

**Zuwanderungsbezogene
Klassenzusammensetzung: Messung sowie direkte
und vermittelte Effekte auf Leistung und
psychosoziale Schülermerkmale**

Dissertation

zur Erlangung des akademischen Grades
Doctor rerum naturalium (Dr. rer. nat.)
im Fach Psychologie

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von Dipl.-Psych. Camilla Rjosk

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz
Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Richard Lucius

Gutachter/Gutachterin: 1. Prof. Dr. Oliver Lüdtke
 2. Prof. Dr. Petra Stanat
 3. Prof. Dr. Manuel C. Völkle

Tag der mündlichen Prüfung: 16.12.2015

Inhaltsverzeichnis

1. Einleitung.....	1
2. Theoretischer Rahmen der Arbeit und Überblick über den Forschungsstand	9
2.1 Zuwanderungsbezogene Disparitäten: Erklärungen durch individuelle Determinanten der Schülerleistung und psychosozialer Merkmale.....	9
2.1.1 Schülerinnen und Schüler mit Zuwanderungshintergrund in Deutschland.....	9
2.1.2 Individuelle und familiäre Determinanten zuwanderungsbezogener Disparitäten in der Schulleistung.....	11
2.1.3 Individuelle und familiäre Determinanten zuwanderungsbezogener Disparitäten in psychosozialen Merkmalen.....	17
2.1.4 Zwischenfazit: Notwendigkeit der Betrachtung kontextueller Determinanten.....	25
2.2 Zuwanderungsbezogene Disparitäten: Erklärungen durch kontextuelle Determinanten der Schülerleistung und psychosozialer Merkmale	26
2.2.1 Theoretische Annahmen und empirische Befunde zur Bedeutung verschiedener Ebenen des Lernumfelds.....	27
2.2.2 Die Ebene der Schulklasse als proximales Entwicklungsumfeld: Klassenzusammensetzung als Anteil Heranwachsender mit Zuwanderungshintergrund und zuwanderungsbezogene Heterogenität.....	35
2.2.3 Allgemeine Modelle der Vermittlung von Effekten der Zusammensetzung und Modelle der Unterrichtsqualität	39
2.2.4 Empirische Befundlage zu zuwanderungsbezogenen, leistungsbezogenen und sozialen Effekten der Zusammensetzung auf Leistung und psychosoziale Schülermerkmale	47
2.2.5 Theoretische Annahmen und empirische Befunde zur Genese von Effekten zuwanderungsbezogener Zusammensetzung auf Schülerleistung	54
2.2.6 Theoretische Annahmen und empirische Befunde zur Genese von Effekten zuwanderungsbezogener Zusammensetzung auf motivationale Schülermerkmale	67
2.2.7 Methodische Überlegungen zur Analyse von Effekten der Klassenzusammensetzung	76
2.3 Zusammenfassung	81
3. Ziele und Fragestellungen der vorliegenden Arbeit.....	86
4. Teilstudie 1: Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality	96
5. Teilstudie 2: Classroom composition and language minority students' motivation in language lessons	122

6. Teilstudie 3: Ethnic composition and heterogeneity in the classroom–Their measurement and relationship with student outcomes	156
7. Zusammenfassung zentraler Befunde der Teilstudien.....	188
8. Gesamtdiskussion und Ausblick	196
8.1 Die Bedeutung zuwanderungsbezogener Klassenzusammensetzung für Schülerleistung und psychosoziale Merkmale: Zu erwartende Richtung und Größe von Effekten.....	196
8.1.1 Effekte des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf Leistungsmaße	196
8.1.2 Effekte der zuwanderungsbezogenen Heterogenität auf Leistungsmaße	199
8.1.3 Effekte der zuwanderungsbezogenen Klassenzusammensetzung auf psychosoziale Schülermerkmale.....	200
8.1.4 Zuwanderungsbezogene Klassenzusammensetzung und die Bedeutung von zusätzlichen Merkmalen der Klassenzusammensetzung und des weiteren Kontexts.....	202
8.2 Unterrichtsqualität als Mediator von Effekten der Klassenzusammensetzung	204
8.3 Die Bedeutung der Klassenzusammensetzung für das Bestehen zuwanderungsbezogener Disparitäten	208
8.4 Methodische Bewertung und Grenzen der Arbeit	210
8.5 Implikationen für die pädagogische Praxis und zukünftige Forschung	216
8.5.1 Implikationen für die pädagogische Praxis	216
8.5.2 Implikationen für zukünftige Forschung.....	221
9. Literaturverzeichnis.....	224
10. Abbildungsverzeichnis	252
11. Anhang	253
11.1 Anhang A1 zu Teilstudie 1.....	253
11.2 Anhang A2 zu Teilstudie 2.....	255
11.3 Anhang A3 zu Teilstudie 3.....	258
11.4 Anhang A4: ergänzender Anhang zur vorliegenden Arbeit	271

Dank

Diese Dissertation hat mich einige Jahre meines Lebens begleitet und ich möchte hiermit allen Personen danken, die auf die eine oder andere Weise diesen Weg mit mir gegangen sind.

Ich danke dem Institut zur Qualitätsentwicklung im Bildungswesen, das mir die Rahmenbedingungen für die Erstellung der Dissertation geboten hat und allen Mitarbeiterinnen und Mitarbeitern des Instituts - insbesondere dem ehemaligen Abi-Englisch-Team und dem Französisch-Sek-I-Team - und den assoziierten Drittmittelprojekten für die angenehme Arbeitsatmosphäre. Ich danke Oliver Lüdtkke, Petra Stanat und Manuel Völkle für die Begutachtung dieser Arbeit sowie Dirk Richter für die ausgezeichnete Betreuung während der gesamten Dissertationszeit. Mein Dank gilt natürlich auch den Co-Autorinnen und Co-Autoren der Zeitschriftenbeiträge der Dissertation, Dirk Richter, Jan Hochweber, Oliver Lüdtkke, Petra Stanat, Eckhardt Klieme und Jacquelynnne Eccles, für den anregenden Austausch sowie dem Deutschen Institut für Internationale Pädagogische Forschung und dem Institut zur Qualitätsentwicklung im Bildungswesen für die Möglichkeit zur Nutzung der Daten für die dieser Arbeit zugrundeliegenden Analysen. Des Weiteren danke ich dem LIFE-Programm des Max-Planck-Instituts für Bildungsforschung, das mir einen Blick über den Tellerrand und lehrreichen Austausch mit Forscherinnen und Forschern verschiedener Disziplinen und Herkunftsländer ermöglichte.

Für Rückmeldungen zum theoretischen Rahmen der Arbeit danke ich insbesondere Malte Jansen sowie Kristin Schotte, Jennifer Paetsch, Sarah Gentrup, Alexandra Marx, Cornelia Gresch und Doreen Schöppe. Und Jakob Simon danke ich für das Korrekturlesen im Eiltempo. Darüber hinaus danke ich insbesondere noch Maike Wäckerle, Birgit Heppt, Anne Ziemke, Anna Lenski, Malte Jansen, Aleksander Kocaj, Sebastian Wurster, Martin Hecht, Nicole Haag, Sebastian Weirich, Karoline Sachse und natürlich noch vielen Weiteren für ein angenehmes und entspannendes Drumherum. Außerhalb von Arbeit und Dissertation danke ich *la familia dos – la venganza* (v.a. Anni, Marco, Käthe, Pia, Patty) für Ablenkung, Feiern, Zeitverbringen, Zuhören, meinen Eltern und Geschwistern für die Unterstützung und Jesus Jimenez für alles (ich kaufe dir eine Weltreise).

Zusammenfassung

Schulische Lernkontexte sind in unterschiedlichem Ausmaß als entwicklungsförderlich einzustufen. Während die Relevanz der Schulart und der leistungsbezogenen und sozialen Zusammensetzung der Schülerschaft für die akademische Entwicklung von Heranwachsenden als empirisch gut bestätigt gilt, ist die Bedeutung der zuwanderungsbezogenen Klassenzusammensetzung nicht eindeutig geklärt. Im Rahmen der Dissertation werden Effekte der zuwanderungsbezogenen Zusammensetzung sowohl auf verschiedene Leistungsmaße als auch auf psychosoziale Schülermerkmale untersucht. Besonderes Augenmerk wird dabei einerseits auf die Messung der zuwanderungsbezogenen Zusammensetzung gelegt. Andererseits wird eine mögliche Vermittlung von Effekten der Zusammensetzung durch Merkmale der Unterrichtsqualität untersucht. Grundlage bilden Datensätze aus verschiedenen Large-Scale-Assessment-Studien, die im Rahmen von drei Zeitschriftenbeiträgen analysiert werden.

In der ersten Teilstudie wurden Effekte der sozialen und zuwanderungsbezogenen Zusammensetzung auf das deutsche Leseverständnis im Verlauf der neunten Klasse sowie deren Vermittlung durch Merkmale der Unterrichtsqualität (strukturierte Klassenführung, schülerorientiertes, unterstützendes Unterrichtsklima und herausfordernder Sprachunterricht) untersucht. Es konnte gezeigt werden, dass der Klassenanteil von Heranwachsenden mit Zuwanderungshintergrund unter Kontrolle des durchschnittlichen sozioökonomischen Status nicht mit der individuellen Leseleistung assoziiert war. Es zeigten sich erste Hinweise einer partiellen Vermittlung des Effekts sozialer Zusammensetzung durch das Ausmaß herausfordernden Sprachunterrichts.

Die zweite Teilstudie beschäftigte sich mit Effekten der zuwanderungsbezogenen Zusammensetzung auf die Lernmotivation im Fach Deutsch im Verlauf der neunten Klasse sowie deren Vermittlung durch das Ausmaß eines schülerorientierten, unterstützenden Unterrichtsklimas. Es konnte gezeigt werden, dass Heranwachsende mit Zuwanderungshintergrund höhere Motivation berichteten, wenn sie Klassen mit einem höheren Anteil von Heranwachsenden mit Zuwanderungshintergrund besuchten; für Heranwachsende ohne Zuwanderungshintergrund lag kein bedeutsamer Effekt vor. Es zeigten sich keine vermittelten Effekte durch Merkmale des Unterrichtsklimas.

In der dritten Teilstudie wurden der Anteil von Heranwachsenden mit Zuwanderungshintergrund und verschiedene Maße zuwanderungsbezogener Heterogenität in ihrer Assoziation mit Mathematik- und Leseleistung sowie dem Verbundenheitsgefühl mit Peers von Kindern der vierten Jahrgangsstufe verglichen. Es zeigten sich stärkere negative Effekte des Anteils von Kindern mit Zuwanderungshintergrund im Vergleich zu Maßen der zuwanderungsbezogenen Heterogenität. Bei simultaner Berücksichtigung des Klassenanteils und eines Heterogenitätsmaßes waren Effekte der Heterogenität größtenteils nicht mehr signifikant. Es zeigten sich jedoch leicht positive Heterogenitätseffekte auf die Mathematikleistung. Es bestanden keine Effektunterschiede zwischen verschiedenen Heterogenitätsmaßen.

Die Befunde der Dissertation verweisen auf die Bedeutung der Schulklasse als Lernumfeld für die leistungsbezogene und psychosoziale Entwicklung Heranwachsender sowie für die Erklärung zuwanderungsbezogener Disparitäten im Bildungssystem.

Abstract

Some school-related learning environments are more prone to support academic development than others. Empirical findings show that students tend to develop higher levels of achievement when they learn in academic tracks and in schools or classrooms with a high average prior achievement level and a high average socioeconomic status. Yet, research is inconclusive whether the classroom composition related to ethnic and language minority students affects student outcomes. The thesis at hand therefore explores effects of the language-related and ethnic makeup of classrooms on several achievement outcomes as well as psychosocial outcomes. Special attention is paid to the measurement of the ethnic makeup and a possible mediating role of instructional quality for classroom composition effects. Analyses are based on data from large-scale-assessment studies in three research articles.

The first research article investigated effects of socioeconomic and language minority classroom composition on ninth graders' reading achievement across two measurement points and explored the mediating role of central features of instructional quality (focus on language, student-oriented, supportive climate, and structured classroom management). The proportion of language minority students had no effect on reading achievement at measurement point two after controlling for the socioeconomic composition. The results also suggest that the effect of the socioeconomic composition on achievement may be mediated partially by the teacher's focus on language during instruction.

The second research article focused on effects of the proportion of language minority students in classrooms on the development of ninth graders' intrinsic motivation in language lessons and the mediating role of a student-oriented, supportive climate. The analyses revealed that language minority students were increasingly motivated in classrooms with a higher proportion of language minority students. There was no effect on language majority students' motivation. No statistical evidence for the prediction that the compositional effect should be mediated by the instructional climate was found.

The third research article explored various measures of the ethnic makeup in a classroom and their relationship with student outcomes. The proportion of ethnic minority students and measures of ethnic diversity were compared in their prediction of fourth graders' achievement in mathematics and reading as well as the feeling of belonging with one's peers. Compared to diversity measures, the proportion of minority students showed stronger negative effects on student outcomes. Including diversity measures and the proportion of minority students, the diversity effects mostly lost significance. However, students showed higher levels of mathematics achievement in more diverse classrooms. The various measures of diversity led to similar results.

The findings of this thesis add to the evidence indicating that the learning environment in the classroom affects students' achievement and psychosocial development and contributes to ethnic disparities in the education system.

Einleitung

1. Einleitung

Globale Entwicklungen, wie solche der Arbeitsmigration und Globalisierung (vgl. Vedder, Horenczyk, Liebkind & Nickmans, 2006), können die Zusammensetzung und Vielfalt einer Gesellschaft beeinflussen, die sich auch innerhalb des Bildungssystems manifestiert: Das deutsche Bildungssystem zeichnet sich seit geraumer Zeit durch eine hohe Anzahl herkunftsbezogen sehr heterogener Schülerinnen und Schüler aus. In Abhängigkeit von der Definition eines „Zuwanderungshintergrunds“ über Staatsangehörigkeit, Geburtsland oder der vorrangig im Elternhaus verwendeten Sprache beträgt der Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund ungefähr 10 bis 26 Prozent (vgl. Autorengruppe Bildungsberichterstattung, 2014; Diefenbach, 2010; Pöhlmann, Haag & Stanat, 2013). Der Prozentanteil liegt in Stadtstaaten wie Berlin oder Hamburg noch bedeutsam über den genannten Werten, in den ostdeutschen Bundesländern ist er deutlich geringer. Die mit Zuwanderung verbundene Heterogenität der Schülerschaft zeigt sich beispielsweise darin, dass in deutschen Grundschulen Kinder unterrichtet werden, die insgesamt bis zu 57 unterschiedliche Sprachen sprechen (Heppt, Haag, Böhme & Stanat, 2014).

Zuwanderungsbezogene Disparitäten sind bereits seit geraumer Zeit Gegenstand der Bildungsforschung, sie rückten aber besonders im letzten Jahrzehnt stärker in das wissenschaftliche und öffentliche Interesse (vgl. Allemann-Ghionda & Pfeiffer, 2008). Maßgebend waren hierbei internationale Studien, die insbesondere dem deutschen Bildungssystem eine starke Kopplung zwischen Zuwanderungshintergrund und Schülerleistung attestierten (vgl. z.B. OECD, 2006, 2010). Schülerinnen und Schüler mit Zuwanderungshintergrund erreichen durchschnittlich geringere Kompetenzen in zentralen Leistungsbereichen als Heranwachsende ohne Zuwanderungshintergrund (z.B. Haag, Böhme & Stanat, 2012; Pöhlmann et al., 2013). Ebenso bestehen zuwanderungsbezogene Disparitäten im Besuch verschiedener Schularten des Bildungssystems: Während eine Überrepräsentation von Lernenden mit Zuwanderungshintergrund an deutschen Haupt- und Förderschulen bzw. nicht-gymnasialen Schularten vorliegt, sind sie an Gymnasien bzw. zum Abitur führenden Schularten deutlich unterrepräsentiert. Sie verlassen öfter das Schulsystem ohne einen Abschluss, haben geringere Chancen auf einen Ausbildungsplatz oder ein Hochschulstudium. Im weiteren Verlauf zeigen sich für diese Personengruppe geringere Chancen auf einen festen Arbeitsplatz (vgl. den 8. Bericht der Beauftragten der Bundesregierung für Migration, Flüchtlinge und Integration, 2010; Becker & Reimer,

2010; Diefenbach, 2010; OECD, 2007). Durch diese Befunde ausgelöst, beschäftigte sich die empirische Bildungsforschung in den letzten Jahren verstärkt mit möglichen Erklärungsansätzen für das Entstehen der nachteiligen Bildungssituation von Kindern und Jugendlichen mit Zuwanderungshintergrund. Zur Erklärung zuwanderungsbezogener Disparitäten wurden vor allem individuelle und familiäre Determinanten von Bildungsergebnissen der Schülerinnen und Schüler hinzugezogen (vgl. z.B. Diefenbach, 2010). Verschiedene Autorinnen und Autoren weisen jedoch darauf hin, dass auch der Lernkontext bei der Entstehung und Aufrechterhaltung von Disparitäten eine Rolle spielen kann, was bislang nicht ausreichend geprüft wurde (Bellin, 2009; Diefenbach, 2010; Stanat, 2006a). Dabei ist davon auszugehen, dass der schulische Lernkontext nicht nur die Leistungsentwicklung beeinflusst, sondern auch die psychosoziale Entwicklung allgemein.

