert werden: Aus HISEI, ISCED und den Bücherangaben der Eltern wird ein gemeinsamer Index gebildet (ESCS). Das Programm Mplus bietet dabei den Vorteil, solche Konstrukte latent modellieren zu können, so dass die drei Einzelitems nicht vorab in einer einzigen Variablen zusammengefasst werden müssen.

Hierfür wurden zunächst die drei zu nutzenden Items z-standardisiert, um sie in derselben Metrik einem möglichen latenten Konstrukt zuordnen zu können. Die konfirmatorische Faktorenanalyse in Mplus ergab einen latenten Faktor („ESCS“), der eine gute interne Konsistenz

von Cronbachs  $\alpha = 0,751$  aufweist<sup>65</sup>. Tabelle 13 gibt einen Überblick über die zugehörigen Einzelitems, ihre Faktorladungen und die Residualvarianz. Dabei handelt es sich um standardisierte Koeffizienten.

**Tabelle 13:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚ESCS‘ aus Mplus

Item	Bezeichnung	Faktorladung	Residualvarianz
zISCED	Schul- und Berufsbildung der Eltern	0,759	0,424
zBOOKS	Bücherbesitz (Elternangaben)	0,548	0,700
zHISEI	Berufliche Stellung der Eltern (höchste Angabe)	0,695	0,517

Es zeigt sich, dass alle drei Einzelitems ausreichend hohe Faktorladungen aufweisen. Die durch die Items in Bezug auf den Faktor ‚ESCS‘ nicht aufgeklärte Varianz liegt für den ISCED und den HISEI im mittleren Bereich (42,4% und 51,7%), für die Büchervariable mit 70% etwas höher. Dies erscheint bei der Unterschiedlichkeit der drei Items, die hier zu einem möglichst ökonomischen (d.h. sparsam modellierten) Index zusammengefasst werden, plausibel. Aussagen über die Passung des Modells können anhand der Fitwerte nicht getroffen werden: Bei der Modellierung eines Konstrukts mit drei Variablen ist das Modell genau passend spezifiziert (saturiertes Modell), so dass keine freien Parameter mehr geschätzt werden müssen. Dies resultiert in einer perfekten Fitstatistik von CFI = 1,000, TLI = 1,000, RMSEA = 0,000 und SRMR<sub>L1</sub> = 0,000. Da die Faktorladungen und die interne Konsistenz jedoch auf ein solides Modell schließen lassen, kann der ESCS als latentes Konstrukt für den sozioökonomischen Hintergrund einer Familie herangezogen werden. In vertiefenden Analysen konnte – analog zu den Ergebnissen von Ehmke und Siegle (2005) – zudem gezeigt werden, dass durch eine latente Modellierung des Indizes mehr Varianz der Schulerfolgsindikatoren Leseleistung und Selbstkonzept im Lesen aufgeklärt werden können als durch die drei Einzelitems (vgl. Tabelle 14):

<sup>65</sup> Für alle der folgend berichteten Messmodelle wird der Cronbachs-Alpha-Wert als Maß für die interne Konsistenz berichtet. Dieser stammt aus den explorativen Analysen aus SPSS, ihm liegt also ein etwas anderer Stichprobenumfang (Auslassen fehlender Werte in SPSS vs. FIML-Methode in Mplus) zugrunde.

**Tabelle 14:** Aufklärung der Varianz in der Leseleistung und in dem Selbstkonzept im Lesen durch sozioökonomische Merkmale in Abhängigkeit ihrer Modellierung, Stichprobengröße N

<b>Index</b>	<b>Aufgeklärte Varianz der Leseleistung</b>	<b>Aufgeklärte Varianz des Selbstkonzepts im Lesen</b>	<b>N</b>
ISCED (manifest)	8,9%	2,7%	3066
HISEI (manifest)	5,9%	1,7%	3090
Bücheranzahl (manifest)	7,3%	2,1%	3362
ISCED + HISEI + Bücher (manifest)	11,6%	3,3%	2815
ESCS (latent)	16,0%	4,8%	4227

Anmerkung: Die Stichprobengröße N liegt in einem Fall mit 4227 über den befragten 4000 Schülerinnen und Schülern, da die durch das Teamteaching vervielfachten Schülerdatensätze für die Analysen durch das Teacherweight zwar heruntergewichtet werden, in die Stichprobenstatistik jedoch ungewichtet einfließen.

Deutlich wird, dass alle drei Einzelindikatoren (ISCED, HISEI, Anzahl der Bücher) jeweils maximal 8,9% der Varianz in der Leseleistung und 2,7% der Varianz des Selbstkonzepts im Lesen aufzuklären vermögen. Selbst ihre gemeinsame Berücksichtigung in einem Modell mit 11,6% (Leseleistung) bzw. 3,3 % (Selbstkonzept im Lesen) Varianzaufklärung bleibt hinter der Varianzaufklärung durch den latenten ESCS (16,0% der Leseleistung bzw. 4,8% des Selbstkonzepts im Lesen). Die latente Modellierung scheint hier also einen empirischen Mehrwert zu haben, wobei sich dieser in Teilen durch den Umgang mit fehlenden Werten erklären lässt: Wie aus der in Tabelle 14 dargestellten Stichprobengröße, die in die Berechnung des jeweiligen Modells einfluss, erkenntlich, werden in den Einzelmodellen bis zu 1161 Schülerangaben (für das Modell mit allen drei Einzelindikatoren sogar 1412 Schülerangaben, dies entspricht etwa einem Drittel der Stichprobe) nicht berücksichtigt – erst die Modellierung eines komplexeren Modells (wie das latente) erlaubt einen angemessenen Umgang mit fehlenden Werten (FIML statt listenweisem Fallausschluss, wie in Kapitel 12.2 beschrieben). Gerade auch unter der für komplexe Modelle geltenden Sparsamkeitsprämisse soll das latente ESCS-Konstrukt für die Analysen herangezogen werden.

Ergänzt wird die Erfassung der sozialen Herkunft einer Familie oftmals durch Angaben zum Migrationshintergrund, da insbesondere hinsichtlich der Bildungsbeteiligung erhebliche Unterschiede zwischen Schülerinnen und Schülern aus Familien mit und ohne (jüngerer) Migrationsgeschichte bestehen. So liegt der relative Anteil an Jugendlichen, die das Gymnasium besuchen und deren Eltern beide im Ausland geboren wurden, bei 15%, während knapp 50% derselben Gruppe die Hauptschule besuchen (vgl. Stanat, 2006; Baumert & Schümer, 2002). Nach Diefenbach (2002) lassen sich bei der Erklärung dieser nachteiligen Position der Schü-

lerinnen und Schüler mit Migrationshintergrund im deutschen Bildungswesen vier Ansätze unterscheiden: *Kulturalistische Erklärungen* beziehen sich auf kulturelle Orientierungen und vergleichen Aspekte wie etwa die traditionelle Familienorientierung in der Zuwanderer- und Aufnahmegesellschaft, die einer erfolgreichen Bildungslaufbahn möglicherweise entgegenstehen könnten. In den Blick genommen werden hier beispielsweise auch die in der Familie gesprochene Sprache und die ethnische Zusammensetzung des Freundeskreises (vgl. auch Stanat, 2006; Alba, Handl & Müller, 1994). *Sozioökonomischen Erklärungen* zufolge ist die Benachteiligung von Schülerinnen und Schülern mit Migrationshintergrund ein Resultat geringer Ressourcen in den zugewanderten Familien, die nur bedingt eine Investition in die Bildungslaufbahn der Kinder erlauben. Eine weitere Erklärung findet sich in den *durch die Migrationssituation verfügbaren Handlungsstrategien*: Insbesondere die Aufenthaltsperspektive bestimmt demnach die Bildungsinvestitionen. Diese könnten gerade in Familien, die ihren Aufenthalt als nur vorübergehend betrachten, gering ausfallen, da die Kinder und Jugendlichen möglichst frühzeitig zum Familieneinkommen beitragen sollen statt einer Bildungslaufbahn nachzugehen, die im Herkunftsland möglicher Weise nicht anerkannt wird bzw. sich nicht auszahlen könnte. Und schließlich lassen sich *Effekte des Bildungssystems* als Erklärung heranziehen. Diese fokussieren insbesondere institutionelle Merkmale, die ungünstige Wirkungen bei Kindern und Jugendlichen mit Migrationshintergrund nach sich ziehen können. Hierunter fallen vor allem Mechanismen direkter und indirekter Diskriminierung, etwa bei Entscheidungen bzw. Empfehlungen an Übergangsstellen in der Schullaufbahn (vgl. z.B. Gomolla & Radtke, 2002; Stanat, 2006).

Aus diesen vier Erklärungsansätzen lassen sich unterschiedlich starke Prädiktoren für die Bildungsbeteiligung und Kompetenzentwicklung von Kindern und Jugendlichen mit Migrationshintergrund ableiten (vgl. Diefenbach, 2002 und andere, zusammenfassend Stanat, 2006). Im Kontext nationaler und internationaler Schulleistungsvergleichsstudien werden oft jedoch nur wenige Indikatoren erfasst, die die unterschiedlichen Erklärungsansätze herkunftsbezogener Disparitäten im deutschen Bildungssystem abbilden – so auch bei IGLU. Herangezogen werden könnten zwar die Angaben zu dem familiären Migrationsstatus (Kind, Eltern und Großeltern in Deutschland bzw. im Ausland geboren), allerdings ist dies nur in der nationalen Erweiterung der Studie verfügbar. Da es hier Ziel ist, ein international anschlussfähiges (Herkunfts-)Modell zu entwickeln, wurde dies nicht berücksichtigt. Dem hingegen findet sich der häusliche Sprachgebrauch, der als Migrationsindikator herangezogen werden

könnte, in den internationalen Daten wieder – jedoch konnten Schwippert et al. (2012) zeigen, dass unter Hinzunahme der sozioökonomischen Stellung der Familie (Buchbesitz, Bildungsniveau, Sozialschicht) und des Migrationshintergrunds (ein bzw. beide Elternteile im Ausland geboren) der häusliche Sprachgebrauch in keinem systematischen Zusammenhang mit dem Leseverständnis der Schülerinnen und Schüler steht. Da die international verfügbaren Daten dem Diskurs um den Einfluss des Migrationshintergrunds auf den Schulerfolg von Kindern und Jugendlichen nicht gerecht werden können, wurde auf die Einbindung dieses Index in das Mehrebenenstrukturgleichungsmodell verzichtet. Folglich wird die sozioökonomische Stellung der Familie latent durch den ESCS modelliert.

### **12.3.2 HLE: Die häusliche Lesesozialisation**

Die zentrale Rolle der Lesesozialisation im Elternhaus (Home Literacy Environment, kurz: HLE) für die Leistungsentwicklung des Kindes wurde in Kapitel 7 beschrieben. Folgend soll geprüft werden, ob sich dies – ähnlich wie der ESCS – in einem latenten Konstrukt abbildet. Herangezogen wird hierfür die international verfügbare Itembatterie „Lesen lernen: Vorbereitende Aktivitäten der Eltern vor Schuleintritt“ (Wendt et al., 2016), da gerade die frühe Förderung vor Beginn der Grundschule eine zentrale Rolle für eine gelungene Lesesozialisation spielt. Hierfür stehen 15 Items zur Verfügung, von denen sechs mathematikbezogene Aktivitäten beschreiben, die für die Modellierung der familiären Lesesozialisation nicht herangezogen werden sollen. Eine explorative Faktorenanalyse mit den übrigen Items ergibt zwei Faktoren, die die Intensität einer Lesesozialisation widerspiegeln: Während dem ersten Faktor Aktivitäten wie Bücher lesen, Geschichten erzählen und sich über das Gelesene unterhalten zugeordnet sind, finden sich speziellere, vertiefende schreib- und lesebezogene Aktivitäten wie Buchstaben und Wörter schreiben, Schilder und Beschriftungen laut lesen und mit Alphabetspielzeug spielen im zweiten Faktor wieder (vgl. Anhang 4.1). Dieser zweite Faktor soll nicht für die nachfolgenden Analysen herangezogen werden, da hier eine sehr spezielle Förderung der lesebezogenen Kompetenzen von Kindern vor der Grundschule im Fokus steht. Solch eine fokussierte Förderung könnte auch in einem negativen Zusammenhang mit der Lesemotivation, dem Leseselbstkonzept und damit auch der späteren Leseleistung des Kindes stehen, nämlich dann, wenn diese vorschulischen Aktivitäten als Zwang empfunden werden und dem Kind die Freude am Lesen nehmen. Herangezogen werden soll stattdessen der erste Faktor, da dieser allgemeinere Aspekte der Lesesozialisation im Home

Literacy Environment fokussiert und somit vermutlich eher für die Beschreibung der Lesekompetenz als zentraler Kulturtechnik genutzt werden kann.

Aus Tabelle 15 können die Faktorladungen und Residualvarianzen entnommen werden. Der Fragestamm der Skala lautet dabei „Wie oft haben Sie oder jemand anders aus Ihrem Haushalt die folgenden Tätigkeiten mit Ihrem Kind durchgeführt, bevor es in die Grundschule kam?“ und verfügt über die (hier invertierten) Ausprägungen 1 = nie oder fast nie, 2 = manchmal und 3 = oft (Wendt et al., 2016).

**Tabelle 15:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚HLE‘ aus Mplus

Item	Bezeichnung	Faktorladung	Residualvarianz
ASBH02Ar	Bücher lesen	0,486	0,763
ASBH02Br	Geschichten erzählen	0,486	0,763
ASBH02Er	Sich über Dinge unterhalten, die Sie gemacht haben	0,486	0,764
ASBH02Fr	Sich darüber unterhalten, was Sie gelesen haben	0,651	0,576

Einige der Faktorladungen liegen mit 0,486 auf mittlerem Niveau (und weisen entsprechend eine höhere Residualvarianz auf). Allerdings zeichnet sich das Modell durch akzeptable bis gute Fitwerte mit einem RMSEA von 0,065, einem CFI von 0,987 und einem TLI von 0,961 aus. Auch die interne Konsistenz liegt mit Cronbachs  $\alpha = 0,620$  gerade noch im akzeptablen Bereich. Da dieses Modell zudem inhaltlich plausibel und für die Ausarbeitungen von theoretischer Bedeutung ist, wird diese Skala bei den Analysen berücksichtigt.

### 12.3.3 Leseleistung

Neben den Herkunftskonstrukten ESCS und HLE wird die Leseleistung der Schülerinnen und Schüler als ein Indikator für den Schulerfolg in das Mehrebenenstrukturgleichungsmodell aufgenommen. Eine Beschreibung ihrer inhaltlichen Konstruktion in IGLU (reading literacy) findet sich in Kapitel 8.1. Methodisch müssen für die Verwendung der Leistungsvariablen einige Besonderheiten berücksichtigt werden. Generell liegen für jeden Schüler und für jede Schülerin mehrere Leistungswerte vor: Zum einen finden sich Variablen für jede einzelne der erfassten Dimensionen des Leseverständnisses (Verstehensprozess und Leseintention) und Variablen mit dem Gesamtscore, also der generellen Leseleistung. Zum anderen liegen für all diese Werte jeweils fünf Replikationen vor. Dies liegt darin begründet, dass die Schülerinnen

und Schüler in einem rotierten Testheftdesign jeweils nur einen Ausschnitt der für die Messung der Leseleistung zur Verfügung stehenden Items bearbeiten, um die Testdauer zu verkürzen. Auf Grundlage der Testergebnisse und ausgesuchten Hintergrundmerkmale der Viertklässlerinnen und Viertklässler wird für jedes befragte Schulkind eine mögliche Leistungsverteilung der Lesekompetenz bestimmt, in der seine ‚wahre‘ (und nicht direkt messbare) Lesekompetenz liegt. Aus dieser latenten Verteilung werden dann fünf *plausible Werte* zufällig gezogen. Solche Mehrfachziehungen reduzieren die Wahrscheinlichkeit, Ausreißer aus der für den jeweiligen Schüler bzw. die jeweilige Schülerin angenommenen Leistungsverteilung zu ziehen. Darüber hinaus dienen sie dazu, die wahren Varianzen der Population besser durch die Stichprobe zu beschreiben.

Ihre Anwendung, wie etwa hier im Rahmen von Mehrebenenstrukturgleichungsmodellen, gelingt anhand einer wiederholten, voneinander getrennten Berechnung der Gesamtmodelle: Hierbei wird das zu schätzende Modell fünf Mal mit je einem Plausible Value (PV) berechnet. Durch ein von Rubin (1987) entwickeltes Verfahren werden im Anschluss an die Mehrfachspezifikation des Modells die Ergebnisse zusammengefasst, die dann erwartungstreue Schätzer der Populationsparameter (Hornberg et al., 2007) darstellen.

Dieses Verfahren kommt entsprechend immer dann zum Einsatz, wenn Leistungswerte in die Analysen einbezogen werden. Wird sich also folgend auf die Lesekompetenz von Viertklässlerinnen und Viertklässlern bezogen, handelt es sich dabei um die nach Rubin (1987) zusammengefassten Plausible Values der Gesamtleseleistung („overall reading score“). Anzumerken sei darüber hinaus, dass die Plausible Values vor der Analyse im Gesamtmodell anhand der gleichen Transformationsformel z-standardisiert (Mittelwert = 0, Standardabweichung = 1) wurden, da sich die ursprüngliche Metrik (Mittelwert  $\approx 541,0$ , Standardabweichung  $\approx 66,4$ ) sehr von der der übrigen Items (i.d.R. drei- bis fünfstufige Metrik) unterscheidet, was in dem Analyseprogramm Mplus zu Schätzproblemen führt.

Da die Verteilung der Leistungswerte im internationalen Vergleich bereits in Kapitel 12.1 beschrieben wurde, soll dies hier nicht weiter ausgeführt werden.

#### 12.3.4 Leseselbstkonzept

Die Konstruktion des Leseselbstkonzepts orientiert sich an dem Vorgehen des ESCS und HLE mit einer Ergänzung: Aufgestellt wird ein Messmodell sowohl für die Individual- als auch für die Kontextebene, da es sowohl als individueller Indikator für den Schulerfolg (L1) herangezogen werden soll als auch als Konstrukt, welches durch die Lehrperson bzw. die individuelle Förderung (L2, vgl. Hypothese 15 in Kapitel 11.5) beeinflusst wird. Genutzt wird hierfür die international verfügbare Skala „Selbstkonzept Lesen (1)“ (Wendt et al., 2016), die aus sieben Items mit dem Fragestamm „Wie gut liest du? Gib an, wie sehr du jeder der folgenden Aussagen zustimmst“ und der Kodierung 1 = stimme stark zu, 2 = stimme einigermaßen zu, 3 = stimme wenig zu und 4 = stimme überhaupt nicht zu besteht. Mittels einer explorativen Faktorenanalyse wurde in SPSS zunächst geprüft, ob die sieben verfügbaren Items auf einen Faktor laden. Das Ergebnis – eine zweifaktorielle Lösung – findet sich im Anhang 4.2. Allerdings konnte diese Faktorstruktur anhand konfirmatorischer Analysen unter Berücksichtigung der Mehrebenenstruktur und mit einer Modellierung des Konstrukts sowohl auf Individual- als auch auf Kontextebene in Mplus nicht bestätigt werden, wie sich an den nicht akzeptablen Fitwerten ablesen lässt (RMSEA = 0,077, CFI = 0,883, TLI = 0,824, SRMR<sub>L1</sub> = 0,055, SRMR<sub>L2</sub> = 0,136).

Für die nachstehenden Analysen herangezogen wird daher ein in der konfirmatorischen Faktorenanalyse in Mplus abgebildeter Faktor, der Items wie „normalerweise bin ich gut im Lesen“ und „lesen fällt mir sehr leicht“ beinhaltet. Die Zuordnung zu einem latenten Konstrukt wird – wie oben dargestellt – in Mplus auf beiden Ebenen (L1 und L2) überprüft. In Tabelle 16 sind die Faktorladungen und Residualvarianzen für beide Ebenen angegeben. Die Items, die mit einem ‚r‘ gekennzeichnet sind, wurden vorab invertiert, so dass hohe Zahlenwerte einem hohen Leseselbstkonzept entsprechen.

**Tabelle 16:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Leseselbstkonzept‘ auf within- und between-Ebene aus Mplus

Item	Bezeichnung	Level 1		Level 2	
		Faktor- ladung	Residual- varianz	Faktor- ladung	Residual- varianz
ASBR08Ar	Normalerweise bin ich gut im Lesen.	0,831	0,309	0,975	0,049
ASBR08Br	Lesen fällt mir sehr leicht.	0,760	0,422	0,943	0,111
ASBR08E	Es fällt mir schwer, Geschichten mit schwierigen Wörtern zu lesen.	0,343	0,883	0,922	0,150
ASBR08Fr	Meine Lehrerin/mein Lehrer sagt, dass ich gut lesen kann.	0,588	0,655	0,428	0,817

Es zeigen sich für drei der vier Items mittlere bis hohe Faktorladungen. Lediglich das Item „Es fällt mir schwer, Geschichten mit schwierigen Wörtern zu lesen“ lädt mit 0,343 etwas geringer auf den Faktor, der hier das Leseselbstkonzept auf L1 widerspiegeln soll (und verfügt entsprechend über eine relativ hohe Residualvarianz). Begründet werden kann dies damit, dass sich die Häufigkeitsverteilungen auf diesem Item von denen der anderen Items unterscheiden: So liegt der Mittelwert hier bei 2,77, während er für die anderen Items durchschnittlich bei 3,37 (im Wertebereich von 1 bis 4) liegt. Das Item ASBR08E ist also etwas ‚normalverteilter‘ als die übrigen drei und fällt daher aus dieser Struktur heraus. Da es sich inhaltlich aber um einen guten Prädiktor für das Leseselbstkonzept handelt, wird das Item weiterhin dem Konstrukt zugeordnet. Zudem ergeben sich für das Messmodell sehr gute Fitwerte mit einem RMSEA von 0,038, einem CFI von 0,991, einem TLI von 0,974 und einem SRMR<sub>L1</sub> von 0,015 bzw. einem SRMR<sub>L2</sub> von 0,033. Die interne Konsistenz ist mit einem Cronbachs  $\alpha$  von 0,685 als akzeptabel zu bewerten, so dass auch dieses Modell in die folgenden Analysen aufgenommen wird.

### 12.3.5 Schüler-Lehrer-Beziehung

Die Schüler-Lehrer-Beziehung sollte ursprünglich ebenfalls latent auf Individual- und Kontextebene modelliert werden – auf beiden Ebenen, da nicht nur angenommen wird, dass eine positive Wahrnehmung der Schüler-Lehrer-Beziehung seitens der Lernenden den Schulerfolg positiv beeinflusst (Individualebene, Hypothese 8 und 9, s. Kapitel 11.2), sondern auch, dass diese durch individuelle Förderung positiv beeinflusst wird (Kontextebene, Hypo-

these 20, s. Kapitel 11.5). Zur Verfügung stehen hier fünf national erhobene Items<sup>66</sup> der Skala „Klassenklima: Schüler-Lehrer-Beziehung“ (Wendt et al., 2016) mit dem Fragestamm „Wie geht es dir mit deiner Klassenlehrerin/deinem Klassenlehrer?“ und der Kodierung 1 = stimmt genau, 2 = stimmt eher, 3 = stimmt eher nicht und 4 = stimmt gar nicht. Anhand einer explorativen Faktorenanalyse in SPSS wurde geprüft, ob diese fünf Items durch einen gemeinsamen Faktor abgebildet werden können. Hierbei ergab sich eine einfaktorielle Lösung (s. Anhang 4.3) mit einer akzeptablen internen Konsistenz in Höhe von Cronbachs  $\alpha = 0,757$ . Die daran angeschlossene konfirmatorische Überprüfung auf L1 und L2 in Mplus ergab jedoch unzureichende Fitwerte in Höhe von RMSEA = 0,100, CFI = 0,914, TLI = 829, SRMR<sub>L1</sub> = 0,048 und SRMR<sub>L2</sub> = 0,045. Solche Unterschiede zwischen SPSS und Mplus können zum einen auf den unterschiedlichen Umgang mit fehlenden Werten (listenweiser Fallausschluss vs. FIML), zum anderen auf die Berücksichtigung der Mehrebenenstruktur in Mplus zurückgeführt werden. In einem empirisch geleiteten Ansatz wurde dann untersucht, ob sich das Konstrukt möglicherweise durch eine Auswahl der fünf Items abbilden lässt. Die beste Lösung ist im Folgenden angegeben (vgl. Tabelle 17).

---

<sup>66</sup> Die sozialen Facetten (Schüler-Lehrer- und Schüler-Schüler-Beziehung) werden in IGLU nur in der nationalen Erweiterung erfasst. Da wie eingangs beschrieben anzunehmen ist, dass sie für die Erklärung von Schulerfolg herangezogen und zudem durch individuelle Förderung beeinflusst werden können, werden sie dennoch in das Mehrebenenstrukturgleichungsmodell aufgenommen. Anders als das herkunftsbezogene ‚Kernmodell‘ (ESCS, HLE, Leseleistung und Leseselbstkonzept) können sie allerdings nicht für an diese Arbeit anschließende internationale Vergleiche genutzt werden.

**Tabelle 17:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Schüler-Lehrer-Beziehung‘ auf within- und between-Ebene aus Mplus

Item	Bezeichnung	Level 1		Level 2	
		Faktorladung	Residualvarianz	Faktorladung	Residualvarianz
BSNG19A	Meine Klassenlehrerin/mein Klassenlehrer wird schnell ungeduldig, wenn sie/er mir etwas erklären muss.	0,296	0,912	0,766	0,413
BSNG19Br	Ich bin nett zu meiner Klassenlehrerin/meinem Klassenlehrer.	0,784	0,386	1,006	-0,011
BSNG19Cr	Ich höre gut zu, wenn meine Klassenlehrerin/mein Klassenlehrer mich etwas fragt.	0,677	0,542	0,822	0,324
BSNG19Dr	Mit meiner Klassenlehrerin/meinem Klassenlehrer komme ich gut aus.	0,671	0,550	0,987	0,026

Unberücksichtigt blieb bei dieser Lösung das Item BSNG19Er („Meine Klassenlehrerin/mein Klassenlehrer ist freundlich zu mir“). Die Fitwerte sind durchgehend mindestens als akzeptabel zu bewerten (RMSEA = 0,062, CFI = 0,977, TLI = 0,930, SRMR<sub>L1</sub> = 0,029, SRMR<sub>L2</sub> = 0,055), allerdings zeigt sich eine sehr geringe Faktorladung des Items BSNG19A auf L1 und eine Ladung außerhalb des gültigen Wertebereichs für Item BSNG19Br auf L2, verbunden mit einer negativen Residualvarianz. Dies sind Hinweise darauf, dass das Modell nicht richtig spezifiziert werden kann. Da das latente Konstrukt auch mit keiner anderen möglichen Kombination der fünf Items (oder einer Auswahl dieser) abgebildet werden konnte, wurde schließlich theoriegeleitet ein Einzelitem ausgewählt, was am ehesten die Wahrnehmung der Schüler-Lehrer-Beziehung mit all ihren Implikationen abbildet und manifest in das Modell aufgenommen. Dabei handelt es sich um das Item BSNG19Dr („Mit meiner Klassenlehrerin/meinem Klassenlehrer komme ich gut aus“), da alle anderen Variablen eher besondere Verhaltensweisen, teilweise in spezifischen Situationen und keine generelle Einschätzung der gegenseitigen Beziehung abbilden.

Die folgende Tabelle 18 gibt einen Überblick über die Häufigkeitsverteilung des Einzelitems.

**Tabelle 18:** Häufigkeitsverteilung des Items „Mit meiner Klassenlehrerin/meinem Klassenlehrer komme ich gut aus“

Ausprägung	Häufigkeit	Prozent
Stimmt gar nicht	19.843	2,7
Stimmt eher nicht	27.161	3,7
Stimmt eher	132.600	17,8
Stimmt genau	465.119	62,6
Fehlend	98.764	13,3
Gesamtsumme	743.487	100,0

Quelle: Eigene Berechnung, nach TCHWGT gewichtet, daraus ergeben sich die dem Populationsumfang entsprechenden Stichprobenzahlen; mehr als 100% ergeben sich durch Rundungen

Deutlich wird, dass die dem Item zugehörigen Antworten schief verteilt sind: Fast zwei Drittel der Viertklässlerinnen und Viertklässler geben an, mit ihrem Klassenlehrer bzw. ihrer Klassenlehrerin gut auszukommen (gegenüber 6,4%, bei denen dies eher oder gar nicht der Fall ist). Dies drückt sich ebenfalls in einer sehr hohen Itemschwierigkeit<sup>67</sup> ( $P_i = 87,3$ ) aus. Allerdings weisen auch die übrigen vier Items, die aus theoretischen Gründen nicht herangezogen wurden, ähnliche Verteilungen auf. Dies wird bei der Interpretation der Ergebnisse entsprechend zu berücksichtigen sein.

### 12.3.6 Schüler-Schüler-Beziehung

Bei der Schüler-Schüler-Beziehung handelt es sich um ein weiteres national verfügbares Konstrukt aus IGLU-E. Für seine Konstruktion wurde die Skala „Einschätzung der eigenen sozialen Integration“ mit dem Fragestamm „Wie wohl fühlst du dich mit deinen Mitschülern?“ herangezogen. Für diese finden sich neun Items mit der Kodierung 1 = stimmt genau, 2 = stimmt fast, 3 = stimmt ein wenig und 4 = stimmt gar nicht. Eine explorative Faktorenanalyse in Mplus ergab eine zweifaktorielle Lösung (vgl. Tabelle 35 im Anhang 4.4) mit jeweils guter interner Konsistenz (Cronbachs  $\alpha$  für Faktor 1 = 0,834, für Faktor 2 = 0,707). Da SPSS jedoch nicht die Möglichkeit bietet, ein Konstrukt auf Kontextebene zu überprüfen, wurde dem eine konfirmatorische Analyse in Mplus nebengestellt. Dabei ergaben sich für beide Konstrukte ungünstige Fitwerte<sup>68</sup>. Da theoriegeleitet keine plausible Einschränkung des Konstrukts vorgenommen werden konnte, wurden die Inter-Item-Korrelationen dahin-

<sup>67</sup>  $P_i = \frac{(\bar{x}_i - 1)}{\max(x_i) - 1} \cdot 100 = \frac{(3,6177 - 1)}{4 - 1} \cdot 100 \approx 87,3$ , mit  $\bar{x}_i$  = Itemmittelwert und  $\max(x_i)$  = maximal erreichbarer Wert  $i$  für ein Item  $x$ .

<sup>68</sup> Faktor 1: RMSEA = 0,121, CFI = 0,897, TLI = 0,794, SRMR<sub>L1</sub> = 0,028, SRMR<sub>L2</sub> = 0,181.  
Faktor 2: RMSEA = 0,130, CFI = 0,876, TLI = 0,629, SRMR<sub>L1</sub> = 0,056, SRMR<sub>L2</sub> = 0,194.

gehend überprüft, ob sich Zusammenhänge zwischen Items aufdecken lassen, aus denen ein reliables Konstrukt abgeleitet werden kann. Die folgende Tabelle 19 gibt einen Überblick über die so dem Faktor ‚Schüler-Schüler-Beziehung‘ zugeordneten Items, ihren Faktorladungen und den Residualvarianzen auf Individual- und Kontextebene. Die Items, die mit einem ‚r‘ gekennzeichnet sind, wurden vorab invertiert, so dass hohe Werte einer hohen sozialen Integration entsprechen.

**Tabelle 19:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Schüler-Schüler-Beziehung‘ auf within- und between-Ebene aus Mplus

Item	Bezeichnung	Level 1		Level 2	
		Faktorladung	Residualvarianz	Faktorladung	Residualvarianz
BSNG18Er	Ich habe wirklich gute Freundinnen oder Freunde in der Klasse.	0,586	0,657	0,506	0,744
BSNG18Fr	Meine Mitschülerinnen und Mitschüler halten viel von mir.	0,716	0,487	0,965	0,068
BSNG18Gr	Meine Mitschülerinnen und Mitschüler mögen mich so, wie ich bin.	0,805	0,352	0,987	0,025
BSNG18Hr	Die anderen arbeiten gern mit mir in Arbeitsgruppen zusammen.	0,802	0,358	0,975	0,049

Für eines der dargestellten Items (BSNG18Er) zeigen sich sowohl auf Individual- als auch auf Kontextebene geringere Faktorladungen als für die übrigen Items. Dies könnte sich inhaltlich begründen lassen: Während die übrigen Items eher die allgemeine Integration im Klassenkontext fokussieren, erfasst das Item BSNG18Er, ob ein Kind „wirklich gute Freunde“ in der Klasse hat. Möglich wäre es, dass ein Kind durchaus sozial integriert ist, seine besten Freunde aber beispielsweise bei Freizeitaktivitäten gefunden hat, diese also nicht dieselbe Klasse besuchen. Dass dieses Item etwas weniger stark dem latenten Konstrukt zugeordnet wird, erscheint daher plausibel. Da die Faktorladungen mit 0,586 (L1) und 0,506 (L2) dennoch als ausreichend hoch zu bewerten sind und die inhaltliche Nähe des Items zum Konstrukt ‚Schüler-Schüler-Beziehung‘ dennoch gegeben ist, wird es in das Messmodell aufgenommen. Letzteres weist mit einem RMSEA von 0,060, einem CFI von 0,987, einem TLI von 0,962 sowie einem SRMR auf L1 von 0,018 und auf L2 von 0,051 eine gute Modellpassung sowie eine hohe interne Konsistenz von Cronbachs  $\alpha = 0,830$  auf. Das Modell wird also wie oben dargestellt in die Mehrebenenstrukturgleichungsanalyse aufgenommen.

### 12.3.7 Klassenzusammensetzung

Wie in den Kapiteln 10 und 11.4 beschrieben, sollen auch Merkmale der Klassenzusammensetzung in die Analysen aufgenommen werden. Dabei handelt es sich um den klassendurchschnittlichen sozioökonomischen Status der Schülerinnen und Schüler (gemessen anhand des in Kapitel 12.3.1 dargestellten ESCS) und den klassendurchschnittlichen häuslichen Sprachgebrauch („Wie oft sprichst du zu Hause Deutsch?“).

Für die Konstruktion des ESCS auf Klassenebene werden die Einzelitems HISEI, ISCED und der Bücherbesitz für jede Klasse gemittelt. Herangezogen werden hierfür jeweils alle für eine Klasse zur Verfügung stehenden Informationen der einzelnen Schülerinnen und Schüler. Mit Hilfe des Aggregate-Befehls in SPSS wird hieraus ein Klassendurchschnitt gebildet, der dann jedem einzelnen Lernenden angespielt wird, so dass alle Schülerinnen und Schüler einer Klasse denselben Wert für das jeweilige Merkmal der Klassenzusammensetzung erhalten. Genutzt werden auch hier die z-standardisierten Werte der drei Einzelindikatoren, wie dies bereits in Kapitel 12.3.1 beschrieben wurde. In der nachstehenden Tabelle 20 werden zusätzlich die nicht-standardisierten Itemanalysen abgebildet, da diese leichter zu interpretieren sind – gerade auch vor dem Hintergrund der in Kapitel 12.3.1 ausführlicher dargestellten Einzelitems.

**Tabelle 20:** Deskriptive Analysen der Klassenaggregate der z-standardisierten und nicht-standardisierten Items HISEI, ISCED und Bücherbesitz

Standardisierung	Item	Wertebereich		Streuung im Wertebereich		Mittelwert	Standardabweichung
		Min	Max	Min	Max		
nicht-stand.	HISEI	22,00	73,00	33,60	71,00	52,27	5,773
	ISCED	1,00	7,00	3,00	7,00	4,64	0,708
	Bücher	1,00	5,00	2,00	5,00	3,54	0,537
z-stand.	HISEI	-2,44	1,65	-1,51	1,48	-0,02	0,462
	ISCED	-2,38	1,53	-1,08	1,52	-0,01	0,461
	Bücher	-2,09	1,19	-1,27	1,19	-0,01	0,440

Verglichen mit dem Gesamtwertebereich der Einzelitems wird hier deutlich, dass mehr Klassen sich durch positive mittlere Hintergrundmerkmale (hoher HISEI, ISCED und Bücherbesitz) auszeichnen als dies im unteren Bereich der Fall ist, was sich auch in leicht überdurchschnittlichen Mittelwerten ausdrückt. Dieser Befund scheint für Deutschland plausibel, wo die unteren Merkmalsausprägungen (kein Schulbesuch, niemals bezahlte Arbeit verrichtet, 0-10 Bücher) nur in wenigen Fällen zutreffen. Dieses für die Individualebene bereits verifi-

zierte latente Modell wurde hier einer Prüfung auf Kontextebene unterzogen. Wenngleich auf die übliche Fitstatistik nicht zurückgegriffen wurde, da das Modell mit null Freiheitsgraden ein saturiertes Modell darstellt, zeigen doch die drei Einzelitems eine hohe Faktorladung auf das latente Konstrukt (vgl. Tabelle 21).

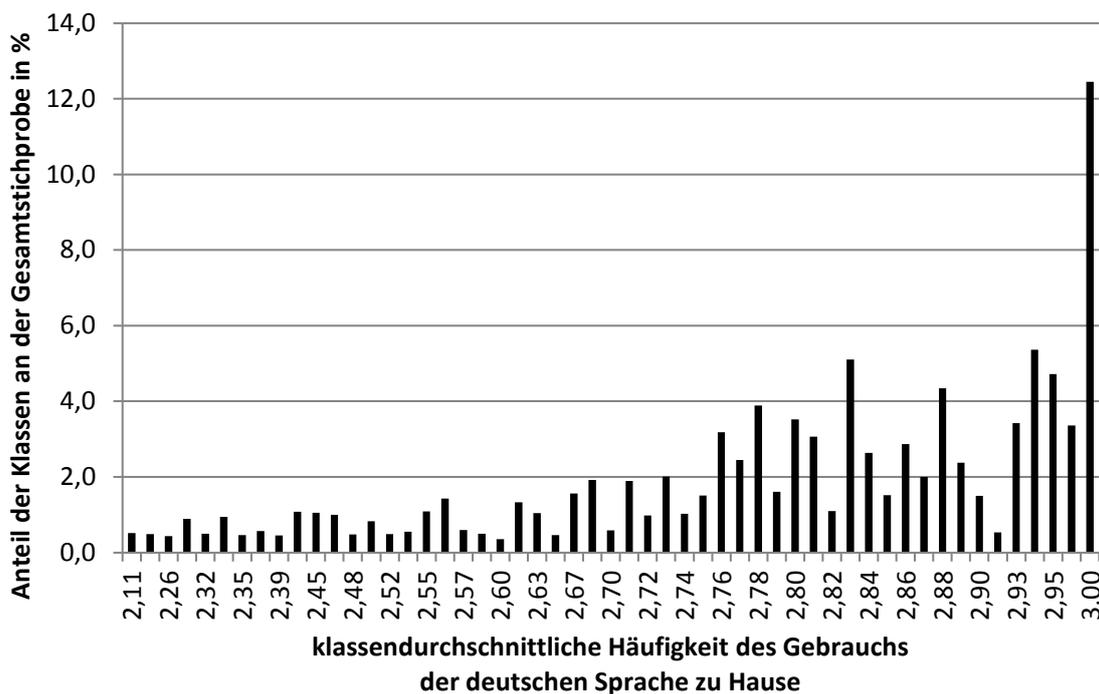
**Tabelle 21:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die auf Klassenebene aggregierten Items des Messmodells ‚Klassenaggregat ESCS‘ auf between-Ebene aus Mplus

Item	Faktorladung	Residualvarianz
HISEI (KA)	0,877	0,230
ISCED (KA)	0,846	0,284
Bücherbesitz (KA)	0,884	0,218

Legende: KA = Klassenaggregat

Dementsprechend kann das latente ESCS-Konstrukt auf Kontextebene als Indikator für die sozioökonomische Klassenzusammensetzung herangezogen werden.