Ein zentrales Ziel des Schulsystems innerhalb einer demokratischen Gesellschaft besteht darin, allen Schülerinnen und Schülern jeglicher Herkunft gleich gute Bildungschancen bereitzustellen (Baumert & Schümer, 2001; Stanat, 2006a). Dies erfordert eine optimale Förderung aller im schulischen Lernkontext bei gleichzeitigem Ausgleich bestehender Disparitäten (vgl. z.B. Berkemeyer et al., 2014; Giesinger, 2007; Hopf, 2011; Stanat, 2006a). Verschiedene empirische Studien zeigen, dass Lernkontexte in unterschiedlichem Ausmaß als entwicklungsförderlich für Schülerinnen und Schüler einzustufen sind (vgl. z.B. Baumert, Stanat & Watermann, 2006; Becker, Lüdtke, Trautwein, Köller & Baumert, 2012; Stanat, 2006a). Der Lernkontext ist dabei vielschichtig und umfasst sowohl strukturelle Merkmale des Bildungssystems, wie dessen Gliederung in Bildungsgänge, als auch organisatorische Merkmale auf Schulebene bis hin zur konkreten Schulklasse als proximales schulbezogenes Entwicklungsumfeld. Der Ebene der Schulklasse kommt für die akademische Entwicklung eine zentrale Bedeutung zu, da die Schülerinnen und Schüler viel Zeit in der Klasse mit ihren Lehrkräften und Klassenkameradinnen und -kameraden verbringen (Hamre & Pianta, 2010). In diesem Lernumfeld finden der Großteil institutionalisierten Unterrichts sowie Interaktionen zwischen Schülerinnen und Schülern sowie Lehrkräften statt. Diese beiden Merkmale auf Klassenebene, d.h. Merkmale der Klassenzusammensetzung und des Unterrichts, werden als zentral für die akademische Entwicklung erachtet (vgl. z.B. Dumont, Neumann, Maaz & Trautwein, 2013; Seidel & Shavelson, 2007).

Empirische Studien belegen die Relevanz der Zusammensetzung der Schülerschaft innerhalb einer Schulklasse für die akademische Entwicklung individueller Schülerinnen

und Schüler. Die Klassenzusammensetzung kann anhand verschiedener Charakteristika beschrieben werden, wie dem Durchschnitt und der Heterogenität von Merkmalsausprägungen in der Klasse. Untersuchungen von Effekten der Klassenzusammensetzung beziehen sich überwiegend auf Effekte eines Klassendurchschnitts, die im Rahmen von Kompositionsanalysen untersucht werden. Kompositionseffekte beschreiben Effekte eines aggregierten Individualmerkmals auf Klassenebene unter Kontrolle des Individualmerkmals (Raudenbush & Bryk, 2002). Weitgehend einheitlich zeigt sich in solchen Studien, dass das aggregierte Vorkenntnisniveau sowie der durchschnittliche sozioökonomische Status in der Schulklasse positiv in Zusammenhang mit der Leistungsentwicklung individueller Schülerinnen und Schüler steht – unabhängig von deren individuellem Hintergrund (vgl. z.B. Dumont, Neumann, Maaz, et al., 2013; Van Ewijk & Slegers, 2010a). Des Weiteren kann die leistungsbezogene und soziale Schulklassenzusammensetzung auch mit psychosozialen Schülermerkmalen wie Lernmotivation, Selbstkonzept oder Wohlbefinden in der Schule in Zusammenhang stehen (vgl. z.B. Baumert et al., 2006; Nagengast & Marsh, 2012). Inwiefern die zuwanderungsbezogene Klassenzusammensetzung in Zusammenhang mit Leistungsmaßen und psychosozialen Schülermerkmalen steht, ist empirisch jedoch nicht eindeutig geklärt. Zuwanderungsbezogene Zusammensetzung umfasst den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund, dessen Effekt in Kompositionsanalysen betrachtet wird. Darüber hinaus umfasst die zuwanderungsbezogene Zusammensetzung auch Informationen über die Heterogenität der Schülergruppe, die mithilfe verschiedener Heterogenitätsmaße abgebildet werden können. Wie diese beiden Charakteristika zuwanderungsbezogener Zusammensetzung mit der akademischen Entwicklung Heranwachsender zusammenhängen, ist bislang nicht ausreichend untersucht. Des Weiteren steht die Untersuchung theoretisch angenommener Vermittlungswege von Effekten der Klassenzusammensetzung noch aus. Diskutiert wird insbesondere die Annahme, dass ein höherer Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund zu einem Lernumfeld führt, in dem die Qualität des Unterrichts beeinträchtigt wird (vgl. z.B. Van Ewijk & Slegers, 2010b). Das unterrichtliche Angebot bietet die Möglichkeit, von institutioneller Seite auf die Leistungsentwicklung von Schülerinnen und Schülern einzuwirken. Die Unterrichtsqualität wird deshalb als relevanter Faktor der Vermittlung von Effekten der Klassenzusammensetzung angenommen, jedoch bislang kaum empirisch überprüft (vgl.

z.B. Baumert et al. 2006; Dumont, Neumann, Maaz et al., 2013; Harker & Tymms, 2004; Schofield, 2006). Erkenntnisse zu Entstehungsbedingungen und Vermittlungswegen von Effekten der Klassenzusammensetzung sind jedoch notwendig, um Hinweise darauf zu gewinnen, wie diese Effekte von institutioneller Seite durch Maßnahmen auf Schul- und Klassenebene reduziert werden können, um allen Schülerinnen und Schülern möglichst gleichermaßen förderliche Bildungschancen bereitzustellen.

Vor diesem Hintergrund setzt sich die vorliegende Arbeit das Ziel, Effekte der zuwanderungsbezogenen Schulklassenzusammensetzung im Sinne eines breiten Verständnisses von Bildungserträgen sowohl auf verschiedene Leistungsmaße als auch auf psychosoziale Schülermerkmale zu untersuchen. Besonderes Augenmerk liegt dabei einerseits auf der Messung zuwanderungsbezogener Klassenzusammensetzung und andererseits auf möglichen Transmissionswegen der Effekte. Grundlage der Untersuchung bilden repräsentative Stichproben verschiedener Large-Scale-Assessment-Studien, die in der vorliegenden Arbeit im Rahmen von drei Teilstudien analysiert werden. Die Datensätze umfassen verschiedene Altersgruppen und sind mehrheitlich längsschnittlich angelegt.

Die Arbeit gliedert sich wie folgt: Im zweiten Kapitel wird der theoretische Rahmen der Arbeit vorgestellt und ein Überblick über den empirischen Forschungsstand gegeben. In Kapitel 2.1 wird zunächst der Beitrag individueller Determinanten von Schülerleistungen und psychosozialen Merkmalen, insbesondere der Lernmotivation, zur Erklärung zuwanderungsbezogener Disparitäten zusammenfassend dargestellt¹. Kapitel 2.2 widmet sich anschließend den kontextuellen Determinanten mit Schwerpunkt auf theoretischen Annahmen und empirischen Befunden zu Effekten der zuwanderungsbezogenen Klassenzusammensetzung sowie ihrer Vermittlung und Messung. Da die zuwanderungsbezogene Zusammensetzung mit der leistungsbezogenen

¹ Die vorliegende Arbeit beschäftigt sich mit Schülerinnen und Schülern mit Zuwanderungshintergrund und entsprechender Klassenzusammensetzung. Unter „Zuwanderungshintergrund“ wird im Rahmen dieser Arbeit eine Zuwanderungserfahrung innerhalb der Familie verstanden, die gewöhnlich durch das Geburtsland von Familienmitgliedern bzw. der Schülerinnen und Schüler selbst oder die Verwendung (auch) einer anderen Sprache als der Verkehrssprache im Elternhaus definiert wird. Bei Darstellung theoretischer Annahmen und empirischer Befunde, vor allem aus US-amerikanischen Studien, wird ebenso im Sinne einer sprachlichen Vereinheitlichung meist der Begriff „Zuwanderungshintergrund“ für die Bezeichnung ethnischer Minderheitengruppen verwendet.

und sozialen Zusammensetzung verbunden ist, wird ebenfalls kurz auf theoretische Annahmen und empirische Befunde zu diesen beiden Merkmalen eingegangen.

In Kapitel 3 werden die Ziele und Fragestellungen der vorliegenden Arbeit erläutert, denen im Rahmen von drei Teilstudien nachgegangen wurde. Diese sind in den Kapiteln 4-6 abgebildet.

Zum Abschluss werden die in der vorliegenden Arbeit gewonnenen Erkenntnisse in Kapitel 7 zusammengefasst und in Kapitel 8 vor dem Hintergrund des bestehenden theoretischen und methodischen Kenntnisstands diskutiert. Die Arbeit schließt mit einer Erörterung von Implikationen für die Schulpraxis und zukünftige Forschungsperspektiven.

2

Theoretischer Rahmen der Arbeit und
Überblick über den Forschungsstand

2. Theoretischer Rahmen der Arbeit und Überblick über den Forschungsstand

2.1 Zuwanderungsbezogene Disparitäten: Erklärungen durch individuelle Determinanten der Schülerleistung und psychosozialer Merkmale

In der Literatur verschiedener Disziplinen wie Psychologie, Erziehungswissenschaft und Soziologie finden sich verschiedene Ansätze zur Erklärung zuwanderungsbezogener Disparitäten. In diesen Ansätzen wird überwiegend eine individuenbezogene Sichtweise vertreten (vgl. Diefenbach, 2010). Die Bildungsdisparitäten werden also über die Merkmale der Schülerinnen und Schüler selbst beziehungsweise ihrer Familien als Determinanten des Erfolgs oder Misserfolgs im Schulsystem erklärt. Diese Merkmale werden in den nachfolgenden Abschnitten theoretisch eingeordnet und die empirische Befundlage zusammenfassend dargestellt. Bezug genommen wird dabei einerseits auf Zusammenhänge mit der Schülerleistung (Abschnitt 2.1.2) und andererseits auf Zusammenhänge mit der Lernmotivation als zentrales psychosoziales Merkmal, welches das Lernverhalten und zukünftige Schülerleistung beeinflusst (Abschnitt 2.1.3). Im nächsten Abschnitt werden zunächst Hintergrundinformationen zur Population von Schülerinnen und Schülern mit Zuwanderungshintergrund im deutschen Bildungssystem skizziert, um eine Grundlage für die Einordnung von familiären Hintergrundmerkmalen von Personen mit Zuwanderungshintergrund in Deutschland zu schaffen.

2.1.1 Schülerinnen und Schüler mit Zuwanderungshintergrund in Deutschland

Die Entstehung der gegenwärtigen zuwanderungsbezogenen Zusammensetzung der Schülerschaft in Deutschland reicht zurück bis in die Mitte der Fünfzigerjahre des letzten Jahrhunderts. In dieser Zeit führte der Arbeitskräftemangel in Deutschland zur Anwerbung von Arbeitskräften im Ausland (z.B. Flam, 2007; Seibert, 2012). Zwischen 1955 und 1967 wurden verschiedene Anwerbeverträge unter anderem mit Italien, Spanien, Türkei, Tunesien und Jugoslawien geschlossen, die insbesondere ab 1961 zu einem starken Anstieg ausländischer Beschäftigter in der damaligen BRD führten. Arbeitskräfte wurden insbesondere für die industrielle Massenfertigung, Schwerindustrie

und Bergbau benötigt – Tätigkeiten, zu deren Ausführung es keiner hohen Qualifikationsanforderungen bedurfte. Entsprechend hatten die angeworbenen Arbeitskräfte ein vergleichsweise niedriges Ausbildungsniveau und wiesen im Herkunftsland einen eher niedrigen sozialen Status auf. Zu diesem Zeitpunkt war ein längerfristiges Beschäftigungsverhältnis nicht vorgesehen; durch die Anwerbung sollte lediglich der Bedarf in einer Hochkonjunkturphase bedient werden. Nach dem Anwerbestopp 1973 wurden ein längerfristiger Aufenthalt und Familiennachzug in die damalige BRD ermöglicht, was vielfach in Anspruch genommen wurde. Im Jahr 1973 betrug die Anzahl ausländischer Arbeitskräfte in der damaligen BRD circa 2,6 Millionen. In den 1980er Jahren nahm die Zuwanderung in die BRD ab, erreichte aber um 1990 einen zweiten Höhepunkt. In diese Zeit fällt auch die verstärkte Aufnahme von (Spät-)Aussiedlerinnen und Aussiedlern aus der Sowjetunion und Osteuropa sowie von Flüchtlingen und Asylbewerberinnen und Asylbewerbern. Nach der Wiedervereinigung ab Mitte/Ende der 1990er Jahre bis zum Jahr 2008 sank die Anzahl von Ausländerinnen und Ausländern in Deutschland von 7,5 auf 7,2 Millionen (Seibert, 2012).

In Berichten über Personen mit Zuwanderungshintergrund gibt es unterschiedliche Möglichkeiten einen Zuwanderungshintergrund operational zu definieren. Während Dokumente des statistischen Bundesamts und des Mikrozensus lange Zeit lediglich zwischen Personen mit deutscher Staatsbürgerschaft und Personen ohne deutsche Staatsbürgerschaft unterschieden, wurden unter anderem in der Bildungsforschung auch Personen ein Zuwanderungshintergrund zugeschrieben, deren Eltern oder Großeltern nicht in Deutschland geboren sind (zweite bzw. dritte Generation), oder die neben Deutsch auch weitere Sprachen im Elternhaus verwenden. Familienmitglieder der letztgenannten Gruppen kamen im Rahmen oben beschriebener Zuwanderungsprozesse in die damalige BRD bzw. in das wiedervereinigte Deutschland. In Abhängigkeit von der Definition des Zuwanderungshintergrunds beläuft sich der Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund im deutschen Bildungssystem auf ungefähr 10 bis 26 Prozent (vgl. Autorengruppe Bildungsberichterstattung, 2014; Diefenbach, 2010; Pöhlmann et al., 2013). Bei der Definition über das Geburtsland der Eltern sind die größten Herkunftsgruppen Personen aus der Türkei, aus der ehemaligen Sowjetunion, aus Polen und aus dem ehemaligen Jugoslawien (vgl. Pöhlmann et al., 2013). In der vorliegenden Arbeit wird ein Zuwanderungshintergrund je nach Fragestellung über das Geburtsland oder die verwendete Sprache im Elternhaus definiert.

2.1.2 Individuelle und familiäre Determinanten zuwanderungsbezogener Disparitäten in der Schulleistung

Schülerinnen und Schüler unterscheiden sich in ihrer Schulleistung, die sich in Form von Noten, Punkten in Klassenarbeiten, aber auch in im Rahmen von Bildungsstudien administrierten Leistungstests manifestiert. Das Lernpotenzial bzw. die kognitiven Grundfähigkeiten eines Individuums stehen mit der Schulleistung in Zusammenhang (vgl. z.B. Helmke & Weinert, 1997; Köller & Baumert, 2002; Wang, Haertel & Walberg, 1993). Darüber hinaus gibt es im deutschen Bildungssystem jedoch eine starke Kopplung von sozialer Herkunft, Zuwanderungshintergrund und Leistung (z.B. Baumert & Maaz, 2012; Kuhl, Siegle & Lenski, 2013).

Theoretische Annahmen, die vornehmlich Merkmale von Schülerinnen und Schülern bzw. Familien mit Zuwanderungshintergrund in den Blick nehmen, um zuwanderungsbezogene Disparitäten zu erklären, umfassen unter anderem (1) sozioökonomische Erklärungen, (2) Bildungsentscheidungen bei Übergangsprozessen, (3) kulturalistische Erklärungen, (4) Erklärungen durch die in der Migrationssituation verfügbaren Handlungsstrategien und (5) die Bedeutung der Sprachkenntnisse für den Schulerfolg (vgl. Baumert & Maaz, 2012; Diefenbach, 2002, 2010; Stanat, 2006b). Diese Annahmen schließen sich nicht gegenseitig aus, sondern beschreiben verschiedene Erklärungsmechanismen, die auch gleichzeitig wirksam sein können.

(1) *Sozioökonomische Erklärungen* beziehen sich auf eine geringere Kapitalausstattung Heranwachsender mit Zuwanderungshintergrund, als die, die für das erfolgreiche Durchlaufen der Schullaufbahn in Deutschland notwendig sei.

Insbesondere seit PISA 2000 wird innerhalb der deutschen Bildungsforschung die enge Kopplung von sozialer Herkunft und Schülerleistung diskutiert (vgl. Kramer & Helsper, 2010). Schülerinnen und Schüler aus Familien mit einem niedrigen sozioökonomischen Status erreichen – unabhängig von einem Zuwanderungshintergrund – durchschnittlich geringere Leistungen in zentralen Kompetenzbereichen als Kinder und Jugendliche aus Familien mit einem höheren sozioökonomischen Status (PISA 2000: Baumert & Schümer, 2001; Schümer, 2004; IQB-Ländervergleich 2011: Richter, Kuhl & Pant, 2012; IQB-Ländervergleich 2012: Kuhl et al., 2013). Diese Befunde werden mit der sozioökonomischen Stellung der Eltern und damit einhergehendem Verfügen über ökonomisches (Becker, 1993), kulturelles (Bourdieu, 1982; Bourdieu & Passeron, 1964) und soziales Kapital (Bourdieu, 1982; Coleman, Hoffer & Kilgore, 1982) erklärt. Die

Kapitalausstattung wirkt über familiäre Prozessmerkmale vermittelt auf die Schulleistung ein (vgl. Baumert, Watermann & Schümer, 2003).

Gemäß dem „Modell des Zusammenhangs von Struktur- und Prozessmerkmalen der familiären Lebensverhältnisse und Bildungsbeteiligung bzw. Kompetenzerwerb“ (Baumert et al., 2003; siehe Abbildung 2.1) beeinflussen Strukturmerkmale der sozialen Herkunft der Familie, wie der sozioökonomische Status und das Bildungsniveau, aber auch der Zuwanderungshintergrund, die Leistungsentwicklung der Schülerinnen und Schüler, sowohl in direkter als auch in vermittelter Weise. Die Strukturmerkmale beeinflussen in der Familie ablaufende Prozesse des konsumptiven Verhaltens, der kulturellen Praxis sowie der kommunikativen und sozialen Praxis, was letztendlich ein mehr oder weniger lern- und leistungsförderliches familiäres Umfeld bedingt und dadurch die Schülerleistung beeinflusst. Diese drei Prozessmerkmale sind näherungsweise den Wirkweisen des ökonomischen, kulturellen und sozialen Kapitals zuzuordnen: Ökonomische Ressourcen beeinflussen das konsumptive Verhalten und bieten die Grundlage für Investitionen in die Bildung der Kinder. Investitionsentscheidungen sind dabei von Kosten-Nutzen-Überlegungen geleitet. Bildungsinvestitionen sind für Familien mit einem niedrigeren sozioökonomischen Status mit höheren Kosten verbunden – indirekte Kosten, wie Einkommensausfall während der Ausbildungszeit, und direkte Kosten, wie Geld zum Bücherkauf, die in Konkurrenz zu anderen Investitionen stehen (vgl. Humankapitaltheorie des Ökonomen Gary S. Becker, 1993). Die kulturelle Praxis steht in Verbindung mit kulturellem Kapital (Bourdieu, 1982; Bourdieu & Passeron, 1964). Es schließt Güter wie Literatur oder Kunstwerke, aber auch Bildungszertifikate oder Titel bis hin zu verinnerlichten Wahrnehmungs-, Deutungs- und Handlungsschemata ein. Die Schemata werden auch als „Habitus“ bezeichnet, der unter anderem in der Familie durch gemeinsame Aktivitäten und durch das elterliche Vorbild vermittelt wird und die kulturelle Passung einer Schülerin bzw. eines Schülers zu einem System beeinflusst (vgl. auch die Darstellung bei Baumert & Schümer, 2001 sowie dazu auch Kramer & Helsper, 2010).

Bedeutsam für die kommunikative und soziale Praxis ist unter anderem das soziale Kapital (Bourdieu, 1982; Coleman et al., 1982). Zu sozialem Kapital gehört das Vorhandensein und die Erreichbarkeit sozialer Netzwerke, die durch ihre Kommunikationspraktiken die Übernahme von sozial anerkannten Werten, Normen und Zielen fördern und unterstützen sowie Informationen bereitstellen (vgl. die Darstellung bei Baumert & Schümer, 2001). In Analysen von Daten der PISA-Studie konnte das

Modell der Struktur- und Prozessmerkmale empirisch untermauert werden (Baumert et al., 2003).

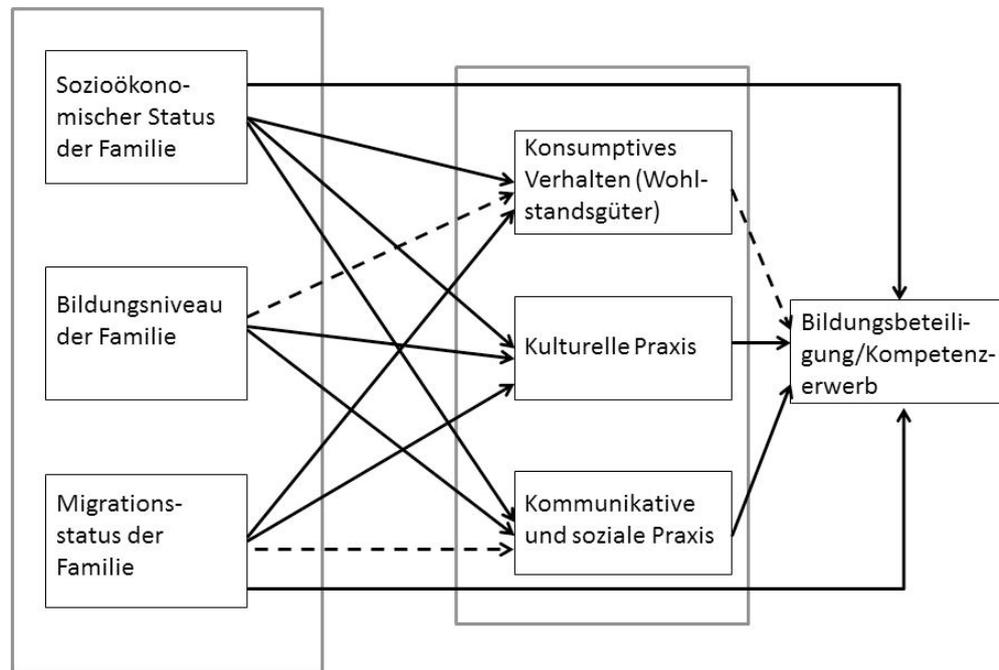


Abbildung 2.1: Modell des Zusammenhangs zwischen Struktur- und Prozessmerkmalen der familiären Lebensverhältnisse und Bildungsbeteiligung bzw. Kompetenzerwerb (aus: Baumert et al, 2003, S. 56)

Da Familien mit Zuwanderungshintergrund in Deutschland im Durchschnitt einen geringen sozialen Status haben (vgl. Abschnitt 2.1.1), sind die eben beschriebenen Annahmen ebenfalls für die Erklärung der Leistung von Heranwachsenden mit Zuwanderungshintergrund bedeutsam. Hierbei wird vor allem eine geringe Bildung und ein geringeres Einkommen von Familien mit Zuwanderungshintergrund im Vergleich zu Familien ohne Zuwanderungshintergrund betont, was ihnen geringere Ressourcen zur Investition in die Bildung ihrer Kinder bietet (vgl. Jungbluth, 1999). Eine zusätzliche Problematik kann sich für die Familien daraus ergeben, dass im Herkunftsland akkumuliertes Kapital in Form von Bildungsabschlüssen und Zertifikaten in der Aufnahmegesellschaft nicht anerkannt oder nachgefragt wird (vgl. Diefenbach, 2010). Die Relevanz der sozialen Herkunft für die Schulleistungen und -abschlüsse von Schülerinnen und Schülern mit Zuwanderungshintergrund in Deutschland ist empirisch gut bestätigt. So zeigt z.B. Walter (2008a) anhand einer Analyse der PISA 2000-, 2003-

und 2006-Daten, dass die Leistungen von Schülerinnen und Schülern mit Zuwanderungshintergrund mit ihrer sozialen Herkunft gemessen über den sozioökonomischen Status und das Bildungsniveau der Eltern sowie die Anzahl kultureller Besitztümer in Zusammenhang stehen (vgl. auch Haag et al., 2012; Kristen, 2008; Pöhlmann et al., 2013). Sozioökonomische Merkmale der Familien können jedoch nicht als alleinige Ursache von zuwanderungsbezogenen Disparitäten gesehen werden, da Unterschiede zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund nach Kontrolle von Merkmalen der sozialen Herkunft weiterhin bestehen (vgl. Bellin, 2009; Diefenbach, 2010)².

(2) Einen weiteren Erklärungsmechanismus bilden die *Entscheidungen bei Übergangsprozessen* innerhalb des Bildungssystems, die mit familiären Merkmalen der sozialen Herkunft und des Zuwanderungshintergrunds in Zusammenhang stehen können. Der Übergangsprozess ist relevant für Leistungsdisparitäten, da der Besuch unterschiedlicher Bildungsgänge sich differentiell auf die Leistungsentwicklung auswirkt (siehe Kapitel 2.2).

Richtungsweisend für Annahmen zu sozialer Herkunft und Übergangsentscheidungen waren die Arbeiten von Raymond Boudon (1974), der *primäre und sekundäre Effekte der sozialen Schichtzugehörigkeit* unterscheidet. Primäre Effekte der sozialen Herkunft beziehen sich auf die eben beschriebenen Kapitalausstattungen von Familien, die unter anderem über Prozesse der Anregung und Unterstützung die Leistungsentwicklung und Schulnoten der Schülerinnen und Schüler beeinflussen (vgl. Maaz, Baumert & Trautwein, 2010). Leistung und Noten bilden die Grundlage für Entscheidungen bei Übergängen im Bildungssystem (vgl. Maaz, Baumert, Gresch & McElvany, 2010). Darüber hinaus können auch sekundäre Herkunftseffekte wirksam sein, die Entscheidungen unabhängig von den individuellen Schülerleistungen durch Erwartungen und Ziele für Bildung und Beruf beeinflussen. Bildungsentscheidungen folgen der Theorie von Boudon (1974) zufolge einem Erwartungs-Wert-Modell (vgl. Maaz, Hausen, McElvany & Baumert, 2006), d.h. dass die Wahl des Bildungsweges als

² So zeigten beispielsweise Analysen im Rahmen des IQB-Ländervergleichs 2012 für das Fach Mathematik, dass Leistungsunterschiede zwischen Schülerinnen und Schülern ohne Zuwanderungshintergrund und Heranwachsenden mit zwei nicht in Deutschland geborenen Elternteilen verschiedener Herkunftsländer nach Kontrolle des sozialen Hintergrunds nicht verschwanden. Wenn die Unterschiede im sozioökonomischen Status und Bildungsniveau der Eltern kontrolliert wurden, nahmen die Effekte vielmehr um circa die Hälfte bis ein Drittel ab (Pöhlmann et al., 2013).

ein Resultat von Investitionsentscheidungen angesehen wird, wobei die Kosten-Nutzen-Abwägungen zwischen den sozialen Schichten variieren. In verschiedenen empirischen Arbeiten konnten Hinweise auf primäre und sekundäre soziale Herkunftseffekte bei Übergangentscheidungen gefunden werden (Maaz, Baumert & Trautwein, 2010; vgl. Dumont, Neumann, Becker, Maaz & Baumert, 2013).

Da der soziale Status systematisch mit dem Zuwanderungshintergrund variiert, können diese Annahmen ebenfalls auf Familien mit Zuwanderungshintergrund zutreffen. Spezifisch für Familien mit Zuwanderungshintergrund kann darüber hinaus eine mangelnde Kenntnis des deutschen Schulsystems Bildungsentscheidungen beeinflussen (vgl. Diefenbach, 2002; Dollmann, 2010). Studien zu primären und sekundären Disparitäten beim Übergang von Kindern mit Zuwanderungshintergrund innerhalb des deutschen Bildungssystems weisen darauf hin, dass sich die geringeren Chancen eines Übergangs auf das Gymnasium von Kindern mit türkischem Zuwanderungshintergrund bzw. Kindern von (Spät-)Aussiedlern im Vergleich zu Kindern ohne Zuwanderungshintergrund vor allem auf primäre Disparitäten zurückführen lassen (z.B. Gresch & Becker, 2010; Kristen & Dollmann, 2009). Der Besuch von Schulen nicht-gymnasialer Bildungsgänge dieser Schülergruppen lässt sich vor allem durch geringere Schulleistungen, Noten und einen niedrigeren sozioökonomischen Status der Familie erklären. Unter Kontrolle von mit primären Disparitäten assoziierten Merkmalen zeigt sich sogar teilweise ein positiver sekundärer Zuwanderungseffekt (vgl. Dollmann, im Druck; Gresch, 2012; Gresch & Becker, 2010; Kristen & Dollmann, 2009). Es wird angenommen, dass diesem Effekt erhöhte Bildungsaspirationen von Familien mit Zuwanderungshintergrund zugrunde liegen (vgl. z.B. Hustinx, 2002; siehe auch Abschnitt 2.1.3.2).

(3) *Kulturalistische Erklärungen* beziehen sich auf Aspekte wie Werte- und Deutungsmuster, die charakteristisch für eine Herkunftskultur sind und sich von der Kultur der Aufnahmegesellschaft unterscheiden. Als mögliche Hindernisse für den Schulerfolg werden unter anderem kulturelle Orientierungen, traditionelle Werte- und Deutungsmuster sowie mangelnde Kenntnisse über das deutsche Schulsystem diskutiert, die es Eltern mit Zuwanderungshintergrund erschweren, ihre Kinder beim erfolgreichen Durchlaufen der Schullaufbahn in der Aufnahmegesellschaft zu unterstützen (vgl. Leenen, Grosch & Kreidt, 1990; Diefenbach, 2010). Empirische Belege für diesen Erklärungsansatz gibt es nur vereinzelt (vgl. Diefenbach, 2010). Wenn die Bevorzugung

von Zeitungen in der Sprache der Aufnahmegesellschaft sowie der Gebrauch der Sprache der Aufnahmegesellschaft als Indikator für kulturelle Orientierungen verwendet werden, lassen sich Hinweise auf positive Zusammenhänge zwischen der Orientierung an der Aufnahmegesellschaft im Elternhaus und dem Bildungserfolg der Kinder finden (vgl. Zusammenfassung bei Stanat, 2006a, 2006b).

(4) Erklärungen durch die in der Migrationssituation verfügbaren

Handlungsstrategien beziehen sich vor allem auf die Aufenthaltsperspektive der Familie. Personen mit unsicherem Aufenthaltsstatus und Bleibeabsichten würden demnach weniger in die Ausbildung ihrer Kinder im Aufnahmeland investieren und es eher bevorzugen, dass die Kinder möglichst schnell zum Familieneinkommen beitragen (vgl. Korte, 1990; Diefenbach, 2002, 2010; Stanat, 2006b). Befunde aus Analysen des Sozioökonomischen Panels (SOEP) konnten allerdings bislang nicht bestätigen, dass die Rückkehrabsicht der Familie mit der Beschulung in unterschiedlichen Bildungsgängen des deutschen Schulsystems in Zusammenhang steht (Alba, Handl & Müller, 1994; Diefenbach, 2002; vgl. Diefenbach, 2010).

(5) Ein weiterer Prädiktor des Schulerfolgs wird in den *Sprachkenntnissen* der Schülerinnen und Schüler gesehen (vgl. z.B. Esser, 2006). Untersuchungen zur Bildungsbeteiligung (vgl. u.a. Alba et al., 1994; Büchel & Wagner, 1996; Herwartz-Emden, 2003) und Ergebnisse aktueller Schulleistungsuntersuchungen (z.B. Haag et al., 2012; Pöhlmann et al., 2013; Walter, 2008a) weisen übereinstimmend darauf hin, dass Kinder und Jugendliche in Deutschland, die Zuhause auch eine andere Sprache sprechen als Deutsch, durchschnittlich geringere Leistungen erreichen als Kinder, die Zuhause in der Verkehrssprache kommunizieren. Insbesondere eine geringere Beherrschung lexikalischer und grammatikalischer Besonderheiten der schulischen Bildungssprache (z.B. Fachwortschatz, Komposita, Präpositionalphrasen) sind mit Leseleistungsunterschieden zwischen ausschließlich deutschsprachigen Heranwachsenden und solchen, die im Elternhaus auch eine andere Sprache als Deutsch sprechen, assoziiert (z.B. Heppt et al., 2014). Sprachliche Kompetenzen in der Verkehrssprache sind demnach auch für Leistungsunterschiede zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund mitverantwortlich.

Die dargestellten Erklärungsansätze zuwanderungsbezogener Disparitäten beziehen sich auf Individualmerkmale der Schülerinnen und Schüler. So spielen insbesondere eine schwächere sozioökonomische Situation und die Beherrschung der Unterrichtssprache für die Leistungsdifferenzen zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund eine Rolle. Werden diese Merkmale bei der Vorhersage von Schulleistung berücksichtigt, bleiben jedoch weiterhin Unterschiede zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund bestehen. Es ist also notwendig, weitere Erklärungsfaktoren hinzuzuziehen (vgl. Bellin, 2009; Diefenbach, 2010).

Zu ergänzen bleibt hierbei noch, dass Unterschiede zwischen Schülerinnen und Schülern unterschiedlicher Herkunftsgruppen im deutschen Bildungssystem bestehen: So sind Kinder und Jugendliche mit Zuwanderungshintergrund, deren Familien aus der Türkei und aus Italien stammen, in Bezug auf Indikatoren der Schulleistung besonders stark benachteiligt. Zugewanderte aus Polen und der ehemaligen UdSSR unterscheiden sich in leistungsrelevanten Merkmalen hingegen weniger deutlich von Schülerinnen und Schülern ohne Zuwanderungshintergrund (Segeritz, Walter & Stanat, 2010; vgl. Diefenbach, 2010). Die Formulierung fundierter herkunftslandspezifischer Erklärungsansätze und deren Untersuchung stehen noch am Anfang. Insgesamt kann allerdings davon ausgegangen werden, dass Schülerinnen und Schüler verschiedener Herkunftsländer im Durchschnitt gegenüber Gleichaltrigen ohne Zuwanderungshintergrund bezüglich sozialer Herkunft und Schulleistung benachteiligt sind (vgl. z.B. Pöhlmann et al., 2013; Walter, 2008a) und der allgemeine Zuwanderungshintergrund ein robustes, unterscheidendes Schülermerkmal darstellt, das in der empirischen Bildungsforschung häufig verwendet wird (Berkemeyer, Bos, Maniti, Hermstein & Khalatbari, 2013).