Der häusliche Sprachgebrauch als weiterer Indikator wurde erfasst anhand der Frage „Wie oft sprichst du zu Hause Deutsch?“ (Schülerfragebogen) mit der hier invertierten Kodierung 1 = Ich spreche zu Hause niemals Deutsch, 2 = Ich spreche zu Hause manchmal Deutsch und manchmal eine andere Sprache, 3 = Ich spreche zu Hause immer oder fast immer Deutsch. Analog zu dem Vorgehen für die Einzelitems des ESCS wurden die Angaben der Schülerinnen und Schüler auf Klassenebene aggregiert. Für die gesamte Stichprobe ergibt sich dabei ein Mittelwert von 2,79 mit einer Standardabweichung von 0,181. Der minimale klassendurchschnittliche Wert liegt bei 2,11 auf der Skala mit dem Wertebereich 1 bis 3, der maximale Wert bei 3. Der folgenden Abbildung 17 kann entnommen werden, dass dies 12,4% der Klassen ausmacht, im Gegensatz zu 0,5% der Klassen mit der niedrigsten Ausprägung, was sich auch in der sehr hohen Itemschwierigkeit  $P_i = 89,7$  ausdrückt.

**Abbildung 17:** Klassendurchschnittliche Häufigkeit des Gebrauchs der deutschen Sprache im familiären Kontext in Prozent

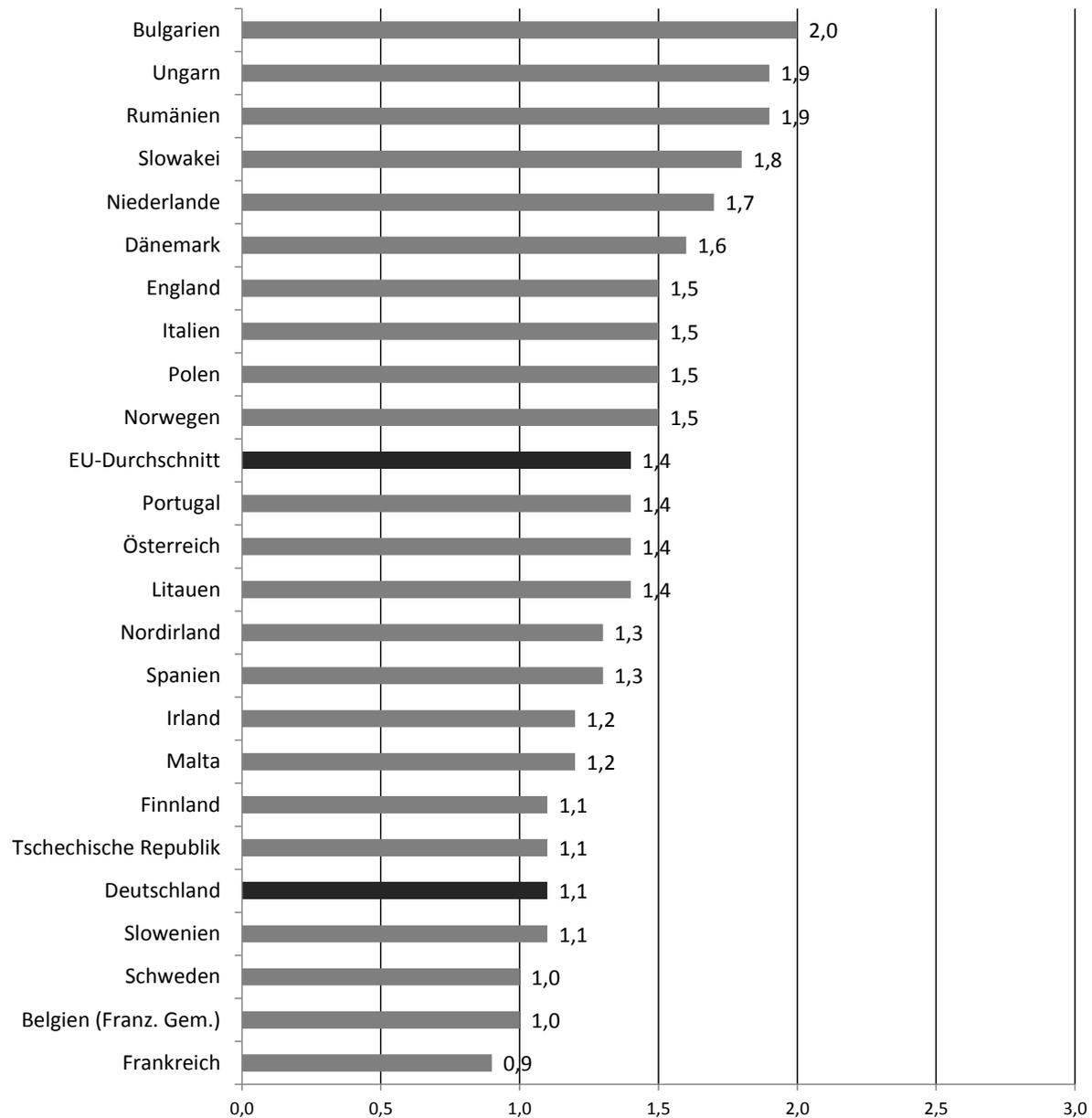
Deutlich wird, dass die Schülerinnen und Schüler in einem Großteil der Klassen überwiegend Deutsch zu Hause sprechen. Dennoch wird dieses auf Klassenebene aggregierte Merkmal – neben dem klassendurchschnittlichen ESCS – für die Analysen herangezogen.

### 12.3.8 Individuelle Förderung

Die Angaben zu der Häufigkeit individueller Förderung im Unterricht finden sich im Lehrer-rahmenfragebogen und stellen daher Selbsteinschätzungen der Lehrkräfte dar. Zwei Skalen lassen sich dabei unterscheiden: Während die international verfügbare Skala „Formen der Differenzierung im Leseunterricht“ und daraus das Item „Ich gebe individuell abgestimmten Leseunterricht“ (ATBR03Dr) mit der Kodierung 1 = nie, 2 = manchmal, 3 = oft und 4 = immer oder fast immer (invertiert) lediglich auf die Häufigkeit individueller Förderung im Leseunterricht rückschließen lässt, finden sich detailliertere Informationen zu dem *didaktischen Vorgehen* der Lehrkraft in der national eingesetzten Skala „Förderung von Schülerinnen und Schülern nach ihren individuellen Lernvoraussetzungen“.

Um einen Überblick darüber zu gewinnen, wie häufig im internationalen (EU-)Kontext Viertklässlerinnen und Viertklässler individualisierten Leseunterricht erfahren, wird das oben beschriebene Einzelitem zunächst für eine länderbezogene Häufigkeitsanalyse herangezogen (vgl. Abbildung 18).

**Abbildung 18:** Häufigkeit des individuellen Leseunterrichts nach Angabe der Lehrkräfte im europäischen Vergleich



Legende: Mittelwerte, Kodierung des Items: 0 = nie, 1 = manchmal, 2 = oft, 3 = immer oder fast immer  
 Datenreihe: Tarelli, Lankes, Drossel & Gegenfurtner, 2012, S. 159. Angaben zum EU-Durchschnitt aus eigenen Berechnungen; eigene Abbildung

Deutlich wird, dass die individuelle Förderung (erfasst anhand des Einzelitems) in Deutschland weit hinter dem EU-Durchschnitt (1,4) bleibt: Mit einer gemittelten Häufigkeit von 1,1 wird so wenig wie in kaum einem anderen europäischen Land im Leseunterricht individuell gefördert<sup>69</sup>. Lediglich in Schweden, Belgien und Frankreich geben Lehrkräfte an, den Leseun-

<sup>69</sup> Anzumerken ist an dieser Stelle, dass es sich dabei um reine Zahlenwertvergleiche handelt aus denen keine Aussagen hinsichtlich der Signifikanz der Unterschiede getroffen werden können.

terricht weniger zu individualisieren<sup>70</sup>. So werden in Deutschland 18,6% der Viertklässlerinnen und Viertklässler nie, 52,5% manchmal, 20,2% oft und lediglich 3,4% immer oder fast immer individuell im Lesen gefördert (Rest: keine Angaben). Der Mittelwert liegt, wie oben dargestellt, bei 1,09, die Standardabweichung bei 0,739. Tarelli et al. (2012) konstatieren, dass sich diese niedrigen Werte für Deutschland in den letzten Jahren nicht verändert hätten.

Dieser Einordnung in den internationalen Kontext muss jedoch mit Vorsicht begegnet werden. Wie Analysen zu dem Zusammenhang zwischen dem oben verwendeten internationalen Einzelitem und der nationalen Skala, die sieben konkrete didaktische Handlungen der Lehrperson abbildet, die Bestandteil individueller Förderung sind, zeigen, korrelieren beide nur in mittlerer Stärke ( $r = 0,444$ ). Angenommen wird hier vor allem, dass Lehrende möglicherweise die Häufigkeit, mit der sie allgemein im Unterricht individuell fördern, unterschätzen, wenngleich die anhand der nationalen Skala erfassten Einzelmerkmale der Individualisierung Bestandteil ihres Unterrichts sind. Bevor auf dies vertiefend eingegangen wird und entsprechende Implikationen abgeleitet werden können, soll nachfolgend zunächst die nationale Skala, die in das Mehrebenenstrukturgleichungsmodell aufgenommen wird, vorgestellt werden.

Bei dieser handelt es sich um die „Förderung von Schülerinnen und Schülern nach ihren individuellen Lernvoraussetzungen“ (Wendt et al., 2016), bestehend aus 14 Items. Ihre explorativ überprüfte Faktorstruktur ergibt dabei eine plausible Strukturierung (vgl. Tabelle 36 in Anhang 4.5): Während eine Großzahl der Items, die inhaltlich die gezielte individuelle Förderung einzelner Schülerinnen und Schüler abdecken, auf einen Faktor laden, sind es vor allem Fragen zu dem Einsatz von Computern für die Leistungsdifferenzierung<sup>71</sup>, zu einer Form des Mastery Learnings (ATNG27Ir) und des offenen Unterrichts (ATNG27Mr) sowie zu binnendifferenzierenden, jedoch nicht zwangsweise für einzelne Lernende geplante Maßnahmen (ATNG27Hr), die – inhaltlich plausibel – den übrigen drei Faktoren zugeordnet sind. Dem-

<sup>70</sup> Zu beachten ist, dass aus den oben dargestellten Befunden nicht geschlossen werden kann, dass in Deutschland und anderen EU-Staaten *allgemein* wenig binnendifferenziert wird: Hierfür finden sich weitere Items in IGLU, die für diese Betrachtung individueller Förderung jedoch nicht relevant sind, deren Verteilung über die EU-Staaten hinweg allerdings von der oben dargestellten abweichen kann.

<sup>71</sup> Items, die explizit die Nutzung von Computersoftware fokussieren, könnten indirekt auch die Medienaffinität der Lehrpersonen oder auch die Ausstattung einer Schule (Vorhandensein von Computern, Erwerb entsprechender Programme etc.) erfassen – dies mag ein Grund dafür sein, dass die beiden Items ATNG27Kr und ATNG27Nr in Abgrenzung zu den übrigen Items auf einen gemeinsamen Faktor laden, wenngleich sie inhaltlich durchaus der individuellen Förderung zugeordnet werden könnten.

entsprechend wurde lediglich der erste Faktor, bezeichnet als ‚Individuelle Förderung‘ einer konfirmatorischen Überprüfung in Mplus unterzogen und konnte bestätigt werden.

In Tabelle 22 finden sich die Items, die das latente Konstrukt abbilden. Der Fragestamm der Skala lautet dabei „Wie häufig machen Sie die folgenden Dinge, um Schüler nach ihren individuellen Lernvoraussetzungen zu fördern?“.

**Tabelle 22:** STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Individuelle Förderung‘ auf between-Ebene aus Mplus

Item	Bezeichnung	Faktorladung	Residualvarianz
ATNG27Ar	Ich gebe Schülern je nach Leistung unterschiedlich schwere Hausaufgaben.	0,652	0,575
ATNG27Br	Ich lasse schnellere Schüler schon zum Nächsten übergehen, während ich mit den langsameren noch übe oder wiederhole.	0,545	0,702
ATNG27Cr	Wenn Schüler etwas nicht verstanden haben, vergebe ich gezielte Zusatzaufgaben.	0,688	0,526
ATNG27Er	Ich gebe schwachen Schülern zusätzliche Unterstützung im Unterricht.	0,487	0,763
ATNG27Fr	Leistungsstarken Schülern gebe ich Extraaufgaben, durch die sie wirklich gefordert werden.	0,675	0,544
ATNG27Gr	Bei der Stillarbeit variiere ich die Aufgabenstellungen, um Schülern unterschiedlicher Leistungsstärke gerecht zu werden.	0,705	0,503
ATNG27Jr	Ich lasse die Schüler regelmäßig in Gruppen oder allein an unterschiedlich schwierigen Aufgaben arbeiten.	0,641	0,590

Die Kodierung wurde für alle Items invertiert, so dass hohe Werte einer häufigen individuellen Förderung entsprechen: 0 = nie, 1 = in 1 von 10 Stunden, 2 = in 2 von 10 Stunden, 3 = in 3 oder 4 von 10 Stunden, 4 = in 5 bis 8 von 10 Stunden und 5 = in fast jeder Stunde. Der Mittelwert liegt bei 3,20, die Standardabweichung bei 0,892. Das latente Konstrukt weist für alle sieben Items ausreichend hohe Faktorladungen auf und zeichnet sich durch einen guten Modellfit und eine hohe innere Konsistenz aus: So liegt der RMSEA bei 0,018, der CFI bei 0,939, der TLI bei 0,908, der SRMR<sub>L2</sub> bei 0,041 und Cronbachs Alpha bei 0,819. Dieses latente Konstrukt stellt also empirisch einen geeigneten Indikator für die allgemeine individuelle Förderung dar. Im Vergleich mit den in Kapitel 4 beschriebenen Elementen individueller Förderung zeigt sich jedoch, dass diese nicht vollständig durch die Skala abgedeckt sind. Wäh-

rend sich zwar die Diagnose der lernbezogenen Stärken und Schwächen der Schülerinnen und Schüler, die Anpassung von Lernaufgaben an ihre Lernvoraussetzungen, die Lernziendifferenzierung, die Lernberatung sowie das Bereitstellen didaktisch unterschiedlicher Materialien meist implizit wiederfinden, fehlt hier die Erfassung lern- und leistungsbezogener, motivierender Rückmeldungen, Angaben zu didaktisch angepassten Formen der Leistungsbeurteilung und -dokumentation sowie zu der Eignung des Klassenzimmers und seiner Ausstattung für individualisierte Settings<sup>72</sup>. Ebenfalls nicht berücksichtigt wird hier die für die individuelle Förderung benötigte adaptive Lehrkompetenz, zu der die Sach-, die diagnostische, die didaktische und die Klassenführungs-kompetenz (s. Beck et al., 2008) gehören. Diesbezügliche Selbsteinschätzungen der Lehrenden anhand einiger weniger Skalen wären möglicherweise ohnehin durch eine hohe soziale Erwünschtheit gekennzeichnet, so dass IGLU für deren Untersuchung nicht den geeigneten Rahmen bietet. Dennoch kann das dargestellte Konstrukt deutlich detaillierter das didaktische Konzept der individuellen Förderung – fachunabhängig – abdecken als das international verwendete Einzelitem für den Leseunterricht. Auf deren Zusammenhang soll im Folgenden eingegangen werden.

Wie oben beschrieben, stehen beide Indikatoren – das international verfügbare Einzelitem zu der Häufigkeit individueller Förderung im Leseunterricht und die deutlich ausdifferenziertere Skala zu konkreten didaktischen Herangehensweisen, die die individuelle Förderung ausmachen, jedoch fachunabhängig erfragt wurden – lediglich in einem mittleren Zusammenhang ( $r = 0,444$ ). Bei der Überprüfung der Häufigkeitsverteilungen fällt auf, dass Lehrende im Allgemeinen scheinbar häufiger individualisieren, dies jedoch deutlich weniger Anwendung im Leseunterricht findet. So liegt der Mittelwert für die allgemeine Dimension bei 3,0 (im Wertebereich von 0 bis 5, gegenüber einem Mittelwert von 1,1 für die lesebezogene Individualisierung im Wertebereich von 0 bis 3). Aufbauend auf diesen Befund wurde untersucht, in welchem Maß sich diese Angaben überschneiden. So werden 5,9% der Viertklässlerinnen und Viertklässler von Lehrenden unterrichtet, die sowohl im Leseunterricht als auch im Allgemeinen häufig individualisieren. 9,2% der Schülerinnen und Schüler werden im Leseunterricht nicht, jedoch im Allgemeinen individuell gefördert, für den umgekehrten Fall (Individualisierung im Leseunterricht und nicht im Allgemeinen) finden sich erwartungsge-

---

<sup>72</sup> Dabei sei anzumerken, dass zu einigen dieser fehlenden Teilaspekte, insbesondere zu den Formen der Leistungsrückmeldung und -beurteilung, Daten in der nationalen Ergänzung von IGLU zu finden sind, diese hier jedoch keine Berücksichtigung finden, da dies vor allem mit einer deutlichen Erweiterung der theoretischen Betrachtungen einherginge, die den Rahmen der vorliegenden Arbeit überschreiten würden.

mäß keine Fälle. 59,8% der Schulkinder erhalten sowohl im Leseunterricht als auch darüber hinaus kaum individuelle Förderung. Abgeleitet werden kann hieraus zunächst die Tendenz, dass ein Teil der Lehrenden Individualisierung als didaktische Maßnahme im Unterricht ergreift, dies jedoch scheinbar weniger im Leseunterricht Anwendung findet. Dies würden auch die Befunde von Lankes (2004), Lankes, Plaßmeier, Bos und Schwippert (2004) und die Darstellung bei Lipowsky et al. (2011) bestätigen, nach welchen das Lesen in der vierten Klasse in Deutschland vor allem lehrergelenkt-klassenbezogen unterrichtet wird und differenzierende Unterrichtsmethoden häufig nur in Form stillen Lesens praktiziert werden.

Andererseits könnte eine weitere Ursache in dem Verständnis individueller Förderung liegen: Möglicherweise interpretieren einige Lehrende die von ihnen gewählten didaktischen Maßnahmen im Unterricht, wie sie anhand der entsprechend differenzierten Skala erfasst werden, nicht als individuelle Förderung und unterschätzen somit systematisch die anhand des Einzelitems erfasste Häufigkeit, mit der sie (den Leseunterricht) individualisieren. Eine Ursache hierfür könnte auch in dem Fehlen unterrichtsnaher Konzepte, wie sie beispielsweise Altrichter et al. (2009) und Solzbacher (2008) konstatieren, liegen: Demnach fehle es möglicherweise an einer Sensibilisierung der Lehrenden alleine für die Kategorisierung ihrer Didaktik (unter dem Deckmantel ‚individuelle Förderung‘), während sie die zugehörigen Maßnahmen wie die Differenzierung der Aufgaben nach den individuellen Lernständen der Schülerinnen und Schüler bereits in den Unterricht integriert haben.

Während beide Erklärungsansätze hier nicht auf ihren empirischen Gehalt hin überprüft werden können, verdeutlichen sie doch, dass das Einzelitem zu der Häufigkeit individueller Förderung im Leseunterricht wenig valide ist um als alleiniger Prädiktor herangezogen werden zu können. Dies schränkt ebenfalls die Einordnung in den oben dargestellten EU-Kontext ein – in jedem Fall auf den Unterrichtsgegenstand Lesen; möglicherweise können auf Grundlage dieses Einzelitems gar keine Aussagen getroffen werden, sollten Lehrende tatsächlich die Häufigkeit, mit der sie im Leseunterricht individuell fördern, systematisch unterschätzen. Dies bleibt in vertiefenden Analysen, die über die Erfassung in IGLU hinausgehen, festzustellen. In jedem Fall kann die national verfügbare Skala, die einzelne Merkmale individueller Förderung erfasst, für die folgenden Analysen sowohl aus theoretischer als auch aus empirischer Sicht herangezogen werden. Eine einzige Einschränkung besteht darin, dass der Unterrichtsgegenstand (wie etwa der Leseunterricht) nicht erfasst wurde und Aussagen nur einen

allgemeinen, fachunabhängigen Charakter haben. Dies sei bei der Interpretation der Befunde berücksichtigt.

Insgesamt konnte in diesem Kapitel 12.3 für jedes für die Überprüfung der theoretisch angenommenen Zusammenhänge benötigte Konstrukt ein plausibles empirisches Messmodell, mindestens jedoch ein Einzelitem identifiziert werden. Im Folgenden sollen die theoriegeleiteten Zusammenhänge zwischen diesen Dimensionen zunächst in Einzelschritten überprüft werden, bevor abschließend das gesamte Mehrebenenstrukturgleichungsmodell abgebildet wird. Dem schließt sich die rein deskriptive Darstellung der Ergebnisse in Kapitel 13 an.

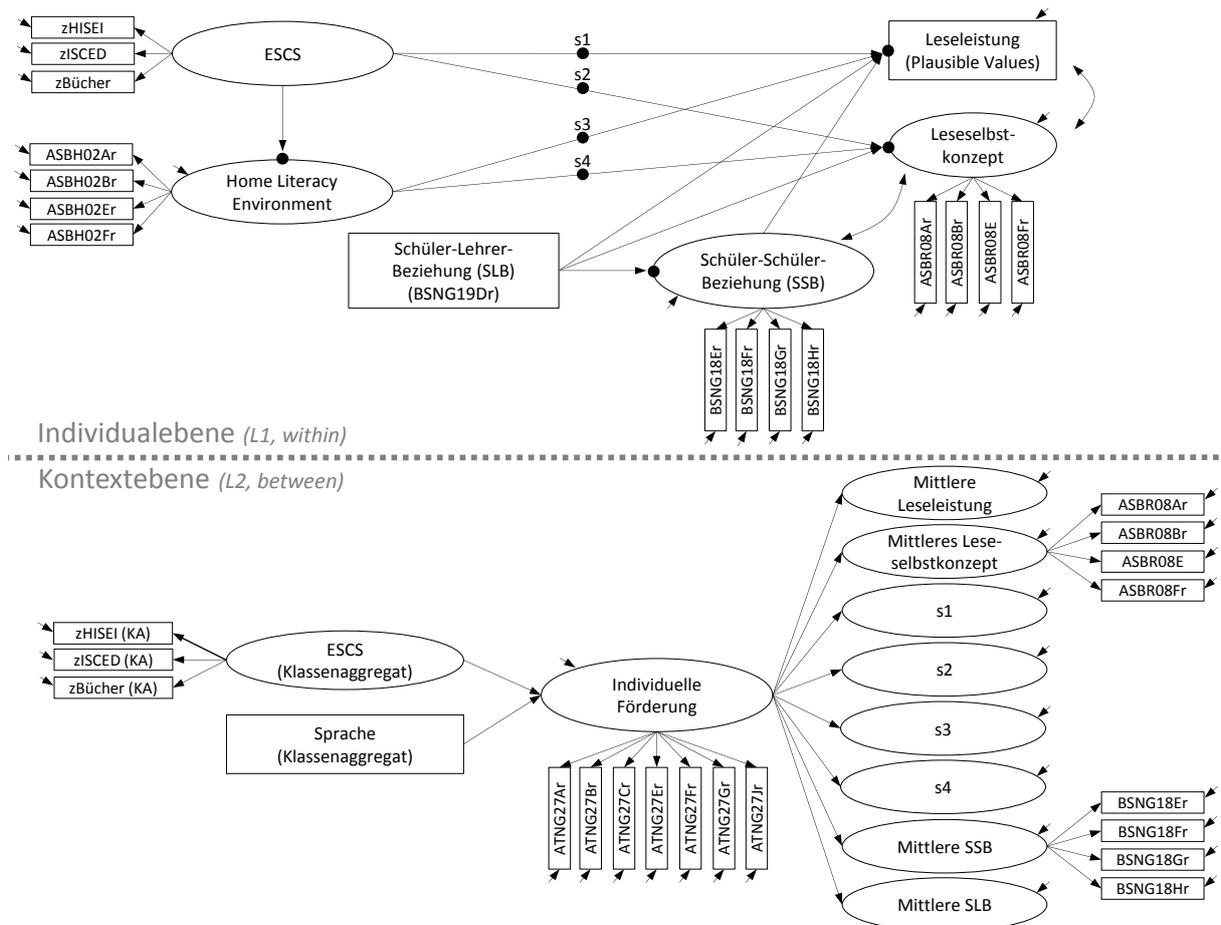
#### **12.4 Prüfung der einzelnen Zusammenhänge**

Bei der empirischen Überprüfung der Beziehungen zwischen den oben dargestellten Konstrukten und Einzelitems ist zu beachten, dass es sich bei Strukturgleichungsmodellen um *strukturtestende* Verfahren handelt. Ihr Ziel ist es nicht, in einem explorativen Ansatz Strukturen zu entdecken; viel eher geht es darum, theoretisch begründete Zusammenhänge simultan in einem Gesamtmodell auf ihre Gültigkeit hin zu testen. Eine zentrale Rolle spielen dabei die Fitwerte (CFI, TLI, RMSEA, SRMR, ggf. auch  $\chi^2$ ): Sie geben an, wie gut die theoretischen Annahmen auf die Daten ‚passen‘ (vgl. Kapitel 12.2). Eine nachträgliche und nicht theoriegeleitete Anpassung der Modellstrukturen an diese Fitwerte hätte zur Folge, dass das modifizierte Modell an Allgemeingültigkeit verliert, da eine Anpassung immer an die Spezifika des zugrunde liegenden Datensatzes erfolgt. Diesem muss entsprechend Rechnung getragen werden, indem derartige Modellanpassungen nicht vorgenommen werden. Ausgenommen hiervon ist der Ausschluss nicht signifikanter Pfade aus dem Gesamtmodell: Dabei handelt es sich um theoretisch begründete Beziehungen, die sich empirisch nicht halten lassen, also im Sinne des konfirmatorischen Charakters der Analyseverfahren nicht bestätigt werden können.

Die Überprüfung der hypothetischen Zusammenhänge, wie sie in Kapitel 11 abgeleitet wurden, erfolgt schrittweise: Die nachfolgende Abbildung 19 gibt einen Überblick über das theoriegeleitete Gesamtmodell mit den oben aufgestellten Messmodellen. Die Beziehungen zwischen diesen Konstrukten werden zunächst einzeln auf ihre Stärke hin geprüft. Hierfür finden sich nachstehend Tabellen, aus denen die entsprechenden Einzelzusammenhänge abgelesen werden können. Nicht signifikante Pfade werden dabei im abschließenden Gesamtmodell zunächst zwar noch einmal aufgenommen (um zu prüfen, ob diese Nichtsignifikanz auch

in der simultanen Analyse aller Zusammenhänge bestehen bleibt und um Suppressions- und ähnliche Effekte auszuschließen), werden bei fortbestehender Nichtsignifikanz jedoch ausgeschlossen. Dieses Verfahren hat sich etabliert, da es das ‚troubleshooting‘ erheblich erleichtert: Da nicht signifikante Pfade die Passung des Gesamtmodells deutlich verzerren, kann aus solchen Voranalysen schneller identifiziert werden, welche theoretischen Annahmen sich empirisch nicht abbilden lassen.

**Abbildung 19:** Theoriegeleitetes Gesamtmodell



Die in Abbildung 19 dargestellten Beziehungen werden folgend in Einzelschritten auf ihre Stärke hin geprüft, bevor in einem finalen Schritt das Gesamtmodell analysiert wird.

### 12.4.1 Nullmodell

Bevor ein Mehrebenenmodell spezifiziert wird, sollte zunächst geprüft werden, wie sich die Varianz der abhängigen Variable(n) auf den Ebenen verteilt. Erst aus dieser Prüfung heraus kann abgeleitet werden, ob und in welchem Umfang bedeutende Unterschiede innerhalb und zwischen den Clustern vorliegen, für deren Erklärung theoriegeleitet Prädiktoren aufgenommen werden können. Mit einer einfaktoriellen Varianzanalyse mit Zufallseffekten in

Mplus („Nullmodell“) wurden die Mittelwerte der abhängigen Variablen (hier Leseleistung und Leseselbstkonzept) und ihre Varianz auf Individual- und Kontextebene modelliert. Aus dieser lässt sich anschließend die Intraclass Correlation (ICC) anhand der Formel  $ICC = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$  mit  $\tau_{00}$  = Varianz der abhängigen Variablen auf L2 und  $\sigma^2$  = Varianz der abhängigen Variablen auf L1 berechnen, aus welcher abgelesen werden kann, wie hoch der Anteil der Gesamtvarianz eines Indikators ist, der auf Unterschiede zwischen den Klassen zurückgeführt werden kann.

**Tabelle 23:** Nicht-standardisierte Varianz, p-Werte und ICC für die abhängige Variable ‚Leseleistung‘ und des abhängige latente Konstrukt ‚Leseselbstkonzept‘ auf Individual- und Kontextebene

Indikator	L1		L2		ICC
	$\sigma^2$	p	$\tau_{00}$	p	
Leseleistung (manifest)	0,768	< 0,001	0,265	< 0,001	0,257
Leseselbstkonzept (latent)	0,354	< 0,001	0,009	0,007	0,025

Aus den Darstellungen in Tabelle 23 wird deutlich, dass 25,7% der Varianz in der Leseleistung auf Unterschiede zwischen den Klassen zurückgeführt werden können. Für das latent modellierte Leseselbstkonzept fällt dies mit 2,5% deutlich geringer aus. Allerdings konstatieren Cohen, Cohen, West und Aiken (2003), dass bereits ICCs in Höhe von 0,01 zu Verzerrungen in Signifikanztests führen können, wenn die Mehrebenenstruktur und mit ihr einhergehend die Ähnlichkeit der Schülerinnen und Schüler innerhalb einer Klasse im Vergleich zu den Unterschieden zwischen Klassen unberücksichtigt bleibt. Allerdings bleibt anzumerken, dass durch den hypothetisch angenommenen Einfluss der individuellen Förderung auf das mittlere Selbstkonzept der Schülerinnen und Schüler einer Klasse (Hypothese 15) maximal die 2,5% der Varianz, die auf Kontextebene festgestellt werden konnte, aufgeklärt werden können – ein vergleichsweise geringer Betrag. Dies soll bei der Prüfung der Hypothese aufgegriffen werden. Dennoch kann die Voraussetzung für Mehrebenenanalysen – eine sich substantiell von Null unterscheidende ICC (vgl. Hox, 2010; Christ & Schlüter, 2010) – hier bestätigt werden.

Im Folgenden wird daher mit der Prüfung der einzelnen Zusammenhänge begonnen. Das finale Mehrebenenstrukturgleichungsmodell findet sich in Kapitel 12.5.

### 12.4.2 Herkunft und Schulerfolg

Begonnen wird dabei mit den Zusammenhängen auf Individualebene und dort mit dem Herkunftseffekt, der zentrales Element der Ausgangsüberlegungen der hier vorgestellten Arbeit ist. Überprüft wird der Einfluss des ESCS auf die Leseleistung (Hypothese 1) und das Leseselbstkonzept (Hypothese 2), die Indikatoren für den Schulerfolg darstellen. Anzumerken sei, dass auf das Angeben der Fit-Indices für diese Vorprüfungen verzichtet wird, dort, wo sie als akzeptabel oder besser zu bewerten sind. Erst für das Gesamtmodell werden diese berichtet.

**Tabelle 24:** STDYX-standardisierte Regressionsparameter  $\beta$ , p-Wert und nicht aufgeklärte Varianz  $R^2$  für den Einfluss des ESCS auf die Leseleistung und das Leseselbstkonzept

Prädiktor	Abhängiges Konstrukt	$\beta$	p	$R^2$
ESCS	Leseleistung	0,410	< 0,001	0,832
	Leseselbstkonzept	0,225	< 0,001	0,947
HLE	Leseleistung	0,295	< 0,001	0,913
	Leseselbstkonzept	0,238	< 0,001	0,943

Es zeigt sich erwartungsgetreu, dass der Economic, Social and Cultural Status (ESCS) einen Einfluss mittlerer Stärke ( $\beta = 0,410$ ) auf die Leseleistung und einen geringeren, dennoch bestehenden Einfluss auf das Leseselbstkonzept ( $\beta = 0,225$ ) aufweist. Für das Home Literacy Environment (HLE) zeigt sich ein ähnlicher Befund: Auch hier ist es die Leseleistung, die stärker ( $\beta = 0,295$ ) als das Leseselbstkonzept ( $\beta = 0,238$ ) von dem häuslichen Anregungsgehalt bzw. der Lesesozialisation abhängt. Allerdings vermag das HLE deutlich weniger Varianz in beiden Indikatoren des Schulerfolgs aufzuklären als der ESCS: So lässt sich 16,8% der Varianz in der Leseleistung der Viertklässlerinnen und Viertklässler durch die unterschiedlichen ESCS' der Elternhäuser erklären, für das HLE liegt dieser Wert bei 8,7%. Die Varianzaufklärungen für das Selbstkonzept liegen für beide Indikatoren deutlich darunter.

Da sich für alle vier theoretisch angenommenen Zusammenhänge signifikante und relevante Regressionsparameter finden lassen, werden diese Pfade im Gesamtmodell entsprechend mitmodelliert.

### 12.4.3 Soziale Beziehungen und Schulerfolg

Folgend dargestellt sind – analog zu den bisherigen Analysen – die Regressions- und Korrelationsparameter für den Einfluss der Schüler-Lehrer-Beziehung auf und den Zusammenhang

der Schüler-Schüler-Beziehung mit der Leseleistung und dem Leseselbstkonzept der Schülerinnen und Schüler (vgl. Tabelle 25 und Tabelle 26).

**Tabelle 25:** STDYX-standardisierte Regressionsparameter  $\beta$ , p-Wert und nicht aufgeklärte Varianz  $R^2$  für den Einfluss der Schüler-Lehrer-Beziehung auf die Leseleistung und das Leseselbstkonzept und der Schüler-Schüler-Beziehung auf die Leseleistung

Prädiktor	Abhängiges Konstrukt	$\beta$	p	$R^2$
Schüler-Lehrer-Beziehung	Leseleistung	0,097	0,019	0,991
	Leseselbstkonzept	0,152	< 0,001	0,977
Schüler-Schüler-Beziehung	Leseleistung	0,069	< 0,001	0,995

**Tabelle 26:** STDYX-standardisierter Korrelationsparameter r, p-Wert und nicht aufgeklärte Varianz  $R^2$  für den Zusammenhang zwischen Schüler-Schüler-Beziehung und Leseselbstkonzept

Korrelation	r	p	$R^2$
Schüler-Schüler-Beziehung mit Leseselbstkonzept	0,207	< 0,001	0,957

Verglichen mit den Einflüssen des ESCS und des HLE auf den Schulerfolg zeigen sich für die sozialen Facetten Schüler-Schüler-Beziehung und Schüler-Lehrer-Beziehung deutlich geringere, wenngleich plausible Zusammenhänge insbesondere mit dem Leseselbstkonzept ( $r = 0,207$  bzw.  $\beta = 0,152$ ). Mit einem p-Wert unter 0,001 bzw. 0,019 handelt es sich bei dem Einfluss der sozialen Dimensionen auf die *Leseleistung* zwar um statistisch signifikant von Null verschiedene Regressions- bzw. Korrelationsparameter, allerdings sind diese mit  $r = 0,069$  und  $\beta = 0,097$  als wenig relevant anzusehen. Um das theoretisch abgeleitete Modell auch empirisch abzubilden, werden sie dennoch in das nationale Gesamtmodell aufgenommen, wenngleich hier schon anzunehmen ist, dass ihr Einfluss marginal ausfallen wird.

#### 12.4.4 Klassenzusammensetzung und individuelle Förderung

Wie in Kapitel 10 beschrieben, soll im Gesamtmodell die Klassenzusammensetzung (bestehend aus dem ESCS und dem Sprachgebrauch, für jede Klasse gemittelt) und ihr Einfluss auf die Häufigkeit der individuellen Förderung modelliert werden. Die folgende Tabelle 27 beschreibt die Einflüsse der einzelnen Klassenkonstrukte auf die Skala ‚Individuelle Förderung‘.

**Tabelle 27:** STDYX-standardisierte Regressionsparameter  $\beta$ , p-Wert und nicht aufgeklärte Varianz  $R^2$  für den Einfluss der Klassenaggregate ESCS und häuslicher Sprachgebrauch Deutsch auf die Häufigkeit der individuellen Förderung

Prädiktor	Abhängiges Konstrukt	$\beta$	p	$R^2$
ESCS (Klassenaggregat)	Individuelle Förderung (Skala)	-0,038	0,642	0,999
Sprache (Klassenaggregat)	Individuelle Förderung (Skala)	-0,130	0,050	0,983

Den p-Werten zu entnehmen ist, dass sich lediglich einer der beiden Klassenaggregat-Effekte signifikant von Null unterscheidet. Dabei handelt es sich um den Einfluss des klassendurchschnittlichen häuslichen Gebrauchs der deutschen Sprache auf die individuelle Förderung (Skala). Mit  $\beta = -0,130$  liegt dieser im negativen Bereich, so dass in Klassen, in denen im Mittel wenige Schülerinnen und Schüler zu Hause Deutsch sprechen, tendenziell häufiger individuell gefördert wird als dies in Klassen mit größeren Anteilen an Kindern, die zu Hause häufig oder immer Deutsch sprechen der Fall ist. Beides wird in den simultanen Analysen des Gesamtmodells zunächst berücksichtigt; allerdings deuten diese Voranalysen bereits stark darauf hin, dass zumindest das ESCS-Klassenaggregat in keinem Zusammenhang mit der Häufigkeit individueller Förderung zu stehen scheint.