2.1.3 Individuelle und familiäre Determinanten zuwanderungsbezogener Disparitäten in psychosozialen Merkmalen

Studien der empirischen Bildungsforschung analysieren als Ergebnis von Bildungsprozessen häufig Schülerleistungen (vgl. Seidel & Shavelson, 2007). Die Leistungsentwicklung von Kindern und Jugendlichen ist jedoch nicht das alleinige pädagogische Ziel; Ziel von schulischer Bildung ist gleichermaßen die psychosoziale Entwicklung der Schülerinnen und Schüler im weiteren Sinne (z.B. Baumert et al., 2006).

Ein zentrales Schülermerkmal im psychosozialen Bereich stellt die Lernmotivation als wichtiger Motor des Lernverhaltens dar (vgl. Heckhausen & Heckhausen, 2006; Krapp, 1993; Rheinberg, 2006). Lernmotivation steht unter anderem in positivem Zusammenhang mit zukünftiger Leistung (z.B. Van der Werf, Opdenakker & Kuyper, 2008) sowie mit psychischer Gesundheit und dem Verbleib im Bildungssystem (OECD, 2004; Roeser, Eccles & Sameroff, 2000).

2.1.3.1 Theoretische Einordnung der Lernmotivation

In der pädagogischen Forschung wird unter motivationalen Merkmalen eine ganze Reihe von Konstrukten wie Lernmotivation, Interesse, Selbstkonzept oder Zielorientierungen subsummiert (Murphy & Alexander, 2000), die in Zusammenhang miteinander stehen.

Bei der (Lern-)Motivation wird zwischen intrinsischen und extrinsischen Aspekten unterschieden (vgl. Schiefele & Köller, 2006). *Intrinsische Lernmotivation* wird als Absicht definiert, eine Lernhandlung auszuführen, weil die Handlung an sich interessant ist beziehungsweise weil der Gegenstand der Handlung (z.B. auch ein bestimmtes Themenfeld) interessant ist. So kann eine Schülerin zum Beispiel ein Buch über mittelalterliche Geschichte lesen, weil sie die Tätigkeit des Lesens an sich gerne ausführt und weil sie daran interessiert ist, mehr über mittelalterliche Geschichte zu erfahren. *Extrinsische Lernmotivation* wird als Absicht definiert, eine Lernhandlung auszuführen, um positive Folgen herbeizuführen bzw. negative Folgen zu vermeiden. Positive Folgen können hier beispielsweise gute Noten oder eine Annäherung an ein Berufsziel sein; die Abwendung negativer Folgen das Verhindern des Klassenwiederholens. Es wird davon ausgegangen, dass Lernmotivation sowohl eher stabile Anteile in Form von Persönlichkeitsmerkmalen oder Dispositionen umfasst als auch aktuelle Anteile, die das Verhalten in einer bestimmten Situation vorhersagen und veränderbar sind (vgl. Schiefele & Köller, 2006; Schiefele & Streblow, 2005).

Eine sehr einflussreiche Theorie, die bisherige Ansätze zur intrinsischen und extrinsischen Motivation umfassend integriert, ist die Selbstbestimmungstheorie (*self-determination theory*) von Deci und Ryan (1985, 2000). Deci und Ryan konzipierten die Selbstbestimmungstheorie als Makrotheorie menschlicher Motivation, die verschiedene Stufen von extrinsischer Motivation unterscheidet sowie Rahmenbedingungen von intrinsischer und extrinsischer Motivation beschreibt. Die Selbstbestimmungstheorie stellt

damit eine erklärungsmächtige Theorie dar, die vielseitig anwendbar und empirisch gut fundiert ist.

In dieser Theorie wird extrinsische Motivation in vier verschiedenen Stufen als Prozess der zunehmenden Internalisierung von Handlungszielen verstanden. Auf der ersten Stufe der „externalen Regulation“ wird Verhalten beschrieben, das ausschließlich ausgeführt wird, um Belohnungen zu erhalten oder Bestrafungen zu vermeiden (z.B. Lernen um gute Noten zu erhalten). Auf der letzten Stufe der „integrierten Regulation“ wird Verhalten beschrieben, das ausgeführt wird, da sich eine Person mit bestimmten Zielen identifiziert und diese vollständig und kohärent in ihr Selbst integriert hat (z.B. Medizin studieren, weil es ermöglicht, anderen zu helfen, und das ein persönlicher Wert dieser Person ist). Insbesondere im schulischen Bereich ist die Verinnerlichung extrinsischer Motivation für die Initialisierung von und Aufrechterhaltung des Lernverhaltens bezüglich Inhalten, die nicht per se interessant sind, von großer Bedeutung für das Vorankommen im Bildungssystem (vgl. Deci & Ryan, 2000). Um externe Werte zu internalisieren und intrinsische Motivation zu entwickeln, müssen laut Deci und Ryan (2000) drei Grundbedürfnisse in der Handlung befriedigt werden: sich selbst als kompetent wahrnehmen (Kompetenzbedürfnis, *need for competence*), sich selbst als autonom handelndes Wesen mit Entscheidungsspielraum wahrnehmen (Autonomiebedürfnis, *need for autonomy*) und sich mit dem sozialen Umfeld verbunden fühlen (Bedürfnis nach interpersonaler Bezogenheit, *need for relatedness*). Insbesondere das Kompetenzbedürfnis und Autonomiebedürfnis sind als notwendige, aber nicht hinreichende Bedingungen der Entwicklung intrinsischer Motivation bzw. Internalisierung extrinsischer Motivation zu sehen, während das Bedürfnis nach interpersonaler Bezogenheit eher für die Aufrechterhaltung intrinsischer Motivation relevant ist (vgl. Deci & Ryan, 2000; Schiefele & Köller, 2006). Innerhalb der Selbstbestimmungstheorie wird somit unmittelbar die Bedeutung des Lernkontexts bei der Befriedigung der Grundbedürfnisse und Entwicklung intrinsischer Lernmotivation deutlich (siehe auch Niemiec & Ryan, 2009; *stage environment fit theory* Eccles et al., 1993). Die Bedeutung des Lernkontexts wird in Kapitel 2.2 der vorliegenden Arbeit dargestellt. In dem nun anschließenden Abschnitt wird zunächst auf individuelle Determinanten der Lernmotivation eingegangen.

2.1.3.2 Individuelle und familiäre Determinanten zuwanderungsbezogener Disparitäten in motivationalen Schülermerkmalen

Es gibt verschiedene theoretische Annahmen, die Motivationsunterschiede zwischen Heranwachsenden mit und ohne Zuwanderungshintergrund postulieren. Hierunter gibt es einerseits Annahmen, die eine geringere Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund im Vergleich zu ihren Peers ohne Zuwanderungshintergrund vorhersagen und andererseits Annahmen, die für eine höhere Lernmotivation von Heranwachsenden mit Zuwanderungshintergrund sprechen.

Annahmen einer geringeren Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund stützen sich unter anderem (1) auf den kulturökologischen Ansatz von Ogbu (1987; Ogbu & Simons, 1998), (2) auf die Wahrnehmung von begrenzten Handlungsalternativen und Konflikten zwischen Familien- und Schulkultur (vgl. Kumar & Maehr, 2010) sowie (3) auf mögliche niedrigere Fähigkeitsselbsteinschätzungen und die damit einhergehende geringere Lernmotivation (vgl. z.B. Roebbers, Mecheril & Schneider, 1998).

(1) Der *kulturökologische Ansatz* von Ogbu (1987, 2004; Ogbu & Simons, 1998) bezieht sich in seinem Ursprung auf Annahmen vor allem zur schulbezogenen Motivation von Afroamerikanern in den USA, die als „*involuntary minorities*“ bezeichnet werden. Demnach würden ihre Geschichte von Diskriminierung und Entmündigung und die Wahrnehmung, dass Bildung ihnen wenig Gewinn bringt, dazu führen, dass afroamerikanische Schülerinnen und Schüler den Wert akademischer Leistungen geringer einschätzen. Diese Schülerinnen und Schüler würden wahrnehmen, dass ihnen ein untergeordneter Status zugeschrieben wird, und sich in der Schule herabgesetzt fühlen. Dies führe letztendlich zu einer Entfremdung von der Schule und ihren Werten einschließlich einer geringeren Motivation, für die Schule zu lernen. Schulversagen würde als (einzige) praktikable Handlungsmöglichkeit im Bildungssystem gesehen, die gleichzeitig eine Demonstration von Solidarität mit der eigenen ethnischen Minderheitengruppe ermögliche.

(2) Ebenso beschreiben Kumar und Maehr (2010) Lernmotivation als Entscheidungsprozess, bei dem sich die Schülerin bzw. der Schüler für eine von mehreren Handlungsmöglichkeiten – in diesem Fall das Lernen für ein Schulfach – entscheidet. Die *Wahrnehmung realisierbarer Handlungsalternativen* werde dabei von den geteilten Normen, Werten, Erwartungen und Vorbildern innerhalb der ethnischen Gruppe und innerhalb der Familie geprägt, die im Gegensatz zu Normen und Werten der Schule

stehen können. Ein Beispiel ist die Orientierung an individualistischen oder kollektivistischen Gruppennormen. Die Orientierung an kollektivistischen Normen kann einer in der Schule betonten Verantwortlichkeit des Individuums für seinen Leistungserfolg entgegenstehen, wenn die ethnische Gruppe dem Schulsystem und schulbezogenem Engagement skeptisch gegenübersteht. Die Wahrnehmung einer Dissonanz zwischen „Schulkultur“ und „Kultur des Elternhauses“, beziehungsweise der Minderheitengruppe, kann sich insgesamt auf unterschiedliche Aspekte, wie Unterschiede in Hautfarbe, finanziellen Mitteln und Verhaltensregeln, beziehen (Kumar, 2006; Okagaki, 2001). Heranwachsende mit Zuwanderungshintergrund können sich also anders als andere Personen der Schule wahrnehmen. Durch diese Unterschiede zwischen Familien- und Schulkultur kann es Heranwachsenden schwerfallen, lernbezogene Entscheidungen zu treffen.

(3) Ein weiterer Aspekt, der mit der Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund in Zusammenhang stehen kann, ist ihr *akademisches Selbstkonzept* in verschiedenen schulischen Bereichen. Die Fähigkeit der Schülerinnen und Schüler in einer Domäne ist unabhängig vom Zuwanderungshintergrund ein zentrales Merkmal, das mit der Lernmotivation in Zusammenhang steht. Die oben vorgestellte Selbstbestimmungstheorie (Deci & Ryan, 1985, 2000) postuliert, dass Menschen dazu neigen, Tätigkeiten intrinsisch motiviert auszuführen, in denen sie sich als kompetent wahrnehmen. Dies bedeutet, dass die Lernmotivation mit der Schulleistung und mit dem akademischen Selbstkonzept einer Schülerin bzw. eines Schülers positiv zusammenhängt (Marsh, 1986; vgl. Abschnitt 2.1.3.1; für empirische Befunde siehe z.B. Artelt, Baumert, Julius-McElvany & Peschar, 2004; Marsh, Trautwein, Lüdtke, Köller & Baumert, 2005; Pekrun & Zirngibl, 2004).

Für Heranwachsende mit Zuwanderungshintergrund kann angenommen werden, dass insbesondere Schülerinnen und Schüler, die gerade in ein Land eingewandert sind, sich in einem neuen Schulsystem zurechtfinden müssen und zunächst ihre Leistungen als geringer einschätzen (Roebbers et al., 1998; vgl. Berry, 1997). Geringere Fähigkeitsselbsteinschätzungen von Schülerinnen und Schülern mit Zuwanderungshintergrund wären aufgrund ihrer potentiellen Schwierigkeiten in der Unterrichtssprache insbesondere in sprachlichen Fächern zu erwarten (vgl. Shajek, Lüdtke & Stanat, 2006). Hierbei könne es möglich sein, dass Schülerinnen und Schüler bei ihrer Fähigkeitseinschätzung ihre eigenen Leistungen in verschiedenen Bereichen vergleichend gegenüberstellen („*internal frame of reference*“, vgl. Marsh, 1986), was bei

einer besonders niedrigen Einschätzung eigener sprachlicher Fähigkeiten zu einer erhöhten Einschätzung eher weniger sprachlicher Fähigkeiten führen könne (vgl. Shajek et al., 2006).

Im Gegensatz zu den eben beschriebenen Annahmen einer niedrigeren Lernmotivation gibt es unter dem Stichwort „*immigrants' aspiration paradox*“ (z.B. Salikutluk, 2013) eine Reihe von Annahmen zu den häufig in empirischen Studien feststellbaren besonders hohen Bildungsaspirationen und Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund bei gleichzeitig geringerer Schulleistung. Diese umfassen (1) *immigrant optimism* (Kao & Tienda, 1995), (2) geringere Kenntnisse über das Bildungssystem und unrealistisch hohe Aspirationen (vgl. Salikutluk, 2013) und (3) Kompensation von Benachteiligungserfahrungen (vgl. Salikutluk, 2013)³.

(1) Kao und Tienda (1995) nehmen mit der „*immigrant optimism hypothesis*“ an, dass es sich bei Familien, die in ein Land einwandern, häufig um eine eher lern- und leistungsfokussierte Gruppe von Personen handelt. Hierbei beziehen sich die Autorinnen auf „*voluntary immigrants*“ im Sinne von Ogbu und Simons (1998). Personen wandern häufig aus, um für sich und ihre Kinder bessere Lebenschancen, zum Beispiel in Bezug auf Beschäftigungs- und Ausbildungsverhältnisse, zu erlangen. Eltern mit Zuwanderungshintergrund sollten demnach besonders hohe Bildungsaspirationen für ihre Kinder hegen, die sie in ihrer familiären Praxis auf die Kinder übertragen.

(2) Zusätzlich können besonders positive Erwartungen von Eltern mit Zuwanderungshintergrund dadurch entstehen, dass sie selbst nicht im Aufnahmeland die Schule besucht haben, über wenig Kenntnisse über das Bildungssystem verfügen und unrealistisch hohe Bildungserwartungen an ihre Kinder setzen (vgl. Salikutluk, 2013).

(3) Eine weitere Möglichkeit für besonders hohe Leistungsansprüche können auch wahrgenommene Benachteiligungen bzw. Diskriminierungen im Aufnahmeland sein, die sie durch besonders starke Anstrengung versuchen zu kompensieren (vgl. Salikutluk, 2013). Diese Annahme steht dabei teilweise im Gegensatz zum kulturökologischen

³ Zudem besteht die Möglichkeit, dass erhöhte Bildungsaspirationen und Lernmotivation von Schülerinnen und Schülern auch vor allem unterschiedliche Antworttendenzen zwischen Schülergruppen abbilden (vgl. Segeritz & Pant, 2013).

Ansatz von Ogbu (1987), der für „*involuntary minorities*“ besonders niedrige Lernmotivation aufgrund von Diskriminierungserfahrungen vorhersagt.

Die empirische Befundlage zu Unterschieden zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund in ihrer Lernmotivation und verbundenen Konzepten wie Selbstkonzept, Interesse oder Lernengagement kann zunächst als uneinheitlich eingestuft werden. Die Mehrzahl der Studien spricht jedoch eher dafür, dass keine Unterschiede zwischen den Schülergruppen oder erhöhte Motivation von Schülerinnen und Schülern mit Zuwanderungshintergrund bestehen, wobei es auch Hinweise auf fachspezifische Effekte gibt.

Shajek et al. (2006) konnten anhand deutscher PISA-2000-E-Daten zeigen, dass Schülerinnen und Schüler mit Zuwanderungshintergrund (definiert über nicht-deutsche Familiensprache) sich von Heranwachsenden ohne Zuwanderungshintergrund in ihrem Selbstkonzept unterschieden. Sie wiesen erwartungskonform ein niedrigeres verbales Selbstkonzept und ein höheres mathematisches Selbstkonzept als Schülerinnen und Schüler ohne Zuwanderungshintergrund auf. Weitere Analysen unter Berücksichtigung der jeweiligen Schulnoten geben Hinweise darauf, dass das erhöhte mathematische Selbstkonzept auf den Kontrast der Noten der Schülerinnen und Schüler in verschiedenen Domänen, also auf die Verwendung eines internen Bezugsrahmens, zurückzuführen sein kann.

Auch in der Studie von Roebbers et al. (1998), in der unter anderem Schülerinnen und Schüler ohne Zuwanderungshintergrund mit Kindern von Spätaussiedlern in der dritten und vierten Klasse längsschnittlich verglichen wurden, zeigten sich Unterschiede im verbalen Selbstkonzept. Schülerinnen und Schüler mit Zuwanderungshintergrund wiesen ein geringeres Niveau des verbalen Selbstkonzepts auf, unterschieden sich jedoch nicht im Entwicklungsverlauf des Selbstkonzepts von Kindern ohne Zuwanderungshintergrund. Es zeigten sich keine Hinweise auf Unterschiede im mathematischen Selbstkonzept.

Walter und Taskinen (2008) untersuchten intrinsische und extrinsische Lernmotivation im Bereich Naturwissenschaften anhand von PISA-2006-Daten und verglichen Schülerinnen und Schüler ohne und mit Zuwanderungshintergrund aus der ehemaligen Sowjetunion und der Türkei der ersten und zweiten Generation. Es zeigte sich eine geringere intrinsische Lernmotivation nur bei Schülerinnen und Schülern der zweiten Generation mit türkischem Zuwanderungshintergrund und erhöhte intrinsische

Lernmotivation nur bei Schülerinnen und Schülern, die selbst aus der ehemaligen Sowjetunion zugewandert sind. Schülerinnen und Schüler der zweiten Zuwanderungsgeneration, deren Eltern nicht aus der Türkei stammten, wiesen eine geringere extrinsische Lernmotivation auf. Es bestanden keine weiteren bedeutsamen Unterschiede zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund.

Demgegenüber fanden Stanat und Christensen (2006) anhand von PISA-2003-Daten erhöhtes Interesse und Freude an Mathematik sowie erhöhte instrumentelle Motivation im Fach Mathematik bei Heranwachsenden mit Zuwanderungshintergrund im Vergleich zu Schülerinnen und Schülern ohne Zuwanderungshintergrund in den meisten OECD-Ländern (vgl. auch Artelt et al., 2004 zu Befunden anhand PISA-2000-Daten). Diese Unterschiede blieben auch nach Kontrolle des sozioökonomischen Status bestehen. Unter Kontrolle der Mathematikleistung fielen die Motivationsunterschiede zugunsten von Schülerinnen und Schülern mit Zuwanderungshintergrund noch höher aus. Vertiefende Analysen von Stanat, Segeritz und Christensen (2010) mit nationalen und internationalen PISA-2003-Daten zeigten, dass das Befundmuster leicht zwischen verschiedenen Herkunftsgruppen variiert, insgesamt aber Schülerinnen und Schüler der ersten und zweiten Zuwanderergeneration erhöhte mathematikbezogene Lernmotivation aufweisen. Des Weiteren fand Stanat (2006b) anhand deutscher PISA-2000-Daten unter Kontrolle kognitiver Grundfähigkeiten erhöhte Anstrengungsbereitschaft, leicht erhöhtes akademisches Selbstkonzept und höhere angestrebte Bildungsabschlüsse bei Hauptschülerinnen und -schülern mit Zuwanderungshintergrund im Vergleich zu Schülerinnen und Schülern ohne Zuwanderungshintergrund (für weitere Befunde zur erhöhten Motivation von Schülerinnen und Schülern mit Zuwanderungshintergrund in den USA siehe z.B. Chen & Stevenson, 1995; Fuligni, 2001; Fuligni, Witkow & Garcia, 2005; Kao & Tienda, 1995).

Insgesamt spricht die Mehrzahl der hier dargestellten Studien eher für eine leicht erhöhte Lernmotivation bei Schülerinnen und Schülern mit Zuwanderungshintergrund, wobei dies insbesondere in weniger sprachlich ausgerichteten Fächern der Fall ist. Bezüglich sprachfokussierteren Fächern bzw. Tätigkeiten (z.B. Deutschunterricht, Lesen und Schreiben) gibt es auch Hinweise auf niedrigere Lernmotivation und niedrigeres Selbstkonzept von Heranwachsenden mit Zuwanderungshintergrund.

2.1.4 Zwischenfazit: Notwendigkeit der Betrachtung kontextueller Determinanten

Bildungsauftrag der Schule ist es, alle Schülerinnen und Schüler in ihrer leistungsbezogenen und psychosozialen Entwicklung möglichst effektiv zu unterstützen. Groß angelegte Vergleichsstudien zeigen jedoch wiederholt geringere Schulleistungen und Bildungsabschlüsse von Schülerinnen und Schülern mit Zuwanderungshintergrund. Ansätze zur Erklärung von zuwanderungsbezogenen Disparitäten nehmen vor allem Merkmale von Schülerinnen und Schülern sowie ihren Familien als ursächlich für Leistungsunterschiede an. Empirisch zeigt sich, dass sich die durchschnittlich geringeren Leistungen von Schülerinnen und Schülern mit Zuwanderungshintergrund weitgehend auf ihre häufig weniger privilegierte soziale Herkunft sowie geringere Kenntnisse in der Unterrichtssprache zurückführen lassen. Auch nach Berücksichtigung von Sprachkenntnissen und Merkmalen der sozialen Herkunft bleiben jedoch weiterhin zuwanderungsbezogene Disparitäten bestehen.

Ein weiteres leistungsrelevantes Merkmal ist die Lernmotivation: Insbesondere die intrinsische Lernmotivation ist für die Initialisierung und Aufrechterhaltung des Lernverhaltens im schulischen Bereich von Bedeutung. Inwiefern sich Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund in ihrer Lernmotivation unterscheiden, ist Gegenstand einer Reihe theoretischer Annahmen. Die empirische Befundlage zu Motivationsunterschieden zwischen Heranwachsenden mit und ohne Zuwanderungshintergrund ist uneinheitlich. Es überwiegen jedoch Befunde, die für eine tendenziell erhöhte Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund – bei gleichzeitig durchschnittlich geringerem Leistungsniveau – sprechen. Zudem gibt es Hinweise auf fachspezifische Besonderheiten, mit erhöhter Lernmotivation in weniger sprachlichen Domänen (z.B. Mathematik) und niedrigerer Lernmotivation bzw. keinen Unterschieden zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund in sprachlichen Domänen (z.B. Lesefähigkeit).

Erklärungen für zuwanderungsbezogene Disparitäten beziehen sich bislang überwiegend auf Merkmale einzelner Schülerinnen und Schüler und ihren familiären Hintergrund. Da ein Großteil der Lernprozesse jedoch im schulischen Rahmen stattfindet, stellt sich die Frage nach der Bedeutung des Lernumfelds für die Leistungs- und Motivationsentwicklung der Schülerinnen und Schüler. Verschiedene Autorinnen (Bellin, 2009; Diefenbach, 2010; Stanat, 2006a, 2006b) weisen insbesondere auf eine mögliche Relevanz der Schulklassenzusammensetzung als Merkmal des Lernumfelds zur Erklärung

von Disparitäten zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund hin.

Die potentielle Relevanz des Lernumfelds und die Notwendigkeit der Analyse seiner Bedeutung für Unterschiede zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund lässt sich aus verschiedenen Perspektiven ableiten: Auftrag des Bildungssystems ist es, Bildungsgerechtigkeit oder Chancengerechtigkeit herbeizuführen, die erfordert, dass das Bildungssystem bestehende Disparitäten aktiv durch die Förderung von Schülerinnen und Schülern mit besonderen Bedürfnissen verringert (vgl. z.B. Darstellung bei Hopf, 2011; Stanat, 2006a). Die Bedeutung des Lernumfelds für eine erfolgreiche akademische Entwicklung zu untersuchen ist hierfür wichtig. Des Weiteren bietet die Ermittlung bedeutsamer Merkmale des Lernumfelds, wie die Klassenzusammensetzung und der Unterricht, Ansatzmöglichkeiten zur institutionell gesteuerten Verringerung von zuwanderungsbezogenen Disparitäten, beispielsweise durch gesteuerte Umstrukturierung der Schülerschaft von Schulen oder Weiterbildungen für Lehrkräfte (Diefenbach, 2010). In dem nachfolgenden Kapitel 2.2 wird dementsprechend auf kontextuelle Determinanten der akademischen Entwicklung von Schülerinnen und Schülern eingegangen.

2.2 Zuwanderungsbezogene Disparitäten: Erklärungen durch kontextuelle Determinanten der Schülerleistung und psychosozialer Merkmale

In Kapitel 2.2 der vorliegenden Arbeit werden kontextuelle Determinanten für die leistungsbezogene und psychosoziale Entwicklung von Schülerinnen und Schülern dargestellt. Fokus der Arbeit ist die zuwanderungsbezogene Klassenzusammensetzung als kontextuelle Determinante. In den folgenden Abschnitten werden zunächst die übergeordnete Bedeutung des Lernkontexts für die leistungsbezogene und psychosoziale Entwicklung einschließlich relevanter Ebenen des Kontexts (2.2.1), eine begriffliche Einordnung der zuwanderungsbezogenen Klassenzusammensetzung und ihre Messung (2.2.2) sowie allgemeine theoretische Modelle der Vermittlung von Effekten der Klassenzusammensetzung und Modelle der Unterrichtsqualität vorgestellt (2.2.3). Daran anschließend werden empirische Befunde zu direkten Effekten der zuwanderungsbezogenen Zusammensetzung zusammenfassend dargestellt (2.2.4) sowie theoretische Annahmen und Befunde zur Genese von Effekten der Zusammensetzung auf

Leistung (2.2.5) und psychosoziale Merkmale (2.2.6) ausgeführt. Da die zuwanderungsbezogene Zusammensetzung häufig mit leistungsbezogener und sozialer Zusammensetzung assoziiert ist, wird auch auf diese beiden Merkmale eingegangen. Kapitel 2.2 schließt mit einer Darstellung methodischer Überlegungen zur Untersuchung von Effekten der zuwanderungsbezogenen Klassenzusammensetzung.

2.2.1 Theoretische Annahmen und empirische Befunde zur Bedeutung verschiedener Ebenen des Lernumfelds

Bereits in den ersten Jahrzehnten des 20. Jahrhunderts setzten sich die Sozialwissenschaften mit dem Verhältnis von Individuum und Kontext auseinander. Meilensteine waren hierbei der humanökologische Forschungsansatz des Soziologen Robert E. Park (Park & Burgess, 1928), die Feldtheorie sozialer Beziehungen des Psychologen Kurt Lewin (1935) und die Begründung einer ökologischen Psychologie (Barker, 1968) sowie eine systematische theoretische Fundierung der ökologischen Sozialisationsforschung durch Urie Bronfenbrenner (1976; vgl. dazu auch Grundmann & Kunze, 2008).

Park und Burgess (1928) bezogen sich vor allem auf die räumliche Struktur der Umwelt in Analysen von Bevölkerungsstrukturen und sozialen Wanderungsbewegungen in den USA, während Lewin (1935, 1963) sich mit der Wahrnehmung der Umwelt durch die in ihr handelnden Akteure und die Bedeutung wechselseitiger Beeinflussung von Personen in sozialen Handlungsbezügen für deren Persönlichkeitsentwicklung auseinandersetzte. Unter Bezugnahme auf Lewin formulierte Urie Bronfenbrenner (1976) schließlich seinen Ansatz der Ökologie menschlicher Entwicklung. Zentrale Annahmen sind hierbei, dass der Kontext eines Individuums mehrere Lebensbereiche und Verbindungen zwischen ihnen im zeitlichen Entwicklungsverlauf umfasst und dass das Individuum nicht nur durch seine Umwelt beeinflusst wird, sondern fortschreitend seine Umwelt aktiv mitgestaltet, was zu einem reziproken Prozess gegenseitiger Anpassung von Individuum und Kontext führt (siehe auch Bronfenbrenner, 1996). In dieser Theorie wird der Kontext als eine ineinander geschachtelte Anordnung von Systemen (Mikro-, Meso-, Exo-, Makro- und Chronosystem) auf unterschiedlichen Ebenen konzipiert. Laut Bronfenbrenner sind hierbei insgesamt die im Mikrosystem ablaufenden proximalen Prozesse bedeutsamer für die Entwicklung eines Individuums als die Strukturen der

distalere Meso- und Exosysteme. Gleichzeitig aber werden die proximalen Prozesse auf Mikroebene von sozialen Rahmenbedingungen auf höherer Ebene beeinflusst.

Bezogen auf die akademische Entwicklung von Schülerinnen und Schülern lässt sich unmittelbar auf eine entwicklungsrelevante Bedeutung des schulischen Umfelds schließen, da junge Menschen in keiner anderen sozialen Einrichtung so viel Zeit verbringen wie in der Schule (vgl. Meece & Schaefer, 2010). Dies steht jedoch im Gegensatz zu dem häufigeren Fokus auf individuenzentrierte Erklärungsansätze in den Sozialwissenschaften (vgl. Battistich, 2010). Erst in den letzten Jahrzehnten wurden Kontextmerkmale zur Erklärung von Entstehung und Aufrechterhaltung von Disparitäten und bei der Untersuchung der akademischen Entwicklung von Kindern und Jugendlichen verstärkt einbezogen. Laut Weinert, Schrader und Helmke (1989) waren insbesondere langjährige Einwände der Bildungssoziologie und Entwicklungspsychologie gegen die Untersuchung von kontextbezogenen Schul-, Lehrkraft- und Unterrichtseinflüssen auf die Leistungsentwicklung für diese Situation verantwortlich. Als Einwand wurde hier meist angeführt, dass individuelle Personenmerkmale für die individuelle Entwicklung entscheidend seien und Merkmale des Kontexts in ihrem Einfluss vernachlässigbar sind (vgl. Weinert et al., 1989). So kamen zum Beispiel Jencks et al. (1972) in ihrer Studie zu Ungleichheiten im Bildungserfolg zu dem Schluss, dass die Schule bzw. die Dauer der Beschulung nicht bedeutsam mit späterem Einkommen zusammenhängt und somit nicht zur Verringerung sozialer Ungleichheit beitragen könne (vgl. Tanner, 1973). Insbesondere mit dem sogenannten Coleman-Report (Coleman et al., 1966) rückte die Bedeutung des Kontexts für die Entwicklung von Schülerinnen und Schülern stärker in das wissenschaftliche Interesse. Der Report zeigte unter anderem auf, dass die soziale Zusammensetzung in Schulen stärker mit den individuellen Schülerleistungen zusammenhängt als irgendein anderer Faktor auf Schulebene. Hierbei lag der Fokus vor allem auf der leistungsbezogenen Entwicklung; die Bedeutung des Kontexts für die psychosoziale Entwicklung, die ebenfalls ein wichtiges Ergebnis von Bildungsprozessen und Ziel eines Ausbildungssystems ist, rückte erst in den letzten Jahren stärker in den Forschungsfokus (vgl. Meece & Schaefer, 2010; Battistich, 2010). Bedeutsam war hierbei insbesondere das Stage-Environment-Fit-Modell von Eccles et al. (Eccles, et al., 1993), das Annahmen des Person-Environment-Fit-Modells (vgl. Hunt, 1975), der Selbstbestimmungstheorie (Deci & Ryan, 1985, 2000; Abschnitt 2.1.3.1) und der ökologischen Sozialisationsforschung von Bronfenbrenner (1976; siehe oben) verband,

um die Bedeutung des sozialen Schulkontexts vor allem für die Lernmotivation von Jugendlichen herauszuarbeiten (siehe auch Abschnitt 2.2.6). Nicht zuletzt trug auch zur Expansion der Forschung zum Individuum im Schulkontext die Weiterentwicklung von Forschungsmethoden, wie der Mehrebenenanalyse, bei (vgl. Meece & Schaefer, 2010).

Die Beschreibung eines Kontexts, der aus mehreren Ebenen besteht, die entwicklungsrelevant sind und sich in der Proximität zum Individuum unterscheiden, findet sich auch in neueren Modellen der schulischen Entwicklung und der Schulqualität im weiteren Sinne wieder (z.B. Eccles & Roeser, 2010; Wilkinson, Paar, Fung, Hattie & Townsend, 2002; vgl. Ditton & Müller, 2011). Zentral für die Systematisierung der Kontextebenen für die vorliegende Arbeit ist das Modell, das Baumert, Blum und Neubrand (2004) in Anlehnung an die Modellannahmen Bronfenbrenners und unter Rückgriff auf das Modell einer Arbeitsgruppe der OECD entwickelten (siehe Abbildung 2.2). In dem Modell wird ein System von Ebenen distaler und proximaler Einflussfaktoren dargestellt, die ein Rahmenkonzept zur Untersuchung von Indikatoren des Lehrens und Lernens bilden (vgl. auch Stanat, 2006a).

Nachfolgend werden nun kurz verschiedene Ebenen und empirische Befunde zu ihrer Relevanz für die Wirkung der Klassenebene und für die akademische Entwicklung von Schülerinnen und Schülern dargestellt. In den nachfolgenden Abschnitten wird dann auf die im Zentrum der vorliegenden Arbeit stehende Ebene der Schulklasse mit ihrer zuwanderungsbezogenen Schülerzusammensetzung detailliert eingegangen.