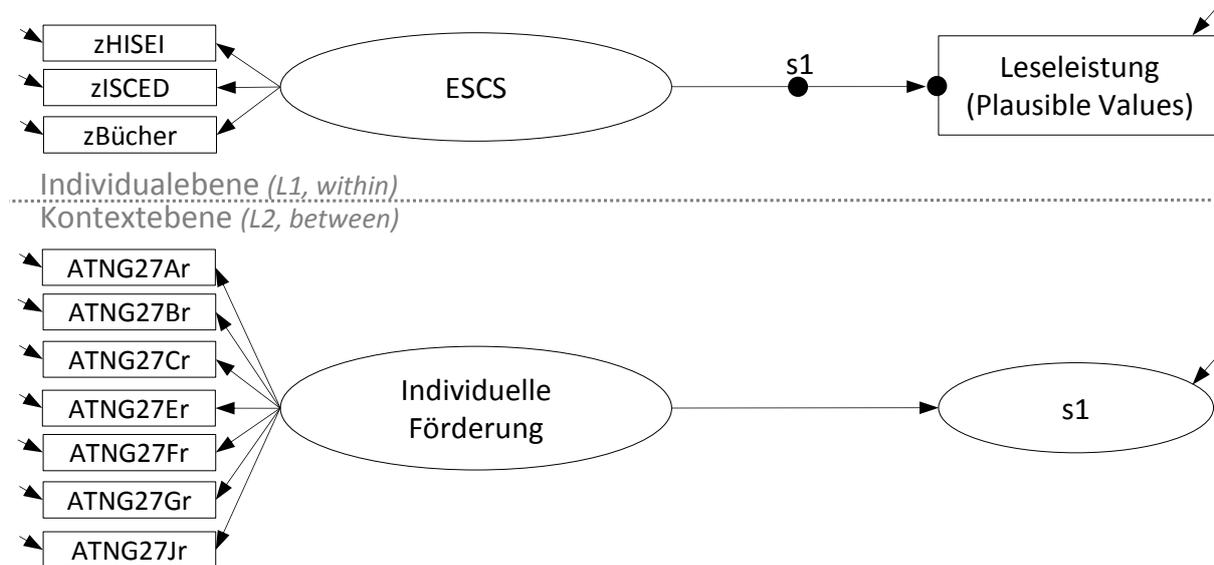
#### 12.4.5 Individuelle Förderung und Herkunft

Ein Vorteil von Mehrebenenmodellen sind die sogenannten Cross-Level-Interaktionseffekte: Sie ermöglichen die Analyse nicht nur von Effekten auf Individual- oder Kontextebene(n), sondern auch, Beziehungen *zwischen* den Ebenen zu modellieren. Dies findet hier Anwendung, wenn die Angaben der Lehrpersonen zu der Häufigkeit individueller Förderung (L2) in Beziehung gesetzt werden mit Zusammenhängen zwischen Schüler- und Elternmerkmalen auf Individualebene (L1). Relevant ist dies entsprechend für den (hypothetischen) Einfluss der individuellen Förderung auf den Herkunftseffekt. Neben diesem Cross-Level-Interaktionseffekt soll auch der Einfluss der individuellen Förderung auf die Schüler-Schüler-, die Schüler-Lehrer-Beziehung sowie auf beide Merkmale des Schulerfolgs – analog zu dem Vorgehen in den bisherigen Unterkapiteln – zunächst in Einzelanalysen betrachtet werden. Daraus abgeleitet werden können dann Hinweise dahingehend, welche Pfade im Anschluss an die simultanen Analysen des Gesamtmodells möglicherweise ausgeschlossen werden sollten, da sie sich in diesen Voranalysen bereits als nicht signifikant erwiesen.

Bei dem Einfluss der individuellen Förderung auf den Herkunftseffekt handelt es sich um das zentrale Element der vorliegenden Arbeit. Dabei stellt sich die Frage, ob Individualisierung

eine Möglichkeit darstellt, den in Deutschland besonders starken Zusammenhang zwischen Herkunft und Schulerfolg zu entkoppeln, mindestens jedoch abzuschwächen. Herangezogen werden hierfür die Slopes, also die Regressionssteigungen, die sich aus (hier: vier) miteinander in Beziehung gesetzten Koordinaten ergeben. Überprüft wird, ob eine unabhängige Variable, hier die individuelle Förderung, dieses Zusammenhangsmaß, in diesem Fall der Zusammenhang zwischen den Herkunftsdimensionen ESCS und HLE und dem Schulerfolg, beeinflussen kann (Moderator- bzw. Cross-Level-Interaktionseffekt). Die folgende Abbildung 20 verdeutlicht dies an einem Beispiel: Auf Individualebene ist der Einfluss des ESCS auf die Leseleistung dargestellt (Herkunftseffekt). Der Slope ( $s_1$ ) wird mit einem Kreis markiert und gibt an, dass der Regressionskoeffizient (von ESCS auf Leseleistung) auf Kontextebene L2 variiert. Auf L2 wird dieser random Slope latent aufgenommen, so dass der Einfluss der individuellen Förderung auf diesen Slope modelliert werden kann (Cross-Level-Interaktionseffekt).

**Abbildung 20:** Beispielhafte Darstellung des Einflusses der individuellen Förderung auf den Slope  $s_1$  als Regressionskonstante zwischen ESCS und Leseleistung



Eine Voraussetzung für die Überprüfung von Cross-Level-Interaktionseffekten bildet entsprechend die Annahme, dass die Slopes zwischen den Klassen variieren – nur dann kann analysiert werden, ob die individuelle Förderung einen Teil dieser Varianz erklären kann, also (theoriekonform) für einen schwächeren Zusammenhang von Herkunft und Schulerfolg in manchen Klassen verantwortlich ist. Nach umfangreichen Betrachtungen kann diese Grundvoraussetzung nicht angenommen werden: Die Slopes variieren in allen Fällen nicht systematisch zwischen den Klassen. Dies soll im Folgenden eingehender erläutert werden.

Da Mplus Slopes, die gering bis gar nicht zwischen den Clustern variieren, nicht für latente Modelle schätzen kann, wurde dies in Einzelitemanalysen der den Konstrukten zugehörigen Variablen berechnet. Die folgende Tabelle 28 gibt entsprechend einen Überblick über den Einfluss der ESCS- und HLE-Items auf beide Facetten des Schulerfolgs. Angegeben sind dabei die Höhe des Zusammenhangs zwischen Herkunftstitem und Schulerfolgskonstrukt ( $\gamma_{10}$ ) und die Varianz zwischen den Klassen ( $\tau_{11}$ ) mit ihrem zugehörigen Signifikanzniveau (für a priori-Festsetzung und Interpretation der Signifikanzniveaus s. Kapitel 12.2).

**Tabelle 28:** Varianzen in den Slopes zwischen den Einzelitems der Konstrukte ESCS und HLE und dem Schulerfolg (Leseleistung und Leseselbstkonzept)

Konstrukt	Item (Prädiktor $X_{ij}$ )	L2-Varianz im Slope ( $Y_{ij}$ = Leseleistung)	L2-Varianz im Slope ( $Y_{ij}$ = Leseselbstkonzept)
ESCS	Bücher	$\tau_{11} = 0,001$ n.s. (bei $\gamma_{10} = 0,263$ )	$\tau_{11} = 0,001$ n.s. (bei $\gamma_{10} = 0,095$ )
	HISEI	$\tau_{11} = 0,005$ n.s. (bei $\gamma_{10} = 0,234$ )	$\tau_{11} = 0,001$ n.s. (bei $\gamma_{10} = 0,087$ )
	ISCED	$\tau_{11} = 0,001$ n.s. (bei $\gamma_{10} = 0,296$ )	$\tau_{11} = 0,001$ n.s. (bei $\gamma_{10} = 0,101$ )
HLE	ASBH02Ar	$\tau_{11} = 0,027$ n.s. (bei $\gamma_{10} = 0,435$ )	$\tau_{11} = 0,001$ n.s. (bei $\gamma_{10} = 0,147$ )
	ASBH02Br	$\tau_{11} = 0,015$ n.s. (bei $\gamma_{10} = 0,158$ )	$\tau_{11} = 0,002$ n.s. (bei $\gamma_{10} = 0,081$ )
	ASBH02Er	$\tau_{11} = 0,006$ n.s. (bei $\gamma_{10} = 0,301$ )	$\tau_{11} = 0,002^{***}$ (bei $\gamma_{10} = 0,069$ )
	ASBH02Fr	$\tau_{11} = 0,020$ n.s. (bei $\gamma_{10} = 0,190$ )	$\tau_{11} = 0,002^{***}$ (bei $\gamma_{10} = 0,090$ )

Deutlich wird, wie oben bereits dargestellt, dass alle Zusammenhänge nicht systematisch zwischen den Klassen variieren<sup>73</sup>. Dies bleibt unberührt von der Höhe des Zusammenhangs, wie er in Kapitel 12.4.2 für die beiden latenten Konstrukte dargestellt wurde. Schlussgefolgert werden muss hieraus, dass die insgesamt vier hier modellierten Herkunftseffekte so stabil sind, dass sie sich zwischen den Klassen nicht unterscheiden. Im Umkehrschluss bedeutet dies, dass sich keine Varianz findet, die durch individuelle Förderung oder ein anderes, sich zwischen Klassen unterscheidendes Konstrukt aufgeklärt werden kann. Ausführlicher wird hierauf in der Diskussion eingegangen.

Für die Modellierung des Gesamtmodells bedeutet dies, dass die Slopes ‚fixiert‘ werden, d.h. ihre Stetigkeit über die Klassen hinweg in das Modell aufgenommen wird. Nicht berücksichtigt wird hingegen der Einfluss der individuellen Förderung auf die vier Slopes der Herkunftseffekte, da hierfür wie gezeigt die Grundvoraussetzung der Varianz zwischen den Klassen

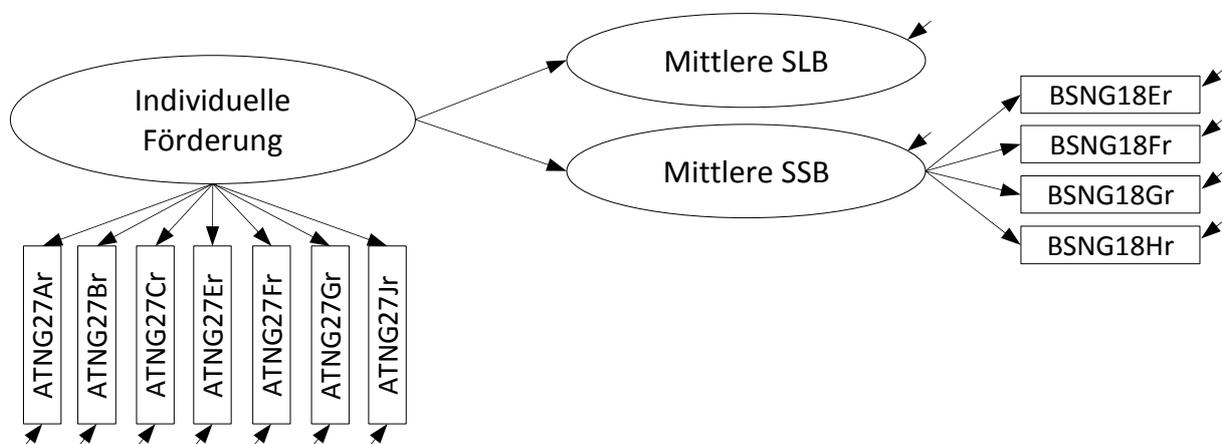
<sup>73</sup> Nur die Between-Class-Varianz für den Zusammenhang zwischen den HLE-Items ASBH02Er und ASBH02Fr und dem Leseselbstkonzept wird signifikant. Da diese mit  $\tau_{11} = 0,002$  jedoch extrem nahe Null liegt, kann auch hier nicht von einer Varianz zwischen den Klassen ausgegangen werden.

verletzt ist, dieser Effekt also nicht vorhanden sein kann. Dementsprechend müssen die Hypothesen 16, 17, 18 und 19 (Einfluss der individuellen Förderung auf den jeweiligen Herkunftseffekt zwischen ESCS und Leseleistung bzw. Selbstkonzept sowie HLE auf Leseleistung bzw. Selbstkonzept) an dieser Stelle verworfen werden.

#### 12.4.6 Individuelle Förderung und soziale Beziehungen

Weiterhin geprüft werden soll in Vorabanalysen, ob sich der Einfluss der individuellen Förderung auf die Schüler-Lehrer- sowie die Schüler-Schüler-Beziehung abbilden lässt. Dabei unterscheidet sich das Vorgehen von dem im vorherigen Kapitel beschriebenen geringfügig: Da es sich bei der angenommenen Wirkung nicht um einen Einfluss auf einen Zusammenhang (Moderatoreffekt), sondern um eine einfache Regression handelt, wird das (klassenbezogene) Mittel der beiden sozialen Konstrukte latent auf L2 modelliert, so dass ihre mögliche Beeinflussung durch das L2-Konstrukt der individuellen Förderung überprüft werden kann. Die folgende Abbildung 21 verdeutlicht dies für die Schüler-Lehrer- sowie Schüler-Schüler-Beziehung und dem Konstrukt der individuellen Förderung.

**Abbildung 21:** Darstellung des Einflusses der individuellen Förderung auf die klassenmittlere Schüler-Lehrer- (SLB) sowie Schüler-Schüler-Beziehung (SSB) auf Kontextebene



Die für dieses Modell berechneten Regressionsparameter finden sich in der nachstehenden Tabelle 29.

**Tabelle 29:** STDYX-standardisierte Regressionsparameter  $\beta$ , p-Wert und nicht aufgeklärte Varianz  $R^2$  für den Einfluss der individuellen Förderung auf die klassenmittlere Wahrnehmung der Schüler-Schüler- und Schüler-Lehrer-Beziehung

Prädiktor	Abhängiges Konstrukt	$\beta$	p	$R^2$
Individuelle Förderung (Skala)	Schüler-Schüler-Beziehung	-0,062	0,702	0,996
	Schüler-Lehrer-Beziehung	-0,015	0,906	1,000

Erkennbar ist, dass sowohl das Einzelitem als auch die Skala zur individuellen Förderung keinen Einfluss auf die Schüler-Schüler- sowie Schüler-Lehrer-Beziehung haben; alle p-Werte der Regressionsparameter liegen deutlich über dem 5%-Niveau. Dementsprechend wird ebenfalls keine Varianz der abhängigen Konstrukte durch die individuelle Förderung erklärt. Dies soll noch einmal im Gesamtmodell simultan zu allen anderen Beziehungen getestet werden. Sollte sich dies auch dort bestätigen, müsste zudem die Annahme, dass sich individuelle Förderung möglicherweise vermittelt durch die Schüler-Schüler- und Schüler-Lehrer-Beziehung indirekt auf die Leseleistung und das Leseselbstkonzept auswirkt, verworfen werden, da hierfür die Haupteffekte (Individuelle Förderung auf SLB bzw. SSB) existent sein müssen.

#### 12.4.7 Individuelle Förderung und Schulerfolg

Bevor folgend das Gesamtmodell aufgestellt werden soll, bleiben noch zwei weitere Beziehungen, die es zu prüfen gilt. Wie in Kapitel 6 angenommen, beeinflusst die individuelle Förderung den Schulerfolg, unter welchem sowohl die Leseleistung als auch das Leseselbstkonzept der Schülerinnen und Schüler verstanden wird (vgl. Hypothesen 14 und 15). Auch hier werden beide Schulerfolgsindikatoren auf L2 modelliert, so dass ihre mögliche Beeinflussung durch die Angaben der Lehrperson überprüft werden können. Tabelle 30 gibt Aufschluss über die entsprechenden Regressionsparameter.

**Tabelle 30:** STDYX-standardisierte Regressionsparameter  $\beta$ , p-Wert und nicht aufgeklärte Varianz  $R^2$  für den Einfluss der individuellen Förderung auf die mittlere Leseleistung und das mittlere Leseselbstkonzept

Prädiktor	Abhängiges Konstrukt	$\beta$	p	$R^2$
Individuelle Förderung (Skala)	Leseleistung	-0,163	0,033	0,973
	Leseselbstkonzept	-0,101	0,530	0,990

Dabei zeigt lediglich der Einfluss der individuellen Förderung auf die Leseleistung einen auf dem 5%-Niveau signifikanten Regressionsparameter von  $\beta = -0,163$  ( $p = 0,033$ ); ihr Einfluss auf das mittlere Leseselbstkonzept unterscheidet sich nicht bedeutend von Null ( $\beta = -0,101$  bei  $p = 0,530$ ). Dabei vermag die individuelle Förderung jedoch lediglich 2,7% der Unterschiede in der Leseleistung zwischen den Klassen aufzuklären, muss also als wenig relevanter Prädiktor eingestuft werden.

Die in diesem und in den vorherigen Kapiteln nachgewiesenen Zusammenhänge sollen folgend in einem Gesamtmodell simultan auf ihre Gültigkeit hin analysiert werden, wobei die

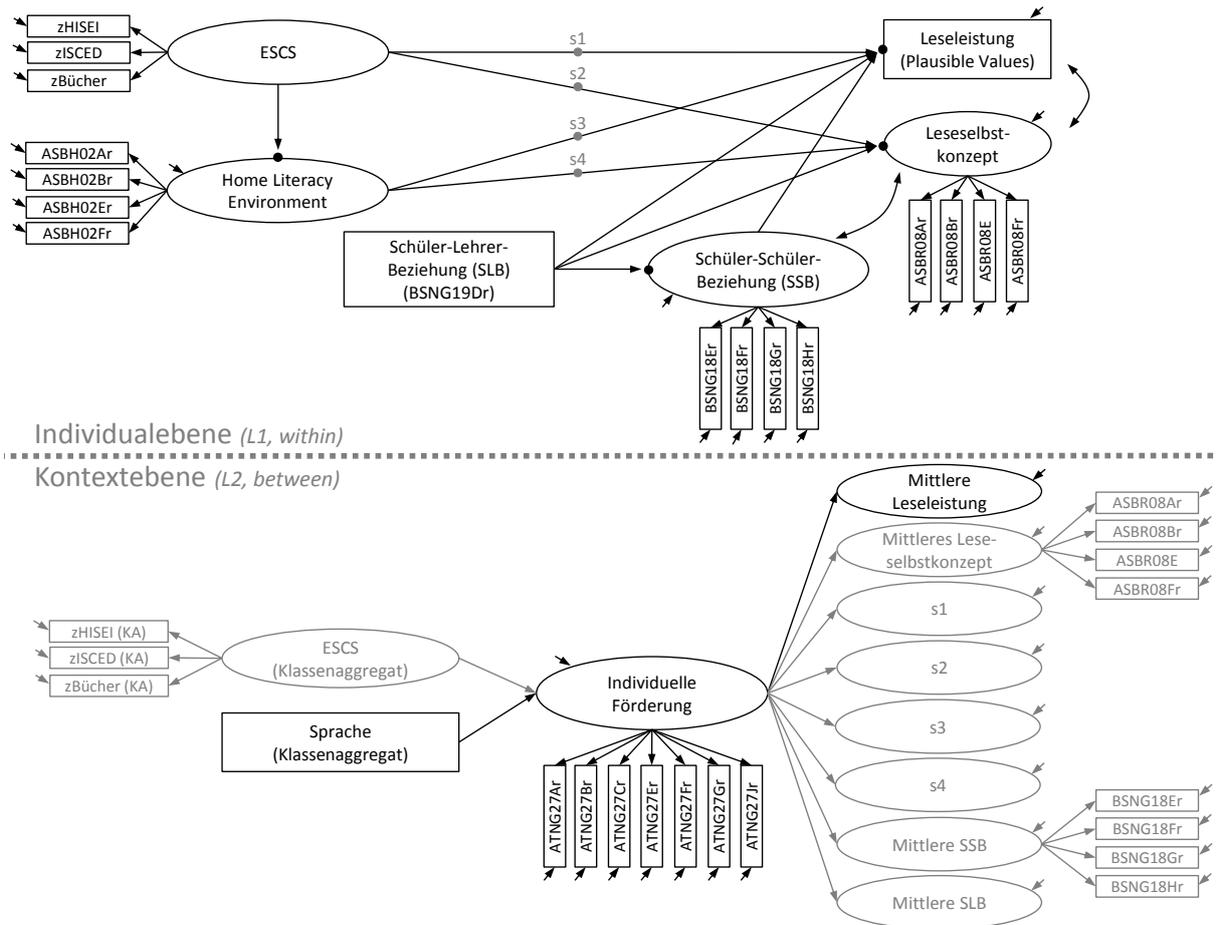
sich in den bivariaten Voranalysen als nicht signifikant erwiesenen Effekte starke Hinweise für eine mögliche Modifikation des Gesamtmodells liefern.

### 12.5 Prüfung des Gesamtmodells

In den vorherigen Unterkapiteln wurden die theoretisch angenommenen Zusammenhänge einzeln auf ihre Gültigkeit hin geprüft. Dabei wurde festgestellt, dass sich nur einer der Effekte der Klassenzusammensetzung auf die Häufigkeit der individuellen Förderung empirisch (bivariat) bestätigen lässt: So zeigt lediglich der Einfluss des klassendurchschnittlichen familiären Gebrauchs der deutschen Sprache einen negativen Einfluss auf die Skala der individuellen Förderung (je weniger die Schülerinnen und Schüler einer Klasse Deutsch zu Hause sprechen, desto eher wird individuell gefördert). Ähnlich verhält es sich mit dem Einfluss der individuellen Förderung auf das Leseselbstkonzept, die Schüler-Lehrer- und die Schüler-Schüler-Beziehung. Diese einzeln betrachteten Regressionsparameter weichen nicht signifikant von Null ab. Lediglich für den Einfluss der individuellen Förderung auf die Leseleistung lässt sich ein signifikanter, wenn auch geringer negativer Zusammenhang finden. Besonders gravierend scheinen die Befunde bezüglich des Herkunftseffekts: Dieser variiert nicht zwischen den Klassen, so dass er durch Maßnahmen individueller Förderung nicht beeinflusst werden kann. Alle Annahmen, die über die Zusammenhänge auf Individualebene getroffen wurden, konnten hingegen in den Einzelanalysen bestätigt werden.

Auf Grundlage dieser bivariaten Voranalysen kann folgend ein Gesamtmodell aufgestellt werden. Aus ihnen kann zudem die Adaption des Modells theoriegeleitet und schrittweise erfolgen: Das Aufstellen des Gesamtmodells *vor* Prüfung der Einzelbeziehungen führt in der Regel zu einem weniger theoriegeleiteten ‚Austesten‘, durch welches Weglassen welcher Pfade der Modellfit verbessert werden kann. Die in dem vorherigen Kapitel 12.4 dargestellten Zusammenhänge, von denen sich einige in bivariaten Analysen nicht nachweisen ließen, liefern jedoch bereits Hinweise darauf, wie sich diese in der simultanen Testung im Mehrebenenstrukturgleichungsmodell abbilden könnten. Abbildung 22 gibt noch einmal einen Überblick über die Pfade, die in den Einzelanalysen nachgewiesen (schwarz dargestellte Pfade) bzw. nicht nachgewiesen (grau dargestellte Pfade) werden konnten.

**Abbildung 22: Theoriegeleitetes Gesamtmodell**



Anmerkung: Pfade, die sich in Voranalysen bereits als nicht verifizierbar herausgestellt haben, sind grau dargestellt, ebenso wie die zugehörigen Konstrukte. Auf diese ausführliche Darstellung, in welcher auch die Items der latenten Konstrukte abgebildet sind, soll im Folgenden verzichtet werden.

In einem Zwischenschritt wurde überprüft, ob für die Pfade, die sich in den bivariaten Voranalysen als nicht signifikant erwiesen, ähnliches in der simultanen Testung im Gesamtmodell festgestellt werden konnte. Dies lässt sich für fast alle Zusammenhänge bestätigen (s. Tabelle 37 bis Tabelle 39 im Anhang 5.1) – lediglich das für die Modellierung der Schüler-Lehrer-Beziehung herangezogene Einzelitem (BSNG19Dr) konnte nicht in das Gesamtmodell aufgenommen werden, da dies zu erheblichen Schätzproblemen in Mplus führte. Möglicherweise liegt dies darin begründet, dass das Item mit der individuellen Förderung auf Kontextebene (Hypothese 20) in keinem Zusammenhang<sup>74</sup> steht und entsprechend keine Varianz in der Schüler-Lehrer-Beziehung aufgeklärt werden kann ( $\beta = -0,015$  bei  $p = 0,906$ ,  $R^2 = 1,000$ , vgl. Kapitel 12.4.6). Da dies so deutlich ausfällt, wird davon ausgegangen, dass sich auch in der simultanen Testung im Gesamtmodell kein gegenläufiger Effekt nachweisen

<sup>74</sup> Dies ist jedoch auch für weitere Zusammenhänge auf Kontextebene der Fall, wenn auch nicht in diesem Ausmaß – möglicherweise bestehen hier also noch andere Ursachen, die zu Schätzproblemen in Mplus führen und nicht identifiziert werden konnten.

lassen würde, so dass an dieser Stelle die Hypothese 20, nach welcher die Häufigkeit individueller Förderung die klassendurchschnittliche Schüler-Lehrer-Beziehung positiv beeinflusst, verworfen werden muss. Aus der im Anhang (vgl. Anhang 5.1, Tabelle 37 bis Tabelle 39) dargestellten Zwischenprüfung des Gesamtmodells mit allen Pfaden können alle übrigen Befunde zu nicht signifikanten Zusammenhängen der Vorabanalysen bestätigt werden. Diese beziehen sich auf den Einfluss des auf Klassenebene aggregierten ESCS auf die Häufigkeit, mit der Lehrerinnen und Lehrer im Unterricht individualisieren (Hypothese 12) sowie auf den Einfluss der Individualisierung auf das klassendurchschnittliche Selbstkonzept im Lesen (Hypothese 15) und die klassendurchschnittliche Schüler-Schüler-Beziehung (Hypothese 21). Da bezüglich der Beeinflussung der klassendurchschnittlichen Einschätzung der sozialen Facetten (Schüler-Schüler- und Schüler-Lehrer-Beziehung) durch Maßnahmen der Individualisierung keine signifikanten Pfade nachgewiesen werden konnten, können sich auch keine darüber indirekt vermittelten Effekte des adaptiven Unterrichtens auf die Leseleistung und das Leseselbstkonzept ergeben. Somit müssen die zugehörigen Hypothesen 22 bis 25 ebenfalls verworfen werden. Darüber hinaus erweist sich jedoch auch der Einfluss des auf Klassenebene aggregierten häuslichen Gebrauchs der deutschen Sprache auf die Häufigkeit der individuellen Förderung (Hypothese 13) als nicht von Null verschieden, der in den bivariaten Analysen noch mit  $\beta = -0,130$  bei  $p = 0,050$  knapp signifikant ausfiel.

Zusammenfassend können hier also neun Hypothesen (12, 13, 15 und 20 bis 25) nicht bestätigt werden. Diese nicht signifikanten Pfade werden in der folgenden Modellierung nicht weiter berücksichtigt, da sie den Modellfit deutlich verschlechtern<sup>75</sup>.

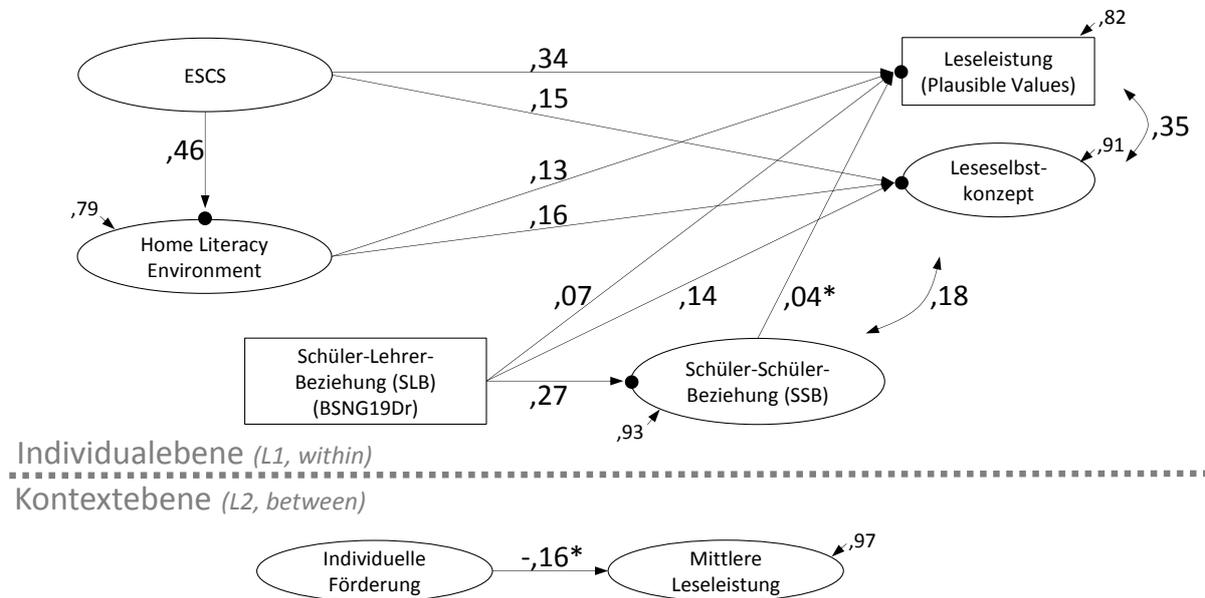
Aus der nachstehenden Abbildung 23 können alle Pfadkoeffizienten und Residualvarianzen abgelesen werden. Nicht aufgeführt sind – wie dies häufig in anderen Darstellungen von Strukturgleichungsmodellen der Fall ist – Asteriske (\*), die eine Irrtumswahrscheinlichkeit von 0,1% (üblicherweise, \*\*\*) angeben; zudem werden alle Parameter der Übersicht halber auf zwei Nachkommastellen gerundet. Ein Überblick über die Faktorladungen der Messmo-

---

<sup>75</sup> Dies ist plausibel: Fitindizes geben an, wie gut theoretische Annahmen auf die Daten passen. Neun theoretische Annahmen konnten anhand der IGLU-2011-Daten nicht nachgewiesen werden; sie ‚passen also nicht auf die Daten‘ – ihre Beibehaltung führt zwangsweise zu einer Verschlechterung der Modellgüte. Dies spiegelt sich in den Fitindizes wider, sie liegen für das Gesamtmodell unter Berücksichtigung aller nicht signifikanten Pfade im nicht akzeptablen Bereich (RMSEA = 0,039, CFI = 0,894, TLI = 0,873, SRMR<sub>L1</sub> = 0,021 und SRMR<sub>L2</sub> = 0,153).

delle und alle Koeffizienten des Strukturmodells findet sich in Tabelle 40 bis Tabelle 43 im Anhang (s. Abschnitte 5.2 und 5.3).

**Abbildung 23:** STDYX-standardisierte Pfadkoeffizienten und nicht erklärte Varianz im Gesamtmodell



Anmerkung: \* = signifikant auf einem 5%-Niveau; keine Angaben = signifikant auf einem 0,1%-Niveau

Die Fitwerte für dieses Modell liegen bei RMSEA = 0,032, CFI = 0,948, TLI = 0,937, SRMR<sub>L1</sub> = 0,026 und SRMR<sub>L2</sub> = 0,118. Abgesehen von dem SRMR für die Kontextebene kann die Modellpassung als gut bewertet werden. Bei der Betrachtung des L2-Modells (individuelle Förderung, mittlere Leseleistung) wird deutlich, dass hier der Regressionskoeffizient lediglich auf dem 5%-Niveau signifikant wird. Für den Einfluss der Häufigkeit der individuellen Förderung auf die mittlere Leseleistung der Schülerinnen und Schüler ergibt sich ein p-Wert von 0,030 (s. Tabelle 41 im Anhang 5.3) bei einem Regressionskoeffizienten von  $\beta = -0,163$ . Durch diesen gering ausgefallenen Zusammenhangskoeffizienten kann nur wenig Varianz in der Leseleistung (2,7%) erklärt werden, was zu einem schlechteren SRMR<sub>L2</sub>-Wert als Indikator für die Modellpassung auf L2 führt.

Durch das Modell auf Individualebene werden 17,6% der Varianz in der Leseleistung und 9% der Varianz im Leseselbstkonzept erklärt<sup>76</sup> (vgl. Tabelle 43 im Anhang 5.3), wobei zwischen den beiden Facetten des Schulerfolgs ein Zusammenhang in Höhe von  $r = 0,35$  besteht (was Hypothese 11 vorläufig bestätigt). Insgesamt ähneln alle Regressions- und Korrelations-

<sup>76</sup> Berücksichtigt werden muss in diesem Zusammenhang, dass die Gesamtvarianz der Leseleistung auf Individualebene 76,8% beträgt, von denen mit diesem Modell 17,6% aufgeklärt werden können (s. Kapitel 12.4.1, die Varianzaufklärung des Leseselbstkonzepts ist analog zu interpretieren).

parameter denen der Vorabprüfungen, wobei die größten Unterschiede ( $\Delta_{\max} = 0,167$ ) hinsichtlich der vier Herkunftseffekte (ESCS und HLE auf je Leseleistung und Leseselbstkonzept) bestehen. Dies verdeutlicht, dass nur durch die simultane Testung von Zusammenhängen, wie dies mithilfe von Strukturgleichungsmodellen gelingt, konfundierten Effekten nachgegangen werden kann.

Insgesamt zeigt sich ein erwartungsgetreuer Effekt des sozioökonomischen Status (ESCS) auf die Leseleistung in Höhe von  $\beta = 0,34$ , der für das Leseselbstkonzept niedriger ausfällt ( $\beta = 0,15$ ). Bei dem Home Literacy Environment (HLE) verläuft dieser Effekt gegenläufig: So steht dies in einem minimal stärkeren Zusammenhang mit dem Leseselbstkonzept ( $\beta = 0,16$ ) als mit der Leseleistung ( $\beta = 0,13$ ). Die Abhängigkeit der häuslichen Lesesozialisation von der sozioökonomischen Stellung der Familie fällt mit  $\beta = 0,46$  theoriekonform stark aus, wobei dadurch über 20% der Varianz im HLE aufgeklärt werden können. Insgesamt können hierdurch die Hypothesen 1 bis 5 vorläufig angenommen werden.

Für die sozialen Dimensionen ergeben sich erwartungsgetreu geringere Werte bezüglich ihres Einflusses auf die kognitive Facette des Schulerfolgs. So beeinflusst die Beziehung zwischen Lehrenden und Lernenden die Leseleistung mit  $\beta = 0,07$  nur marginal, bei der Schüler-Schüler-Beziehung liegt dieser Effekt bei  $\beta = 0,04$  (mit einer höheren Irrtumswahrscheinlichkeit von 3,9%). Etwas stärker zeichnen sich die Zusammenhänge mit dem Leseselbstkonzept ab (Schüler-Lehrer-Beziehung:  $\beta = 0,14$ , Schüler-Schüler-Beziehung:  $r = 0,18$ ), was gerade bezüglich der Bedeutung der sozialen Bezugsgruppen für das Selbstkonzept plausibel erscheint. Der Einfluss der durch die Schülerinnen und Schüler wahrgenommenen Beziehung zu ihrer Lehrkraft auf die Einschätzung ihrer sozialen Integration in den Klassenkontext liegt bei  $\beta = 0,27$ , wodurch etwa 7% der Varianz in der Schüler-Schüler-Beziehung erklärt werden können. Hierdurch werden die Hypothesen 6 bis 10 vorläufig bestätigt.

Für den Zusammenhang auf Kontextebene zeigt sich ein mit den Voranalysen identischer Regressionsparameter für den Einfluss der Häufigkeit, mit der im Unterricht individualisiert wird, auf die klassenmittlere Leseleistung ( $\beta = -0,16$ , mit einer höheren Irrtumswahrscheinlichkeit von 3%, Varianzaufklärung in der mittleren Leseleistung: 2,7%). Dies fällt entsprechend nicht erwartungsgetreu aus, so dass Hypothese 14, die einen positiven Zusammenhang postulierte, zurückgewiesen wird. Allerdings bestätigt dies die Befunde von Gruehn (1995 und 2000) und Seidel und Shavelson (2007), nach welchen diese möglicher-

weise in umgekehrter Kausalrichtung zu interpretieren sind. Auf diesen Effekt wird in der Diskussion (Kapitel 13.2) nochmals gesondert eingegangen.

### **13. Zusammenfassung und Diskussion der Ergebnisse**

In den folgenden Unterkapiteln sollen die zentralen Befunde der vorliegenden Arbeit zunächst zusammengefasst (Kapitel 13.1) und anschließend diskutiert werden (Kapitel 13.2). Dem schließt sich eine Darstellung der Grenzen der empirischen Analysen auf Grundlage der Daten der Internationalen Grundschul-Lese-Untersuchung an (Kapitel 13.3). Daraus abgeleitete Implikationen für anschließende Forschungsarbeiten, die schulische Praxis und den bildungspolitischen Diskurs um individuelle Förderung im Unterricht sind Gegenstand des abschließenden Kapitels 14.

#### **13.1 Zusammenfassung der Ergebnisse**

In den vorangegangenen Abschnitten wurden alle theoretischen Zusammenhänge, wie sie in Kapitel 11 zusammengefasst wurden, auf ihren empirischen Gehalt hin geprüft. Dabei ergab sich vor allem bezüglich des Kernelements der vorliegenden Arbeit ein überraschender Befund: Der Herkunftseffekt, wie er zwischen sozioökonomischem Status und häuslicher Lesesozialisation auf der einen und der Leseleistung und dem Leseselbstkonzept als Indikatoren des Schulerfolgs auf der anderen Seite modelliert wurde, zeigt keine Varianz zwischen den Klassen (bzw. Schulen). Da Moderatoreffekte, wie sie hier für die Wirkung der individuellen Förderung auf den Herkunftseffekt angenommen wurden, zum Ziel haben, bestehende Varianz (zwischen Individuen und/oder Klassen) zu erklären, muss hier folglich davon ausgegangen werden, dass individuelle Förderung keinen Einfluss auf den für diese Analysen spezifischen Herkunftseffekt haben kann. Mögliche Ursachen hierfür sollen im folgenden Kapitel 13.2 erörtert werden.