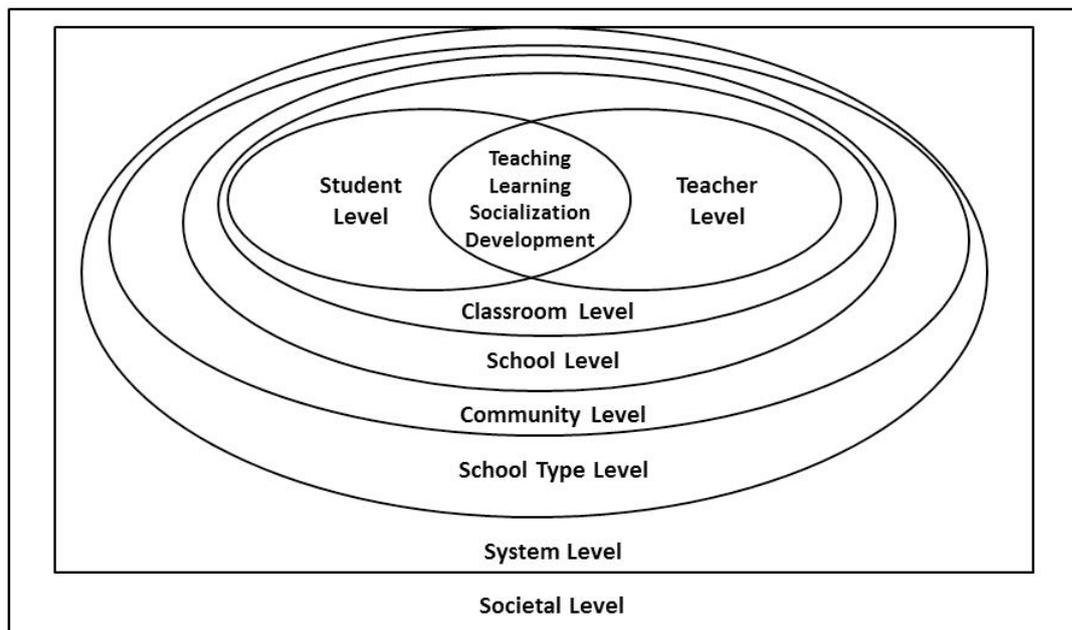


Abbildung 2.2: Rahmenkonzept für die Analyse von Indikatoren des Lehrens und Lernens
(aus: Baumert et al., 2004, S. 146)

Innerhalb der gesellschaftlichen Ebene liegt die Ebene des Bildungssystems eines Landes oder Bundeslandes. Charakteristisch für das System eines Landes sind unter anderem die inhaltliche Ausgestaltung und zeitliche Organisation von Curricula, Richtlinien zur Lehrerbildung, politische Vorgaben zur Finanzierung von Schulen, zur Struktur des Bildungssystems und zu Schulungszeiten (vgl. Dreeben & Barr, 1988; Eccles & Roeser, 2010, 2011; Stanat, 2006a; Weinert et al., 1989). Empirische Studien stützen die Annahme, dass diese Ebene für Effekte auf hierarchieniedrigeren Ebenen bedeutsam ist. So zeigten beispielsweise Braun, Wang, Jenkins und Weinbaum (2006) anhand von Analysen der Daten von zehn Staaten der USA, dass Unterschiede in bildungspolitischen Reformen in Zusammenhang mit Unterschieden in der Leistungssteigerung von Schülerinnen und Schülern mit afroamerikanischem Hintergrund standen. Chudgar und Luschei (2009) zeigten anhand einer Analyse von TIMSS-Daten von Viertklässlerinnen und Viertklässlern in 25 Teilnehmerländern, dass der Zusammenhang zwischen Merkmalen der Schule und der Schülerleistung stärker ausgeprägt ist in Ländern, in denen allgemein starke Einkommensunterschiede zwischen Bevölkerungsgruppen bestehen. Für das deutsche Bildungssystem weisen Analysen von Baumert et al. (2006) darauf hin, dass der Zusammenhang zwischen dem Besuch von

Schulen unterschiedlicher Bildungsgänge und Schülerleistung zwischen Bundesländern variiert. Diese Zusammenhangsunterschiede gehen mit der unterschiedlichen schulstrukturellen Ausdifferenzierung in Bildungsgänge in den Bundesländern einher (für Unterschiede zwischen Bundesländern in Deutschland in zuwanderungsbezogenen Disparitäten siehe z.B. auch Haag et al., 2012; Pöhlmann et al., 2013).

Die Bildungsgänge oder Schularten innerhalb eines Bildungssystems bilden eine weitere Ebene des schulischen Entwicklungskontexts. Die Bildungsgänge können sich bezüglich ihrer unterrichtlichen Rahmenbedingungen zum Beispiel durch Lehrpläne und Merkmale der Schülerschaft unterscheiden. Verschiedene Bildungsgänge eines Schulsystems und Arten der Leistungsgruppierung können allgemein mit der leistungsbezogenen und psychosozialen Entwicklung von Schülerinnen und Schülern in Zusammenhang stehen (vgl. z.B. Baumert et al., 2006; Becker et al., 2012; Diefenbach, 2010; Gamoran, 1992; Knigge, 2009; Köller, Schütte, Zimmermann, Retelsdorf & Leucht, 2013; Maaz, Baumert & Trautwein, 2010). Leistungsgruppierungen in Bildungsgängen oder innerhalb von Schulen werden mit dem Ziel einer leistungsbezogenen Homogenisierung der Schülerschaft und einer optimierten Förderung der Schülerinnen und Schüler durch erleichterte Anpassungsmöglichkeiten des unterrichtlichen Angebots an die Schülervoraussetzungen durchgeführt. Da die Schülerleistung mit der sozialen Herkunft kovariiert, kann die Leistungsgruppierung jedoch auch sozial segregierte Schulen mit sich bringen (vgl. Baumert et al., 2006; Maaz et al., 2010). Die verschiedenen Schularten werden auch als „differenzielle Lern- und Entwicklungsmilieus“ bezeichnet, die sich in der Zusammensetzung ihrer Schülerschaft (z.B. durch eine Konzentration von Risikoschülern an bestimmten Schularten) sowie in institutionellen Merkmalen wie Stundentafeln, Lehrplänen, Unterrichtskulturen und Lehrerbildung unterscheiden können. Für das deutsche Bildungssystem zeigten beispielsweise Baumert et al. (2006), dass 72 Prozent der Zwischenschulvarianz in der Leseleistung von Neuntklässlerinnen und Neuntklässlern allein durch die Schularten erklärt wird. Bei Varianzzerlegung in spezifische und konfundierte Komponenten anhand einer Kommunalitätenanalyse zeigte sich, dass ein Großteil der Zwischenschulvarianz auf eine konfundierte Komponente von Schulart, kognitivem Fähigkeitsniveau und der sozialen Zusammensetzung der Schülerschaft zurückzuführen ist. Schulart, kognitives Fähigkeitsniveau und soziale Zusammensetzung wiesen jedoch auch je eine spezifische, wenn auch geringe, Erklärungskraft auf. Köller et al. (2013) analysierten zwei Kohorten

aus dem EIKA-Projekt in Bremen in einem längsschnittlichen Design zur Vorhersage der Mathematik- und Leseleistung von Neuntklässlerinnen und Neuntklässlern. Sie konnten zeigen, dass unter Kontrolle von Individualmerkmalen die Schulart signifikant mit der Leistungsentwicklung in Zusammenhang stand. Unter Kontrolle der Leistung in der siebten Klasse und der kognitiven Grundfähigkeiten sowie auf kognitive Grundfähigkeiten bezogene Klassenkomposition erreichten Schülerinnen und Schüler des Gymnasiums einen höheren Leistungszuwachs in Mathematik als Heranwachsende nicht-gymnasialer Schularten. Die Leseleistungsentwicklung der Schülerinnen und Schüler war jedoch nicht bedeutsam von Schulart und Schülerzusammensetzung beeinflusst, was die Autorinnen und Autoren unter anderem auf eine geringe Bedeutung systematischer Instruktion im Leseverstehen innerhalb der Sekundarstufe I zurückführen. Eine Vielzahl von Studien zeigt insgesamt, dass sich die akademische Entwicklung von Kindern und Jugendlichen in verschiedenen Schularten voneinander unterscheidet (z.B. Becker, 2009; Becker et al., 2012; Köller et al., 2013; Retelsdorf, Becker, Köller & Möller, 2012). Inwiefern die Schul- oder Klassenzusammensetzung der Schülerschaft auch nach Kontrolle der Schulart für die leistungsbezogene und psychosoziale Entwicklung von Schülerinnen und Schülern bedeutsam ist, wird unter anderem in Abschnitten 2.2.4 bis 2.2.6 ausführlicher dargestellt.

Ein weiterer Bereich, der zwar nicht institutioneller Bestandteil des Bildungssystems ist, von dem jedoch angenommen werden kann, dass er die Entwicklung von Schülerinnen und Schülern beeinflusst, ist die Ebene der Gemeinde oder des Wohnumfelds mit Merkmalen wie Stadt- oder Landbevölkerung und Aspekten der soziokulturellen Zusammensetzung der Bevölkerungsstruktur in einer Nachbarschaft. Unter dem Stichwort der ethnischen und sozialen Segregation setzt sich die Forschung mit Zusammenhängen zwischen der soziokulturellen Zusammensetzung einer Nachbarschaft und der Zusammensetzung von Schulen, Schulklassen und der Schülerleistung auseinander (vgl. z.B. Baumert, Carstensen & Siegle, 2005; Friedrichs, 2008; Helbig, 2010; Karsten et al., 2006; Kristen, 2008; Maggi, Hetzman, Kohen & D'Angiulli, 2004). So zeigte beispielsweise Helbig (2010) anhand der Analyse von Daten der ELEMENT-Studie mit Berliner Grundschülerinnen und Grundschülern der vierten bis sechsten Klassenstufe (Lehmann, 2008), dass soziostrukturelle Merkmale der Nachbarschaft mit der Kompetenzentwicklung der Schülerinnen und Schüler zusammenhängen. Eine besonders positive Kompetenzentwicklung zeigte sich im

Wohnumfeld mit einer günstigen Sozialstruktur unter Kontrolle individueller Schülermerkmale. Es ergaben sich jedoch keine Hinweise darauf, dass sich sozial benachteiligte Nachbarschaften negativ auf die Kompetenzentwicklung auswirken.

Kristen (2008) ging der Frage nach, ob das elterliche Schulwahlverhalten ethnische Segregation an Schulen beeinflusst. Die Autorin zeigte anhand der Daten von in Essen wohnhaften Familien, dass Familien mit türkischem Hintergrund bei der Grundschulwahl bedingt durch geringeres Wissen über das Bildungssystem häufig nur eine Schule in Betracht ziehen, die dann meist in dem direkten Wohnumfeld liegt, anstatt verschiedene potentielle Schulen für ihr Kind zu vergleichen. Elternentscheidungen beeinflussen also die Verbindung von segregierten Nachbarschaften und segregierten Schulen. Dies kann dazu führen, dass Effekte der Zusammensetzung einer Nachbarschaft und der Zusammensetzung einer Schule oder Schulklasse miteinander konfundiert sind (siehe auch Abschnitt 8.4).

Merkmale der Ebene der Schule beeinflussen ebenfalls hierarchieniedrigere Ebenen. Bedeutsam ist hier die Zusammensetzung der Schülerschaft einer Schule unter anderem bezogen auf Leistung, soziale Herkunft oder Zuwanderungshintergrund, die innerhalb der nachfolgenden Abschnitte näher beleuchtet wird. Des Weiteren wird auf dieser Ebene insbesondere die Ressourcenausstattung von Schulen als zentrales Merkmal unter anderem zur Erklärung von durchschnittlich geringeren Schulleistungen von Heranwachsenden mit Zuwanderungshintergrund und/oder niedrigem sozioökonomischem Status diskutiert (vgl. Dumont, Neumann, Maaz, et al., 2013; Eccles & Roeser, 2011; Hanushek, Kain & Rivkin, 2009; Harker & Tymms, 2004; Hattie, 2002; Palardy, 2013; Raudenbush, Fotiu & Cheong, 1998; Rumberger & Palardy, 2005; Wilkinson, Hattie, Parr, & Townsend, 2000). Als ursächlich für eine Verbindung von Schulressourcenausstattungen und Schülermerkmalen werden die Verfügbarkeit in Wohnortnähe, systematische Unterschiede in Elternentscheidungen bei der Schulwahl sowie Schulwahlverhalten der Lehrkräfte diskutiert. Empirisch konnte gezeigt werden, dass Schülerinnen und Schüler mit Zuwanderungshintergrund und/oder niedrigem sozioökonomischem Status häufiger an Schulen und auch Schularten vertreten sind, die sich durch weniger günstige organisationale und strukturelle Merkmale wie höhere Schul- und Klassengröße, weniger qualifizierte Lehrkräfte und weniger zusätzliche Förderprogramme auszeichnen (Baumert & Schümer, 2001; Hanushek et al., 2009; Raudenbush et al., 1998; Roscigno, 2000).

Insgesamt sind verschiedene Charakteristika der Schulebene bedeutsam, die den Rahmen für Lernprozesse auf hierarchieniedrigerer Ebene der Schulklasse bilden: zum Beispiel Normen und Praktiken einer Schule im Sinne einer Schulkultur, Merkmale des Lehrerkollegiums einer Schule, wie etwa Kooperationsbereitschaft, sowie auch formale Schulcharakteristika, wie staatliche oder private Trägerschaft, finanzielle Ausstattung, Schulgröße oder Beschaffenheiten des Schulgebäudes (vgl. z.B. Benner, Graham & Mistry, 2008; Eccles & Roeser, 2010, 2011; Evans, Yoo & Sipple, 2010; Klieme, Steinert & Hochweber, 2010; Lubienski & Lubienski, 2006; Opdenakker & Van Damme, 2007; Palardy, 2013; Rumberger & Palardy, 2005). So zeigten beispielsweise Opdenakker und Van Damme (2007) mittels Ergebnissen von Regressionsanalysen auf Schulebene, dass die Kooperation im Lehrerkollegium positiv mit der mittleren Mathematikleistung der Schule assoziiert ist. Lubienski und Lubienski (2006) kamen in ihren Analysen von Schulen verschiedener Trägerschaften in den USA zu dem Schluss, dass Heranwachsende in Privatschulen höhere Leistungsergebnisse erzielen, was sich jedoch durch die Selektivität der Privatschulen erklären ließ. Für das deutsche Bildungssystem konnten zum Beispiel Klieme et al. (2010) anhand der DESI-Daten in einem Drei-Ebenen-Modell zeigen, dass die Kooperation im Fachkollegium und die wahrgenommene Relevanz von Kompetenzziele auf Seiten der Lehrkräfte mit einigen Unterrichtsmerkmalen auf Klassenebene (Strukturiertheit, Häufigkeit sprachbezogener Lerngelegenheiten), jedoch nicht mit dem individuellen Leistungs- und Motivationszuwachs der Schülerinnen und Schüler in Zusammenhang standen. Dies unterstreicht die allgemeine Annahme, dass distale Kontextebenen einen Rahmen für das proximale Entwicklungsumfeld bilden. Die in dem Beitrag berichteten Varianzanteile von Unterrichtsmerkmalen sowie Leistungs- und Motivationszuwachsen im Verlaufe eines Schuljahrs auf den drei Ebenen (Individuum, Klasse, Schule/Schulart) in einem Nullmodell zeigten eine durchgängig höhere Ausprägung dieser Varianzanteile auf Klassenebene als auf Schulebene, was noch einmal die Bedeutung der proximalen Umwelt hervorhebt (vgl. z.B. De Fraine, Van Damme, Van Landeghem, Opdenakker & Onghena, 2003; Hamre & Pianta, 2010; Wilkinson et al., 2002). Auf die Ebene der Schulklasse mit ihrer Schülerzusammensetzung und den in ihr stattfindenden Prozessen, wird in den nächsten Abschnitten (2.2.2 bis 2.2.6) eingegangen. Im nun direkt anschließenden Abschnitt wird das Konstrukt Klassenzusammensetzung begrifflich eingeordnet und ihre Operationalisierung diskutiert. Daran schließt sich eine Ausführung theoretischer

Modellannahmen zur Genese von Effekten der Klassenzusammensetzung auf die leistungsbezogene und psychosoziale Entwicklung von Schülerinnen und Schülern an.

2.2.2 Die Ebene der Schulklasse als proximales Entwicklungsumfeld: Klassenzusammensetzung als Anteil Heranwachsender mit Zuwanderungshintergrund und zugewanderungsbezogene Heterogenität

2.2.2.1 Begriffliche Einordnung

Laut Hamre und Pianta (2010) verbringen Schülerinnen und Schüler bis zum Ende ihrer Schullaufbahn mindestens 15.000 Stunden im Klassenraum mit ihren Lehrkräften und Klassenkameradinnen und -kameraden. Klassen sind also ein bedeutsames Lernumfeld, in dem der Großteil institutionalisierten Unterrichts stattfindet und in dem Schülerinnen und Schüler mit den Lehrkräften und untereinander interagieren. Mit wem eine Schülerin bzw. ein Schüler eine Klasse besucht, kann leistungsbezogene und psychosoziale Lernergebnisse von Heranwachsenden beeinflussen. Solche Effekte der Zusammensetzung der Schülerschaft wurden bereits durch Veröffentlichungen unter anderem zur sozialen Schulzusammensetzung Anfang der 1960er Jahre bekannt (Blau, 1960; Wilson, 1959; vgl. Dreeben & Barr, 1988) und erreichten spätestens mit dem sogenannten Coleman-Report bzw. der Studie „*Equality of Educational Opportunity*“ (Coleman et al., 1966) weitreichendes wissenschaftliches und öffentliches Interesse. Seitdem setzen sich verschiedene Disziplinen wie Psychologie, Erziehungswissenschaften, Soziologie oder Ökonomie mit der Erforschung des Effekts auseinander. Auch in aktuellen internationalen Studien wird diese Thematik diskutiert (Agirdag, Van Houtte & Van Avermaet, 2012; Goldsmith, 2011; Benner & Crosnoe, 2011; Stanat, Schwippert & Gröhlich, 2010; Klieme et al., 2010; Palardy, 2013). Als Ansatz zur Erklärung von zugewanderungsbezogenen Disparitäten wurden Effekte der Zusammensetzung der Schülerschaft jedoch bislang kaum einbezogen (vgl. Bellin, 2009; Diefenbach, 2010; Stanat, 2006a, 2006b).