Für das über diesen Herkunftseffekt hinausgehende Modell zeigten sich generell erwartungsgetreue Effekte: So beeinflusst der sozioökonomische Status einer Familie die individuelle Leseleistung der Schülerinnen und Schüler ebenso wie (etwas schwächer) das Leseselbstkonzept. Von ihm abhängig ist ebenfalls die häusliche Lesesozialisation, die wiederum in stärkerem Zusammenhang mit dem Leseselbstkonzept, aber auch mit der Leseleistung der Viertklässlerinnen und Viertklässler steht. Diese beiden Facetten des Schulerfolgs werden darüber hinaus durch Merkmale der schulischen Lernumwelt der Kinder beeinflusst, wobei

die Effekte der individuell wahrgenommenen Schüler-Schüler- und Schüler-Lehrer-Beziehung insbesondere bezüglich der kognitiven Dimension (Leseleistung) vergleichsweise schwach ausfallen. Für die Beeinflussung der Häufigkeit, mit der Lehrerinnen und Lehrer ihren Unterricht individualisieren, durch Merkmale der Klassenzusammensetzung (klassendurchschnittlicher sozioökonomischer Status, durchschnittlicher Gebrauch der deutschen Sprache zu Hause) ergeben sich keine bedeutsam von Null verschiedenen Effekte. Nachgewiesen werden konnte für diese allerdings ein negativer Zusammenhang mit der klassendurchschnittlichen Leseleistung der Schülerinnen und Schüler, während ein Einfluss auf das mittlere Leseselbstkonzept sowie auf die gemittelte Schüler-Schüler- und Schüler-Lehrer-Beziehung nicht nachgewiesen werden konnte – wodurch auch die Prüfung indirekter Effekte obsolet wurde. Diese Befunde werden im Folgenden interpretiert.

### **13.2 Diskussion der Ergebnisse**

Unter Berücksichtigung der wenigen bisher vorliegenden Forschungsergebnisse, die der individuellen Förderung zumeist einen positiven Einfluss auf kognitive und affektiv-motivationale Schülercharakteristika nachweisen, sowie aktueller, vorrangig bildungspolitischer Diskussionen war es nicht nur Ziel dieser Arbeit, dieser Befundlage für die vierten Klassen in Deutschland nachzugehen, sondern darüber hinaus zu prüfen, ob individuelle Förderung dem Herkunftseffekt ausgleichend begegnen kann. Dessen Brisanz für den Schulerfolg von Kindern und Jugendlichen hatte sich insbesondere im Rahmen der groß angelegten Schulleistungsvergleichsstudien für Deutschland eindrucksvoll gezeigt: So sind es sozioökonomische Merkmale der Familie, zu denen vor allem kulturelle, ökonomische und soziale Ressourcen gezählt werden und die wiederum weitere Charakteristika wie etwa den häuslichen Anregungsgehalt und die kulturelle Sozialisation beeinflussen, denen eine starke prognostische Bedeutsamkeit für den Schulerfolg ihrer Kinder zukommt. Diese Effekte, die sich unter dem sog. Herkunftseffekt subsumieren, manifestieren sich in unterschiedlichen Eingangsvoraussetzungen zu Schulbeginn und später in einem lern- und leistungsbezogenen Vorsprung von Schülerinnen und Schülern aus höheren sozialen Schichten, der dem Leistungszuwachs eines ganzen Schuljahres entsprechen kann (Bos & Lankes et al., 2003; Wendt et al., 2012). Als Konsequenz hieraus muss angenommen werden, dass die Klassenzusammensetzungen und mit ihnen die Lernbedingungen der einzelnen Schülerinnen und Schüler gerade bezüglich solcher Herkunftsmerkmale als heterogen zu beschreiben sind, denen auch mit ‚äußeren‘ Homogenisierungsbetreibungen wie Klassenwiederholungen und

leistungsbezogenen Schulformuntergliederungen nur in geringem Ausmaß beizuwohnen ist (Lehmann et al., 1995; Artelt et al., 2001; Arnold et al., 2007; Belser & Küsel, 1976; Tietze & Roßbach, 1998). Zudem postulieren aktuellere lernpsychologische und neurodidaktische Befunde, dass jegliches Lernen ohnehin individuell und klassenbezogen entsprechend heterogen ist (Roth, 2009; Herrmann, 2010), so dass ein Unterricht, der sich möglicherweise an einem (fiktiven) Durchschnittsschüler (Weinert, 1997) orientiert und diese lern- und leistungsbezogenen Unterschiede zwischen den Kindern und Jugendlichen vernachlässigt, dazu beitragen könnte, dass Potenziale nicht ausgeschöpft und herkunftsbedingte Disparitäten – so die Annahme – deutlich stärker ins Gewicht fallen, als dies in einem an die Voraussetzungen, Neigungen und Interessen der Lernenden angepassten, individualisierten Unterricht der Fall sein könnte.

Übertragen auf die hier durchgeführten empirischen Analysen anhand der Daten der Internationalen Grundschul-Lese-Untersuchung aus 2011 zeigten sich zunächst theoriekonforme Zusammenhänge auf Individualebene: So konnte etwa nachgewiesen werden, dass das Ausmaß der häuslichen Lesesozialisation in hohem Maß bedingt ist durch die sozioökonomische Stellung einer Familie – je ausgeprägter deren kulturelles und ökonomisches Kapital, desto eher werden lesebezogene Tätigkeiten (z.B. Bücher lesen, Geschichten erzählen, sich über Gelesenes unterhalten) unternommen. Und auch die Effekte der sozioökonomischen Stellung und der häuslichen Lesesozialisation auf den Schulerfolg (Herkunftseffekt) fielen erwartungsgetreu positiv aus. Dies erscheint plausibel, da sich eine gehobene Sprachkultur, die Möglichkeit der Finanzierung von zusätzlichen Lernangeboten, ggf. auch die hohe Bildungsaspiration der Eltern, die einhergeht mit der Unterstützung und Förderung ihrer Kinder ebenso wie die Effekte einer kulturellen, hier insbesondere einer lesebezogenen Sozialisation, wie sie sozioökonomisch gut gestellte Haushalte charakterisieren, positiv auf die (hier lesebezogene) Kompetenzentwicklung der Schülerinnen und Schüler auswirkt. Dies führt gleichzeitig dazu, dass sich diese Kinder solche Leistungen eher zutrauen, als Kinder, die im häuslichen Kontext deutlich weniger Förderung und Unterstützung erfahren. Diese empirisch auch in der vorliegenden Arbeit nachgewiesene starke Abhängigkeit des Schulerfolgs von Herkunftsmerkmalen sollte zum Anlass genommen werden zu untersuchen, ob eine individuelle Förderung der Schülerinnen und Schüler diesen Zusammenhang mindern könne. Wie in Kapitel 12.4.5 dargestellt, beruht die Prüfung solcher Effekte auf der Varianz zwischen Clustern: So werden Klassen, die sich durch eine geringe

Vorhersage des Schulerfolgs durch familiäre Merkmale auszeichnen, dahingehend geprüft, ob in diesen häufiger individuell gefördert wird als in Klassen, in denen die sozioökonomische Stellung der Eltern in einem starken Zusammenhang steht mit dem Schulerfolg ihrer Kinder. Ist dies strukturelles Merkmal der Gesamtstichprobe, kann der in einigen Klassen schwächere Zusammenhang auf die Individualisierung zurückgeführt werden – was im Umkehrschluss die Hypothesen 16 bis 19 bestätigen würde, nach welchen individuelle Förderung den Herkunftseffekt moderiert. Allerdings zeigte sich für alle vier untersuchten Herkunftseffekte keine systematische Varianz zwischen den Klassen.

Unterschieden werden muss dies zunächst von dem allgemeinen Leistungsniveau (und durchschnittlichen Selbstkonzept) einer Klasse: So ist anzunehmen, dass diese zwischen den Klassen bzw. Schulen<sup>77</sup> variieren, wobei ihre schulinterne Streuung möglicherweise geringer ausfällt: Da in Deutschland regional starke sozioökonomische Unterschiede zwischen Wohngebieten vorherrschen (Dumont, Neumann, Maaz & Trautwein, 2013; Bonsen et al., 2010; Tillmann & Weishaupt, 2015), unterscheiden sich die Schülerinnen und Schüler einer Schule in einem regionalen Einzugsgebiet nur marginal in ihren herkunftsbezogenen Merkmalen (geringe Varianz innerhalb von Schulen), während Unterschiede mit Schulen anderer Wohngebiete stark ausfallen können (hohe Varianz zwischen Schulen). Entsprechend ist anzunehmen, dass an Schulen in bildungsbenachteiligten Regionen das durchschnittliche Leistungsniveau deutlich niedriger ausfällt als an Schulen in bildungsprivilegierten Einzugsgebieten. Von diesen leistungs- und selbstkonzeptbezogenen Mittelwertunterschieden zwischen Schulen kann jedoch nicht auf die Zusammenhänge zwischen Herkunft und Schulerfolg geschlossen werden: Letztere erwiesen sich in den hier vorgestellten empirischen Analysen als derart robust, dass sie über alle Klassen hinweg und damit unabhängig von sämtlichen Kontextmerkmalen den Schulerfolg von Viertklässlerinnen und Viertklässlern beeinflussen. Dieser Befund fällt nicht nur erwartungswidrig aus – er überrascht auch in seiner Robustheit, bedeutet er doch, dass der klassendurchschnittliche sozioökonomische Status und die klassendurchschnittliche häusliche Lesesozialisation sich *immer in etwa gleichem Ausmaß* auf die Leseleistung und das Leseselbstkonzept auswirken (unabhängig davon, wie hoch oder niedrig der klassendurchschnittliche sozioökonomische Status und die häusliche Lesesozialisation ausfallen). Die Brisanz dieses Befundes zeigt sich

---

<sup>77</sup> Anzumerken sei hier noch einmal, dass in IGLU meist eine Klasse pro Schule erhoben wurde, so dass gerade für den Zusammenhang mit regionalen Merkmalen eher von Schul- statt Klassenunterschieden gesprochen werden kann.

nicht nur in seiner Robustheit, sondern auch in seiner Stärke: So lassen sich etwa ein Sechstel der Unterschiede in der Lesekompetenz zwischen Viertklässlerinnen und Viertklässlern auf zwei Merkmale ihrer Familien – deren sozioökonomische Stellung und die teilweise mit ihr einhergehende Lesesozialisation – zurückführen<sup>78</sup>. Ihnen kommt also eine enorme Bedeutung für den lesebezogenen Schulerfolg von Kindern zu. Allerdings besteht Grund zu der Annahme, dass sich dieser Effekt bezüglich der hier untersuchten Lesekompetenz deutlich stärker abzeichnet als dies möglicherweise in anderen Domänen der Fall sein könnte: Die Lesekompetenz ist eine zentrale Kulturtechnik, für deren Erlernen außerschulische Gelegenheiten womöglich relevanter sind als zum Beispiel für die naturwissenschaftliche oder mathematische Kompetenz. Auch ist es nicht Ziel von IGLU, die Erfassung der Lesekompetenz explizit an dem im Schulfach Deutsch durchgeführten Leseunterricht auszurichten, was den Einfluss familiärer Lernsituationen – im Vergleich zu schulischen Gelegenheiten – noch verstärken könnte. Überprüft werden sollte daher, wie sich der Herkunftseffekt und seine Varianz zwischen den Klassen abzeichnet, wenn andere leistungsbezogene Kompetenzen curriculumnäher erfasst werden. Dies wird in Kapitel 14.1 noch einmal aufgegriffen.

Wenngleich individuelle Förderung also nicht mit dem Herkunftseffekt in Verbindung gebracht werden konnte, können für sie andere Implikationen aus den theoretischen und empirischen Darstellungen abgeleitet werden. Sie alle beruhen auf dem Befund, dass individueller Förderung sehr heterogene Verständnisse zugrunde liegen. Dies wird vor allem dann deutlich, vergleicht man die Definitionen, auf welchen die unterschiedlichen (wenngleich wenigen) Forschungsarbeiten zu diesem Themengebiet beruhen: Neben einigen Studien, für die das Konstrukt nicht differenzierter betrachtet werden konnte (beispielsweise „individuelle fachliche Unterstützung“ oder „individualisierter Unterricht“ bei Helmke & Schrader, 1998; Weinert et al., 1989 und Krätzschar, 2010 oder „individualized instruction“ bei Hattie, 2009), werden die individuelle Bezugsnormorientierung (Gruehn, 2000), binnendifferenzierende Maßnahmen (ebd.), Merkmale des offenen Unterrichts (Giaconia & Hedges, 1982), Feedback (Seidel & Shavelson, 2007), Diagnostik (Giaconia & Hedges, 1982; Helmke & Weinert, 1997a) oder auch Kleingruppenunterricht (Helmke & Schrader, 1990) unter individueller Förderung subsumiert. Auch findet sich verstärkt ein Verständnis,

---

<sup>78</sup> Für das Leseselbstkonzept fällt dieser Effekt mit einer Varianzaufklärung von 7,4% (das entspricht etwa einem Sechszehntel) nicht ganz so stark aus.

welches in etwa postuliert, dass Schülerinnen und Schüler in Settings individueller Förderung an unterschiedlich schweren Materialien arbeiten und sich diese mitunter selbstständig aussuchen (etwa bei Gruehn, 2000; Giaconia & Hedges, 1982; Lipowsky et al., 2011 und Reis et al., 2011). Gerade letzteres entspricht nicht unbedingt dem theoriegeleiteten Verständnis, wie es hier in Kapitel 3 vorgestellt wurde und deren Kernelement die *Planung der Passung der Lernmaterialien* an die Voraussetzungen und Interessen der jeweiligen Schülerinnen und Schüler ist, wenn sich diese eigenständig Aufgaben aus einem Pool unterschiedlich schwerer Übungen heraussuchen, die der oder die Lehrende nicht explizit für die individuellen Kinder bzw. Jugendlichen vorbereitet hat.

Und auch in der bildungspolitischen Diskussion um die Implementation individueller Förderung im Unterricht zeigen sich unterschiedliche Verständnisse insbesondere bezüglich der *Zielgruppen* individueller Förderung, die die ohnehin differierenden Definitionen in Theorie und Empirie noch ergänzen. In der in Kapitel 5 vorgestellten Systematisierung der aktuellen Schulgesetze der sechzehn deutschen Bundesländer durch Fischer (2014) werden solche Settings insbesondere für Schülerinnen und Schüler mit Entwicklungsverzögerungen oder Lernbehinderungen, für besonders begabte Lernende und für Kinder und Jugendliche, die Bedarf im Erwerb der deutschen Sprache haben, festgehalten. Dabei werden unter individueller Förderung meist sonderpädagogische Maßnahmen oder zusätzlicher Förderunterricht verstanden (vgl. Fischer, 2014; Kunze & Solzbacher, 2008; Solzbacher, 2008), häufig verbunden mit der defizitären Perspektive, dass lernbezogene Missstände behoben werden sollen – nicht so sehr hingegen, dass der Unterricht und die Unterrichtsmaterialien an die individuellen Interessen, Lernstile und Lernvoraussetzungen *aller* Schülerinnen und Schüler angepasst werden.

Zuletzt muss aus den Untersuchungen, die explizit die Einstellungen der Lehrkräfte gegenüber individueller Förderung erfassen (vgl. etwa Kunze & Solzbacher, 2008; Solzbacher, 2008), aber auch aus theoretischen Betrachtungen, wie sie empirischen Analysen vorweggehen (vgl. etwa Altrichter et al., 2009; Klieme & Warwas, 2011; Bräu, 2007; Fischer, 2014), geschlussfolgert werden, dass es bislang kein theoretisch, praktisch und empirisch erprobtes Konzept für die Umsetzung individueller Förderung im Unterricht gibt, auf das sich Lehrende, Forschende und Bildungspolitikern und -politiker stützen können. Es ist anzunehmen, dass hierin die Ursache für die didaktischen, empirischen und

bildungspolitischen Herausforderungen in der Implementierung und Untersuchung der individuellen Förderung liegt. Erklären könnte dies auch die von Solzbacher (2008) beschriebene Unsicherheit der Lehrenden, dies im Unterricht anzuwenden. Allerdings zeigte sich auch, dass einzelne Facetten individueller Förderung durchaus in den Unterricht einfließen, jedoch mit dem Ziel, diesen schüleraktivierend und abwechslungsreich zu gestalten (vgl. Wischer, 2008) – und nicht unbedingt, um ihn an die individuellen Voraussetzungen der Lernenden anzupassen mit dem Ziel, die einzelnen Schülerinnen und Schüler nach ihren Interessen, Lernständen und Neigungen zu fördern. Möglicherweise lässt sich damit auch erklären, warum das hier nicht verwendete internationale IGLU-Item „Ich gebe individuell abgestimmten Leseunterricht“ und die in Kapitel 12.3.8 dargestellte, nur national verfügbare Skala ‚Individuelle Förderung‘ erwartungswidrig lediglich in mittlerem Ausmaß in Zusammenhang stehen: Angenommen werden könnte, dass Lehrende ihren Unterricht nicht als individuell fördernd einschätzen, wenngleich sie entsprechende Teilelemente wie etwa die individuelle Unterstützung einzelner Lernender und die Vergabe gezielter Zusatzaufgaben durchaus anwenden. Letzteres geschieht eventuell jedoch nicht derart, dass allen Schülerinnen und Schülern solch eine Förderung zugute kommt, sondern dass insbesondere leistungsschwache Lernende vermehrte Aufmerksamkeit, mitunter in individualisierten Settings, erhalten bzw. dass, wie von Wischer (2008) konstatiert, der Unterricht hierdurch vor allem abwechslungsreich gestaltet werden soll. Bezüglich des internationalen Einzelitems bleibt daraus ableitend also unklar, ob dieses eine korrekte Einschätzung der Lehrenden bezüglich der hier zugrunde liegenden Definition der individuellen Förderung als Planung der Passung der Unterrichtsangebote an die Lernvoraussetzungen *aller* Schülerinnen und Schüler darstellt – oder möglicherweise lediglich abbildet, dass Lehrende der individuellen Förderung unsicher gegenüber stehen, so dass diese angeben, kaum (ihren Leseunterricht) zu individualisieren. Eine weitere Erklärung hierfür könnte allerdings auch sein, dass entsprechende Maßnahmen nicht im Leseunterricht (wie es das Einzelitem erfasst) ergriffen werden, da dies möglicherweise eher über Stillarbeiten oder das gemeinsame laute (Vor-)Lesen unterrichtet wird, wie es beispielsweise Lankes (2004), Lankes et al. (2004) und Lipowsky et al. (2011) nahelegen. Solche Erörterungen müssen jedoch Gegenstand vertiefender, über IGLU hinausgehender Analysen sein, auf die im Kapitel 14.1 noch einmal ausführlicher eingegangen werden soll. Für einen internationalen Vergleich, der zum Ziel hatte, die in Deutschland noch nicht weit

implementierte individuelle Förderung bezüglich ihres Potenzials, den Herkunftseffekt zu mindern, mit einem Land zu vergleichen, in welchem das Konzept bereits langjährige Anwendung findet und in welchem zudem der Herkunftseffekt vergleichbar mit dem deutschen ist (z.B. Norwegen), konnte das Einzelitem daher nicht ohne Einschränkungen herangezogen werden (vgl. Schulz-Heidorf & Solheim, 2016).

Aus den nationalen Analysen zeigt sich, dass individuelle Förderung mit dem klassendurchschnittlichen Leistungsniveau in einem leicht negativen Zusammenhang steht. Da es, wie bereits dargestellt, anhand von Querschnittsdaten nicht möglich ist, kausale *Wirkrichtungen* festzustellen, verbleibt die Zuschreibung von Ursachen hier auf interpretativer Ebene. So könnte sich individuelle Förderung negativ auf die Lesekompetenz auswirken, da ihre Umsetzung zeitintensiv sein kann: Anders als bei didaktischen Settings, die auf das gleichzeitige Unterrichten der gesamten Klasse ausgerichtet sind, bedarf es hier der individuellen Instruktion. Diese geht zudem vermehrt mit einer individuellen Lernberatung, einer aufwändigeren Leistungsdokumentation (gerade wenn Schulkinder an unterschiedlichen Materialien arbeiten) und einer individualisierten Rückmeldung einher – wofür möglicherweise deutlich mehr Zeit aufgewendet wird, die in anderen Settings bereits für das Lernen genutzt werden könnte (dies legt etwa auch Gruehn, 1995, 2000 nahe). Ist dies dauerhaft der Fall, könnte sich daraus ein Kompetenzrückstand ergeben, der dann auch nicht mehr durch die Qualität des Lernens (wie sie vermeintlich mit individuell angepasster Förderung einhergeht) ausgeglichen werden kann. Eine zweite Ursache für den negativen Zusammenhang könnte in der lernbezogenen Selbstregulation (vgl. etwa Pintrich, 2000; Klieme & Warwas, 2011; Otto et al., 2011) liegen, die bei Viertklässlerinnen und Viertklässlern möglicherweise noch nicht stark ausgeprägt ist: Gehen individualisierte Settings damit einher, dass Schülerinnen und Schüler über einen längeren Zeitraum selbstständig lernen, müssen diese nicht nur über entsprechende Lernstrategien verfügen, sondern beispielsweise auch in der Lage sein, Lernprozesse zu strukturieren und zu reflektieren sowie die Lernmotivation eigenständig aufrecht zu erhalten. Dies ist gerade dann relevant, wenn individuelle Förderung verstanden wird als Bereitstellen unterschiedlich schwerer Aufgaben, die die Schülerinnen und Schüler eigenständig auswählen und bearbeiten (wie oben bereits ausführlicher diskutiert wurde). Rückbeziehend auf die in IGLU eingesetzten Items, die hier verwendet wurden, kann nicht überprüft werden, ob solche Phasen eigenständigen Lernens (sowie freier Wahl des Lernmaterials) Element des

Verständnisses individualisierten Unterrichts sind, da dies nicht explizit erfasst wird. Allerdings findet sich ein deutlicher Fokus auf eine starke Begleitung der Lernprozesse durch die Lehrkraft. So heißt es dort verkürzt z.B. ‚Ich gebe (Zusatz-)Aufgaben‘, ‚Ich variiere die Aufgabenstellungen‘ und ‚Ich übe bzw. wiederhole mit den langsameren Lernenden‘ (ausführlicher siehe Kapitel 12.3.8), so dass anzunehmen ist, dass für das hier abgebildete Verständnis individueller Förderung selbstregulative Kompetenzen bei den Lernenden weniger bedeutsam sind, als würden diese die Lernmaterialien selbstständig auswählen und über einen längeren Zeitraum eigenständig bearbeiten.

Eine weitere Ursache für den negativen Zusammenhang zwischen individueller Förderung und Lesekompetenz kann aus der inversen Interpretation der angenommenen Wirkrichtung abgeleitet werden: Möglicherweise individualisieren Lehrende ihren Unterricht gerade dann, wenn das durchschnittliche Leistungsniveau der Klasse gering ist, sich unter Umständen also viele Schülerinnen und Schüler in der Klasse befinden, die über ungünstige Lernvoraussetzungen verfügen. Solch eine Annahme würde dafür sprechen, in künftigen Analysen auf Basis von Längsschnittdaten das klassendurchschnittliche Leistungsniveau zum ersten Messzeitpunkt als Prädiktor für die Häufigkeit, mit der Lehrerinnen und Lehrer individualisieren, aufzunehmen. Nicht bedingt wird letztere hingegen durch den klassendurchschnittlichen (aggregierten) sozioökonomischen Familienstand und den auf Klassenebene aggregierten häuslichen Gebrauch der deutschen Sprache. Es muss also angenommen werden, dass andere als diese herkunftsbezogenen Merkmale Lehrende dazu veranlassen, Maßnahmen der individuellen Förderung zu ergreifen – möglicherweise das oben beschriebene klassendurchschnittliche Leistungsniveau, vielleicht aber auch dessen Streuung innerhalb der Klasse (wie es etwa bei Scharenberg, 2012 aufgegriffen wird), die ein einheitliches, nicht-adaptives Unterrichten erschwert. Auf dies soll in Kapitel 14.1, welches sich möglicher, dieser Arbeit anschließender empirischer Untersuchungen widmet, vertiefend eingegangen werden.

Für alle anderen theoretisch angenommenen Wirkungen individueller Förderung ergaben sich ebenfalls keine signifikanten Befunde. Zu diesen gehört der Effekt individueller Förderung auf das klassenmittlere Selbstkonzept im Lesen. Dessen Nichtvorhandensein lässt sich möglicherweise auf die verwendeten Items individueller Förderung zurückführen, die lediglich eine *leistungsbezogene* Differenzierung der Lehrkraft erfassen, nicht jedoch eine

Anpassung des Lernmaterials an die Interessen und Neigungen der einzelnen Schülerinnen und Schüler. Dabei wäre es gerade diese Passung, die das Leseengagement, die Leselust und mit ihnen das Vertrauen der Schülerinnen und Schüler in ihre Fähigkeiten steigern könnte (vgl. etwa Reis et al., 2011). Angenommen wird zudem, dass es den Lernenden gelingt, das Niveau, auf welchem sie und die Klassenkameradinnen und -kameraden durch die Lehrkraft eingestuft wurden, zu identifizieren, was dann soziale Vergleichsmomente wie in herkömmlichen Unterrichtsettings nach sich zieht. Dabei wäre es vor allem die Anpassung des Unterrichts an die individuellen Neigungen und Interessen der Schulkinder, die den Vergleich zwischen den Lernenden nichtig machen würde: Wenn sich das Lernmaterial beispielsweise inhaltlich und bezüglich der didaktischen Herangehensweise unterscheidet, erlaubt der Vergleich unter den Schülerinnen und Schülern einer Klasse keine Rückschlüsse auf das Leistungsniveau, auf dem sich der oder die Lernende befindet bzw. durch die Lehrkraft eingestuft wurde – woraus dann keine Rückschlüsse für das eigene Fähigkeitsselbstbild geschlossen werden könnten. Diesen theoretischen Annahmen kann also anhand der in IGLU verwendeten Daten nicht nachgegangen werden; ihre Überprüfung muss Gegenstand vertiefter empirischer Analysen bleiben.

Ein weiterer untersuchter Effekt der individuellen Förderung, der ebenso nicht nachgewiesen werden konnte, steht ebenfalls in einem Zusammenhang mit sozialen Vergleichssituationen innerhalb der Klasse: Wie oben dargestellt wird angenommen, dass individualisierte Lernsituationen zu weniger leistungsbezogenen Vergleichen zwischen den Schülerinnen und Schülern und damit zu weniger Konkurrenzverhalten führen. Dies könnte zu einer positiveren Wahrnehmung der eigenen sozialen Integration innerhalb des Klassenverbands („Schüler-Schüler-Beziehung“) führen, da Leistungsmerkmale dann weniger relevant für deren Gelingen sein könnten. Vermutlich würde dies vor allem besonders guten und schlechten Schülerinnen und Schülern zugute kommen, die – so die Annahme – in anderen, stärker an einem Vergleich orientierten Settings eher Neid oder Mobbing erfahren und folglich möglicherweise die eigene soziale Integration als weniger gelungen wahrnehmen. Dass sich solch ein Effekt im Klassenmittel nicht zeigt, kann auf die Fachspezifität der individuellen Förderung zurückgeführt werden: In IGLU befragt wurden nur Lehrende, die die Viertklässlerinnen und Viertklässler *im Fach Deutsch* unterrichten (und diesen Unterricht mehr oder minder individualisieren). Die Bewertung der eigenen sozialen Integration durch die Schülerinnen und Schüler hingegen erfolgt allgemein für den

Klassenverbund, so wie er über viele (wenn nicht alle) Schulfächer, in denen die Klasse unterrichtet wird, besteht. Vermutet werden kann, dass der Unterricht in anderen Fächern selten in gleichem Ausmaß individualisiert wird wie der Leseunterricht, so dass soziale Vergleiche – selbst wenn sie im Fach Deutsch nicht bestehen – in anderen Fächern möglich sind und die Wahrnehmung der sozialen Integration in anderem Maß beeinflussen. Da die Schüler-Schüler-Beziehung also ein fachübergreifendes Konstrukt darstellt, erscheint es plausibel, dass das didaktische Setting *eines* Unterrichtsfachs diese nicht zu beeinflussen vermag.

Der dritte hier untersuchte Effekt der (leistungsbezogenen) individuellen Förderung auf die klassendurchschnittliche Wahrnehmung der Schüler-Lehrer-Beziehung fiel ebenfalls erwartungswidrig aus, da er nicht nachgewiesen werden konnte. Dabei war es vor allem die Annahme, dass individualisierte Unterrichtssettings die Interaktion zwischen Lehrenden und Lernenden persönlicher gestalten würden, in welcher die je unterschiedlichen Lernbedingungen der einzelnen Schülerinnen und Schüler stärkere Beachtung finden könnten als wenn der bzw. die Lehrende die Klasse gemeinsam, d.h. nicht differenziert unterrichtet. Dass sich dies nicht nachweisen ließ, hat seine Ursache möglicherweise darin, dass anhand der hier verwendeten IGLU-Items fast ausschließlich Situationen erfasst werden, in denen unterschiedlich schwere Aufgaben an Schülerinnen und Schüler unterschiedlicher Leistungsniveaus verteilt werden – Elemente individueller Interaktionen finden sich in Ansätzen lediglich in den Items „Ich lasse schnellere Schüler schon zum Nächsten übergehen, während ich mit den langsameren noch übe oder wiederhole“ (ATNG27Br) und „Ich gebe schwachen Schülern zusätzliche Unterstützung im Unterricht“ (ATNG27Er). Hier bleibt allerdings unklar, wie solch eine Förderung der Leistungsschwächsten aussieht – möglich wäre beispielsweise ein Setting, in dem sich der bzw. die Lehrende mit der Gruppe der schwächsten Lernenden zusammensetzt und diese gemeinsam unterrichtet, wobei es unter Umständen nur zu wenigen individuellen Interaktionen kommt. Individuelle Interaktionen mit Lernenden anderer Leistungsniveaus (bzw. individuelle Interaktionen mit allen Lernenden, unabhängig von deren Leistung) sind hingegen nicht Bestandteil der für die Messung der Häufigkeit individueller Förderung herangezogenen Skala. Es erscheint daher auch hier nicht überraschend, dass sich anhand der Items, die kaum individuelle Interaktionen abdecken (von denen angenommen wurde, dass sie sich

positiv auf die Wahrnehmung der Schüler-Lehrer-Beziehung auswirken könnten), keine Beeinflussung durch die individuellen Förderung nachweisen lässt.

Auf Individualebene zeigen sich für die beiden Konstrukte der Wahrnehmung der sozialen Lernumwelt andere Effekte. Dabei fallen vor allem die Zusammenhänge des Selbstkonzepts im Lesen mit der durch den Einzelschüler bzw. die Einzelschülerin wahrgenommenen Beziehung zu der Lehrkraft und der eigenen sozialen Integration in den Klassenverbund vergleichsweise stark aus. Angenommen wurde hier zum einen, dass freundschaftliche Beziehungen innerhalb der Klasse als ‚Puffer‘ für negativ ausfallende Leistungsvergleiche dienen und bei leistungsstarken Schülerinnen und Schülern möglicherweise das Erleben der eigenen Kompetenz stärken – etwa dann, wenn die Klassenkameradinnen und Klassenkameraden häufig um Unterstützung in schulischen Dingen, z.B. Hausaufgaben bitten. Für stark integrierte Kinder fiel dann die Selbstkonzeptentwicklung positiver (oder weniger negativ) aus als für weniger integrierte Kinder. Zum anderen könnten dauerhaft negativ ausfallende Vergleiche mit den Mitschülerinnen und Mitschülern das Gefühl verstärken, nicht zu den übrigen – leistungsstärkeren – Schülerinnen und Schülern dazu zu gehören, was die soziale Integration auf Dauer belasten würde. Dass dieser Zusammenhang besteht und dabei nicht unwesentlich hoch ausfällt, konnte in den hier vorgestellten Analysen gezeigt werden. Dies untermauert, dass die positive Wahrnehmung der sozialen, auf die Mitschülerinnen und Mitschüler bezogenen Lernumwelt in seiner Bedeutung für die Selbstkonzeptentwicklung nicht zu unterschätzen ist. Andererseits zeigte sich auch, dass die Ausprägung des Selbstkonzepts nicht nur in Zusammenhang mit der Leistung steht, sondern auch soziale Folgen, insbesondere für die Integration im Klassenkontext haben kann.

Und auch für die Wirkung einer positiv eingeschätzten Schüler-Lehrer-Beziehung auf das Selbstkonzept im Lesen konnte ein vergleichsweise starker Effekt festgestellt werden. Möglicherweise haben Lernende, die angeben, ‚mit ihrem Klassenlehrer bzw. ihrer Klassenlehrerin gut auszukommen‘, mehr Freude, dem Unterricht zu folgen, sind motivierter und haben eventuell eher das Gefühl, dass ihnen die Lehrkraft höhere Leistungen zutraut als Kinder, die angeben, nicht mit ihrem Lehrer bzw. ihrer Lehrerin auszukommen, was sich dann in einem entsprechend ausgeprägteren (bzw. geminderten) Selbstkonzept niederschlägt. Allerdings könnte dieser Effekt auch in anderer Wirkrichtung interpretiert werden: Ein Item der Leseselbstkonzept-Skala lautet „Meine Lehrerin/mein Lehrer sagt, dass

ich gut lesen kann“. Wird ein Kind häufig durch die Lehrperson gelobt (was sich hier in einer höheren Ausprägung des Leseselbstkonzepts manifestieren würde), nimmt es die Beziehung zu dem Lehrer bzw. der Lehrerin möglicherweise positiver wahr. Allerdings wird durch das Einzelitem der Schüler-Lehrer-Beziehung explizit die Wahrnehmung hinsichtlich des *Klassenlehrers* bzw. der *Klassenlehrerin* erfasst – bei der Lehrperson, die die Schülerinnen und Schüler im Lesen unterrichtet und die in IGLU befragt wurde, muss es sich jedoch nicht um die bzw. den *Klassenlehrenden* handeln. Anzunehmen ist, dass der Zusammenhang zwischen Selbstkonzept und Schüler-Lehrer-Beziehung sogar noch stärker ausfallen würde, würden sich beide Konstrukte auf dieselbe Dimension (hier: die Lehrperson, die Lesen unterrichtet) beziehen. In jedem Fall wird deutlich, dass der Beziehung zwischen Lehrenden und Lernenden, wie sie durch die Schülerinnen und Schüler wahrgenommen wird, eine nicht zu unterschätzende Bedeutung für die Ausprägung des Leseselbstkonzepts zukommt. Deren Bedeutsamkeit lässt sich auch – wie angenommen – für die Wahrnehmung der eigenen sozialen Integration im Klassenverbund nachweisen. Angenommen wurde, dass ein soziales, empathisches Lehrerverhalten vor allem die Konkurrenzorientierung unter den Schülerinnen und Schülern mindert und sich mit ihr auf die sozialen Vergleiche zwischen ihnen auswirkt – was dann die Wahrnehmung der sozialen Integration stärkt. Daraus abgeleitet werden kann, dass solch eine ‚Beziehungsarbeit‘ zur Stärkung der Schüler-Lehrer-Beziehung durchaus von Vorteil ist – nicht nur für die Erfüllung des Bedürfnisses nach sozialer Eingebundenheit sondern auch für die Entwicklung eines gestärkten Selbstkonzepts, was hier als ein Indikator von Schulerfolg verstanden wird.

Dahingegen erwies sich der Einfluss beider Facetten der sozialen Lernumwelt auf die Leseleistung als kaum bedeutsam. Angenommen wurde hier zum einen, dass sich eine Lernumgebung, in der die Schülerinnen und Schüler ihren Klassenkameradinnen und -kameraden vertrauen können, Hilfe und Unterstützung leisten und erhalten, Rücksicht erfahren und mit ihnen nicht in Konkurrenz stehen, leistungsförderlich auswirke – insbesondere, weil durch eine als positiv wahrgenommene Schüler-Schüler-Beziehung das Bedürfnis nach sozialer Eingebundenheit (als eines der von Deci & Ryan, 1993 konstatierten drei psychologischen Grundbedürfnisse) erfüllt wäre, was dann wiederum – so die Annahme – Kapazitäten für das Verfolgen des Unterrichts und das Lernen schaffen würde. Diesem kommt jedoch, wie in den Analysen gezeigt, kaum eine Relevanz zu, zumindest nicht für die Lesekompetenz. Allerdings werden Facetten wie Vertrauen, Unterstützung, Rücksichtnahme

und Konkurrenzverhalten nicht durch die hier verwendeten Items der Schüler-Schüler-Beziehung (dargestellt in Kapitel 12.3.6) abgedeckt. Zumindest für Viertklässlerinnen und Viertklässler kann abgeleitet werden, dass das Vorhandensein von Freundschaften und die soziale Akzeptanz unter den Mitschülerinnen und Mitschülern nicht mit einer gesteigerten Lesekompetenz einhergeht.

Und auch für den Einfluss der Beziehung zu dem Klassenlehrer bzw. der Klassenlehrerin ergibt sich ein zwar vorhandener, jedoch kaum bedeutsamerer Effekt auf die Leseleistung. Angenommen wurde hier, dass eine stärkere, positive Beziehung zu der Lehrkraft die Motivation, dem Unterricht zu folgen, steigern würde, was in mehr lernbezogenem Engagement und dadurch in einer besseren Leseleistung münden könnte. Dass dieser Effekt sehr gering ausfällt, lässt sich zum einen dadurch erklären, dass (wie oben bereits beschrieben) das Einzelitem der Schüler-Lehrer-Beziehung sich auf den *Klassenlehrer* bzw. die *Klassenlehrerin* bezieht, der oder die nicht zwangsweise die befragten Schülerinnen und Schüler im *Lesen* unterrichten muss – ist dies nicht der Fall, wird hier untersucht, wie sich die Beziehung zu einer Lehrkraft anderer Fächer auf die Lesekompetenz auswirkt. Andererseits wurde aus den theoretischen Betrachtungen abgeleitet, dass sich die Schüler-Lehrer-Beziehung vor allem über motivationale Faktoren auf die Lesekompetenz auswirkt, was so im Modell nicht berücksichtigt wurde. Dies stellt eine wichtige Perspektive für Folgeuntersuchungen dar, wie sie in Kapitel 14.1 noch einmal aufgegriffen werden sollen.