In der vorliegenden Arbeit wird der Begriff „Klassenzusammensetzung“ verwendet, um übergeordnet auf verschiedene Charakteristika der Zusammensetzung, wie dem Niveau oder der Streuung eines Merkmals in einer Gruppe, Bezug zu nehmen. Klassischerweise wird unter der Zusammensetzung die durchschnittliche Merkmalsausprägung einer Klasse verstanden, deren Effekt in Kompositionsanalysen untersucht wird. Die verschiedenen Disziplinen verwenden hier teils unterschiedliche

Begrifflichkeiten, um dasselbe Phänomen zu beschreiben. So wird beispielsweise in der Ökonomie vor allem der Begriff *peer effects* verwendet (vgl. Angrist, 2014; siehe dazu auch Van Ewijk & Slegers, 2010a), während in anderen Disziplinen vor allem Begriffe wie „*contextual effects*“, „*school-mix effects*“ oder „*composition effect*“ bzw. „*compositional effect*“, also Kompositionseffekt, zu finden sind (vgl. Harker & Tymms, 2004). Letzterer hat sich mittlerweile auch aufgrund seiner Eindeutigkeit etabliert (vgl. z.B. auch Dumont, Neumann, Maaz et al., 2013; Harker & Tymms, 2004). Zusätzlich variieren Begrifflichkeiten teilweise auch mit der Art der betrachteten abhängigen Variablen: Bei Effekten auf Schülerleistung ist meist von Kompositionseffekten die Rede. In Untersuchungen von Effekten der Klassenzusammensetzung zum Beispiel auf das Schülerselbstkonzept werden sie eher als Referenzgruppeneffekte oder Big-Fish-Little-Pond-Effekte bezeichnet (vgl. Abschnitt 2.2.6).

Kompositionseffekte beziehen sich also auf Effekte von Gruppenmerkmalen, die durch das Aggregieren von Individualmerkmalen gewonnen wurden. Sie stellen die Effektschätzung des Aggregatmerkmals unter Kontrolle des Individualmerkmals dar (z.B. Harker & Tymms, 2004; Lüdtke, Marsh, Robitzsch & Trautwein, 2011; Raudenbush & Bryk, 2002; für methodische Ausführungen zu Kompositionseffekten siehe Abschnitt 2.2.7). Das bedeutet, dass man mit Untersuchungen von Kompositionseffekten Informationen darüber gewinnen kann, welchen Erklärungsbeitrag die Klassenzusammensetzung zu Leistungsunterschieden von Schülerinnen und Schülern leistet, die bezüglich der Hintergrundmerkmale vergleichbar sind. Die Untersuchung von Effekten der Klassenzusammensetzung auf Leistung und psychosoziale Schülermerkmale kann sich auf eine Vielzahl von Merkmalen der Klassenzusammensetzung beziehen, wie Zuwanderungshintergrund, sozioökonomischen Status, Vorkenntnisse, Geschlecht, Klassenwiederholer oder sonderpädagogischen Förderbedarf.

2.2.2.2 Operationalisierung der Klassenzusammensetzung

„Klassenzusammensetzung“ umfasst also inhaltlich verschiedene Charakteristika, die eine unterschiedliche Operationalisierung erfordern. In den meisten Studien werden das Niveau eines Merkmals (z.B. Vorkenntnisniveau) oder ein Prozentanteil (z.B. Anteil von Mädchen) betrachtet. Seltener werden zur Abbildung der Klassenzusammensetzung auch Maße, die die Heterogenität bezüglich eines Merkmals beschreiben (z.B. Leistungsstreuung), verwendet. Dies trifft auch auf die zugewanderungsbezogene

Zusammensetzung zu: Unter dem Konstrukt „zuwanderungsbezogene Komposition“ wird üblicherweise der Prozentanteil von Kindern und Jugendlichen mit Zuwanderungshintergrund oder der Prozentanteil von Schülerinnen und Schülern eines bestimmten landesspezifischen Hintergrunds gefasst (z.B. Prozentanteil von Heranwachsenden mit türkischem Hintergrund). In Zusammenhang mit Klassenzusammensetzung wird häufig auch eine zuwanderungsbezogene Heterogenität (*racial/ethnic/cultural diversity, heterogeneity*) diskutiert und teilweise implizit inhaltlich mit dem Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund gleichgesetzt (siehe z.B. Konan, Chatard, Selimbegović & Mugny, 2010). Der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund beinhaltet jedoch kaum Informationen über Heterogenität, was die Notwendigkeit der Verwendung weiterer Maße verdeutlicht. So folgern Benner und Crosnoe (2011) in ihrer Studie zur ethnischen Zusammensetzung von Grundschulen: *“Taken as a whole, findings from the current study suggest that, as a field, we need to adopt a more nuanced view of school racial/ethnic composition and its relation to children’s developmental competencies”* (S. 642).

Heterogenität an sich ist ein relevantes Konzept in einer Vielzahl von Disziplinen (vgl. Budescu & Budescu, 2012; McDonald & Dimmick, 2003; Stirling, 2007), im bildungswissenschaftlichen Kontext wird unter dem Schlagwort „Heterogenität“ jedoch bislang vor allem die Leistungsheterogenität der Schülerschaft anhand von Streuungsmaßen betrachtet (vgl. Bos, Lankes, Pläßmeier & Schwippert, 2004; Scharenberg, 2012). Nach Harrison und Klein (2007) lassen sich allgemein drei verschiedene Arten von Heterogenität bzw. *diversity* unterscheiden: Heterogenität in Einstellungen oder Überzeugungen von Personen einer Gruppe, die sich lateral in einer Rangreihe von einem Pol zum anderen anordnen lassen (z.B. Einstellungen zur Integration von Ausländern; *separation diversity* bei Harrison & Klein, 2007); Heterogenität in der Verteilung von hierarchischen Aspekten wie Gütern oder Einkommen (*disparity diversity*); und Heterogenität in kategorialen Merkmalen von Personen, wobei starke Heterogenität sich in einer Vielzahl vertretener Merkmale äußern sollte (*variety diversity*).

Bei der Betrachtung zuwanderungsbezogener Heterogenität handelt es sich somit um die *variety diversity* (siehe auch Lieberson, 1969). Zuwanderungsbezogene Heterogenität ist gegeben, wenn eine große Anzahl verschiedener Herkunftsgruppen und von jeder Herkunftsgruppe eine gleiche Schüleranzahl in einer Klasse vertreten ist (vgl.

z.B. Stirling, 2007). In der Literatur verschiedener Disziplinen finden sich Beschreibungen von verschiedenen Heterogenitätsindizes. Die Indizes unterscheiden sich dabei in der Anzahl von Informationseinheiten, die sie berücksichtigen (*single*, *dual* und *threefold concept measures*; vgl. Junge, 1994; McDonald & Dimmick, 2003; Stirling, 2007). Einbezogen werden die Kategorienanzahl, die Verteilung von Elementen über Kategorien und ein numerisches Distanzmaß, das die Ähnlichkeit der Kategorien zueinander abbildet. Diese Informationseinheiten werden innerhalb der verschiedenen Indizes anhand von Häufigkeiten oder logarithmierten Häufigkeiten kombiniert. Beispiel für ein Maß, das nur eine Informationseinheit berücksichtigt, ist das Zählen vorhandener Kategorien innerhalb einer Gruppe (siehe MacArthur, 1965). Die meisten Indizes berücksichtigen zwei Informationseinheiten, nämlich die Kategorienanzahl und die Verteilung von Elementen über Kategorien⁴. Beispiele für häufigkeitsbasierte diesbezügliche Indizes sind Simpsons D (1949), Junges H (1994), Kvålseths OD (1991) und Fagers S (1972). Das bekannteste Maß, Simpsons D , errechnet sich aus der Summe der quadrierten relativen Häufigkeiten der einzelnen Kategorien, die von dem Wert eins subtrahiert wird. Simpsons D gibt damit die Wahrscheinlichkeit an, dass zwei zufällig gezogene Elemente unterschiedlichen Kategorien angehören. Ein hoher Wert für Simpsons D entspricht also einer großen Heterogenität. Beispiele für Indizes, die Logarithmierungen von Häufigkeiten beinhalten, sind Shannons H (Shannon & Weaver, 1949) und Gleasons D (Sanders, 1968). Die Berechnung von Shannons H sieht beispielsweise vor, dass jede relative Häufigkeit einer Kategorie mit dem Logarithmus naturalis dieser Häufigkeit multipliziert wird. Diese einzelnen Produkte werden dann aufsummiert und vom Wert eins subtrahiert. Erneut bilden hohe Werte für Shannons H eine große Heterogenität ab.

Zusammenfassend lässt sich festhalten, dass zugewanderungsbezogene Klassenzusammensetzung unterschiedliche Charakteristika einschließt, die miteinander assoziiert sind. Meistens findet der Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund als Maß der Zusammensetzung Verwendung. Seine Effekte werden als Kompositionseffekte bezeichnet. Die konzeptuell darüber hinausgehende Heterogenität wird seltener in Studien zugewanderungsbezogener Klassenzusammensetzung untersucht, Heterogenitätsindizes verschiedener Disziplinen können jedoch auch für die

⁴ In der vorliegenden Arbeit wird nicht auf *threefold concept measures* Bezug genommen, da numerische Distanzmaße zur Ähnlichkeit von Herkunftsländern nicht den Fokus der Arbeit darstellen und keine geeigneten Distanzwerte zur Verfügung stehen.

Untersuchung der Klassenzusammensetzung verwendet werden. Wie sehr diese beiden Charakteristika zuwanderungsbezogener Zusammensetzung zur Vorhersage von Leistung und psychosozialen Merkmalen der Schülerinnen und Schüler beitragen, stellt ein Forschungsdesiderat dar, auf das in späteren Abschnitten eingegangen wird (vgl. Abschnitt 2.2.4 und 6.2). Im nachfolgenden Abschnitt werden zunächst allgemeine theoretische Annahmen zur Vermittlung von Effekten der Zusammensetzung vorgestellt.

2.2.3 Allgemeine Modelle der Vermittlung von Effekten der Zusammensetzung und Modelle der Unterrichtsqualität

Die Formulierung von Annahmen zur Genese von Effekten der Zusammensetzung und ihre empirische Überprüfung sind zentral für das Verständnis darüber, wie etwaige negative Konsequenzen der Zusammensetzung reduziert werden könnten (vgl. Schofield, 2006; Walter & Stanat, 2008). Hierbei ist davon auszugehen, dass die Effekte vor allem durch Prozesse auf proximaler Klassenebene vermittelt werden (vgl. Bellin, 2009).

„Vermittlung“ oder „Mediation“ bedeutet hier, dass Effekte der Klassenzusammensetzung auf Lernergebnisse von Schülerinnen und Schülern durch andere Klassenmerkmale indirekt geleitet werden und durch sie (teilweise) erklärbar sind. Beispielsweise könnte ein positiver Effekt zwischen Klassenleistungsniveau am Anfang des Schuljahres auf die individuelle Schülerleistung am Ende des Schuljahres durch das Anspruchsniveau des Unterrichts teilweise vermittelt werden: Klassen mit einem hohen Leistungsniveau erhalten einen anspruchsvolleren Unterricht und anspruchsvollerer Unterricht führt zu höherer individueller Schülerleistung am Ende des Schuljahres. Das Anspruchsniveau erklärt teilweise den Effekt der Klassenzusammensetzung auf Individualleistung.

Wie bereits in Abschnitt 2.2.1 ausgeführt, wirken auch Charakteristika und Prozesse höherer Kontextebenen auf die Klassenebene und individuelle Schülerinnen und Schüler. Dies betrifft insbesondere die Zusammensetzung einer Schule, die in theoretischen Annahmen häufig gemeinsam mit der Klassenzusammensetzung behandelt wird. Es werden nun zunächst prominente Modelle zur Schul- und Klassenzusammensetzung vorgestellt. Diese beziehen sich allgemein auf die Vermittlung von Effekten verschiedener Merkmale wie zuwanderungsbezogene, soziale und leistungsbezogene Zusammensetzung.

2.2.3.1 Modelle zur Wirkweise der Schul- und Klassenzusammensetzung

Ein erstes prominentes Modell der Vermittlung von Effekten der Zusammensetzung wurde von Dreeben und Barr (1988; vgl. Barr & Dreeben, 1983) formuliert. Darin werden drei Hauptvermittlungswege angenommen: organisatorische, curriculare und didaktische Arrangements in der Klasse, die normative Kultur der Schülergruppe und Vergleichsprozesse zwischen Schülerinnen und Schülern. Die Autoren setzten sich vor allem mit der leistungsbezogenen Zusammensetzung auseinander und betonen die vermittelnde Rolle des unterrichtlichen Angebots für leistungsbezogene Kompositionseffekte auf die individuelle Schülerleistung. Auf diesem Modell aufbauend entwickelten Baumert et al. (2006) ihr „Vermittlungsmodell für Kontexteffekte“ (siehe Abbildung 2.3). Den Kontext bilden hier Merkmale der Schulzusammensetzung und Merkmale der Institution wie curriculare und didaktische Tradition von Schularten sowie schulartspezifische Lehrerbildung. In dem Modell werden verschiedene Merkmale der Schulzusammensetzung (z.B. zugewanderungsbezogene Zusammensetzung, soziale und leistungsbezogene Zusammensetzung) beschrieben, die sich über verschiedene Prozesse auf die individuelle Leistung und motivationale Merkmale der Schülerinnen und Schüler wie Selbstbewertung und Aspirationen auswirken. Diese Prozesse sind den Bereichen normative Kultur in der Elternschaft, normative Kultur in der Schülergruppe, Vergleichsprozesse auf Schülerebene sowie Curriculum und Unterricht zugeordnet. Die Kultur der Elternschaft ist eher distal für Prozesse auf der Klassenebene (vgl. auch Wilkinson et al., 2002), und Vergleichsprozesse zwischen Schülerinnen und Schülern sind insbesondere für motivationale Merkmale von Relevanz (vgl. Abschnitt 2.2.6.4). Als vermittelnde Faktoren für die Schülerleistung und psychosoziale Merkmale werden von den Autoren neben der normativen Peerkultur insbesondere Lehrererwartungen, curriculare Angebote und Adaptation des Unterrichts betont.

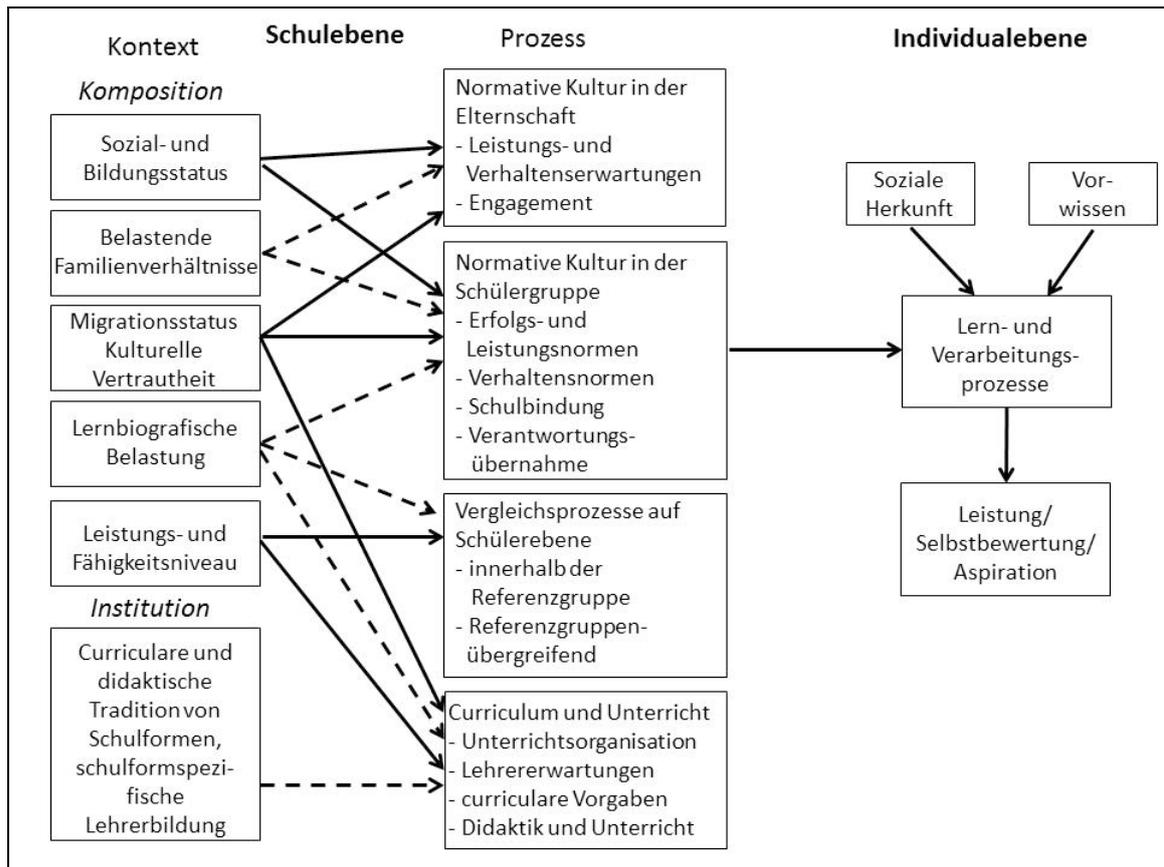


Abbildung 2.3: Vermittlungsmodell für Kontexteffekte von Baumert et al. (2006, S. 126)

Auch im englischsprachigen Raum entwickelte Modelle verweisen auf die Bedeutung des Unterrichts für die Vermittlung von Effekten der Zusammensetzung. Wilkinson et al. (2002; vgl. auch Wilkinson et al., 2000) konzeptualisieren in ihrem auf Grundlage eines Reviews von Forschungsbefunden entwickelten „*Conceptual model of peer influences on learning*“ (siehe Abbildung 2.4) unterschiedliche lernrelevante Ebenen, wobei den hierarchieniedrigeren Ebenen eine größere Bedeutung für das Lernen des Individuums beigemessen wird als den hierarchiehöheren. Curriculum und Unterricht umfassen in dem Modell ein Lernangebot mit für die Schülerschaft angemessenem Anspruchsniveau und die Kompetenz der Lehrkräfte die Lernerfahrungen für die Schülerschaft adäquat zu gestalten. Sie sind dabei auf allen Ebenen relevant, ihre Bedeutung steigt aber auf den hierarchieniedereren, lernrelevanteren Ebenen.