Insgesamt bestätigt sich für die Wirkzusammenhänge der sozialen Facetten der Lernumwelt der empirische Befund, dass sich diese vor allem auf motivational-affektive Schülercharakteristika, nicht so sehr jedoch auf kognitive Merkmale auswirken. Auf Individualebene ließ sich außerdem der Herkunftseffekt theoriekonform modellieren – allerdings nicht in Zusammenhang mit der Häufigkeit individueller Förderung bringen. Für diese kann aus den hier dargestellten theoretischen Betrachtungen und empirischen Befunden abgeleitet werden, dass ihre Umsetzung möglicherweise defizitorientiert ist, insgesamt jedoch selten erfolgt. Dies könnte sich darauf zurückführen lassen, dass insbesondere das Vorbereiten von individuell angepassten Lernmaterialien mit einem hohen Aufwand einhergeht, der eine ausführliche, individuelle Diagnostik des Lernstands und der Interessen der einzelnen Schülerinnen und Schüler erfordert, gepaart mit einer umfangreichen Dokumentation der Leistungsfortschritte und einer individuellen Form der Rückmeldung. Dass Lehrende, wie

dies in Kapitel 12.3.8 gezeigt wurde, unter solchen Umständen Elemente der individuellen Förderung wenig und dann vor allem in Form von leistungsbezogen differierenden Materialien einsetzen, verwundert wenig und wird in den folgenden Kapiteln noch einmal aufgegriffen.

### **13.3 Grenzen der empirischen Untersuchung**

Die sekundäre Nutzung von Daten aus Large-Scale-Assessments, wie sie für die hier vorliegenden Analysen auf Grundlage der Internationalen Grundschul-Lese-Untersuchung herangezogen wurden, geht mit einer Reihe von Vorteilen, aber auch mit Einschränkungen einher. Dabei sind es vor allem die Repräsentativität der Daten für Viertklässlerinnen und Viertklässler und deren Lesekompetenz in Deutschland sowie die Möglichkeit eines direkten internationalen Vergleichs, die die Nutzung solcher Daten attraktiv machen. Andererseits können für Sekundäranalysen nur Variablen bzw. Konstrukte herangezogen werden, die auch in den Studien erfasst werden und möglicherweise nur einen Teil der jeweiligen theoretischen – ex post an die Studien herangetragenen – Annahmen abzudecken vermögen. Im Rahmen der vorliegenden Arbeit betrifft dies vor allem die Erfassung der individuellen Förderung: Hierfür herangezogen werden können sowohl ein international eingesetztes Einzelitem als auch eine Skala der nationalen Erweiterung (vgl. Kapitel 12.3.8), die – so die theoretische Annahme – inhaltlich das gleiche Konstrukt abbilden sollten, jedoch nur in mittlerem Maß miteinander korrelieren. Zurückgeführt wurde dieser erwartungswidrig ausfallende Zusammenhang zum einen auf die ihnen zugrunde liegende Fachspezifität: So wird anhand des Einzelitems erfasst, wie häufig Lehrende individuell abgestimmten Leseunterricht geben, während die sieben Items, die anhand von Faktorenanalysen dem Faktor ‚Individuelle Förderung‘ zugeordnet wurden, fachunabhängig erfassen, wodurch und wie häufig Schülerinnen und Schüler nach ihren individuellen Lernvoraussetzungen gefördert werden. Von den Lehrpersonen, die angeben, ihren Unterricht häufig zu individualisieren, geben lediglich knapp 40% an, dies sowohl im Leseunterricht (Einzelitem) als auch fachübergreifend (Skala) zu tun – für etwa 60% finden sich lediglich Angaben zu allgemeinen, fachunabhängigen Maßnahmen der individuellen Förderung. Möglicherweise werden also andere didaktische Methoden als die individuelle Förderung herangezogen, um Schülerinnen und Schüler *im Lesen* zu unterrichten. Eine weitere Erklärung für die mittlere Korrelation findet sich vermutlich in einem heterogenen Verständnis dessen, was individuelle Förderung ausmacht – möglich ist es, dass Lehrende

zwar Maßnahmen individueller Förderung, wie sie anhand der Skala erfasst werden, ergreifen, diese jedoch nicht der individuellen Förderung (wie sie, nicht weiter differenziert, durch das Einzelitem erfasst wird) zuordnen. Da anhand der sieben Items der nationalen Skala konkrete didaktische Maßnahmen der individuellen Förderung erfasst werden, ist anzunehmen, dass hierdurch inhaltlich validere Aussagen zu der Häufigkeit, mit der Viertklässlerinnen und Viertklässler im Unterricht individuell gefördert werden, getroffen werden können als durch das (wenig spezifische) Einzelitem. Untermauert wird dies durch rezente Analysen auf Grundlage dieses Einzelitems, welches (als ‚kleinster möglicher Nenner‘<sup>79</sup>) für einen internationalen Vergleich herangezogen wurde und welches (anders als die hier verwendete nationale Skala) in keinem Zusammenhang mit der Lesekompetenz der Viertklässlerinnen und Viertklässler steht (vgl. Schulz-Heidorf & Solheim, 2016).

Auf Grundlage von IGLU lassen sich solche Annahmen jedoch nicht überprüfen. Nötig wäre hier vor allem eine Erfassung des Verständnisses individueller Förderung der Lehrerinnen und Lehrer, die Rückschlüsse darauf zulassen würde, ob insbesondere die über das Einzelitem („Ich gebe individuell abgestimmten Leseunterricht“) erfassten Angaben dem hier dargestellten Verständnis individueller Förderung als Planung der Passung der Lernangebote an die Lernvoraussetzungen, Interessen und Fähigkeiten des einzelnen Schulkindes entsprechen. Erst dann könnten die oben aufgestellten Vermutungen über eine möglicherweise eingeschränkte Validität des Items überprüft werden. Ebenfalls unberücksichtigt bleiben muss eine Verzerrung des Antwortverhaltens beispielsweise durch soziale Erwünschtheit – ebenso wie die Wahrnehmung des Lehrerverhaltens durch die Schülerinnen und Schüler. Letzterem käme im Kontext der vorliegenden Arbeit möglicherweise eine größere Bedeutung zu: Während Lehrende lediglich angeben können, die Anpassung des Unterrichtsangebots an die Voraussetzungen, Fähigkeiten und Interessen der Schülerinnen und Schüler *anzustreben*, könnte eine Wahrnehmung dessen durch die Lernenden selbst diese Passung bestätigen (oder negieren). Solch eine Gegenüberstellung scheint gerade im Anbetracht von Forschungsarbeiten, in denen Schülerurteilen eine entsprechende Validität nachgewiesen werden konnte (etwa Clausen, 2002; Grewe, Strietholt & Schwippert, 2007), sinnvoll. In einer entsprechenden Erweiterung des Schülerfragebogens liegt also möglicher-

---

<sup>79</sup> Da nur das Einzelitem in dem internationalen Datensatz verfügbar ist, konnte auch nur dieses für den Ländervergleich herangezogen werden.

weise eine (zudem ökonomische) Chance, dem oben geschilderten Problem der Erfassung individueller Förderung im Rahmen von IGLU zu begegnen.

Eine weitere Grenze der vorgestellten empirischen Analysen ergibt sich ebenfalls aus der Verwendung der Daten der Internationalen Grundschul-Lese-Untersuchung. Mit IGLU erfasst wird die Lesekompetenz von Viertklässlerinnen und Viertklässlern, die hier (neben dem Leseselbstkonzept) das abhängige, zu erklärende Merkmal des Schulerfolgs von Schülerinnen und Schülern darstellt. Auf sie – und den ihr zugehörigen Unterricht – bezieht sich ein Großteil der in den Rahmenfragebögen erfassten Merkmale der Unterrichtsqualität, der Lernbedingungen und der Wahrnehmung der schulischen Umwelt. Mit IGLU ebenfalls einher geht eine eingeschränkte Verallgemeinerbarkeit der Lehrerangaben: Da die Stichprobe der Lehrerinnen und Lehrer nicht zufällig, sondern immer in Relation zu den zufällig ausgewählten Schülerinnen und Schülern und ihren Klassen gezogen wurde, können diese auch nur für das Unterrichten von Viertklässlerinnen und Viertklässlern verallgemeinert werden. Werden hier also Aussagen beispielsweise zu der Häufigkeit individueller Förderung und ihrem Einfluss auf die Schulleistung getroffen, gelten diese folglich für die Lesekompetenz von Viertklässlerinnen und Viertklässlern in Deutschland, die von Lehrenden mit bestimmten Merkmalen (beispielsweise der Häufigkeit individueller Förderung) im Fach Deutsch<sup>80</sup> unterrichtet werden. Dies ist gerade dann wichtig, wird wie in Kapitel 13.2 dargestellt angenommen, dass sich das Kernmodell der hier vorgestellten Analysen, der Herkunftseffekt, in anderen Domänen wie beispielsweise in Mathematik oder Naturwissenschaften anders darstellen könnte.

Darüber hinaus lässt die Betrachtung von Effekten im nicht-experimentellen Querschnitt keine Aussagen hinsichtlich ihrer Kausalrichtungen zu: Erst unter der Berücksichtigung von Eingangsniveaus, wie es Designs mit mehreren Messwiederholungen erlauben, können Ursache-Wirk-Beziehungen ausgemacht werden. Besonders deutlich zeigt sich dies in dem in den vorliegenden Analysen invertierten Zusammenhang zwischen individueller Förderung und dem klassendurchschnittlichen Leistungsniveau. Würden dieser Untersuchung mehrere Messzeitpunkte zugrunde liegen, hätte überprüft werden können, ob Lehrende insbesondere dann zu Maßnahmen individueller Förderung greifen, wenn das

---

<sup>80</sup> Möglich ist es hierbei allerdings, dass diese Schülerinnen und Schüler durch die befragte Lehrperson nicht nur in Deutsch, sondern auch in anderen Schulfächern unterrichtet werden, die jedoch nicht im Fokus von IGLU stehen.

Leistungsniveau einer Klasse gering ist oder ob diese (i.d.R. zeitintensiven) didaktischen Settings zu einer geringeren klassendurchschnittlichen Leseleistung führen. Aus diesen Analysen auf Basis eines Querschnitts ergeben sich also Perspektiven für zukünftige Untersuchungen, die gerade solche Abhängigkeitsstrukturen im Längsschnitt fokussieren könnten.

## **14. Ausblick**

Aus den hier vorgestellten Sekundäranalysen der IGLU-Daten ergeben sich eine Vielzahl von Implikationen für den Anwendungsbereich Schule, die sich sowohl auf dessen Erforschung als auch auf dessen Praxis – im Klassenzimmer und auf übergeordneter administrativer Ebene – beziehen. Diese sollen in dem hier abschließenden Kapitel 14 diskutiert werden.

### **14.1 Implikationen für zukünftige Forschungsarbeiten**

Die sich aus den Ergebnissen dieser Arbeit ableitenden Implikationen für Anschlussforschungsarbeiten beziehen sich im Wesentlichen auf eine Erweiterung des Modells, auf die Verwendung über IGLU hinausgehender (Längsschnitt-)Daten und die vertiefende Betrachtung der Wirkung individueller Förderung.

Ein Element, das sich sowohl bei der Beschreibung von Wirkzusammenhängen der sozialen Facetten der schulischen Lernumwelt, aber auch bei der Darstellung bisher nachgewiesener Effekte individueller Förderung findet, ist die Lernmotivation<sup>81</sup>. So stellen etwa Reis et al. (2011) fest, dass in Klassen, deren Leseunterricht durch das ‚Schoolwide Enrichment Model-Reading‘ (SEM-R) individualisiert wurde, ein gesteigertes Lernengagement und eine gesteigerte Lesemotivation zu finden sei, was die Autoren auf die Anpassung des Unterrichts an die Interessen und das Leseniveau der einzelnen Schülerinnen und Schüler zurückführten. Nach Gerecht (2010; auch Ladd, 1990; Ladd et al., 1997; Buhs & Ladd, 2001) wirkt sich zudem eine gelungene soziale Integration positiv auf die Lernmotivation aus, ebenso wie die positive Ausgestaltung der Schüler-Lehrer-Beziehung (Den Brok et al., 2004; Furrer & Skinner, 2003; Kunter & Voss, 2011; Opdenakker et al., 2012; Osterman, 2000; Lipowsky, 2015). Dass die Lernmotivation einen Einfluss auf die Qualität des Lernens und die Leistung hat, konnte in einer Vielzahl von Studien nachgewiesen werden (vgl. etwa Deci & Ryan,

---

<sup>81</sup> Anzumerken sei jedoch, dass auf die Darstellung von Zusammenhängen bezüglich der Lernmotivation in der theoretischen Rahmung dieser Arbeit kein Fokus gelegt wurde, so dass diese hier nur unvollständig wiedergegeben werden kann.

1993; Fend, 1997; Graham & Weiner, 1996; Jerusalem, 2006; Jerusalem & Pekrun, 1999; Krapp, Schiefele & Schreyer, 1993; zusammenfassend auch Lohrmann & Hartinger, 2014 und Schiefele & Köller, 2010). Intrinsisch motivierte, lernzielorientierte und interessierte Schülerinnen und Schüler, die Lern- und Leistungsaufgaben mit Begeisterung begegnen, „bemühen sich intensiv darum, den Stoff zu verstehen, indem sie z.B. verstärkt versuchen, neue Inhalte mit ihrem bisherigen Wissen in Verbindung zu bringen. An der Sache interessierte Schüler(innen) setzen tiefer gehender [sic] Lernstrategien ein und ziehen adäquatere Lösungsstrategien heran. Sie geben bei Schwierigkeiten nicht so schnell auf und zeigen insgesamt mehr Ausdauer“ (Lohrmann & Hartinger, 2014, S. 277). Wenn wie hier angenommen wird, dass durch Maßnahmen individueller Förderung – also der Anpassung der Lerngegenstände, -materialien und -situationen an die Voraussetzungen, Leistungsniveaus, Interessen und Neigungen der einzelnen Schülerinnen und Schüler – das Lernen für den einzelnen Lerner bzw. die einzelne Lernerin bedeutsamer gemacht wird, könnte dies mit einer gesteigerten Lernmotivation einhergehen, die möglicherweise als vermittelndes (mediierendes) Konstrukt zwischen der individuellen Förderung und dem Schulerfolg von Kindern und Jugendlichen angesiedelt ist. Interessant wäre es in diesem Kontext zu untersuchen, ob ein individualisierter Unterricht ebenfalls zu Veränderungen in der motivationalen Zielorientierung führen kann, also ob Kinder und Jugendliche unter adaptiven Unterrichtsettings eher zu einer Lernzielorientierung (Ausrichtung der Lernbemühungen auf die Entwicklung von Fähigkeiten und Kompetenzen, unabhängig davon, ob diese bspw. für Tests gebraucht werden, vgl. Job, Sevincer & Oettingen, 2014) denn zu einer Leistungszielorientierung (Ausrichten der Lernbemühungen auf die Demonstration von Kompetenzen, auch im Vergleich zu anderen, vgl. Job, Oettingen & Sevincer, 2014) tendieren, wenn sie dem Lerngegenstand eine größere persönliche (und damit evtl. eine größere außerschulische) Relevanz beimessen. Würde sich solch ein Effekt nachweisen lassen, würde der individuellen Förderung eine enorme Bedeutsamkeit für das schulische Lernen beikommen, welches dann nicht nur ausgerichtet wäre auf ein gutes Abschneiden in Leistungsevaluationen, sondern auf den Kompetenzerwerb als solchen.

Ein weiterer Teilaspekt der individuellen Förderung könnte ebenfalls das Modell auf Grundlage der IGLU-Daten ergänzen: Die individuelle Bezugsnormorientierung stellt ein Element der Leistungsrückmeldung bzw. Leistungsbeurteilung dar, welches sich für individualisierte Settings in besonderem Maß eignet (Altrichter et al., 2009; Weinert, 1997;

Hofmann, 2000; Zeinz & Köller, 2006). Bei der Bezugsnormorientierung handelt es sich im Allgemeinen um „einen Standard, mit dem ein Resultat verglichen wird, wenn man es als Leistung wahrnehmen und bewerten will“ (Rheinberg & Fries, 2010, S. 61); bei der *individuellen* Bezugsnorm wird die Leistung eines Schülers bzw. einer Schülerin mit seinen bzw. ihren vorherigen Leistungen verglichen<sup>82</sup>. Dies erscheint vor allem dann sinnvoll, wenn die Lernenden einer Klasse in adaptiven Settings an höchst unterschiedlichen Materialien in unterschiedlichem Tempo und mitunter mit unterschiedlichen Lernzielen arbeiten (vgl. Zeinz & Köller, 2006). Dabei konnte der individuellen Bezugsnorm bereits vielfach ein leistungsförderlicher und selbstkonzeptsteigernder Effekt nachgewiesen werden, der auch mit einer Minderung sozialer Vergleichsmomente innerhalb der Klassen – wie sie Teil der hier vorgestellten theoretischen Betrachtungen sind, auch zu dem Zusammenhang mit sozialen Merkmalen der Lernumwelt – einhergeht (vgl. etwa Lüdtko & Köller, 2002; Köller, 2004a; Lüdtko, Köller, Marsh & Trautwein, 2005; Zeinz & Köller, 2006; Fischbach, Brunner, Krauss & Baumert, 2015). Eine Überprüfung dieser Effekte im Rahmen des hier vorgestellten Modells hätte dabei zunächst nur replizierenden Charakter. Allerdings könnte durch die Aufnahme dieser Facette (möglicherweise in Abgrenzung zu der sozialen und kriterialen Bezugsnorm) dezidierter überprüft werden, wie individuelle Förderung (IGLU: im Leseunterricht in der vierten Klasse) im Unterricht umgesetzt wird – ob diese also beispielsweise vor allem die leistungsbezogene Unterscheidung von Unterrichtsmaterialien (wie sie die hier verwendete Skala ‚Individuelle Förderung‘ vor allem abdeckt) fokussiert oder auch mit Formen des individuellen Feedbacks bzw. der Leistungsbeurteilung einhergeht. Gerade hinsichtlich dessen, dass bisher wenig über die Umsetzung individueller Förderung im Unterricht bekannt ist (Fischer, 2014; Solzbacher, 2008; Altrichter et al., 2009; Bräu, 2007, vgl. auch Kapitel 4 und 5), unterliegt solch vertiefenden Analysen eine große bildungspolitische und mitunter schulpraktische Bedeutung.

Eine weitere mögliche Ergänzung des Modells ergibt sich vor allem aus der umgekehrten Kausalinterpretation des negativen Zusammenhangs zwischen der Häufigkeit, mit der individualisiert wird und der klassendurchschnittlichen Leseleistung der Schülerinnen und Schüler. Wie in Kapitel 13.2 beschrieben, könnte unter anderem daraus abgeleitet werden,

---

<sup>82</sup> Dem stehen die soziale Bezugsnorm (Vergleich der Leistungen eines Schulkindes bspw. mit den klassendurchschnittlichen Leistungen) und die sachliche bzw. kriteriale Bezugsnorm (Vergleich der Leistungen eines Schulkindes mit vorab festgelegten Standards, beispielsweise Kompetenz- oder Bildungsstandards) gegenüber.

dass Lehrerinnen und Lehrer insbesondere dann zu Maßnahmen des adaptiven Unterrichtens tendieren, wenn das klassendurchschnittliche Leistungsniveau gering ist<sup>83</sup> – in der individuellen Förderung also auch eine Reaktionsform auf die (leistungsbezogene) Klassenzusammensetzung zu sehen wäre. Möglich wäre darüber hinaus, dass Lehrende in Klassen, die eine sehr breite Leistungs*verteilung* zeigen, die einzelnen Lernenden individuell fördern, da das gemeinsame Unterrichten einer sehr heterogenen Schülerschaft als nicht zielführend eingestuft wird. Beide Annahmen wurden bisher lediglich als Ursache für den negativen Zusammenhang zwischen individueller Förderung und Leistung *interpretiert*, ihre empirische Überprüfung steht noch aus. Dies müsste idealerweise anhand von Längsschnittuntersuchungen erfolgen, da nur hier für Eingangsniveaus der Klassenzusammensetzung kontrolliert und solche Ursache-Wirk-Beziehungen nachgewiesen werden können. Solch eine Analyse würde dann einen Rückschluss dahingehend erlauben, unter welchen Bedingungen Maßnahmen des adaptiven Unterrichtens durch Lehrkräfte eingesetzt werden. Gerade hinsichtlich des hier bestätigten Eindrucks, dass solche didaktischen Settings bisher wenig in Deutschland etabliert sind (vgl. etwa Altrichter et al., 2009; Klieme & Warwas, 2011; Bräu, 2007; Fischer, 2014), gleichzeitig aber als „Schlüssel zur Lösung bildungspolitischer und pädagogischer Probleme“ (Klieme & Warwas, 2011, S. 805) verstanden werden, könnte hiermit ein wichtiger, wenn auch nur einen Teilaspekt betreffender Beitrag für die Umsetzung individueller Förderung im Unterricht geleistet werden.

Darüber hinaus wurde der negative Zusammenhang zwischen der individuellen Förderung und der Leistung der Schülerinnen und Schüler in Anlehnung an Gruehn (1995, 2000) dahingehend interpretiert, dass das adaptive Unterrichten mit einem hohen zeitlichen Aufwand, insbesondere bezüglich der Instruktion der einzelnen Schülerinnen und Schüler einhergeht, diese dann – verglichen mit Kindern in anderen didaktischen Settings – weniger Zeit für das Lernen und Bearbeiten von Aufgaben haben, was sich (über längere Zeit) in einer geringeren Leistung widerspiegeln könnte. Dabei wurde der sogenannten ‚Time on Task‘, also der „vom Schüler tatsächlich aufgewendete[n] Zeit für das Erreichen der angestrebten Ziele“ (deutsch: „echte Lernzeit“, beides Meyer, 2005, S. 40) ein deutlicher Zusammenhang

---

<sup>83</sup> Wie in Kapitel 13.2 beschrieben, kann dieser Effekt auch dahingehend interpretiert werden, dass individuelle Förderung mit einem großen zeitlichen Aufwand einhergeht – Zeit, die möglicherweise eher für Instruktionen und nicht für das Lernen genutzt wird (wenig ‚Time on Task‘), was sich in geringeren Leistungen widerspiegeln kann.

mit der Leistungsentwicklung nachgewiesen<sup>84</sup> (Meyer, 2005; Anderson, 1995; Fisher, 1995; Helmke, 2003). Ob solch ein Effekt möglicherweise die leistungsbezogene Wirkung der individuellen Förderung überlagert, könnte in einem Studiendesign, in welchem für die Zeit, die Schülerinnen und Schüler für das Bearbeiten von Aufgaben bzw. dem Erwerb von Fähigkeiten aufwenden, kontrolliert werden. Sollte sich unter solch zeit-standardisierten Settings ein positiver Effekt der individuellen Förderung zeigen, könnten daraus Maßnahmen abgeleitet werden, die das Ziel verfolgen, die ‚echte Lernzeit‘ im individualisierten Unterricht zu erhöhen.

Im Allgemeinen zeigt sich in den hier vorgestellten Analysen und den theoretischen Darstellungen ein verhaltenes Bild hinsichtlich der Wirkung individueller Förderung auf die Lesekompetenz und ihren Einsatz im Unterricht. So konnte in Kapitel 6 anhand bisher zu diesem Thema veröffentlichter Studien gezeigt werden, dass diese im Fach Deutsch (bzw. der jeweiligen Landessprache) eher mit schreibbezogenen, weniger jedoch mit lesebezogenen Fähigkeiten der Schülerinnen und Schüler in Zusammenhang steht. Zurückgeführt wurde dies auf den Umstand, dass es sich bei der Lesekompetenz um eine zentrale Kulturtechnik handelt, für deren Erwerb möglicherweise außerschulische, vor allem auf das familiäre Umfeld bezogene Lernsituationen von größerer Bedeutung sind als beispielsweise für die mathematische oder naturwissenschaftliche Kompetenzen. Dies könnte darüber hinaus auch eine Ursache dafür sein, dass der Zusammenhang zwischen Merkmalen der familiären Herkunft und dem Schulerfolg der Viertklässlerinnen und Viertklässler so extrem robust ausfällt, dass er nicht systematisch zwischen den Klassen variiert. Zudem wurde, wie in Kapitel 13.2 dargestellt, angenommen, dass der Leseunterricht möglicherweise nicht die Domäne darstellt, in der Formen der individuellen Förderung vermehrt Anwendung finden. All dies sollte als Anlass genommen werden zu prüfen, wie sich dasselbe Modell für den Erwerb einer anderen Kompetenz, beispielsweise der mathematischen, wie sie in TIMSS ebenfalls für Viertklässlerinnen und Viertklässler erfasst wurde, abbildet. Anzunehmen wäre, dass insbesondere der Herkunftseffekt weniger konstant ausfällt, so dass er dann mit der Häufigkeit individueller Förderung in Verbindung gebracht werden kann.

---

<sup>84</sup> Anzumerken sei, dass dieser sonst positiv lineare Effekt ab einem gewissen Zeitpunkt nur noch eine minimale Steigung erfährt, zusätzliche Unterrichtszeit dann also kaum noch zu einer Leistungssteigerung führt (vgl. Helmke, 2003).

In der vorliegenden Arbeit wurde individuelle Förderung – theoretisch – verstanden als Möglichkeit, *alle* Schülerinnen und Schüler einer Klasse dahingehend zu fördern, dass sie das für sie persönlich jeweils maximale Lern- und Leistungsniveau erreichen (vgl. Kapitel 3). Nicht untersucht wurde, ob einzelne Schülergruppen möglicherweise in besonderem Ausmaß von adaptiven Unterrichtssettings profitieren. Auf einen diesbezüglichen Aspekt wird bei Schulz-Heidorf und Solheim (2016) hingewiesen: In Anlehnung an Konzepte und Theorien der Sprachkompetenz (vgl. v.a. Bernstein, 1975 und Cummins, 1984, 2000) könnten Kinder und Jugendliche aus bildungsferneren Elternhäusern gerade in sprachbezogenen Unterrichtsfächern einen deutlichen Nachteil aufweisen, da sie nicht so sehr mit der in der Schule und in bildungsnahen Familien vorhandenen Bildungssprache vertraut sind. Ähnlich wie Bernstein (elaborierter vs. restringierter Sprachcode, vgl. Bernstein, 1975) unterscheidet Cummins (1984, 2000) dabei die Sprachfähigkeiten nach „basic interpersonal communication skills“ (BICS), also basale zwischenmenschliche Kommunikationsfähigkeiten, die auch als ‚Alltags-sprache‘ bezeichnet werden und vor allem in kontextuell stark eingebetteten und kognitiv wenig anspruchsvollen Situationen Anwendung finden und die „cognitive academic language proficiency“ (CALP). Bei letzterer handelt es sich um die sogenannte Bildungssprache, die auch als akademische Sprachfähigkeit bezeichnet wird und in kontextreduzierten Situationen, in denen kognitiv anspruchsvolle Inhalte vermittelt werden sollen (bspw. in der Schule), herangezogen wird (für einen Überblick s. auch Berendes, Dragon, Weinert, Heppt & Stanat, 2013). Diese Einteilung ermöglicht es zu begründen, warum Kindern aus bildungsfernen oder anderssprachigen Familien auf den Schulerfolg bezogene Rückstände aufweisen: So reiche die in diesen Elternhäusern vorrangig gesprochene Alltagssprache nicht aus, um der fach- und bildungssprachlich orientierten Sprache der Schule zu folgen (vgl. u.a. Dehn, 2011; Baumert & Schümer, 2001; Berendes et al., 2013; Gantefort, 2013; Gogolin, 2006, 2007, 2009). Da Kinder aus bildungsnahen Elternhäusern mit der in der Schule gebräuchlichen Bildungssprache bereits vertraut sind, weisen sie schon zu Schulbeginn einen (insbesondere auf die sprachlichen Komponenten schulischer Lernsituationen bezogenen) Vorteil gegenüber Kindern aus bildungsfernen Haushalten auf. Wie Schulz-Heidorf und Solheim (2016) argumentieren, könnte dies dazu führen, dass Lehrende zwar angeben, ihren Unterricht auf den Einzelschüler bzw. die Einzelschülerin abzustimmen – diese Passung jedoch dadurch reduziert wird, da ihr unterliegende Kommunikationsmechanismen dazu

führen, dass einigen Lernenden bereits die Dechiffrierung der Bildungssprache misslingt<sup>85</sup>. Möglich wäre es sogar, dass durch eine Kombination aus Unkenntnis hinsichtlich der Verwendung und Wirkung des akademischen Sprachgebrauchs in der Schule und der Individualisierung im Unterricht bestehende soziale Disparitäten noch ausgeweitet werden, wenn einige Kinder in besonderem Maß von solchen Settings profitieren, andere hingegen bereits an dem sprachlichen Verständnis beispielsweise der Instruktion, aber auch des Lernmaterials scheitern.

Möglicherweise zeigen sich aber auch divergierende Effekte individueller Förderung bei Schülerinnen und Schülern, die in unterschiedlichem Ausmaß über die insbesondere für das eigenständige Lernen benötigten Kompetenzen (lernbezogene Selbstregulation, Lernstrategien, Aufrechterhalten der Lernmotivation etc., vgl. Darstellungen in Kapitel 4) verfügen. Da bisher wenig über solch differenzielle Effekte individueller Förderung bekannt ist, sollten diese Gegenstand vertiefender Forschungsarbeiten sein – gerade auch, da ihr mittlerweile eine große Bedeutung für die Begegnung von Heterogenität im Schulalltag beigemessen wird.

#### **14.2 Implikationen für die schulische Praxis und den bildungspolitischen Diskurs um individuelle Förderung im Unterricht**

In den hier vorgestellten Analysen wurde der individuellen Förderung ein negativer Zusammenhang mit der Leseleistung der Schülerinnen und Schüler nachgewiesen. Abgeleitet werden kann daraus jedoch nicht, dass Lehrerinnen und Lehrer ihren Unterricht per se nicht an die individuellen Lernvoraussetzungen, Leistungsstände, Begabungen, Neigungen und Interessen der einzelnen Kinder und Jugendlichen anpassen sollten, um eine ungünstige Leistungsentwicklung zu verhindern. Dies hat verschiedene Gründe: Zum einen gelten die hier vorgestellten Befunde lediglich für den Zusammenhang mit der Lesekompetenz. Und auch die anhand der IGLU-Skala erfasste Form der Individualisierung deckt fast ausschließlich Merkmale der Anpassung des Lernmaterials an die *Leistungsstände* der einzelnen Lernenden ab – individuelle Förderung beinhaltet jedoch noch weitere Facetten wie die Planung der Passung des Unterrichtsmaterials und der Lehr- und Lernformen an die Lernvoraus-

---

<sup>85</sup> Dies fasst beispielsweise auch Koch (2008) zusammen: So sei in der Schulpraxis lange nicht erkannt worden, „dass Migrantenkinder bei der Einschulung häufig nicht in der Lage sind, dem Unterricht sprachlich zu folgen: Da sie auf dem Kommunikationsniveau von BICS mithalten konnten, ging man davon aus, dass ihre Deutschkenntnisse ausreichen, um auch den komplexeren Ansprüchen der Schulsprache entsprechen zu können“ (Koch, 2008, S. 44, nach Berendes, Dragon, Weinert, Heppt & Stanat, 2013, S. 22).

setzungen, Lernstile, Interessen und Begabungen der Schülerinnen und Schüler (vgl. Kapitel 3), die hier (aufgrund der verfügbaren Daten) nicht Gegenstand der Analysen sein konnten. Zudem erlaubt die Datenstruktur (lediglich ein Messzeitpunkt, Querschnitt; keine Unterrichtsstudie mit Prozessdaten) keine Rückschlüsse auf Ursache-Wirk-Beziehungen. So konnte nicht überprüft werden, wie sich die Anpassung der Lernmaterialien an die Leistungsniveaus der einzelnen Viertklässlerinnen und Viertklässler auf die *Entwicklung* der Lesekompetenz auswirkt – oder ob das klassendurchschnittliche Leistungsniveau im Lesen möglicherweise dazu führt, dass Lehrpersonen ihren Unterricht häufiger individualisieren. Wenngleich Maßnahmen der individuellen Förderung, wie in Kapitel 5 gezeigt wurde, nur selten Bestandteil des Leseunterrichts sind, könnte die inverse Interpretation ihres Effekts auf die Lesekompetenz als ein Hinweis dahingehend verstanden werden, dass der bildungspolitisch geforderten individuellen Förderung vor allem von Kindern, die leistungsbezogene Defizite aufweisen, im Leseunterricht tendenziell nachgekommen wird.

Dass sie jedoch selten Anwendung findet, steht möglicherweise in Zusammenhang mit ihrer bislang weitestgehend fehlenden konzeptionellen Ausgestaltung (vgl. Altrichter et al., 2009; Klieme & Warwas, 2011; Bräu, 2007; Fischer, 2014). In einem Interview bezüglich der Befunde der Studie „Individuelle Förderung in Ganztagschulen – inwiefern gelingt sie bei Kindern in schwierigen Lebenssituationen? Studie zu Chancen und Problematiken besonderer erzieherischer Förderung in Ganztagschulen“ konstatiert Prof. Dr. W. Böttcher, dass „es [...] aber völlig unklar [sei], was das [die individuelle Förderung, Anm. KSH] genau bedeutet und wie man es umsetzen kann [...]. Die Professionellen im Bereich der Bildung und Erziehung brauchen konkrete Vorstellungen davon, was wirkt, was gemeint ist und wie man eine gute Idee auch tatsächlich umsetzen kann“ (Bundesministerium für Bildung und Forschung, 2012, S. 25). Es erscheint daher nicht verwunderlich, dass Lehrende über eine entsprechende Unsicherheit berichten, ihren Unterricht zu individualisieren (vgl. Solzbacher, 2008). Begegnet werden könnte dies zum einen mit der Ausgestaltung solcher Konzepte und Förderinstrumente, die den Lehrerinnen und Lehrern einen didaktischen Handlungsrahmen eröffnen. Andererseits müsste dies – sobald solche Konzepte erarbeitet und praktisch wie empirisch überprüft wurden – in die Aus- und Fortbildung von Lehrenden integriert werden, um diese Unsicherheit zu reduzieren und die Implementation der individuellen Förderung im Unterricht zu verstärken.

Hinsichtlich der möglichen Beeinflussung des Herkunftseffekts seitens der Lehrperson ergeben sich bezüglich der hier für die Lesekompetenz überprüften Effekte gravierende Befunde: Da der Einfluss der sozioökonomischen Stellung der Eltern und der häuslichen Lesesozialisation auf die Lesekompetenz und das Leseselbstkonzept nicht systematisch zwischen den Klassen variiert, kann dieser (lesespezifische) Herkunftseffekt folglich zwischen Klassen bzw. Schulen auch nicht durch Maßnahmen der individuellen Förderung beeinflusst werden. Das heißt im Umkehrschluss, dass – unabhängig von der besuchten Klasse bzw. Schule – den Eltern und dem häuslichen, auf das Lesen bezogenen Anregungsgehalt eine gleichbleibend hohe Bedeutung für die Lesekompetenzentwicklung ihrer Kinder zukommt. Dies konstatieren ebenfalls Hertel, Jude und Naumann (2010), nach welchen „[d]ie häusliche Lernumgebung und die Leseförderung im Elternhaus vor und während der Schulzeit [...] eine wesentliche Rolle in der kindlichen Schriftsprachkompetenz [einnehmen]“, da Kinder dort „die Erfahrungen [machen], dass Sprache, Schrift und Lesen dem alltäglichen Leben zugehörig sowie gesamtgesellschaftlich und kulturell bedeutsam sind“ (beides Hertel et al., 2010, S. 255). Gezeigt werden konnte in der vorliegenden Arbeit, dass etwa ein Sechstel der Varianz in der Leseleistung zwischen den Schülerinnen und Schülern in den Klassen auf die sozioökonomische Stellung und mit ihr die Lesesozialisation innerhalb von Familien zurückgeführt werden kann. Unter diesen Bedingungen scheint es ratsam, Eltern möglichst frühzeitig aktiv in die Leseförderung der Kinder einzubinden, sie hierfür zu sensibilisieren und praktische Hinweise beispielsweise für das Vorlesen und die Auswahl von geeigneten Büchern zu geben (wie es beispielsweise das Teilprojekt ‚Family Literacy‘ des BLK-Programms FörMig fokussiert; weitere Ausführungen auch bei Nickel, 2007). Die flächendeckende Umsetzung solcher Konzepte könnte sich dann, frühzeitig eingesetzt, günstig, d.h. mindernd auf den lesebezogenen Herkunftseffekt auswirken.

In Anbetracht der Bedeutung, die Eltern für den lesebezogenen Kompetenzerwerb ihrer Kinder zukommt, scheint ihre Sensibilisierung dahingehend also sinnvoll. Andererseits muss ebenfalls darauf geachtet werden, dass sich ihr Einfluss auf schulbezogene Lernsituationen nicht maximiert: Beispielsweise wird durch die Vergabe von Hausaufgaben ein Teil der Förderung von Schülerinnen und Schülern auf den privaten Bereich übertragen – deren Betreuung und Ausgestaltung (wie eingangs gezeigt) mitunter von sozioökonomischen Merkmalen der Familie abhängt, also den Ausbau sozialer Disparitäten weiter fördert. Dies betrifft in gleicher Weise zusätzliche Unterstützungsangebote bei Leistungsdefiziten, etwa in

Form von privat organisierten (und finanzierten) Nachhilfestunden statt schulintern organisiertem Förderunterricht. Eine Möglichkeit, dem zu begegnen, können Ganztagsangebote sein. Mit ihnen einher geht die Annahme, dass familiären Lernsituationen weniger prognostische Relevanz für den Schulerfolg zukommt, da diese beispielsweise durch eine Hausaufgabenbetreuung, wie sie charakterisierendes Merkmal fast aller Ganztagsangebote ist (vgl. Holtappels, Kamski & Schnetzer, 2007), vermehrt in der Schule stattfinden. Ob in solchen oder anderen, auf die zusätzliche Förderung der Schülerinnen und Schüler ausgelegten Settings: Die hier vorgestellten Analysen zeigen, dass die vermehrte Entkopplung von schulischen Lernsituationen (bspw. Hausaufgaben) und der (alleinigen) elterlichen Unterstützung in deren Bewältigung nicht zu vernachlässigen ist, um herkunftsbedingten Unterschieden weniger Bedeutung für den lesebezogenen Schulerfolg zukommen zu lassen.

Unabhängig von diesen herkunftsbedingten Effekten zeigt sich für die praktische Bedeutung der individuellen Förderung im Unterricht ein verhaltenes Bild: Sie wird in Deutschland bisher wenig im lesebezogenen Unterricht der vierten Klassen eingesetzt, zeigt – anders als erwartet – keinen (positiven) Zusammenhang mit dem Leseselbstkonzept und verändert nicht die Wahrnehmung der Schüler-Lehrer-Beziehung und der sozialen Integration der Schülerinnen und Schüler in den Klassenverbund. Zudem steht sie in einem negativen Zusammenhang mit der Lesekompetenz der Viertklässlerinnen und Viertklässler. Zwar konnte hier nicht abschließend beurteilt werden, ob individuelle Förderung tatsächlich zu geringeren Leistungen führt (beispielsweise durch die mit den längeren, individualisierten Instruktionen einhergehende verminderte ‚echten Lernzeit‘) oder ob eine geringe klassendurchschnittliche Lesekompetenz dazu führt, dass Lehrende vermehrt Maßnahmen des adaptiven Unterrichtens ergreifen. Sollte sich diese zweite Annahme bestätigen, würde dies bedeuten, dass individuelle Förderung eher herangezogen wird, um lernbezogenen Defiziten der Schülerinnen und Schüler zu begegnen und nicht – ihrer hier vorgestellten Definition folgend – um den Unterricht an die Voraussetzungen, Fähigkeiten und Interessen *aller* Schülerinnen und Schüler anzupassen, ihn damit für den einzelnen Lerner, die einzelne Lernerin bedeutsamer zu machen und so den individuellen Lernvoraussetzungen, wie sie vor allem durch familiäre Merkmale bedingt sind, weniger prognostische Relevanz für den Schulerfolg zukommen zu lassen.

In Verbindung mit der Feststellung, dass sich praktisch erprobte und empirisch auf ihre Wirksamkeit hin bestätigte Konzepte der individuellen Förderung bisher kaum finden lassen (Altrichter et al., 2009; Klieme & Warwas, 2011; Bräu, 2007; Fischer, 2014), scheinen die sich aus dieser Arbeit ableitenden Implikationen zuallererst bildungsadministrativer Natur zu sein: So bedarf es zunächst der Ausgestaltung eines solchen Konzepts, entsprechender Förderinstrumente und didaktischer Handlungsrahmen, denen ein homogenes Verständnis individueller Förderung zugrunde liegt. Dies würde dann eine einheitliche Operationalisierung, gerade auch im Rahmen von Large-Scale-Assessments wie IGLU erlauben. Solch ein Konzept müsste anschließend einer praktischen und empirischen Überprüfung unterzogen werden, insbesondere hinsichtlich der Wirkung individueller Förderung auf die verschiedenen Maße des Schulerfolgs, differenziert nach Schulfächern und Jahrgängen. Vertieft werden könnten solche Analysen dann dahingehend, unter welchen Bedingungen Lehrende Maßnahmen der individuellen Förderung ergreifen (und nicht ergreifen), warum sie (zumindest im Leseunterricht der vierten Klassen) bisher wenig Anwendung findet und mit was für Hürden ihre konsequente Umsetzung verbunden ist. Erst dann kann geurteilt werden, ob individuelle Förderung das Potenzial besitzt, „Schlüssel zur Lösung bildungspolitischer und pädagogischer Probleme“ (Klieme & Warwas, 2011, S. 805) zu sein – und konsequenterweise Eingang in die Aus- und Fortbildung von Lehrerinnen und Lehrern finden sollte.

## Verzeichnisse

### Literaturverzeichnis

- Abs, H. J. (2005). LehrerbildnerInnen im Rahmen des Vorbereitungsdienstes. *Journal für Lehrerinnen- und Lehrerbildung*, 5, 21–31.
- Abt Gürber, N. (2011). *Bereichsspezifische Selbstkonzepte bei Kindern in der Schuleingangsstufe: Zusammenhänge mit Leistungen und Wohlbefinden in der Schule*. Herdern: Selbstverlag.
- Aiello, N. C. & Wolfle, L. M. (1980). *A meta-analysis of individualized instruction in science*. Boston: Paper presented at the Annual Meeting of the American Educational Research Association, April 1980.
- Alba, R. D., Handl, J. & Müller, W. (1994). Ethnische Ungleichheiten im deutschen Bildungssystem. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 46, 209–237.
- Altrichter, H., Trautmann, M., Wischer, B., Sommerauer, S. & Doppler, B. (2009). Unterrichten in heterogenen Gruppen: Das Qualitätspotenzial von Individualisierung, Differenzierung und Klassenschülerzahl. In W. Specht (Hrsg.), *Nationaler Bildungsbericht Österreich 2009* (S. 341–360). Graz: Leykam.
- Anderson, G. J. (1973). *The assessment of learning environments: A manual for the Learning Environment Inventory and the My Class Inventory*. Halifax: Atlantic Institute of Education.
- Anderson, L. W. (1995). Time. Allocated and instructional. In L. W. Anderson (Hrsg.), *International Encyclopedia of Teaching and Teacher Education* (2. Aufl., S. 204–207). Oxford: Pergamon.
- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T. C. (2007). Schullaufbahnpräferenzen am Ende der vierten Klassenstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes et al. (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 272–297). Münster: Waxmann.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–131). Opladen: Leske + Budrich.
- Atash, M. N. & Dawson, G. O. (1986). Some effects of the ISCS program: A meta-analysis. *Journal of Research in Science Teaching*, 23 (5), 377–385.
- Atkinson, J. W. (1974). Motivational determinants of intellectual performance and cumulative achievement. In J. W. Atkinson & J. Raynor (Hrsg.), *Achievement and performance* (S. 389–410). Washington, D. C.: Winston.
- Autorengruppe Bildungsberichterstattung (Hrsg.) (2010). *Bildung in Deutschland 2010. Ein indikatorengestützter Bericht mit einer Analyse zu Perspektiven des Bildungswesens im demographischen Wandel. Im Auftrag der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland und des Bundesministeriums für Bildung und Forschung*. Bielefeld: Bertelsmann.
- Backhaus, K. (2006). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (11. Aufl.). Berlin: Springer-Verlag.

- Bangert, R. L., Kulik, J. A. & Kulik, C.-L. C. (1983). Individualized systems of instruction in secondary schools. *Review of Educational Research*, 53 (2), 143–158.
- Baumert, J. & Maaz, K. (2006). Das theoretische und methodische Konzept von PISA zur Erfassung sozialer und kultureller Ressourcen der Herkunftsfamilie: Internationale und nationale Rahmenkonzeption. 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. 11–29). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–410). Opladen: Leske + Budrich.
- Baumert, J. & Schümer, G. (2002). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb im nationalen Vergleich. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele et al. (Hrsg.), *PISA 2000 – die Länder der Bundesrepublik Deutschland im Vergleich* (S. 159–201). Opladen: Leske + Budrich.
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 15–68). Opladen: Leske + Budrich.
- Beck, E., Baer, M., Guldemann, T., Bischoff, B., Brühwiler, C., Müller, P. et al. (2008). *Adaptive Lehrkompetenz. Analyse und Struktur, Veränderbarkeit und Wirkung handlungssteuernden Lehrerwissens*. Münster: Waxmann.
- Becker, R. (2010). Soziale Ungleichheit von Bildungschancen und Chancengerechtigkeit – eine Reanalyse mit bildungspolitischen Implikationen. In R. Becker (Hrsg.), *Bildung als Privileg. Erklärungen und Befunde zu den Ursachen der Bildungsungleichheit* (S. 161–189). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Beckmann, H.-K. (1978). *Leistung in der Schule*. Braunschweig: Westermann.
- Belser, H. & Küsel, G. (1976). Zum Sitzenbleiber-Problem an Volksschulen. Empirische Untersuchungen an Volksschulen in Hamburg. In R. Biermann (Hrsg.), *Schulische Selektion in der Diskussion* (S. 101–115). Bad Heilbrunn: Klinkhardt.
- Bennett, N. (1976). *Teaching styles and pupil progress*. London: Open Books.
- Berendes, K., Dragon, N., Weinert, S., Heppt, B. & Stanat, P. (2013). Hürde Bildungssprache? Eine Annäherung an das Konzept "Bildungssprache" unter Einbezug aktueller empirischer Forschungsergebnisse. In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik. Interdisziplinäre Perspektiven*. Münster: Waxmann.
- Bernstein, B. (1975). *Sprachliche Kodes und soziale Kontrolle*. Düsseldorf: Schwann.
- Bessoth, R. (1989). *Verbesserung des Unterrichtsklimas*. Neuwied: Luchterhand.
- Betz, T. (2006a). *Zur Bildungsbedeutsamkeit von Familie: Informelle Bildung und Schulerfolg*. Arbeitspapiere des Zentrums für sozialpädagogische Forschung der Universität Trier. Arbeitspapier II/19. Zugriff am 30.10.2014. Verfügbar unter [http://www-neu.uni-trier.de/fileadmin/fb1/prof/PAD/SP2/Arbeitspapiere/Arbeitspapier19\\_betz.pdf](http://www-neu.uni-trier.de/fileadmin/fb1/prof/PAD/SP2/Arbeitspapiere/Arbeitspapier19_betz.pdf)

- Betz, T. (2006b). Ungleiche Kindheit – Ein (erziehungswissenschaftlicher) Blick auf die Verschränkung von Herkunft und Bildung. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 26 (1), 52–68.
- Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment*, 1 (2), 1–12.
- Bloom, B. S. (1971). *Mastery learning*. New York: Holt, Rinehart, & Winston.
- Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.
- Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research*, 31 (6), 445–457.
- Bohl, T., Batzel, A. & Richey, P. (2011). Öffnung-Differenzierung-Adaptivität: Charakteristika, didaktische Implikationen und Forschungsbefunde verwandter Unterrichtskonzepte zum Umgang mit Heterogenität. *Schulpädagogik heute*, 2 (4), 1–23.
- Bonfadelli, H. & Fritz, A. (1993). Lesen im Alltag von Jugendlichen. In H. Bonfadelli, A. Fritz & R. Köcher (Hrsg.), *Leseerfahrung und Lesekarrieren* (S. 7–213). Gütersloh: Bertelsmann.
- Bonsen, M., Bos, W., Gröhlich, C., Harney, B., Imhäuser, K., Makles, A. et al. (2010). *Zur Konstruktion von Sozialindizes. Ein Beitrag zur Analyse sozialräumlicher Benachteiligung von Schulen als Voraussetzung für qualitative Schulentwicklung* (Bildungsforschung, Bd. 31). Bonn: BMBF.
- Bortz, J. & Döring, N. (2009). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Aufl.). Heidelberg: Springer-Medizin-Verlag.
- Bos, W., Bonsen, M., Gröhlich, C., Guill, K., May, P., Rau, A. et al. (2007). *KESS 7. Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 7*. Hamburg: Behörde für Bildung und Sport.
- Bos, W., Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Lesekompetenzen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 91–135). Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.) (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Valtin, R. & Walther, G. (2002). *IGLU/PIRLS: Information*. Zugriff am 22.07.2014. Verfügbar unter <http://www.grundschulpaedagogik.uni-bremen.de/lehre/2001ws/gspaed/pisa/iglu2001infobroschuere.pdf>
- Bos, W. & Scharenberg, K. (2010). Lernentwicklung in leistungshomogenen und -heterogenen Schulklassen. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (S. 173–194). Münster: Waxmann.
- Bos, W., Stubbe, T. C. & Buddeberg, M. (2010). Gibt es eine armutsbedingte Bildungsbenachteiligung? Die Operationalisierung verschiedener Indikatoren der sozialen Herkunft in der empirischen Bildungsforschung. In D. H. Rost (Hrsg.), *Intelligenz, Hochbegabung, Vorschulerziehung, Bildungsbenachteiligung* (S. 165–208). Münster: Waxmann.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects of western society*. New York: Wiley.

- Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In *Soziale Ungleichheiten. Sonderband 2 der Zeitschrift Soziale Welt* (S. 183–198). Göttingen: Schwarz.
- Bourdieu, P. (1989). Antworten auf einige Einwände. In K. Eder (Hrsg.), *Klassenlage, Lebensstil und kulturelle Praxis. Beiträge zur Auseinandersetzung mit Pierre Bourdieus Klassentheorie* (S. 395–411). Frankfurt am Main: Suhrkamp.
- Bourdieu, P. (1992). *Die verborgenen Mechanismen der Macht*. Hamburg: VSA-Verlag.
- Bradley, R. & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53 (1), 371–399.
- Bräu, K. (2007). Die Betreuung der Schüler im individualisierenden Unterricht der Sekundarstufe. Strategien und Handlungsmuster der Lehrenden. In K. Rabenstein & S. Reh (Hrsg.), *Kooperatives und selbstständiges Arbeiten von Schülern. Zur Qualitätsentwicklung von Unterricht* (S. 173–195). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bremerich-Vos, A., Tarelli, I. & Valtin, R. (2012). Das Konzept von Lesekompetenz in IGLU 2011. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 69–89). Münster: Waxmann.
- Brügelmann, H. (2000). Wie verbreitet ist offener Unterricht? In O. Jaumann-Graumann & W. Köhnlein (Hrsg.), *Lehrerprofessionalität – Lehrerprofessionalisierung* (Jahrbuch Grundschulforschung, Bd. 3, S. 133–143). Bad Heilbrunn: Klinkhardt.
- Bruns, J. (2014). *Adaptive Förderung in einer elementarpädagogischen Praxis. Empirische Studie zum didaktischen Handeln von Erzieherinnen und Erziehern im Bereich Mathematik*. Münster: Waxmann.
- Büchner, P. (2003). Stichwort: Bildung und soziale Ungleichheit. *Zeitschrift für Erziehungswissenschaft* (1), 5–24.
- Buhs, E. S. & Ladd, G. W. (2001). Peer rejection as antecedent of young children's school adjustment: An examination of mediating processes. *Developmental Psychology*, 37 (4), 550–560.
- Bundesministerium für Bildung und Forschung (2012). *Ganztägig bilden. Eine Forschungsbilanz*. Zugriff am 25.07.2016. Verfügbar unter [http://www.ganztagschulen.org/\\_media/121206\\_BMBF\\_GTS-Forschungsbilanz\\_bf\\_df.pdf](http://www.ganztagschulen.org/_media/121206_BMBF_GTS-Forschungsbilanz_bf_df.pdf)
- Burgess, S., Hecht, S. & Lonigan, C. (2002). Relations of the home literacy environment (HLE) to the development of reading-related abilities: A one-year longitudinal study. *Reading Research Quarterly*, 37 (4), 408–426.
- Bus, A. G., Van Ijzendoorn, M. H. & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65 (1), 1–21.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS and SIMPLIS. Basic concepts, applications and programming*. Mahwah: Lawrence Erlbaum Associates.
- Caro, D. H. & Cortés, D. (2012). Measuring family socioeconomic status: An illustration using data from PIRLS 2006. *Issues and Methodologies in Large-Scale Assessments*, 5, 9–33.
- Carroll, J. B. (1963). A model of school learning. *Teacher College Record*, 64, 723–733.

- Carter, D. R., Chard, D. J. & Pool, J. L. (2009). A family strengths approach to early language and literacy development. *Early Childhood Education Journal*, 36 (6), 519–526.
- Christ, O. & Schlüter, E. (2010). *Strukturgleichungsmodelle mit Mplus. Eine praktische Einführung*. München: Oldenbourg-Verlag.
- Clausen, M. (2002). *Unterrichtsqualität: eine Frage der Perspektive? Empirische Analysen zur Übereinstimmung, Konstrukt- und Kriteriumsvalidität* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 29). Münster: Waxmann.
- Cochran-Smith, M. (1984). *The making of a reader*. Nordwood, NJ: Ablex.
- Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah: Erlbaum.
- Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, 77 (1), 113–143.
- Corno, L. & Snow, R. E. (1986). Adapting teaching to individual differences among learners. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (3. Aufl., S. 605–629). New York: Macmillan.
- Creemers, B. P. M., Scheerens, J. & Reynolds, D. (2000). Theory development in SER. In C. Teddlie & D. Reynolds (Hrsg.), *The international handbook of school effectiveness research* (S. 283–298). London: Falmer Press.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods. A handbook for research on interactions*. New York: Irvington Publishers.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Clevedon: Multilingual Matters.
- Cummins, J. (2000). *Language, power and pedagogy. Bilingual children in the crossfire*. Clevedon: Multilingual Matters.
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, 27 (2), 223–238.
- Dehn, M. (2011). Elementare Schriftkultur und Bildungssprache. In S. Fürstenau & M. Gommolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit* (S. 129–152). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Den Brok, P., Brekelmans, M. & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcome. *School Effectiveness and School Improvement*, 15 (3–4), 407–442.
- DeVellis, R. (2012). *Scale development: Theory and applications* (3. Aufl.). Los Angeles: Sage.
- Diamantopoulos, A. & Siguaw, J. A. (2000). *Introducing LISREL*. London: Sage.
- Dickhäuser, O. (2003). Überprüfung des erweiterten Modells des internal/external frame of reference. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 35 (4), 200–207.

- Dickhäuser, O. (2006). Fähigkeitsselbstkonzepte: Entstehung, Auswirkung, Förderung. *Zeitschrift für Pädagogische Psychologie*, 20, 5–8.
- Diefenbach, H. (2002). Bildungsbeteiligung und Berufseinmündung von Kindern und Jugendlichen aus Migrantenfamilien. Eine Fortschreibung der Daten des Sozio-Ökonomischen Panels (SOEP). In Sachverständigenkommission 11. Kinder- und Jugendbericht (Hrsg.), *Migration und die europäische Integration. Herausforderungen für die Kinder- und Jugendhilfe* (S. 9–70). München: Verlag Deutsches Jugendinstitut.
- 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* (Zeitschrift für Pädagogik, Beiheft 41, S. 73–92). Weinheim: Beltz.
- Ditton, H., Krüsken, J. & Schauenberg, M. (2005). Bildungsungleichheit – Der Beitrag von Familie und Schule. *Zeitschrift für Erziehungswissenschaft*, 8 (2), 285–304.
- Dreesmann, H. (1979). Zusammenhänge zwischen Unterrichtsklima, kognitiven Prozessen bei Schülern und deren Leistungsverhalten. *Zeitschrift für Empirische Pädagogik*, 3, 121–133.
- Dreesmann, H. (1982). *Unterrichtsklima. Wie Schüler den Unterricht wahrnehmen*. Weinheim: Beltz.
- Dreesmann, H., Eder, F., Fend, H., Pekrun, R., Saldern, M. von & Wolf, B. (1992). Schulklima. In K. Ingenkamp & R. S. Jäger (Hrsg.), *Empirische Pädagogik 1970 – 1990. Eine Bestandsaufnahme der Forschung in der Bundesrepublik Deutschland* (S. 655–682). Weinheim: Deutscher Studien-Verlag.
- Du Bois-Reymond, M. (2000). Jugendkulturelles Kapital in Wissensgesellschaften. In H.-H. Krüger & H. Wenzel (Hrsg.), *Schule zwischen Effektivität und sozialer Verantwortung* (Studien zur Schul- und Bildungsforschung, Bd. 9, S. 235–254). Opladen: Leske + Budrich.
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schülerleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60, 163–183.
- Eder, F. (1996). *Schul- und Klassenklima. Ausprägung, Determinanten und Wirkungen des Klimas an höheren Schulen*. Innsbruck: Studien-Verlag.
- Eder, F. (1998). *Linzer Fragebogen zum Schul- und Klassenklima für die 8.-13. Klasse (LSK 8-13)*. Göttingen: Hogrefe.
- Eder, F. (2010). Schul- und Klassenklima. In D. H. Rost (Hrsg.), *Handwörterbuch pädagogische Psychologie* (4. Aufl., S. 694–703). Weinheim: Beltz.
- Ehmke, T., Drechsel, B. & Carstensen, C. H. (2008). Klassenwiederholen in PISA-I-Plus. Was lernen Sitzenbleiber in Mathematik dazu? *Zeitschrift für Erziehungswissenschaft*, 11 (3), 368–387.
- Ehmke, T. & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. *Zeitschrift für Erziehungswissenschaft*, 8 (4), 521–539.
- 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.

- Erikson, R., Goldthorpe, J. H. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, 30, 341–415.
- Fend, H. (1977). *Schulklima: Soziale Einflußprozesse in der Schule*. Weinheim: Beltz.
- Fend, H. (1990). *Vom Kind zum Jugendlichen: Der Übergang und seine Risiken*. Bern: Huber.
- Fend, H. (1997). *Der Umgang mit Schule in der Adoleszenz. Aufbau und Verlust von Lernmotivation, Selbstachtung und Empathie*. Bern: Huber.
- Filipp, S.-H. (Hrsg.) (1979). *Selbstkonzept-Forschung. Probleme, Befunde, Perspektiven*. Stuttgart: Klett-Cotta.
- Fischbach, A., Brunner, M., Krauss, S. & Baumert, J. (2015). Die Bezugsnormorientierung von Mathematiklehrkräften am Ende der Sekundarstufe I: Konvergenz verschiedener Messverfahren und Wirkung auf motivational-affektive Aspekte des Mathematiklernens und Leistung. *Journal for Educational Research Online*, 7 (3), 3–27.
- Fischer, C. (2014). *Individuelle Förderung als schulische Herausforderung*. Berlin: Friedrich-Ebert-Stiftung.
- Fisher, C. W. (1995). Academic learning time. In L. W. Anderson (Hrsg.), *International Encyclopedia of Teaching and Teacher Education* (2. Aufl., S. 430–434). Oxford: Pergamon.
- Fraser, B. J. (1981). *Validity and use of individualized classroom environment questionnaire*. Zugriff am 21.02.2016. Verfügbar unter <http://files.eric.ed.gov/fulltext/ED204351.pdf>
- Fraser, B. J. & Fisher, D. L. (1982). Predicting students' outcomes from their perceptions of classroom psychosocial environment. *American Educational Research Journal*, 19, 498–518.
- Freie und Hansestadt Hamburg & Behörde für Schule und Berufsbildung (2011). *Bildungsplan Grundschule. Deutsch*. Hamburg: Behörde für Schule und Berufsbildung.
- Furrer, C. & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology*, 95 (1), 148–162.
- Gantefort, C. (2013). ‚Bildungssprache‘ - Merkmale und Fähigkeiten im sprachtheoretischen Kontext. In I. Gogolin (Hrsg.), *Herausforderung Bildungssprache – und wie man sie meistert* (FörMig-Edition, Bd. 9, S. 71–105). Münster: Waxmann.
- Ganzeboom, H. B. G., Graaf, P. M. D., Treiman, D. J. & Leeuw, J. de (1992). A Standard International Socio-Economic Index of Occupational Status. *Social Science Research*, 21, 1–56.
- Geiser, C. (2011). *Datenanalyse mit Mplus. Eine anwendungsorientierte Einführung* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gerecht, M. (2010). *Schul- und Unterrichtsqualität und ihre erzieherischen Wirkungen. Eine Sekundäranalyse auf der Grundlage der Pädagogischen Entwicklungsbilanzen* (Empirische Erziehungswissenschaft, Bd. 27). Münster: Waxmann.
- Gerecht, M., Steinert, B., Klieme, E. & Döbrich, P. (2007). *Skalen zur Schulqualität: Dokumentation der Erhebungsinstrumente. Pädagogische Entwicklungsbilanzen mit Schulen (PEB)* (Materialien zur Bildungsforschung, Bd. 17, 2., überarbeitete Aufl.). Frankfurt am Main: GFFP.

- Getzels, J. W. & Thelen, H. (1960). The classroom group as a unique social system. In N. B. Henry (Hrsg.), *The dynamics of instructional groups. Sociopsychological aspects of teaching and learning* (S. 53–82). Chicago: University Press.
- Giaconia, R. M. & Hedges, L. V. (1982). Identifying features of effective open education. *Review of Educational Research*, 52, 579–602.
- Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education*, 5 (1), 351–379.
- Gogolin, I. (2006). Bilingualität und die Bildungssprache der Schule. In P. Mecheril & T. Quehl (Hrsg.), *Die Macht der Sprachen: Englische Perspektiven auf die mehrsprachige Schule* (S. 79–85). Münster: Waxmann.
- Gogolin, I. (2007). Herausforderung Bildungssprache. Textkompetenz aus der Perspektive interkultureller Bildungsforschung. In K.-R. Bausch, E. Burwitz-Melzer, F. G. Königs & H.-J. Krumm (Hrsg.), *Textkompetenzen. Arbeitspapiere der 27. Frühjahrskonferenz zur Erforschung des Fremdsprachenunterrichts* (S. 73–80). Tübingen: Narr.
- Gogolin, I. (2009). „Bildungssprache“ – The importance of teaching language in every school subject. In T. Tajmel & K. Starl (Hrsg.), *Science education unlimited. Approaches to equal opportunities in learning science* (S. 91–102). Münster: Waxmann.
- Gomolla, M. & Radtke, F.-O. (2002). *Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule*. Opladen: Leske + Budrich.
- Graham, S. & Weiner, B. (1996). Theories and principles of motivation. In D. C. Berliner (Hrsg.), *Handbook of educational psychology* (S. 63–84). New York: Macmillan.
- Graumann, O. (2008). Förderung und Heterogenität: Die Perspektive der Schulpädagogik. In K.-H. Arnold, O. Graumann & A. Rakhkockhine (Hrsg.), *Handbuch Förderung* (S. 16–25). Weinheim: Beltz.
- Green, J., Nelson, G., Martin, A. J. & Marsh, H. (2006). The causal ordering of self-concept and academic motivation and its effect on academic achievement. *International Education Journal*, 7 (4), 534–546.
- Grewe, M., Strietholt, R. & Schwippert, K. (2007). Unterrichtsqualität aus Schülersicht. In K. Möller (Hrsg.), *Qualität von Grundschulunterricht. Entwickeln, erfassen und bewerten* (Jahrbuch Grundschulforschung, Bd. 11, S. 179–182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gröhlich, C., Scharenberg, K. & Bos, W. (2009). Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus? *Journal for Educational Research Online*, 1 (1), 86–105.
- Gruehn, S. (1995). Die Vereinbarkeit kognitiver und nichtkognitiver Ziele im Unterricht. *Zeitschrift für Pädagogik*, 41 (4), 531–553.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen. Schüler als Quellen der Unterrichtsbeschreibung*. Münster: Waxmann.
- Guay, F., Marsh, H. W. & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, 95 (1), 124–136.

- Haag, L. & Streber, D. (2014). *Individuelle Förderung* (BildungsWissen Lehramt, Bd. 124). Weinheim: Beltz.
- Hamre, B. K. & Pianta, R. C. (2001). Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Development*, 72 (2), 625–638.
- Hanke, P. (2001). Offener Unterricht in der Grundschule – erforscht? Zum Stand der Forschung zu einem umstrittenen pädagogisch-didaktischen Ansatz. *Erziehung und Unterricht*, 151 (1–2), 200–208.
- Hanke, P. (2005). *Öffnung des Unterrichts in der Grundschule. Lehr-Lernkulturen und orthographische Lernprozesse im Grundschulbereich* (Internationale Hochschulschriften, Bd. 451). Münster: Waxmann.
- Hansford, B. D. & Hattie, J. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, 52, 123–142.
- Hardy, I., Jonen, A., Möller, K. & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology*, 98 (2), 307–326.
- Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. *Studies in Science Education*, 36 (1), 79–103.
- Harnishfeger, A. & Wiley, D. E. (1975). The teaching learning process in elementary schools: A synoptic view. *Curriculum Inquiry*, 6, 5–43.
- Hartinger, A. (2005). Verschiedene Formen der Öffnung von Unterricht und ihre Auswirkung auf das Selbstbestimmungsempfinden von Grundschulkindern. *Zeitschrift für Pädagogik*, 51 (3), 397–414.
- Hartley, S. S. (1977). *Meta-analysis of the effects of individually paced instruction in mathematics*. University of Colorado at Boulder, CO (Unpublished Ph.D.).
- Hattie, J. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement* (Reprinted). London: Routledge.
- Hattie, J. (2013). *Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von „Visible Learning“*. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Heller, K. A. (Hrsg.) (1984). *Leistungsdiagnostik in der Schule* (4. Aufl.). Bern: Huber.
- Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen*. Göttingen: Hogrefe.
- Helmke, A. (2003). *Unterrichtsqualität. Erfassen, bewerten, verbessern*. Seelze: Kallmeyer.
- Helmke, A. (2006). Was wissen wir über guten Unterricht? Über die Notwendigkeit einer Rückbesinnung auf den Unterricht als dem "Kerngeschäft" der Schule. *Pädagogik* (2), 42–45.
- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze-Velber: Kallmeyer.
- Helmke, A. (2014). Forschung zur Lernwirksamkeit des Lehrerhandelns. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 807–821). Münster: Waxmann.
- Helmke, A. & Schrader, F.-W. (1990). Zur Kompatibilität kognitiver, affektiver und motivationaler Zielkriterien des Schulunterrichts – clusteranalytische Studien. In M. Knopf & W.

- Schneider (Hrsg.), *Entwicklung: allgemeine Verläufe – individuelle Unterschiede, pädagogische Konsequenzen. Festschrift zum 60. Geburtstag von Franz Emanuel Weinert* (S. 180–200). Göttingen: Verlag für Psychologie Hogrefe.
- Helmke, A. & Schrader, F.-W. (1998). Entwicklung im Grundschulalter. Die Münchner Studie SCHOLASTIK. *Pädagogik*, 50 (6), 24–28.
- Helmke, A. & Weinert, F. E. (1997a). Unterrichtsqualität und Leistungsentwicklung: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter* (S. 241–251). Weinheim: Beltz.
- Helmke, A. & Weinert, F. E. (1997b). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 71–176). Göttingen: Hogrefe.
- Herber, H.-J. (1983). *Innere Differenzierung im Unterricht*. Stuttgart: Kohlhammer.
- Herder, J. G. (1970). Vitae, non scholae discendum. In W. Flitner (Hrsg.), *Die Erziehung. Pädagogen und Philosophen über die Erziehung und ihre Probleme* (6. Aufl., S. 223–232). Bremen: Schünemann.
- Herrmann, U. (2010). "Freilich thut es dis auf seine Art.". Innere Differenzierung im Lichte reformpädagogischer Erfahrungen und neurowissenschaftlicher Bestätigungen. In S. Bolter & R. Lau (Hrsg.), *Innere Differenzierung in der Sekundarstufe II. Ein Praxishandbuch für Lehrer/innen* (S. 148–157). Weinheim: Beltz.
- Hertel, S., Jude, N. & Naumann, J. (2010). Leseförderung im Elternhaus. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 255–275). Münster: Waxmann.
- Hodges, E. V. E., Boivin, M., Vitaro, F. & Bukowski, W. M. (1999). The power of friendship: Protection against an escalating cycle of peer victimization. *Developmental Psychology*, 35, 94–101.
- Hofmann, F. (2000). *Aufbau von Lernkompetenzen fördern. Neue Wege zur Realisierung eines bedeutsamen pädagogischen Ziels*. Innsbruck: Studien-Verlag.
- Holtappels, H. G., Kamski, I. & Schnetzer, T. (2007). *Ganztagschule im Spiegel der Forschung. Zentrale Ergebnisse der Ausgangserhebung der "Studie zur Entwicklung von Ganztagschulen" (StEG)*. Berlin: DKJS.
- Hong, G. & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27 (3), 205–224.
- Hooper, D., Coughlan, J. & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, 6 (1), 53–60.
- Horak, V. M. (1981). A meta-analysis of research findings on individualized instruction in mathematics. *Journal of Educational Research*, 74 (4), 249–253.
- Hornberg, S., Bos, W., Buddeberg, I., Potthoff, B. & Stubbe, T. C. (2007). Anlage und Durchführung von IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lanke et al. (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 21–46). Münster: Waxmann.

- Hox, J. J. (2010). *Multilevel analysis. Techniques and applications* (2. Aufl.). New York: Routledge.
- Hu, L. T. & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6 (1), 1–55.
- Hurrelmann, K. (2013). Das Schulsystem in Deutschland: Das "Zwei-Wege-Modell" setzt sich durch. *Zeitschrift für Pädagogik*, 59 (4), 455–468.
- Ingenkamp, K. (1972). *Zur Problematik der Jahrgangsklasse*. Weinheim: Beltz.
- International Labour Office (2012). *International standard classification of occupations 2008 (ISCO-08). Structure, group definitions and correspondence tables*. Geneva: International Labour Office.
- International Labour Organization (1969). *International standard classification of occupations. Revised edition 1968*. Geneva: ILO.
- James, W. (1890). *The principles of psychology*. New York: Holt and Macmillan.
- Janke, N. (2006). *Soziales Klima an Schulen aus Lehrer-, Schulleiter- und Schülerperspektive. Eine Sekundäranalyse der Studie "Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 4 (KESS 4)"*. Münster: Waxmann.
- Jerusalem, M. (1985). Selbstkonzeptentwicklung von Kindern und Jugendlichen und der Einfluß perzipierten Lehrerverhaltens. In D. Liepmann & H.-A. Stiksrud (Hrsg.), *Entwicklungsaufgaben und Bewältigungsprobleme in der Adoleszenz. Sozial- und entwicklungspsychologische Perspektiven*. Göttingen: Verlag für Psychologie Hogrefe.
- Jerusalem, M. (2006). Motivationale und volitionale Voraussetzungen des Unterrichts. In K.-H. Arnold (Hrsg.), *Handbuch Unterricht* (S. 575–579). Bad Heilbrunn: Klinkhardt.
- Jerusalem, M. & Klein-Heßling, J. (2002). Soziale Kompetenz. Entwicklungstrends und Förderung in der Schule. *Zeitschrift für Psychologie*, 210 (4), 164–174.
- Jerusalem, M. & Pekrun, R. (Hrsg.) (1999). *Emotion, Motivation und Leistung*. Göttingen: Hogrefe.
- Jerusalem, M. & Schwarzer, R. (1991). Entwicklung des Selbstkonzepts in verschiedenen Lernumwelten. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung* (S. 115–128). Stuttgart: Enke.
- Job, V., Oettingen, G. & Sevincer, A. (2014). Leistungszielorientierung. In M. Wirtz (Hrsg.), *Dorsch - Lexikon der Psychologie* (17. Aufl., S. 993). Bern: Huber.
- Job, V., Sevincer, A. & Oettingen, G. (2014). Lernzielorientierung. In M. Wirtz (Hrsg.), *Dorsch - Lexikon der Psychologie* (17. Aufl., S. 1015). Bern: Huber.
- Johnson, D. W., Johnson, R. T., Buckman, L. A. & Richards, P. S. (1985). The effect of prolonged implementation of cooperative learning on social support within the classroom. *The Journal of Psychology*, 119 (5), 405–411.
- Kahl, T., Buchmann, M. & Witte, E. H. (1977). Ein Fragebogen zur Schülerwahrnehmung unterrichtlicher Lernsituationen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 9, 277–285.

- Kirschhock, E.-M. (2003). *Die Entwicklung schriftsprachlicher Kompetenzen im ersten Schuljahr. Untersuchung im Rahmen eines Vergleichs verschiedener Unterrichtskonzeptionen*. Nürnberg: Friedrich-Alexander-Universität Erlangen-Nürnberg. Zugriff am 01.12.2015. Verfügbar unter <https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/49>
- Klafki, W. (2007). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik* (6. Aufl.). Weinheim: Beltz.
- Klemm, K. (2009). *Klassenwiederholungen – teuer und unwirksam. Eine Studie zu den Ausgaben für Klassenwiederholungen in Deutschland*. Gütersloh: Bertelsmann Stiftung.
- Klieme, E. & Warwas, J. (2011). Konzepte der Individuellen Förderung. *Zeitschrift für Pädagogik*, 57 (6), 805–818.
- Kluczniok, K., Große, C. & Roßbach, H.-G. (2014). Heterogene Lerngruppen in der Grundschule. In W. Einsiedler, M. Götz, A. Hartinger & F. Heinzl (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (4. überarbeitete Aufl., S. 194–200). Stuttgart: UTB.
- Koch, K. (2008). *Zweitspracherwerb von Grundschulkindern nichtdeutscher Herkunftssprache im Kontext institutioneller Unterstützungsleistungen*. Unveröffentlichte Habilitationsschrift, Georg-August-Universität Göttingen.
- Köller, O. (2000). *Leistungsgruppierungen, soziale Vergleiche und selbstbezogene Fähigkeitskognitionen in der Schule. Habilitationsschrift zur Erlangung der venia legendi für das Fach Psychologie an der Universität Potsdam*. Potsdam.
- Köller, O. (2004a). *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- Köller, O. (2004b). *Bildungsstandards in der Bundesrepublik und ihre Bedeutung für die Schul- und Unterrichtsentwicklung*, Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK). Zugriff am 21.07.2014. Verfügbar unter [http://www.dvlfb.de/cms/images/stories/fachtagungen/04fachtagung/prof\\_dr\\_koeller.pdf](http://www.dvlfb.de/cms/images/stories/fachtagungen/04fachtagung/prof_dr_koeller.pdf)
- Köller, O. (2012). What Works Best in School? Hatties Befunde zu Effekten von Schul- und Unterrichtsvariablen auf Schulleistungen. *Psychologie in Erziehung und Unterricht*, 59 (1), 72–78.
- Köller, O., Klemmert, H., Möller, J. & Baumert, J. (1999). Eine längsschnittliche Überprüfung des Modells des Internal/External Frame of Reference. *Zeitschrift für Pädagogische Psychologie*, 13, 128–134.
- Köller, O. & Möller, J. (2000). Spontaneous and reactive attributions following academic achievement. *Social Psychology of Education*, 4 (1), 67–86.
- Köller, O. & Möller, J. (2001). Selbstbezogene Fähigkeitskognitionen im Kontext Schule: Die Rolle unterschiedlicher Referenzrahmen. In R. K. Silbereisen & M. Reitzle (Hrsg.), *Bericht über den 42. Kongress der Deutschen Gesellschaft für Psychologie in Jena 2000* (S. 380–390). Lengerich: Pabst.
- König, J. (2006). *Sekundarschulen als differenzielle Entwicklungsmilieus? Der Umgang mit Schule unter schulorganisatorischen Bedingungen am Beispiel des Klassenklimas, des Selbstkonzepts und der Attribuierung bei Schülerinnen und Schülern in Klasse 7 und 8*. Duisburg: WiKu-Verlag.
- Krammer, K. (2009). *Individuelle Lernunterstützung in Schülerarbeitsphasen. Eine videobasierte Analyse des Unterstützungsverhaltens von Lehrpersonen im Mathematikunterricht*. Münster: Waxmann.

- Krapp, A. (1976). Bedingungsfaktoren der Schulleistung. *Psychologie in Erziehung und Unterricht*, 23 (2), 91–109.
- Krapp, A. (1984). Forschungsergebnisse zur Bedingungsstruktur der Schulleistung. In K. A. Heller (Hrsg.), *Leistungsdiagnostik in der Schule* (4. Aufl., S. 46–62). Bern: Huber.
- Krapp, A., Schiefele, U. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 10 (2), 120–148.
- Krappmann, L. (2002). Kinder und ihre Freunde. In LBS-Initiative Junge Familie (Hrsg.), *Was Kinder wünschen, hoffen und befürchten* (S. 257–274). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krätzschmar, M. (2010). *Selbstkonzepte in altersgemischten Lerngruppen. Eine Längsschnittstudie mit Kontrollgruppen in der Sekundarstufe*. Münster: Waxmann.
- Krüger, H.-H. & Deppe, U. (2009). Peers und Schule – positiver oder negativer Einfluss von Freunden auf schulische Bildungsbiografien? In O. Böhm-Kasper, M. Haring, C. Palentien & C. Rohlf (Hrsg.), *Freundschaften, Cliques und Jugendkulturen. Peer Groups als Bildungs- und Sozialisationsinstanzen* (S. 223–241). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kristen, C. (1999). *Bildungsentscheidungen und Bildungsungleichheit – ein Überblick über den Forschungsstand*. Mannheim: Mannheimer Zentrum für Europäische Sozialforschung.
- Kroesbergen, E. H. & Van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs – a meta-analysis. *Remedial and Special Education*, 24 (2), 97–114.
- Kultusministerium des Landes Nordrhein-Westfalen (Hrsg.) (1989). *Richtlinien und Lehrpläne für die Hauptschule in Nordrhein-Westfalen. Mathematik*. Düsseldorf: Kultusministerium des Landes Nordrhein-Westfalen.
- Kultusministerium des Landes Nordrhein-Westfalen (1993). *Richtlinien und Lehrpläne für das Gymnasium – Sekundarstufe I in Nordrhein-Westfalen. Mathematik*. Düsseldorf: Kultusministerium des Landes Nordrhein-Westfalen.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 51). Münster: Waxmann.
- Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multi-kriteriale 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.
- Kunze, I. & Solzbacher, C. (2008). *Individuelle Förderung in der Sekundarstufe I und II*. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment? *Child Development*, 61 (4), 1081–1100.
- Ladd, G. W. & Coleman, C. C. (1997). Children's classroom peer relationship and early school attitudes: Concurrent and longitudinal associations. *Early Education and Development*, 8, 51–66.

- Ladd, G. W., Kochenderfer, B. J. & Coleman, C. C. (1997). Classroom peer acceptance, friendship, and victimization: Distinct relational systems that contribute uniquely to children's school adjustment? *Child Development*, 68 (6), 1181–1197.
- Ladd, G. W. & Troop-Gordon, W. (2003). The role of chronic peer difficulties in the development of children's psychological adjustment problems. *Child Development*, 74 (5), 1344–1367.
- Lange, B., Kuffner, H. & Schwarzer, R. (1983). *Schulangst und Schulverdrossenheit. Eine Längsschnittanalyse von schulischen Sozialisierungseffekten*. Opladen: Westdeutsche Verlagsgesellschaft.
- Lankes, E.-M. (2004). Leseunterricht in der Grundschule. Unterschiede zwischen Lehrkräften im internationalen Vergleich. *Zeitschrift für Erziehungswissenschaft*, 7 (4), 551–568.
- Lankes, E.-M., Plassmeier, N., Bos, W. & Schwippert, K. (2004). Lehr- und Lernbedingungen in einigen Ländern der Bundesrepublik Deutschland und im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 21–48). Münster: Waxmann.
- Leary, T. (1957). *Interpersonal diagnosis of personality. A functional theory and methodology for personality evaluation*. New York: Ronald Press Company.
- Lee, V. E. & Burkam, D. T. (2002). *Inequality at the starting gate. Social background differences in achievement as children begin school*. Washington, D.C.: Economic Policy Institute.
- Lehmann, R. H., Peek, R., Pieper, I. & Stritzky, R. von (1995). *Leseverständnis und Lesegewohnheiten deutscher Schüler und Schülerinnen*. Weinheim: Beltz.
- Lewin, K. (1963). *Feldtheorie in den Sozialwissenschaften*. Bern: Huber.
- Liedtke-Schöbel, M. (2012). *Grundlagen für die schulinterne Konzeptentwicklung „Individuelle Förderung statt Klassenwiederholung“*. Handreichung des Landesinstituts für Lehrerbildung und Schulentwicklung Hamburg. Hamburg: Landesinstitut für Lehrerbildung und Schulentwicklung.
- Lin, N. & Dumin, M. (1986). Access to occupations through social ties. *Social Networks*, 8, 365–385.
- Lipowsky, F. (2002). Zur Qualität offener Lernsituationen im Spiegel empirischer Forschung. In U. Drews & W. Wallrabenstein (Hrsg.), *Freiarbeit in der Grundschule* (S. 126–159). Frankfurt am Main: Grundschulverband.
- Lipowsky, F. (2015). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (2. Aufl., S. 69–105). Berlin: Springer-Verlag.
- Lipowsky, F., Kastens, C., Lotz, M. & Faust, G. (2011). Aufgabenbezogene Differenzierung und Entwicklung des verbalen Selbstkonzepts im Anfangsunterricht. *Zeitschrift für Pädagogik*, 57 (6), 868–884.
- Lohrmann, K. & Hartinger, A. (2014). Lernemotionen, Lernmotivation und Interesse. In W. Einsiedler, M. Götz, A. Hartinger & F. Heinzel (Hrsg.), *Handbuch Grundschulpädagogik und Grundschuldidaktik* (4. überarbeitete Aufl., S. 275–279). Stuttgart: UTB.

- Lörz, M. & Schindler, S. (2011). Bildungsexpansion und soziale Ungleichheit: Zunahme, Abnahme oder Persistenz ungleicher Chancenverhältnisse – eine Frage der Perspektive? *Zeitschrift für Soziologie*, 40 (6), 458–477.
- Lüdtke, O. & Köller, O. (2002). Individuelle Bezugsnormorientierung und soziale Vergleiche im Mathematikunterricht. Einfluss unterschiedlicher Referenzrahmen auf das fachspezifische Selbstkonzept der Begabung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 34 (3), 156–166.
- Lüdtke, O., Köller, O., Marsh, H. W. & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology*, 30 (3), 263–285.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58 (2), 103–117.
- Maaz, K. & Nagy, G. (2010). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems. Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft, Sonderheft 12, S. 153–182). Wiesbaden: Springer Fachmedien.
- Marks, G., Cresswell, J. & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation*, 12 (2), 105–128.
- Markus, H. & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38.
- Marsh, H. W. (1984). Self-concept: The application of a frame of reference model to explain paradoxical results. *Australian Journal of Education*, 28 (2), 165–181.
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23 (1), 129–149.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79, 280–295.
- Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology*, 82 (4), 646–656.
- Marsh, H. W. (2003). *A reciprocal effects model of the causal ordering of academic self-concept and achievement*. NZARE AARE, Auckland, New Zealand. Zugriff am 10.03.2016. Verfügbar unter <http://www.aare.edu.au/03pap/mar03755.pdf>
- Marsh, H. W., Byrne, B. M. & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80, 366–380.
- Marsh, H. W. & Köller, O. (2004). Unification of theoretical models of academic self-concept/achievement relations: Reunification of East and West German school systems after the fall of the Berlin Wall. *Contemporary Educational Psychology*, 29, 264–282.
- Marsh, H. W., Parker, J. W. & Smith, I. D. (1983). Preadolescent self-concept: Its relation to self-concept as inferred by teachers and to academic ability. *British Journal of Educational Psychology*, 53, 60–78.

- Marsh, H. W. & Yeung, A. S. (1998). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology*, 75 (2), 509–527.
- Martin, M. O., Mullis, I. V. S. & Kennedy, A. M. (Hrsg.) (2007). *PIRLS 2006 technical report*. Chestnut Hill: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- McCallister, L. & Fischer, C. S. (1978). A procedure for surveying personal networks. *Sociological Methods & Research*, 7, 131–148.
- McElvany, N. (2008). *Förderung von Lesekompetenz im Kontext der Familie*. Münster: Waxmann.
- Meyer, H. (2005). *Was ist guter Unterricht?* (3. Aufl.). Frankfurt am Main: Scriptor.
- Möller, J. & Köller, O. (2001). Frame of reference effects following the announcement of exam results. *Contemporary Educational Psychology*, 26, 277–287.
- Möller, J. & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 180–203). Berlin: Springer-Verlag.
- Möller, J. & Trautwein, U. (2015). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (2. Aufl., S. 177–199). Berlin: Springer-Verlag.
- Moos, R. H. (1979). *Evaluating educational environments: Measures, procedures, findings and policy implications*. San Francisco: Jossey-Bass.
- Moos, R. H. & Trickett, E. J. (1974). *Classroom Environment Scale manual*. Palo Alto, CA: Consulting Psychologists Press.
- Muijs, R. D. (1997). Predictors of academic achievement and academic self-concept: A longitudinal perspective. *British Journal of Educational Psychology*, 67 (3), 263–277.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L. & Sainsbury, M. (Hrsg.) (2009). *PIRLS 2011 assessment framework*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Mullis, I. V. S., Martin, M. O., Minnich, C. A., Drucker, K. T. & Ragan, M. A. (2012). *PIRLS 2011 encyclopedia. Education policy and curriculum in reading*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- Muthén, B. & Muthén, L. (2012). *Mplus. Statistical analysis with latent variables. User's guide*. Los Angeles: Muthén & Muthén.
- Neugebauer, M. (2010). Bildungsungleichheit und Grundschulempfehlung beim Übergang auf das Gymnasium: Eine Dekomposition primärer und sekundärer Herkunftseffekte. *Zeitschrift für Soziologie*, 39 (3), 202–214.
- Nickel, S. (2007). Family Literacy in Deutschland. In M. Elfert & G. Rabkin (Hrsg.), *Gemeinsam in der Sprache baden: Family Literacy. Internationale Konzepte zur familienorientierten Schriftsprachförderung* (S. 65–84). Stuttgart: Klett Sprachen.
- Nikolova, R. (2011). *Grundschulen als differenzielle Entwicklungsmilieus. Objektive und subjektive Kontextmerkmale der Schülerzusammensetzung und deren Auswirkung auf die*

- Mathematik- und Leseleistungen* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 81). Münster: Waxmann.
- Opendakker, M., Maulana, R. & Den Brok, P. (2012). Teacher-student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23 (1), 95–119.
- Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung (1998). *Grundqualifikationen, Wirtschaft und Gesellschaft. Ergebnisse der ersten internationalen Untersuchung von Grundqualifikationen Erwachsener*. Paris: OECD Publishing.
- Organisation für Wirtschaftliche Zusammenarbeit und Entwicklung (2001). *Lernen für das Leben. Erste Ergebnisse der Internationalen Schulleistungsstudie PISA 2000*. Paris: OECD Publishing.
- Oser, F. & Baeriswyl, F. J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of Research on Teaching* (4. Aufl., S. 1031–1065). Washington: American Educational Research Association.
- Osterman, K. F. (2000). Students' need for belonging in the school community. *Review of Educational Research*, 70 (3), 323–367.
- Otto, B., Perels, F. & Schmitz, B. (2011). Selbstreguliertes Lernen. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (S. 33–44). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Pedersen, S., Vitaro, F., Barker, E. D. & Sorge, A. I. (2007). The timing of middle-childhood peer rejection and friendship: Linking early behavior to early-adolescent adjustment. *Child Development*, 78, 1037–1051.
- Pekrun, R. (1983). *Schulische Persönlichkeitsentwicklung: Theorieentwicklungen und empirische Erhebungen zur Persönlichkeitsentwicklung von Schülern der 5. bis 10. Klassenstufe*. Frankfurt am Main: Lang.
- Pekrun, R. (1985). Schulklima. In W. Twellmann (Hrsg.), *Handbuch Schule und Unterricht* (S. 524–547). Düsseldorf: Schwann.
- Pekrun, R. (1998). Schüleremotionen und ihre Förderung: Ein blinder Fleck in der Unterrichtsforschung. *Psychologie in Erziehung und Unterricht*, 45, 230–248.
- Pekrun, R. (2001). Familie, Schule und Entwicklung. In S. Walper & R. Pekrun (Hrsg.), *Familie und Entwicklung. Aktuelle Perspektiven der Familienpsychologie* (S. 84–105). Göttingen: Hogrefe.
- Pietsch, M. & Stubbe, T. C. (2007). Inequality in the transition from primary to secondary school: School choices and educational disparities in Germany. *European Educational Research Journal*, 6 (4), 424–445.
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Hrsg.), *Handbook of self-regulated learning* (S. 451–502). San Diego: Academic Press.
- Power, C. N. & Tisher, R. P. (1979). A self-paced environment. In H. J. Walberg (Hrsg.), *Educational environments and effects* (S. 200–217). Berkeley: McCutchan.

- Psychometrica – Institut für psychologische Diagnostik (2015). *Berechnung von Effektstärken*. Zugriff am 01.03.2016. Verfügbar unter <http://www.psychometrica.de/effektstaerke.html>
- Puntambekar, S. & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist*, 40 (1), 1–12.
- Quack, L. (1979). Zur Bedingungsanalyse der Schulleistung: Der Beitrag kognitiver und nicht-kognitiver Merkmale der Schülerpersönlichkeit. In K. J. Klauer & H.-J. Kornadt (Hrsg.), *Jahrbuch für Empirische Erziehungswissenschaft 1979* (S. 93–116). Düsseldorf: Schwann.
- Rakoczy, K. (2006). Motivationsunterstützung im Mathematikunterricht. *Zeitschrift für Pädagogik*, 52 (6), 822–843.
- Rashid, F., Morris, R. & Sevcik, R. (2005). Relationship between home literacy environment and reading achievement in children with reading disabilities. *Journal of Learning Disabilities*, 38 (1), 2–11.
- Rauer, W. & Schuck, K. D. (2003). *FEESS 3.-4. Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen*. Göttingen: Beltz Test.
- Rauin, U. (1987). Differenzierender Unterricht: Empirische Studien im Überblick. In U. Steffens & H. Bargel (Hrsg.), *Untersuchungen zur Qualität des Unterrichts* (S. 111–137). Wiesbaden: Hessisches Institut für Bildungsplanung und Schulentwicklung.
- Reis, S. M., McCoach, D. B., Little, C. M., Muller, L. M. & Kaniskan, R. B. (2011). The effects of differentiated instruction and enrichment pedagogy on reading achievement in five elementary schools. *American Educational Research Journal*, 48 (2), 462–501.
- Rheinberg, F. & Fries, S. (2010). Bezugsnormorientierung. In D. H. Rost (Hrsg.), *Handwörterbuch pädagogische Psychologie* (Anwendung Psychologie, 4., überarb. und erw. Aufl., S. 61–68). Weinheim: Beltz.
- Roberts, J. K., Jurgens, J. & Burchinal, M. (2005). The role of home literacy practices in pre-school children's language and emergent literacy skills. *Journal of Speech, Language, and Hearing Research*, 48 (2), 345–359.
- Rohlf's, C. (2009). Freundschaft und Zugehörigkeit – Grundbedürfnis und Entwicklungsaufgabe. In O. Böhm-Kasper, M. Harring, C. Palentien & C. Rohlf's (Hrsg.), *Freundschaften, Cliques und Jugendkulturen. Peer Groups als Bildungs- und Sozialisationsinstanzen* (S. 61–71). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L. & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: A meta-analytic approach. *Review of Educational Research*, 81 (4), 493–529.
- Rosebrock, C. (2006). Literarische Sozialisation. Basisartikel. In H.-J. Kliewer & J. Pohl (Hrsg.), *Lexikon Deutschdidaktik*. Baltmannsweiler: Schneider-Verlag Hohengehren.
- Rosenholtz, S. J. & Simpson, C. (1984). The formation of ability conceptions: Developmental trend or social construction? *Review of Educational Research*, 54 (1), 31–63.
- Roßbach, H.-G. & Tietze, W. (1996). *Schullaufbahnen in der Primarstufe. Eine empirische Untersuchung zu Integration und Segregation von Grundschulern*. Münster: Waxmann.

- Roth, G. (2009). Warum sind Lehren und Lernen so schwierig? In U. Herrmann (Hrsg.), *Neurodidaktik. Grundlagen und Vorschläge für gehirngerechtes Lehren und Lernen* (2. Aufl., S. 58–68). Weinheim: Beltz.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581–592.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Ryan, R. M., Stiller, J. D. & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *Journal of Early Adolescence*, 14 (2), 226–249.
- Saldern, M. von (1987). *Sozialklima von Schulklassen. Überlegungen und mehrebenenanalytische Untersuchungen zur subjektiven Wahrnehmung von Lernumwelten*. Frankfurt am Main: Lang.
- Saldern, M. von & Littig, K. E. (1985). Die Konstruktion der Landauer Skalen zum Sozialklima. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 17 (2), 138–149.
- Saldern, M. von & Littig, K. E. (1987). *Landauer Skalen zum Sozialklima*. Weinheim: Beltz.
- Saldern, M. von, Littig, K. E. & Ingenkamp, K. (1986). *Landauer Skalen zum Sozialklima für 4. bis 13. Klassen (LASSO 4-13)*. Weinheim: Beltz.
- Salzmann, C. G. (1784). *Noch etwas über die Erziehung nebst Ankündigung einer Erziehungsanstalt*. Leipzig: Crusius.
- Satow, L. (2000). *Klassenklima und Selbstwirksamkeitsentwicklung. Eine Längsschnittstudie in der Sekundarstufe I*. Berlin: Freie Universität Berlin.
- Sauer, J. & Gattringer, H. (1986). Zur Aufklärung der Schulleistung durch spezifische und gemeinsame Varianzanteile von Intelligenz und Motivation. In K. Daumenlang & J. Sauer (Hrsg.), *Aspekte psychologischer Forschung. Festschrift zum 60. Geburtstag von Erwin Roth* (S. 237–257). Göttingen: Hogrefe.
- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie* (Empirische Erziehungswissenschaft, Bd. 36). Münster: Waxmann.
- Scheerens, J. (1990). School effectiveness and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1, 61–80.
- Schicke, M. & Fagan, T. K. (1994). Contributions of self-concept and intelligence to the prediction of academic achievement among grade 4, 6, and 8 students. *Canadian Journal of School Psychology*, 10, 62–69.
- Schiefele, U. & Köller, O. (2010). Intrinsische und extrinsische Motivation. In D. H. Rost (Hrsg.), *Handwörterbuch pädagogische Psychologie* (4. Aufl., S. 336–344). Weinheim: Beltz.
- Schönwälder, K., Petermann, S. & Vertovec, S. (Max-Planck-Gesellschaft, Hrsg.). (o.J.). *Vielfalt, soziale Interaktion und Solidarität. Forschungsperspektiven 2010+*. Zugriff am 10.03.2016. Verfügbar unter [https://www.mpg.de/99379/hm05\\_Vielfalt.pdf](https://www.mpg.de/99379/hm05_Vielfalt.pdf)
- Schreiner, G. (1973). *Schule als sozialer Erfahrungsraum: Überlegungen und Untersuchungen zum Phänomen des Schulklimas*. Frankfurt am Main: Athenäum Fischer Taschenbuch Verlag.

- Schulz-Heidorf, K. (2012). *Elternverhalten und Klimawahrnehmung von Viertklässlern. Eine Re-Analyse aus IGLU-E 2006*. Unveröffentlichte Masterarbeit, Universität Hamburg.
- Schulz-Heidorf, K. & Schwippert, K. (2014). Die soziale Integration im Klassenkontext – abhängig vom elterlichen Erziehungs- und Unterstützungsverhalten? *Erziehung und Unterricht* (7–8), 593–603.
- Schulz-Heidorf, K. & Solheim, O. J. (2016). Adapted teaching: A chance to reduce the effect of social origin? A comparison between Germany and Norway, using PIRLS 2011. *Tertium Comparationis*, 22 (2) 230–259.
- Schwippert, K., Wendt, H. & Tarelli, I. (2012). Lesekompetenz von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 191–207). Münster: Waxmann.
- Seidel, T. (2003). *Lehr-Lernskripts im Unterricht*. Münster: Waxmann.
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77 (4), 454–499.
- Senechal, M. & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, 73 (2), 445–460.
- Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46 (3), 407–441.
- Siegler, R. S., DeLoache, J. S. & Eisenberg, N. (2011). *Entwicklungspsychologie im Kindes- und Jugendalter*. Heidelberg: Spektrum Akademischer Verlag.
- Siewert, J. (2013). *Herkunftsspezifische Unterschiede in der Kompetenzentwicklung: Weil die Schule versagt? Untersuchungen zum Ferieneffekt in Deutschland*. Münster: Waxmann.
- Skaalvik, E. M. & Hagtvet, K. A. (1990). Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. *Journal of Personality and Social Psychology*, 58 (2), 292–307.
- Skaalvik, E. M. & Valas, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *Journal of Experimental Education*, 67, 135–149.
- Skinner, E. A., Kindermann, T. A. & Furrer, C. J. (2009). A motivational perspective on engagement and disaffection: Conceptualization and assessment of children's behavioral and emotional participation in academic activities in the classroom. *Educational and Psychological Measurement*, 69 (3), 493–525.
- Slavin, R. E. (1996). *Education for all*. Lisse: Swets & Zeitlinger Publishers.
- Snijders, T. A. B. (1999). Prologue to the measurement of social capital. *The Tocqueville Review*, 20, 27–44.
- Solzbacher, C. (2008). Was denken Lehrerinnen und Lehrer über individuelle Förderung? *Pädagogik*, 60, 38–42.
- Specht, W. (Hrsg.) (2009). *Nationaler Bildungsbericht Österreich 2009*. Graz: Leykam.

- Speck-Hamdan, A. (2009). Heterogenität und Integration. In H. Bartnitzky, H. Brügelmann, U. Hecker, F. Heinzl, G. Schönknecht & A. Speck-Hamdan (Hrsg.), *Kursbuch Grundschule* (S. 258–289). Frankfurt am Main: Grundschulverband.
- Stanat, P. (2006). Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft. 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. 189–219). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (1997). *Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland – Konstanzer Beschluss. Beschluss der Kultusministerkonferenz vom 24.10.1997*. Zugriff am 17.07.2014. Verfügbar unter [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/1997/1997\\_10\\_24-Konstanzer-Beschluss.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf)
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2003a). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (Jahrgangsstufe 10)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2003b). *Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss (Jahrgangsstufe 10)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2003c). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss (Jahrgangsstufe 10)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004a). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004b). *Bildungsstandards im Fach Mathematik für den Hauptschulabschluss (Jahrgangsstufe 9)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004c). *Bildungsstandards im Fach Deutsch für den Primarbereich (Jahrgangsstufe 4)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004d). *Bildungsstandards im Fach Deutsch für den Hauptschulabschluss (Jahrgangsstufe*

- 9), Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2004e). *Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Hauptschulabschluss (Jahrgangsstufe 9)*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2010). *Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe)*. Informationsschrift des Sekretariats der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2010/2010\\_10\\_18-Uebergang-Grundschule-S\\_el1-Orientierungsstufe.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_10_18-Uebergang-Grundschule-S_el1-Orientierungsstufe.pdf)
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2012a). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2012b). *Bildungsstandards im Fach Deutsch für die Allgemeine Hochschulreife*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2012c). *Bildungsstandards für die fortgeführte Fremdsprache (Englisch/Französisch) für die Allgemeine Hochschulreife*, Beschlüsse der Kultusministerkonferenz. Zugriff am 30.10.2014. Verfügbar unter <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/dokumente.html>
- Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2014). *Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004 in der Fassung vom 12.06.2014*. Zugriff am 20.02.2016. Verfügbar unter [http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf](http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf)
- Statistik Austria (o. J.a). *Klassifikationsdatenbank. ISCO 08 - Struktur*. Zugriff am 11.11.2014. Verfügbar unter [http://www.statistik.at/KDBWeb/kdb\\_VersionAuswahl.do](http://www.statistik.at/KDBWeb/kdb_VersionAuswahl.do)
- Statistik Austria (o.J.b). *ISCO 08 – Element 4132, Datenerfasser*. Zugriff am 11.11.2014. Verfügbar unter [http://www.statistik.at/KDBWeb/kdb\\_Erlaeuterungen.do?KDBtoken=null&&elementID=5462356](http://www.statistik.at/KDBWeb/kdb_Erlaeuterungen.do?KDBtoken=null&&elementID=5462356)
- Stephenson, A. K., Parrila, K. R., Georgiou, K. G. & Kirby, R. J. (2008). Effects of home literacy, parents' beliefs, and children's task-focused behavior on emergent literacy and word reading skills. *Scientific Studies of Reading*, 12 (1), 24–50.
- Stern, G. G. (1970). *People in context. Measuring person-environment-congruence in education and industry*. New York: Wiley.

- Stone, C. A. (1998). The metaphor of scaffolding: Its utility for the field of learning disabilities. *Journal of Learning Disabilities*, 31 (4), 344–364.
- Streblov, L. (2004). *Bezugsrahmen und Selbstkonzeptgenese* (Pädagogische Psychologie und Entwicklungspsychologie, Bd. 42). Münster: Waxmann.
- Stringfield, S. A. (1994). A model of elementary school effects. In D. Reynolds (Hrsg.), *Advances in school effectiveness research and practice* (S. 153–187). New York: Pergamon.
- Stubbe, T. C. (2010). *Die Analyse sozialer Disparitäten in der empirischen Bildungsforschung. Zusammenfassung einer kumulativen Habilitation*. Dortmund: Technische Universität Dortmund.
- Stubbe, T. C., Bos, W. & Euen, B. (2012). Der Übergang von der Primar- in die Sekundarstufe. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 209–226). Münster: Waxmann.
- Stubbe, T. C., Buddeberg, I., Hornberg, S. & McElvany, N. (2007). Lesesozialisation im Elternhaus im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes et al. (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 299–327). Münster: Waxmann.
- Swanson, H. L. & Lussier, C. M. (2001). A selective synthesis of the experimental literature on dynamic assessment. *Review of Educational Research*, 71 (2), 321–363.
- Tarelli, I., Lankes, E.-M., Drossel, K. & Gegenfurtner, A. (2012). Lehr- und Lernbedingungen an Grundschulen im internationalen Vergleich. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 137–174). Münster: Waxmann.
- Tarelli, I., Valtin, R., Bos, W., Bremerich-Vos, A. & Schwippert, K. (2012). IGLU 2011: Wichtige Ergebnisse im Überblick. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 11–25). Münster: Waxmann.
- Tarelli, I., Wendt, H., Bos, W. & Zylowski, A. (2012). Ziele, Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2011). In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–67). Münster: Waxmann.
- Terhart, E. (2006). Kompetenzen von Grundschullehrerinnen und -lehrern: Kontext, Entwicklung, Beurteilung. In P. Hanke (Hrsg.), *Grundschule in Entwicklung. Herausforderungen und Perspektiven für die Grundschule heute* (S. 233–248). Münster: Waxmann.
- Thillmann, K. (2012). *Schulentwicklung und Schulorganisation. Eine empirische Untersuchung schulischer Organisationsgestaltung vor dem Hintergrund der Neuen Steuerung im Bildungssystem*. Berlin: Freie Universität Berlin.
- Tietze, W. & Roßbach, H.-G. (1998). Sitzenbleiben. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 465–469). Weinheim: Beltz.
- Tillmann, K. & Weishaupt, H. (2015). Ansätze bedarfsorientierter Ressourcenausstattung von sozial belasteten Schulen in Deutschland. Eine Situationsanalyse. *Zeitschrift für Bildungswissenschaften*, 31 (2), 5–26.

- Trapmann, S., Hell, B., Weigand, S. & Schuler, H. (2007). Die Validität von Schulnoten zur Vorhersage des Studienerfolgs – eine Metaanalyse. *Zeitschrift für Pädagogische Psychologie*, 21 (1), 11–27.
- Treiman, D. J. (1977). *Occupational prestige in comparative perspective*. New York: Academic Press.
- UNESCO Institute for Statistics (2012). *International Standard Classification of Education. ISCED 2011*. Montreal: UNESCO Institute for Statistics.
- Valentine, J. C., DuBois, D. L. & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39, 111–133.
- Van de Pol, J., Volman, M. & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22 (3), 271–296.
- Van Der Gaag, M. & Snijders, T. A. B. (2005). The Resource Generator: social capital quantification with concrete items. *Social Networks*, 27, 1–29.
- Van Essen, F. (2013). *Soziale Ungleichheit, Bildung und Habitus. Möglichkeitsräume ehemaliger Förderschüler*. Wiesbaden: Springer VS.
- Van Steensel, R. (2006). Relations between socio-cultural factors, the home literacy environment and children's literacy development in the first years of primary education. *Journal of Research in Reading*, 29 (4), 367–382.
- Wagner, H. (1982). Bezugsnormspezifische Lehrerunterschiede im Urteil von Schülern. In F. Rheinberg (Hrsg.), *Jahrbuch für empirische Erziehungswissenschaft* (S. 173–191). Düsseldorf: Schwann.
- Walberg, H. J. & Anderson, G. J. (1972). Properties of the achieving urban classes. *Journal of Educational Psychology*, 63, 381–385.
- Walberg, H. J., Singh, R. & Rasher, S. P. (1977). Predictive validity of student perceptions: A cross-cultural replication. *American Educational Research Journal*, 14, 45–49.
- Warwas, J., Hertel, S. & Labuhn, A. S. (2011). Bedingungsfaktoren des Einsatzes von adaptiven Unterrichtsformen im Grundschulunterricht. *Zeitschrift für Pädagogik*, 57 (6), 854–867.
- Watzlawick, P., Beavin, J. & Jackson, D. D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. New York: Norton.
- Waxman, H. C., Wang, M. C., Anderson, K. A. & Walberg, H. J. (1985). Synthesis of research on the effects of adaptive education. *Educational Leadership*, 43 (1), 26–29.
- Weber, H. S. (2012). *Die relative Bedeutung kognitiver und non-kognitiver Bedingungsfaktoren bei der Vorhersage des Schulerfolgs. Vergleich der prädiktiven Validität von Arbeitsgedächtnis, Intelligenz, Motivation und SES*. Saarbrücken: Saarländische Universitäts- und Landesbibliothek.
- Weinert, F. E. (1997). Notwendige Methodenvielfalt: Unterschiedliche Lernfähigkeit der Schüler erfordern variable Unterrichtsmethoden des Lehrers. In M. A. Meyer & U. Rampillon (Hrsg.), *Lernmethoden, Lehrmethoden: Wege zur Selbstständigkeit* (S. 50–52). Seelze: Friedrich.

- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Weinheim: Beltz.
- Weinert, F. E. & Helmke, A. (1996). Der gute Lehrer: Person, Funktion oder Fiktion? In A. Leschinsky (Hrsg.), *Die Institutionalisierung von Lehren und Lernen. Beiträge zu einer Theorie der Schule* (Zeitschrift für Pädagogik, Beiheft 34, S. 223–234). Weinheim: Beltz.
- Weinert, F. E., Schrader, F.-W. & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, 13 (8), 895–914.
- Weinhold, S. (2006). Entwicklungsverläufe im Lesen- und Schreibenlernen in Abhängigkeit verschiedener didaktischer Konzepte. Eine Longitudinalstudie in Klasse 1–4. In S. Weinhold (Hrsg.), *Schriftspracherwerb empirisch. Konzepte – Diagnostik – Entwicklung* (S. 120–151). Baltmannsweiler: Schneider-Verlag Hohengehren.
- Weißeno, G. & Eck, V. (2013). *Wissen, Selbstkonzept und Fachinteresse: Ergebnisse einer Interventionsstudie zur Politikkompetenz*. Münster: Waxmann.
- Wendt, H., Bos, W., Tarelli, I., Vaskova, A. & Walzebug, A. (2016). *IGLU/TIMSS 2011 – Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen*. Münster: Waxmann.
- Wendt, H., Stubbe, T. C. & Schwippert, K. (2012). Soziale Herkunft und Lesekompetenzen von Schülerinnen und Schülern. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 175–190). Münster: Waxmann.
- Wenning, N. (2007). Heterogenität als Dilemma für Bildungseinrichtungen. In S. Boller, E. Rosowski & T. Stroot (Hrsg.), *Heterogenität in Schule und Unterricht. Handlungsansätze zum pädagogischen Umgang mit Vielfalt* (S. 21–31). Weinheim: Beltz.
- Wentzel, K. R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, 89 (3), 411–419.
- Wigfield, A. & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, 26 (3–4), 233–261.
- Wild, E. & Lorenz, F. (2010). *Elternhaus und Schule*. Stuttgart: UTB.
- Wilkinson, I. A. G. & Fung, I. Y. Y. (2002). Small-group composition and peer effects. *International Journal of Educational Research*, 37 (5), 425–447.
- Wischer, B. (2008). Binnendifferenzierung ist ein Wort für das schlechte Gewissen des Lehrers. *Erziehung und Unterricht*, 158 (9–10), 714–722.
- Wubbels, T., Créton, H. A. & Hoymayers, H. (1985). *Discipline problems of beginning teachers, interactional teacher behaviour mapped out*. Paper presented at the 69. Annual Meeting of the American Educational Research Association (AERA), Chicago.
- Zeinz, H. & Köller, O. (2006). Noten, soziale Vergleiche und Selbstkonzepte in der Grundschule. In A. Schröder-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang* (S. 177–190). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Zinnecker, J., Behnken, I., Maschke, S. & Stecher, L. (2003). *Null zoff & voll busy. Die erste Jugendgeneration des neuen Jahrhunderts*. Opladen: Leske + Budrich.

## Tabellenverzeichnis

Tabelle 1:	Leseverständnis von Achtklässlerinnen und Achtklässlern in west- und ostdeutschen Bundesländern im Schuljahr 1990/1991 nach Schulform.....	8
Tabelle 2:	Effektstärken (Cohens d) des Einflusses der Lehrerfacette ‚Regulation and Monitoring‘ auf die verschiedenen Aspekte des Schulerfolgs in unterschiedlichen Dimensionen.....	40
Tabelle 3:	Prozentwerte der Studien mit großen (L) und kleinen (S) Effektgrößen für die Outcome-Variablen Selbstkonzept, Lese-, Mathematik- und Sprachleistung, die die Teilaspekte ‚Diagnostische Evaluation‘ und ‚Individualisierte Instruktion‘ aufweisen.....	43
Tabelle 4:	Komponenten wichtiger Klima-Instrumente .....	81
Tabelle 5:	Effektstärken der Fixed Effects Modelle der Zusammenhänge zwischen einer positiven und negativen Ausprägung der Schüler-Lehrer-Beziehung, der Leistung und den schulbezogenen Einstellungen in der Gesamtstichprobe und grundschulbezogenen Untersuchungen.....	93
Tabelle 6:	Effektstärken des Einflusses der Charakteristika des Lehrerverhaltens auf kognitive und affektive Merkmale der Schülerinnen und Schüler .....	95
Tabelle 7:	Beispielhafte Kodierung der Klassenzugehörigkeit für zwei Klassen mit jeweils drei Schülern, ohne und mit Teamteaching.....	125
Tabelle 8:	Häufigkeitsverteilung des HISEI in Deutschland .....	136
Tabelle 9:	Internationale Klassifizierung von Berufsabschlüssen nach ISCED.....	137
Tabelle 10:	IGLU-Kodierung und Häufigkeitsverteilung des ISCED für Deutschland.....	138
Tabelle 11:	Häufigkeitsverteilung der Büchervariablen (Elternangaben) für Deutschland.....	140
Tabelle 12:	Varianzaufklärung in der mathematischen Kompetenz in PISA durch verschiedene Prädiktoren, differenziert für Deutschland und den OECD-Durchschnitt.....	141
Tabelle 13:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚ESCS‘ aus Mplus.....	142
Tabelle 14:	Aufklärung der Varianz in der Leseleistung und in dem Selbstkonzept im Lesen durch sozioökonomische Merkmale in Abhängigkeit ihrer Modellierung, Stichprobengröße N .....	143

Tabelle 15:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚HLE‘ aus Mplus .....	146
Tabelle 16:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Leseselbstkonzept‘ auf within- und between-Ebene aus Mplus .....	149
Tabelle 17:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Schüler-Lehrer-Beziehung‘ auf within- und between-Ebene aus Mplus .....	151
Tabelle 18:	Häufigkeitsverteilung des Items „Mit meiner Klassenlehrerin/meinem Klassenlehrer komme ich gut aus“ .....	152
Tabelle 19:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Schüler-Schüler-Beziehung‘ auf within- und between-Ebene aus Mplus.....	153
Tabelle 20:	Deskriptive Analysen der Klassenaggregate der z-standardisierten und nicht-standardisierten Items HISEI, ISCED und Bücherbesitz.....	154
Tabelle 21:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die auf Klassenebene aggregierten Items des Messmodells ‚Klassenaggregat ESCS‘ auf between-Ebene aus Mplus.....	155
Tabelle 22:	STDYX-standardisierte Faktorladungen und nicht erklärte Varianz für die Items des Messmodells ‚Individuelle Förderung‘ auf between-Ebene aus Mplus.....	159
Tabelle 23:	Nicht-standardisierte Varianz, p-Werte und ICC für die abhängige Variable ‚Leseleistung‘ und des abhängige latente Konstrukt ‚Leseselbstkonzept‘ auf Individual- und Kontextebene.....	164
Tabelle 24:	STDYX-standardisierte Regressionsparameter $\beta$ , p-Wert und nicht aufgeklärte Varianz $R^2$ für den Einfluss des ESCS auf die Leseleistung und das Leseselbstkonzept .....	165
Tabelle 25:	STDYX-standardisierte Regressionsparameter $\beta$ , p-Wert und nicht aufgeklärte Varianz $R^2$ für den Einfluss der Schüler-Lehrer-Beziehung auf die Leseleistung und das Leseselbstkonzept und der Schüler-Schüler-Beziehung auf die Leseleistung.....	166

Tabelle 26:	STDYX-standardisierter Korrelationsparameter $r$ , $p$ -Wert und nicht aufgeklärte Varianz $R^2$ für den Zusammenhang zwischen Schüler-Schüler-Beziehung und Leseselbstkonzept .....	166
Tabelle 27:	STDYX-standardisierte Regressionsparameter $\beta$ , $p$ -Wert und nicht aufgeklärte Varianz $R^2$ für den Einfluss der Klassenaggregate ESCS und häuslicher Sprachgebrauch Deutsch auf die Häufigkeit der individuellen Förderung .....	167
Tabelle 28:	Varianzen in den Slopes zwischen den Einzelitems der Konstrukte ESCS und HLE und dem Schulerfolg (Leseleistung und Leseselbstkonzept) .....	169
Tabelle 29:	STDYX-standardisierte Regressionsparameter $\beta$ , $p$ -Wert und nicht aufgeklärte Varianz $R^2$ für den Einfluss der individuellen Förderung auf die klassenmittlere Wahrnehmung der Schüler-Schüler- und Schüler-Lehrer-Beziehung .....	170
Tabelle 30:	STDYX-standardisierte Regressionsparameter $\beta$ , $p$ -Wert und nicht aufgeklärte Varianz $R^2$ für den Einfluss der individuellen Förderung auf die mittlere Leseleistung und das mittlere Leseselbstkonzept.....	171
Tabelle 31:	Kodierung der Items ASBH20A und ASBH20B (von Mutter bzw. Vater ausgeübte Berufe) in IGLU 2011 .....	238
Tabelle 32:	Explorative Faktorenanalyse für die lesebezogenen Items der Skala „Lesen lernen: Vorbereitende Aktivitäten der Eltern vor Schuleintritt“ .....	239
Tabelle 33:	Explorative Faktorenanalyse für die Items der Skala „Selbstkonzept Lesen (1)“ .....	240
Tabelle 34:	Explorative Faktorenanalyse für die Items der Skala „Schüler-Lehrer-Beziehung“ .....	241
Tabelle 35:	Explorative Faktorenanalyse für die Items der Skala „Schüler-Schüler-Beziehung“ .....	242
Tabelle 36:	Explorative Faktorenanalyse für eine Auswahl der Items der Skala „Förderung von Schülerinnen und Schülern nach ihren individuellen Lernvoraussetzungen“ .....	243
Tabelle 37:	STDYX-standardisierte $\beta$ -Parameter und $p$ -Werte der Regressionen im Gesamtstrukturmodell unter Berücksichtigung nicht signifikanter Pfade ..	245

Tabelle 38:	STDYX-standardisierte $r$ -Parameter und $p$ -Werte der Korrelationen im Gesamtstrukturmodell unter Berücksichtigung nicht signifikanter Pfade ..	245
Tabelle 39:	STDYX-standardisierte Residualvarianzen $R^2$ und $p$ -Werte der latenten Konstrukte des Gesamtmodells unter Berücksichtigung nicht signifikanter Pfade.....	245
Tabelle 40:	STDYX-standardisierte Faktorladungen $\lambda$ , $p$ -Werte und Residualvarianz $R^2$ der latenten Konstrukte des Gesamtmodells .....	246
Tabelle 41:	STDYX-standardisierte $\beta$ -Parameter und $p$ -Werte der Regressionen im Gesamtstrukturmodell .....	247
Tabelle 42:	STDYX-standardisierte $r$ -Parameter und $p$ -Werte der Korrelationen im Gesamtstrukturmodell .....	247
Tabelle 43:	STDYX-standardisierte Residualvarianzen $R^2$ und $p$ -Werte der latenten Konstrukte des Gesamtmodells .....	247

## Abbildungsverzeichnis

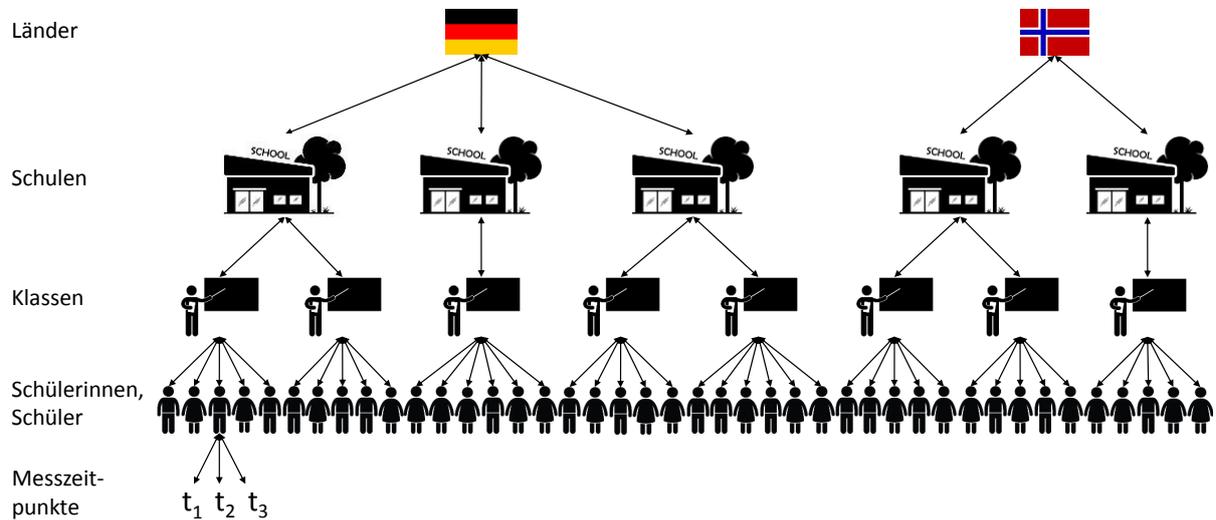
Abbildung 1:	Verteilung der Leistungen auf der Gesamtskala Lesen nach Bildungsgang in PISA 2000.....	7
Abbildung 2:	Testleistungen im Leseverständnis (Gesamtskala) nach Schullaufbahnpräferenzen der Lehrkräfte differenziert, IGLU 2006 .....	9
Abbildung 3:	Dimensionen- und Kriterienraster (Ordnungs- und Suchraster) zur Inneren Differenzierung.....	18
Abbildung 4:	Hierarchisches Modell des Selbstkonzepts nach Shavelson et al. (1976).....	66
Abbildung 5:	Struktur des schulischen Selbstkonzepts im revidierten Modell.....	67
Abbildung 6:	Typisches Ergebnismuster in Pfadanalysen zu der Untersuchung des internalen frame-of-reference-Modells.....	68
Abbildung 7:	Strukturgleichungsmodell der Classic Causal Ordering Study von Marsh (1990) zu vier Messzeitpunkten .....	69
Abbildung 8:	Erweitertes Angebot-Nutzungs-Modell des Unterrichts .....	73
Abbildung 9:	IGLU-Rahmenmodell für den Zusammenhang zwischen Schülerleistungen und deren Bedingungen .....	74
Abbildung 10:	Modell des sozialen Lehrerverhaltens .....	94
Abbildung 11:	Leistungsentwicklung in Klassen unterschiedlicher Eingangsvoraussetzungen .....	109
Abbildung 12:	PIRLS-2011-Testleistungen der Schülerinnen und Schüler im internationalen Vergleich – Gesamtskala Lesen .....	118
Abbildung 13:	Hierarchische Schachtelung von Schülerdaten auf Individualebene in Klassen auf der Kontextebene.....	121
Abbildung 14:	Beispiel der ISCO-08-Kodierung für Bürokräfte und verwandte Berufe, Auszüge .....	131
Abbildung 15:	Sozioökonomischer Status (ISEI) und Berufsprestige (SIOPS) für ausgewählte Berufe .....	133
Abbildung 16:	Anordnung der ursprünglichen Kodierung des von Mutter bzw. Vater ausgeübten Berufs auf der ISEI-Skala.....	135
Abbildung 17:	Klassendurchschnittliche Häufigkeit des Gebrauchs der deutschen Sprache im familiären Kontext in Prozent .....	156

Abbildung 18: Häufigkeit des individuellen Leseunterrichts nach Angabe der Lehrkräfte im europäischen Vergleich .....	157
Abbildung 19: Theoriegeleitetes Gesamtmodell .....	163
Abbildung 20: Beispielhafte Darstellung des Einflusses der individuellen Förderung auf den Slope $s_1$ als Regressionskonstante zwischen ESCS und Leseleistung .....	168
Abbildung 21: Darstellung des Einflusses der individuellen Förderung auf die klassenmittlere Schüler-Lehrer- (SLB) sowie Schüler-Schüler-Beziehung (SSB) auf Kontextebene .....	170
Abbildung 22: Theoriegeleitetes Gesamtmodell .....	173
Abbildung 23: STDYX-standardisierte Pfadkoeffizienten und nicht erklärte Varianz im Gesamtmodell .....	175
Abbildung 24: Hierarchische Schachtelung von Daten einer Analyseebene in vier weiteren Analyseebenen .....	236

## Anhang

### Anhang 1: Mehrebenenstruktur: Ein Beispiel

**Abbildung 24:** Hierarchische Schachtelung von Daten einer Analyseebene in vier weiteren Analyseebenen



Anmerkung: Die verschiedenen Messzeitpunkte wurden der Übersicht halber nur für eine Person dargestellt

Quelle: eigene Darstellung; Piktogramme: Khoon Lay Gan/123rf.de

## **Anhang 2: ISCO-08-Kodierung Bürokräfte und verwandte Berufe am Beispiel des Datenerfassers**

Berufsgattung: **4132** Titel: **Datenerfasser**

### **Erläuterungen:**

Datenerfasser geben mithilfe einer Tastatur, Maus oder eines optischen Scanners, Spracherkennungs-Software oder anderen Dateneingabewerkzeugen codierte, statistische, finanzielle oder sonstige numerische Daten in elektronische Geräte, EDV-gestützte Datenbanken, Tabellenkalkulationsprogramme oder andere Datenverzeichnisse ein. Sie geben Daten in mechanische und elektronische Geräte ein, um mathematische Berechnungen durchzuführen.

Aufgaben umfassen:

- Entgegennahme und Eintragung von Rechnungen, Formularen, Aufzeichnungen und anderen Dokumenten zur Datenerfassung;
- Eingabe numerischer Daten, Codes und Texte aus Quellmaterial in computerkompatible Speicher- und Verarbeitungsanlagen;
- Überprüfung der Fehlerfreiheit und Vollständigkeit von Daten und bei Bedarf Korrektur eingegebener Daten;
- Bedienung von Buchhaltungs- und Rechenmaschinen;
- Datenimport und -export zwischen verschiedenen Datenbanksystemen und Softwareprogrammen.

Beispiele für hier zugeordnete Berufe:

- Dateneingabekraft
- Datenerfasser
- Erfasser von Zahlungseingängen

Quelle: Statistik Austria, o.J.b

**Anhang 3: Kodierung der Items ASBH20A und ASBH20B****Tabelle 31:** Kodierung der Items ASBH20A und ASBH20B (von Mutter bzw. Vater ausgeübte Berufe) in IGLU 2011

<b>Kodierung</b>	<b>Label</b>
1	Hat niemals bezahlte Arbeit verrichtet
2	Eigentümer eines kleinen Betriebes – <i>Zum Beispiel Betriebe mit weniger als 25 Beschäftigten, wie etwa Einzelhandelsgeschäfte, Dienstleistungsbetriebe, Restaurants</i>
3	Angestellter – <i>Zum Beispiel Büroangestellte, Sekretäre, EDV-Fachkräfte, Angestellte im Kundenservice</i>
4	Angestellter im Bereich Dienstleistung oder Verkauf – <i>Zum Beispiel Reisebegleiter, Servicepersonal im Restaurant, Pflegepersonal, Sicherheitspersonal, nieder-rangige Angehörige von Militär oder Polizei, Verkäufer, Straßenverkäufer</i>
5	Fachkraft in Landwirtschaft oder Fischerei – <i>Zum Beispiel Landwirte, Beschäftigte in der Forstwirtschaft, Beschäftigte in der Fischerei</i>
6	Handwerker – <i>Zum Beispiel Maurer, Tischler, Klempner, Elektriker usw., Metallarbeiter, Maschinenschlosser, auch Kunsthandwerker</i>
7	Anlagen- oder Maschinenbediener – <i>Zum Beispiel Anlagen- und Maschinenbediener, Montagebandarbeiter, Kraftfahrer</i>
8	Angelernter Arbeiter oder Hilfsarbeitskraft – <i>Zum Beispiel Reinigungspersonal in Privathaushalten, Betrieben und öffentlichen Gebäuden, Haushaltspersonal, Boten, Pförtner und Gepäckträger, Arbeiter in Landwirtschaft, Fischerei und Baugewerbe</i>
9	Leiter eines großen Unternehmens oder leitender Bediensteter – <i>Zum Beispiel Leiter einer großen Firma (mindestens 25 Angestellte) oder einer Abteilung in einer großen Firma, Abgeordnete oder hohe Regierungsbeamte, leitende Angestellte oder Vorstandsmitglieder in Vereinen und Verbänden, Offiziere</i>
10	Wissenschaftler oder in einem verwandten Beruf tätig – <i>Zum Beispiel Wissenschaftler, Mathematiker, Informatiker, Architekten, Ingenieure, Mediziner und Ärzte, Lehrkräfte, Juristen, Sozialwissenschaftler, Schriftsteller und Künstler, Priester und Pfarrer</i>
11	Techniker oder in einem gleichrangigen nicht technischen Beruf tätig – <i>Zum Beispiel wissenschaftlicher oder technischer Mitarbeiter in den Bereichen Wissenschaft, Ingenieurwesen und Informatik, biowissenschaftliche und Gesundheitsfachkräfte, nicht wissenschaftliche Lehrkräfte, Finanz- und Verkaufsfachkräfte, Vermittler gewerblicher Dienstleistungen, Verwaltungsfachkräfte</i>
12	Nichts trifft zu
Anmerkung:	Die Kodierungen sind für beide Items einheitlich.
Quelle:	Wendt et al., 2016

**Anhang 4: Explorative Faktorenanalysen****Anhang 4.1: HLE: Die häusliche Lesesozialisation****Tabelle 32:** Explorative Faktorenanalyse für die lesebezogenen Items der Skala „Lesen lernen: Vorbereitende Aktivitäten der Eltern vor Schuleintritt“

Item	Bezeichnung	Faktorladungen	
		Faktor 1	Faktor 2
ASBH02Ar	Bücher lesen	0,646	0,049
ASBH02Br	Geschichten erzählen	0,624	0,185
ASBH02Cr	Lieder singen	0,609	0,025
ASBH02Dr	Mit Alphabetspielzeug spielen (z.B. Holzklötze mit Buchstaben)	0,062	0,724
ASBH02Er	Sich über Dinge unterhalten, die Sie gemacht haben	0,625	-0,007
ASBH02Fr	Sich darüber unterhalten, was Sie gelesen haben	0,669	0,181
ASBH02Gr	Wortspiele spielen	0,469	0,519
ASBH02Hr	Buchstaben oder Wörter schreiben	-0,046	0,809
ASBH02Ir	Schilder und Beschriftungen laut lesen	0,171	0,685

Quelle: Elternfragebogen, Wendt et al., 2016

**Anhang 4.2: Leseselbstkonzept****Tabelle 33:** Explorative Faktorenanalyse für die Items der Skala „Selbstkonzept Lesen (1)“

Item	Bezeichnung	Faktorladungen	
		Faktor 1	Faktor 2
ASBR08Ar	Normalerweise bin ich gut im Lesen.	0,487	0,649
ASBR08Br	Lesen fällt mir sehr leicht.	0,590	0,554
ASBR08C	Lesen fällt mir schwerer als vielen meiner Mitschüler.	0,783	0,190
ASBR08Dr	Wenn ein Buch interessant ist, ist es mir egal, wie schwierig es zu lesen ist.	-0,156	0,757
ASBR08E	Es fällt mir schwer, Geschichten mit schwierigen Wörtern zu lesen.	0,683	-0,001
ASBR08Fr	Meine Lehrerin/mein Lehrer sagt, dass ich gut lesen kann.	0,293	0,659
ASBR08G	Lesen fällt mir schwerer als alle anderen Fächer.	0,760	0,170
Quelle:	Schülerfragebogen, Wendt et al., 2016		

**Anhang 4.3: Schüler-Lehrer-Beziehung****Tabelle 34:** Explorative Faktorenanalyse für die Items der Skala „Schüler-Lehrer-Beziehung“

<b>Item</b>	<b>Bezeichnung</b>	<b>Faktorladung</b>
BSNG19A	Meine Klassenlehrerin/mein Klassenlehrer wird schnell ungeduldig, wenn sie/er mir etwas erklären muss.	0,458
BSNG19Br	Ich bin nett zu meiner Klassenlehrerin/meinem Klassenlehrer.	0,800
BSNG19Cr	Ich höre gut zu, wenn meine Klassenlehrerin/mein Klassenlehrer mich etwas fragt.	0,721
BSNG19Dr	Mit meiner Klassenlehrerin/meinem Klassenlehrer komme ich gut aus.	0,858
BSNG19Er	Meine Klassenlehrerin/mein Klassenlehrer ist freundlich zu mir.	0,826

Quelle: Schülerfragebogen, Wendt et al., 2016

**Anhang 4.4: Schüler-Schüler-Beziehung****Tabelle 35:** Explorative Faktorenanalyse für die Items der Skala „Schüler-Schüler-Beziehung“

Item	Bezeichnung	Faktorladung	
		Faktor 1	Faktor 2
BSNG18A	An unserer Schule ist es schwer, Freundinnen und Freunde zu finden.	-0,078	0,791
BSNG18Br	Ich bin beliebt.	0,710	0,232
BSNG18Cr	Andere kommen mit ihren Problemen zu mir.	0,785	-0,147
BSNG18Dr	Mir wird in der Schule das Gefühl gegeben, wichtig zu sein.	0,795	0,067
BSNG18Er	Ich habe wirklich gute Freundinnen oder Freunde in der Klasse.	0,331	0,628
BSNG18Fr	Meine Mitschülerinnen und Mitschüler halten viel von mir.	0,796	0,269
BSNG18Gr	Meine Mitschülerinnen und Mitschüler mögen mich so, wie ich bin.	0,558	0,572
BSNG18Hr	Die anderen arbeiten gern mit mir in Arbeitsgruppen zusammen.	0,641	0,492
BSNG18I	Ich fühle mich in der Klasse ein bisschen als Außenseiterin/Außenseiter.	0,086	0,774

Quelle: Schülerfragebogen, Wendt et al., 2016

**Anhang 4.5: Individuelle Förderung****Tabelle 36:** Explorative Faktorenanalyse für eine Auswahl der Items der Skala „Förderung von Schülerinnen und Schülern nach ihren individuellen Lernvoraussetzungen“

Item	Bezeichnung	Faktorladung				
		F1	F2	F3	F4	F5
ATNG27Ar	Ich gebe Schülern je nach Leistung unterschiedlich schwere Hausaufgaben.	0,625	0,154	0,461	-0,161	-0,157
ATNG27Br	Ich lasse schnellere Schüler schon zum Nächsten übergehen, während ich mit den langsameren noch übe oder wiederhole.	0,697	0,107	-0,019	0,174	-0,209
ATNG27Cr	Wenn Schüler etwas nicht verstanden haben, vergebe ich gezielte Zusatzaufgaben.	0,692	-0,160	0,217	0,137	0,100
ATNG27D	In meiner Unterrichtsgestaltung orientiere ich mich am Leistungsdurchschnitt.	0,058	-0,166	0,039	0,737	-0,317
ATNG27Er	Ich gebe schwachen Schülern zusätzliche Unterstützung im Unterricht.	0,629	0,090	-0,221	-0,031	0,430
ATNG27Fr	Leistungsstarken Schülern gebe ich Extraaufgaben, durch die sie wirklich gefordert werden.	0,728	0,033	0,171	0,049	0,013
ATNG27Gr	Bei der Stillarbeit variiere ich die Aufgabenstellungen, um Schülern unterschiedlicher Leistungsstärke gerecht zu werden.	0,594	0,004	0,502	-0,067	0,070
ATNG27Hr	Bei Gruppenarbeit unterscheide ich verschiedene Leistungsgruppen, die jeweils gesonderte Aufgaben erhalten.	0,267	0,043	0,823	0,062	0,102
ATNG27Ir	Ich achte darauf, dass alle Schüler den Unterrichtsstoff verstanden haben, bevor ich ein neues Thema beginne.	0,038	-0,092	0,132	-0,075	0,854
ATNG27Jr	Ich lasse die Schüler regelmäßig in Gruppen oder allein an unterschiedlich schwierigen Aufgaben arbeiten.	0,614	0,211	0,141	0,151	0,290

*Fortsetzung*

Item	Bezeichnung	Faktorladung				
		F1	F2	F3	F4	F5
ATNG27Kr	Ich setze Computerprogramme zur Leistungsdifferenzierung ein.	0,282	0,764	0,102	-0,110	-0,011
ATNG27Lr	Ich gebe Mädchen und Jungen gemäß ihren geschlechtsspezifischen Interessen unterschiedliche Lesetexte.	-0,185	0,425	0,434	0,456	0,184
ATNG27Mr	Ich lasse die einzelnen Schüler ihre Lesetexte frei wählen (z.B. aus der Klassenbücherei, von zu Hause).	0,232	0,222	-0,036	0,683	0,122
ARNG27Nr	Ich setze insbesondere zur Leseförderung von Jungen Computerprogramme ein.	-0,061	0,813	-0,016	0,157	-0,072

Quelle: Lehrerfragebogen, Wendt et al., 2016

**Anhang 5: Parameter des Gesamtmodells****Anhang 5.1: Parameter des Strukturmodells unter Berücksichtigung nicht signifikanter Pfade**

Legende: HLE = Home Literacy Environment  
 ESCS = Economic, Social and Cultural Status  
 SSB = Schüler-Schüler-Beziehung  
 SLB = Schüler-Lehrer-Beziehung  
 SK-L = Leseselbstkonzept

**Tabelle 37:** STDYX-standardisierte  $\beta$ -Parameter und p-Werte der Regressionen im Gesamtstrukturmodell unter Berücksichtigung nicht signifikanter Pfade

<b>Ebene</b>	<b>Beziehung</b>		<b><math>\beta</math></b>	<b>p</b>
L1	HLE	on ESCS	0,459	< 0,001
	Leseselbstkonzept	on ESCS	0,150	< 0,001
	Leseselbstkonzept	on HLE	0,163	< 0,001
	Leseselbstkonzept	on SLB	0,149	< 0,001
	SSB	on SLB	0,257	< 0,001
	Leseleistung	on ESCS	0,343	< 0,001
	Leseleistung	on HLE	0,135	< 0,001
	Leseleistung	on SSB	0,045	0,020
	Leseleistung	on SLB	0,060	0,001
L2	individuelle Förderung	on Klassenaggregat Sprache	-0,127	0,088
	individuelle Förderung	on Klassenaggregat ESCS	0,104	0,334
	individuelle Förderung	on Klassenaggregat Leistung	-0,212	0,059
	mittlere Leseleistung	on individuelle Förderung	-0,231	0,033
	mittleres Leseselbstkonzept	on individuelle Förderung	-0,146	0,343
	mittlere SSB	on individuelle Förderung	-0,113	0,467

**Tabelle 38:** STDYX-standardisierte r-Parameter und p-Werte der Korrelationen im Gesamtstrukturmodell unter Berücksichtigung nicht signifikanter Pfade

<b>Ebene</b>	<b>Beziehung</b>		<b>r</b>	<b>p</b>
L1	Leseselbstkonzept	with SSB	0,195	< 0,001
	Leseleistung	with SK-L	0,326	< 0,001

**Tabelle 39:** STDYX-standardisierte Residualvarianzen  $R^2$  und p-Werte der latenten Konstrukte des Gesamtmodells unter Berücksichtigung nicht signifikanter Pfade

<b>Ebene</b>	<b>Konstrukt</b>	<b><math>R^2</math></b>	<b>p</b>
L1	Leseleistung	0,815	< 0,001
	HLE	0,789	< 0,001
	SSB	0,934	< 0,001
	Leseselbstkonzept	0,906	< 0,001
L2	mittlere Leseleistung	0,947	< 0,001
	individuelle Förderung	0,911	< 0,001

**Anhang 5.2: Faktorladungen der Messmodelle****Tabelle 40:** STDYX-standardisierte Faktorladungen  $\lambda$ , p-Werte und Residualvarianz  $R^2$  der latenten Konstrukte des Gesamtmodells

<b>Ebene</b>	<b>Faktor</b>	<b>Item</b>	$\lambda$	p von $\lambda$	$R^2$	p von $R^2$
L1	Home Literacy Environment (HLE)	ASBH02Ar	0,556	< 0,001	0,691	< 0,001
		ASBH02Br	0,468	< 0,001	0,781	< 0,001
		ASBH02Er	0,484	< 0,001	0,766	< 0,001
		ASBH02Fr	0,574	< 0,001	0,671	< 0,001
	Economic, Social and Cultural Status (ESCS)	ZISCED	0,733	< 0,001	0,462	< 0,001
		ZBücher	0,589	< 0,001	0,653	< 0,001
		ZHISEI	0,670	< 0,001	0,551	< 0,001
	Schüler-Schüler-Beziehung	BSNG18Er	0,588	< 0,001	0,654	< 0,001
		BSNG18Fr	0,711	< 0,001	0,495	< 0,001
		BSNG18Gr	0,813	< 0,001	0,340	< 0,001
		BSNG18Hr	0,796	< 0,001	0,367	< 0,001
	Leseselbstkonzept	ASBR08Ar	0,808	< 0,001	0,347	< 0,001
		ASBR08Br	0,777	< 0,001	0,396	< 0,001
		ASBR08E	0,376	< 0,001	0,859	< 0,001
		ASBR08Fr	0,598	< 0,001	0,643	< 0,001
	L2	Individuelle Förderung	ATNG27Ar	0,685	< 0,001	0,531
ATNG27Br			0,558	< 0,001	0,688	< 0,001
ATNG27Cr			0,698	< 0,001	0,513	< 0,001
ATNG27Er			0,487	< 0,001	0,763	< 0,001
ATNG27Fr			0,658	< 0,001	0,567	< 0,001
ATNG27Gr			0,716	< 0,001	0,488	< 0,001
ATNG27Jr			0,613	< 0,001	0,624	< 0,001

**Anhang 5.3: Parameter des Strukturmodells**

Legende: HLE = Home Literacy Environment  
 ESCS = Economic, Social and Cultural Status  
 SSB = Schüler-Schüler-Beziehung  
 SLB = Schüler-Lehrer-Beziehung  
 SK-L = Leseselbstkonzept

**Tabelle 41:** STDYX-standardisierte  $\beta$ -Parameter und p-Werte der Regressionen im Gesamtstrukturmodell

<b>Ebene</b>	<b>Beziehung</b>		<b><math>\beta</math></b>	<b>p</b>
L1	HLE	on ESCS	0,457	< 0,001
	Leseselbstkonzept	on ESCS	0,151	< 0,001
	Leseselbstkonzept	on HLE	0,158	< 0,001
	Leseselbstkonzept	on SLB	0,143	< 0,001
	SSB	on SLB	0,268	< 0,001
	Leseleistung	on ESCS	0,335	< 0,001
	Leseleistung	on HLE	0,128	< 0,001
	Leseleistung	on SSB	0,039	0,039
	Leseleistung	on SLB	0,073	< 0,001
L2	mittlere Leseleistung	on Individuelle Förderung	-0,163	0,030

**Tabelle 42:** STDYX-standardisierte r-Parameter und p-Werte der Korrelationen im Gesamtstrukturmodell

<b>Ebene</b>	<b>Beziehung</b>		<b>r</b>	<b>p</b>
L1	Leseselbstkonzept	with SSB	0,180	< 0,001
	Leseleistung	with SK-L	0,347	< 0,001

**Tabelle 43:** STDYX-standardisierte Residualvarianzen  $R^2$  und p-Werte der latenten Konstrukte des Gesamtmodells

<b>Ebene</b>	<b>Konstrukt</b>	<b><math>R^2</math></b>	<b>p</b>
L1	Leseleistung	0,824	< 0,001
	HLE	0,792	< 0,001
	SSB	0,928	< 0,001
	Leseselbstkonzept	0,910	< 0,001
L2	mittlere Leseleistung	0,973	< 0,001

## **Anhang 6: Zusammenfassung der Dissertation**

### **Anhang 6.1: Deutschsprachige Zusammenfassung**

In der Diskussion um Bedingungen des schulischen Lernerfolgs von Kindern und Jugendlichen konnte insbesondere der Familie als primärer Sozialisationsinstanz eine entscheidende Rolle nachgewiesen werden (vgl. etwa Boudon, 1974; Wild & Lorenz, 2010; Becker, 2010). Die enge Kopplung von sozialer Herkunft und Schulerfolg hat weitreichende Folgen: Wie Bos und Lankes et al. (2003; aktueller auch Wendt et al., 2012) zeigen konnten, liegen Viertklässlerinnen und Viertklässler aus Elternhäusern der untersten Sozialschicht im Leseverständnis etwa eine halbe Standardabweichung hinter ihren Mitschülerinnen und Mitschülern der obersten Sozialschicht, was in etwa dem Leistungszuwachs eines Schuljahres entspricht. Dies wurde zum Anlass genommen zu analysieren, ob und wie Schule diesem Herkunftseffekt ausgleichend begegnen kann. Dabei wurde der Fokus auf das Konzept der individuellen Förderung im Unterricht gelegt, welches spätestens seit der Veröffentlichung der PISA-Ergebnisse 2001 starke Popularität erfährt und als „Schlüssel zur Lösung bildungspolitischer und pädagogischer Probleme“ (Klieme & Warwas, 2011, S. 805) Eingang in fast alle Schulgesetze der sechzehn deutschen Bundesländer erfuhr (Fischer, 2014). Ihr liegt die Annahme zugrunde, dass der (unter anderem herkunftsbedingt) heterogenen Zusammensetzung der Schülerschaft durch die Anpassung der Lernsituationen an ihre individuellen Lernvoraussetzungen, Leistungsniveaus, aber auch an ihre Interessen und Begabungen, wie sie die individuelle Förderung charakterisieren, angemessen begegnet werden kann (Altrichter et al., 2009; Bräu, 2007; Klieme & Warwas, 2011; Beck et al., 2008 u.a.).

In den wenigen bisher zu diesem Themengebiet vorliegenden Forschungsarbeiten zeigt sich ein uneinheitliches Bild der Wirkung individueller Förderung, welches vor allem darauf zurückgeführt werden kann, dass ihr unterschiedliche Definitionen unterliegen. Eine Tendenz zeichnet sich dahingehend ab, dass vor allem die affektiv-motivationale Entwicklung der Schülerinnen und Schüler von adaptivem Unterricht profitiert (z.B. bei Seidel & Shavelson, 2007; Giaconia & Hedges, 1982; Reis et al., 2011 und Lipowsky et al., 2011), für kognitive Charakteristika des Schulerfolgs konnte dies weniger deutlich festgestellt werden (Seidel & Shavelson, 2007; Giaconia & Hedges, 1982; Gruehn, 2000 u.a.). Offen bleibt bisher, ob individuelle Förderung ebenfalls eine Möglichkeit darstellt, dem Herkunftseffekt ausgleichend zu begegnen. Dabei wird hier auf Grundlage insbesondere von lernpsychologischen und neuro-

wissenschaftlichen Befunden argumentiert, dass neu zu Lernendes an bestehende Wissensbestände anknüpfen und für den Lernenden von Bedeutung sein sollte, damit eine aktive Auseinandersetzung mit dem Lerngegenstand initiiert und Lernprozesse aufrechterhalten werden (Roth, 2009; Herrmann, 2010). Erreicht werden könnte dies durch die Anpassung des Unterrichts an die Lernvoraussetzungen, Fähigkeiten und Interessen der Schülerinnen und Schüler, wie es in individualisierten Settings der Fall ist. Dann – so die Annahme – könnte herkunftsbezogenen Lernbedingungen von Schülerinnen und Schülern weniger prognostische Relevanz für das Lernen und mit ihnen den Schulerfolg beikommen als dies bisher der Fall ist.

Ziel dieser Untersuchung war es demzufolge, in einem Mehrebenenstrukturgleichungsmodell die Moderation des Herkunftseffekts durch die Häufigkeit individueller Förderung im Unterricht, ebenso wie ihr direkter Einfluss auf Merkmale des Schulerfolgs zu untersuchen. Herangezogen wurden hierfür die Daten der Internationalen Grundschul-Lese-Untersuchung aus 2011. Erweitert wurde das Modell um zwei Facetten der Wahrnehmung der schulischen Lernumwelt, da in Anlehnung an bekannte Rahmenmodelle des Schulerfolgs (v.a. Helmke, 2006, auch Ditton, 2000; Slavin, 1996) angenommen wurde, dass sich Lehrangebote nicht direkt in Lernergebnisse bei den Schulkindern niederschlagen.

Insbesondere hinsichtlich des Herkunftseffekts ergab sich ein unerwarteter Befund: Dieser variiert – zumindest hinsichtlich des Einflusses der sozioökonomischen Stellung der Eltern und der häuslichen Lesesozialisation auf die Leseleistung und das Leseselbstkonzept – nicht zwischen den Klassen, kann also durch Merkmale, die sich wie die individuelle Förderung zwischen den Klassen unterscheiden, nicht moderiert werden. Mit der Leseleistung der Schülerinnen und Schüler steht das adaptive Unterrichten hingegen in einem leicht negativen Zusammenhang, alle anderen diesbezüglichen Effekte unterschieden sich nicht bedeutend von Null. Interpretiert wurde dieser negative Effekt unter anderem dahingehend, dass Maßnahmen der individuellen Förderung möglicherweise gerade dann ergriffen werden, wenn das durchschnittliche Leistungsniveau einer Klasse gering ist. Allerdings können solche Ursache-Wirk-Zuschreibungen anhand der vorliegenden Querschnittdaten nicht überprüft werden. Abgeleitet wurde daraus zunächst ein deutlicher Bedarf an vertiefenden Untersuchungen auf Basis von Längsschnittdaten. Darüber hinaus zeigte sich in den theoretischen wie empirischen Betrachtungen dieser Arbeit, dass individuelle Förderung bislang kaum im

Unterricht implementiert ist, was vor allem auf das Fehlen eines entsprechenden Konzepts sowie Förderinstrumenten zurückgeführt wurde (vgl. etwa auch Altrichter et al., 2009; Klieme & Warwas, 2011; Bräu, 2007; Fischer, 2014). Entsprechende bildungsadministrative wie schulpraktische Implikationen dieser Arbeit beziehen sich daher vordergründig auf die Entwicklung eines solchen Konzepts und entsprechender didaktischer Handlungsrahmen für deren Implementation im Unterricht. Diesen müsste dann eine empirische Überprüfung ihrer Wirksamkeit und – bei entsprechendem Nachweis – eine dahingehende Anpassung der Lehreraus- und -fortbildung folgen, um den berichteten Unsicherheiten der Lehrerinnen und Lehrer, Maßnahmen individueller Förderung zu ergreifen, zu begegnen und der bildungspolitischen Forderung nach adaptivem Unterricht nachzukommen.

## **Anhang 6.2: Englischsprachige Zusammenfassung**

In the discussion on conditions of school achievement it is evident that in Germany, achievement and achievement-related attitudes of students are highly influenced by family characteristics such as their socio-economic situation. One predominant finding shows that, in reading literacy, fourth graders with the lowest social background rank about half a standard deviation below children with the highest social background – which equals the learning and achievement gains of one whole year of schooling (cf. Bos & Lankes et al., 2003, also Wendt et al., 2012). Such findings elucidate the significance of the so-called effect of social origin. Hence, the focus of this analysis was put on if and how schools can compensate this effect. With regards to recent discussions that derived from the publication of the PISA results in 2001, the focal point was put on adapted teaching. This didactical approach has gained strong popularity as the “key for the solution of educational as well as pedagogical problems” (Klieme & Warwas, 2011, S. 805) and was implemented into nearly all of the sixteen federal school laws in Germany (Fischer, 2014). It is based on the assumption that by adapting the teaching to the learning dispositions, abilities, learning conditions as well as interests and talents of the individual learner, the heterogeneous learning conditions and skill levels that teachers face in classrooms can be met (cf. Altrichter et al., 2009; Bräu, 2007; Klieme & Warwas, 2011; Beck et al., 2008).

Research studies that focus on the effects of adapted teaching are scarcely available and show a mixed picture, which can be explained by the varying definitions that underlie the didactical setting. Two tendencies show: Adapted teaching seems to have a slight positive effect on affective-motivational student characteristics (cf. Seidel & Shavelson, 2007; Giaconia & Hedges, 1982; Reis et al., 2011 and Lipowsky et al., 2011); for cognitive aspects, the effect does not show as clearly (Seidel & Shavelson, 2007; Giaconia & Hedges, 1982; Gruehn, 2000). Whether adapted teaching can also influence the effect of social origin remained open so far and was subject to the work presented here. Based on psychological and neuroscientific research findings it is argued that the matters to be learned need to be linked to already existing knowledge and of importance to the individual learner for him or her to actively engage with the matter and to sustain learning processes (Roth, 2009; Herrmann, 2010). This can be achieved by adapting the teaching to the individual learning dispositions, abilities and interests of the students. In theory, such settings might then reduce the

prognostic relevance that the social origin has for learning and school achievements of children and young adults.

The aim of this study was to analyse the moderating influence of adapted teaching on the effect of social origin as well as its direct influence on characteristics of academic success. Such relationships were tested by multilevel structural equation models on the basis of data from the Progress in International Reading Literacy Study (PIRLS) and its national extension (IGLU-E). Since established frame models of school achievement (cf. Helmke, 2006, also Dittton, 2000; Slavin, 1996) show that teaching does not directly result in learning outcomes among students, two facets of the perception of the school learning environment were also added to the model.

Particularly regarding the effect of social origin, the analysis resulted in unexpected findings: The effect does not differ between classes (at least not for the influence of the socio-economic status and the home literacy environment of families on reading literacy and the self-concept in reading of fourth graders) and as a consequence can't be moderated by school-related characteristics that vary between classes, such as adapted teaching. In addition, adapted teaching shows a negative relationship with the reading achievement of the students; all other related effects do not differ significantly from zero. The negative regression coefficient was (amongst other possible explanations) inversely interpreted: Measures of adapted teaching might be taken especially when the average reading ability of a class is low. However, such cause- and outcome-attributions can't be tested on the basis of cross-sectional data as they are used in the analysis presented in this work; further in-depth studies, especially with longitudinal data are needed. In the theoretical framework as well as in the empirical analysis it furthermore showed that adapted teaching is not yet fully implemented in the classroom. This was attributed mainly to the lack of a corresponding concept and instruments of adapted teaching (cf. Altrichter et al., 2009; Klieme & Warwas, 2011; Bräu, 2007; Fischer, 2014). Therefore, administrative and practical implications of this study focus on the development of such a concept and related didactical frameworks for its implementation in the classroom. This needs to be followed by an empirical evaluation of its effectiveness and – after its verification – its inclusion in teacher education and training. This, in turn, could not only help to meet the call by educational policy makers but could also

support teachers in reducing the insecurity they report when applying measures of adapted teaching.