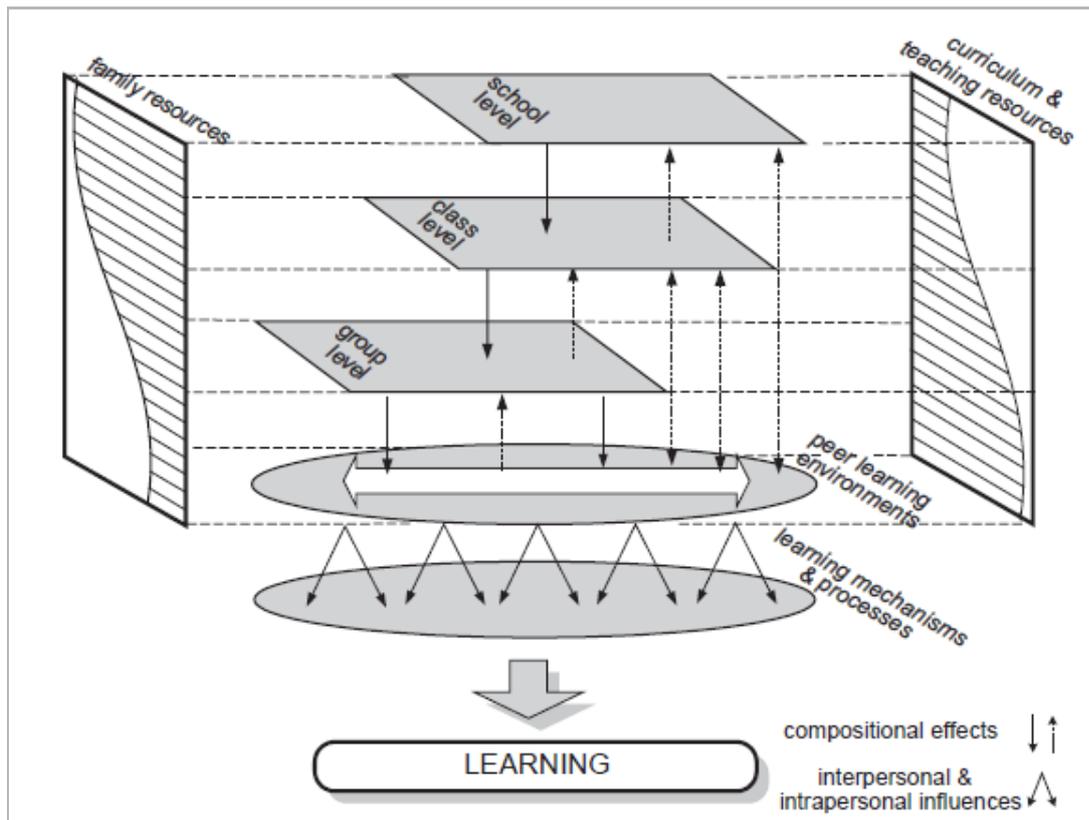


Abbildung 2.4: Conceptual model of peer influences on learning (entnommen aus: Wilkinson et al. 2002, S. 524)

Ebenso beschreibt Palardy (2013) in seinem „*Conceptual framework for mechanisms through which socioeconomic composition impacts educational attainment*“ (siehe Abbildung 2.5) bedeutsame Faktoren für Bildungsergebnisse von Schülerinnen und Schülern – hier am Beispiel des Übergangsverhaltens in verschiedene Schularten und des Schulabschlusses. Neben gegenseitiger Beeinflussung der Peers untereinander nennt er die Unterrichtsqualität einschließlich Anspruchsniveau und Unterstützung von Seiten der Lehrkraft sowie das akademische Klima einschließlich des Curriculums als relevante Faktoren.

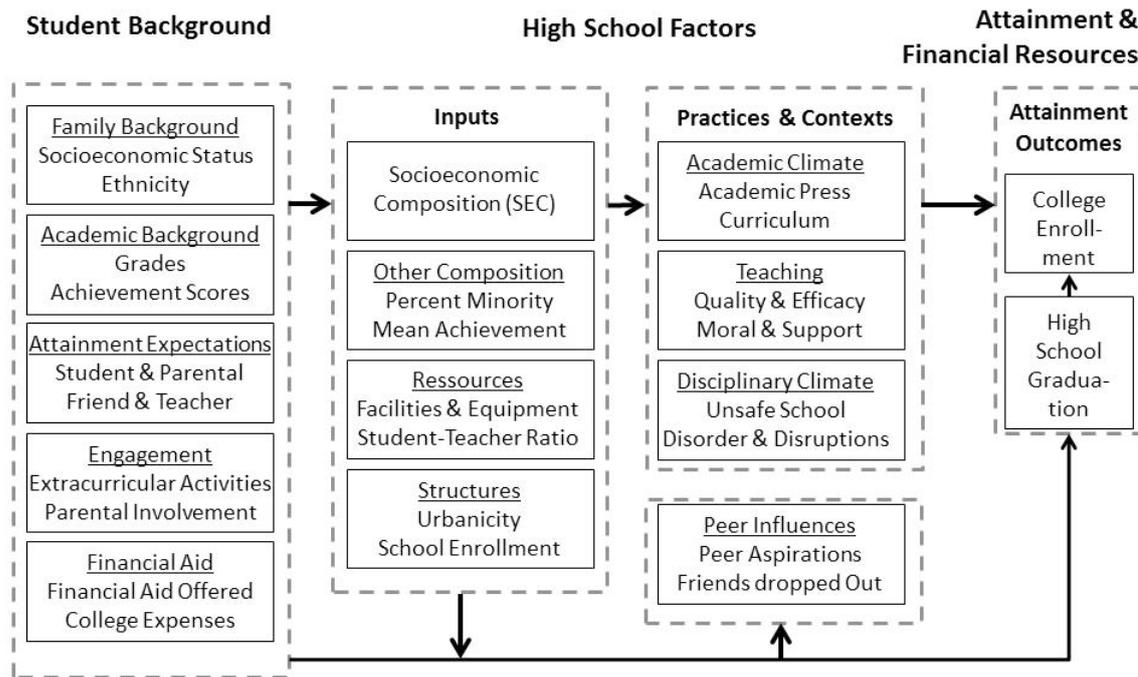


Abbildung 2.5: Conceptual framework for mechanisms through which socioeconomic composition impacts educational attainment von Palardy (2013, S. 724)

Zusammenfassend lässt sich festhalten, dass die beschriebenen Modelle zur Wirkweise von Effekten der Schul- und Klassenzusammensetzung insgesamt die zentrale Bedeutung des Unterrichts für die Vermittlung dieser Effekte hervorheben. Weitere Annahmen beziehen sich auf Prozesse innerhalb der Peergruppe im Sinne von geteilten Normen und Werten sowie Vergleichsprozessen und Unterstützungsmöglichkeiten. Nachfolgend werden nun allgemeine Modelle der Unterrichtsqualität beschrieben, bevor in späteren Abschnitten (2.2.5 und 2.2.6) detailliert theoretische Annahmen und empirische Befunde zu verschiedenen Vermittlungswegen von Effekten der Zusammensetzung dargestellt werden.

2.2.3.2 Modelle der Unterrichtsqualität

Wodurch sich qualitativ hochwertiger Unterricht auszeichnet, wird in einer Reihe von theoretischen Abhandlungen und Modellen beschrieben (vgl. z.B. Baumert et al., 2010; Brophy, 1986, 2000; Hamre & Pianta, 2010; Helmke, 2010; Helmke & Klieme, 2008; Klieme, Lipowsky, Rakoczy & Ratzka, 2006; Klieme & Rakoczy, 2008; Klieme, Schümer, & Knoll, 2001; Kunter, Baumert & Köller, 2007; Pianta & Hamre, 2009; Weinert et al., 1989). Mittlerweile haben sich sowohl in der internationalen als auch in

der deutschsprachigen Forschung Modelle durchgesetzt, die für drei Hauptdimensionen von Unterrichtsqualität sprechen (vgl. z.B. Klieme et al., 2006; Kunter et al., 2007; Pianta & Hamre, 2009) und die sich auch empirisch bestätigt haben (siehe z.B. Fauth, Decristan, Rieser, Klieme & Büttner, 2013; Wagner, Göllner, Helmke, Trautwein & Lüdtke, 2013). Mit teils leicht variierenden Bezeichnungen umfassen die drei Hauptdimensionen der Unterrichtsqualität kognitive Aktivierung, strukturierte Klassenführung und unterstützendes, schülerorientiertes Unterrichtsklima. So beschreiben beispielsweise Klieme und Rakoczy (2008) die drei Grunddimensionen der Unterrichtsqualität und deren vermutete Wirkung in dem in Abbildung 2.6 dargestellten Modell als unterrichtliches Angebot, das abhängig von der Nutzung des Angebots durch die Schülerinnen und Schüler deren Bildungsergebnisse beeinflusst (vgl. Helmke, 2010). Das Modell entwickelten Klieme et al. (2001) unter Rückgriff auf die TIMSS-Video-Studie sowie die Selbstbestimmungstheorie von Deci und Ryan (1985, 2000; vgl. Abschnitt 2.1.3.1). „Guter“ Unterricht fördert demnach das konzeptuelle Verständnis der Schülerinnen und Schüler und ihre Lernmotivation. Im anglo-amerikanischen Sprachraum gibt es entsprechend das Modell von Pianta und Hamre (2009) mit den drei Hauptbereichen *emotional support*, *classroom organization* und *instructional support*.

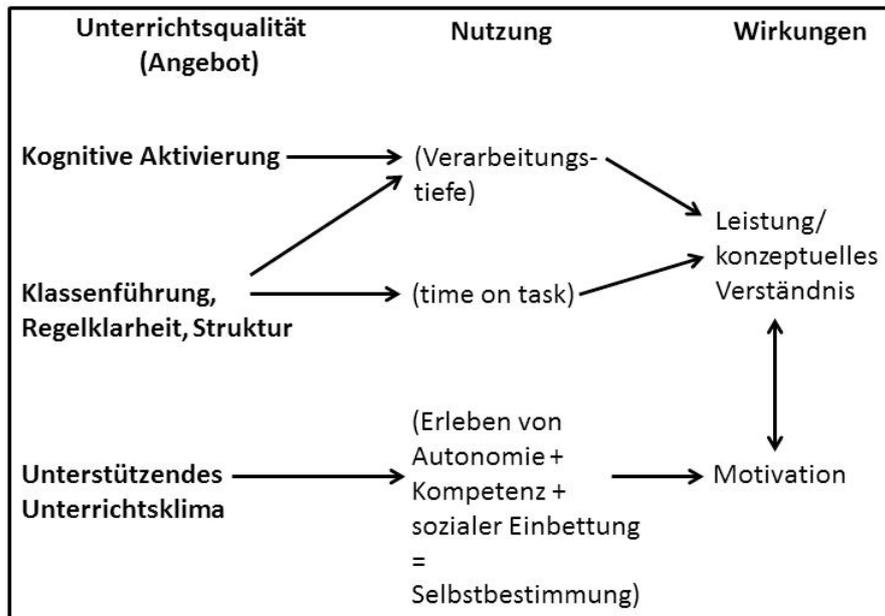


Abbildung 2.6: Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung (entnommen aus Klieme & Rakoczy, 2008, S. 228)

Während die Ausprägungen des Unterrichtsklimas und der Klassenführung weitgehend unabhängig vom Schulfach sein sollten, benötigt die Umsetzung der kognitiven Aktivierung im Unterricht eine fachspezifischere Ausrichtung. Als allgemeine Hauptdimension der Unterrichtsqualität beschreibt die *kognitive Aktivierung* zum Beispiel einen Unterricht, der selbstgesteuertes Lernen und Verwenden verschiedener Lösungsansätze der Schülerinnen und Schüler unterstützt. Ein kognitiv aktivierendes Unterrichtsangebot ermöglicht es, Lerninhalte tief zu durchdringen, anspruchsvolle Lernstrategien bei komplexen Aufgabenstellungen einzusetzen und bringt Schülerinnen und Schüler dazu, ihre bisherigen Überzeugungen zu hinterfragen (vgl. z.B. Klieme et al., 2001; Kunter et al., 2007). Wichtig für das Konzept der kognitiven Aktivierung ist, dass dieses nicht mit Aufgabenschwierigkeit gleichzusetzen ist. Kognitive Aktivierung kann auf unterschiedlichen Schwierigkeitsniveaus stattfinden (Kunter & Baumert, 2006a).

Eine *strukturierte Klassenführung* zeichnet sich beispielsweise dadurch aus, dass die Lehrkraft einen strukturierten Rahmen für die Lehr- und Lernaktivitäten schafft und die aktive Lernzeit steuert, dass sie für die jeweilige Unterrichtsstunde zentral zu erwerbende Kompetenzen deutlich macht, dass sie ihre Leistungserwartungen an die Schülerinnen und Schüler transparent macht und durch Störungsprävention dafür sorgt, dass die Schülerinnen und Schüler dem Unterricht aufmerksam folgen können (vgl. z.B.

Klieme et al., 2001; Kunter et al., 2007). Eine strukturierte Klassenführung ermöglicht insbesondere eine effektive Ausnutzung der Lernzeit im Klassenraum und kann somit förderlich für die Schülerleistung sein.

Ein *unterstützendes, schülerorientiertes Unterrichtsklima* zeichnet sich beispielsweise dadurch aus, dass die Lehrkraft einzelnen Schülerinnen und Schülern Hilfestellungen bietet, konstruktiv mit Fehlern und unzutreffenden Vorstellungen der Schülerinnen und Schüler umgeht, sich bei der Unterrichtsthemenauswahl an Interessen der Schülerinnen und Schüler orientiert und Aufgaben und Arbeitsaufträge verständlich gestaltet. Solch ein Klima schließt also mit ein, dass Lehrkräfte den Lernfortschritt und mögliche Schwierigkeiten der Schülerinnen und Schüler überwachen und gleichzeitig Autonomie der Lerner fördern, dass sie Rückmeldungen an die Schülerinnen und Schüler geben und dass sie eine Atmosphäre vertrauensvollen, angstfreien Umgangs miteinander im Klassenraum schaffen (vgl. z.B. Klieme et al., 2001; Kunter et al., 2007; Weinert et al., 1989). Diese Unterrichtsdimension sollte vor allem auf die Lernmotivation, schulisches Wohlbefinden und das Lernverhalten der Schülerinnen und Schüler wirken, was letztendlich auch für die Schülerleistung förderlich sein sollte.

Zur Beschreibung eines motivationsfördernden Unterrichts wird in dem „Modell zu Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung“ von Klieme et al. (2001; siehe Abbildung 2.6) auf die Selbstbestimmungstheorie der Motivation zurückgegriffen (Deci & Ryan, 1985, 2000; siehe Abschnitt 2.1.3.1). Wie auch Niemiec und Ryan (2009) zeigen, verdeutlichen Ausführungen der Selbstbestimmungstheorie mit den in ihr beschriebenen Grundbedürfnissen von Kompetenzerleben, Autonomie und Verbundenheitsgefühl die Relevanz des unterrichtlichen Lernkontexts für die Motivationsentwicklung der Schülerinnen und Schüler. Die eben beschriebenen Merkmale der Unterrichtsqualität hängen mit den drei Grundbedürfnissen zusammen. Ein für das *Kompetenzerleben* der Schülerinnen und Schüler förderlicher Unterricht zeichnet sich beispielsweise durch angemessen herausfordernde Aufgaben, Unterstützung und Leistungsrückmeldungen sowie transparente Leistungserwartungen der Lehrkraft aus (vgl. Eccles & Roeser, 2011; Niemiec & Ryan, 2009; Roeser et al., 2000; Wang & Eccles, 2013). Ein für das *Autonomieerleben* der Schülerinnen und Schüler förderlicher Unterricht zeichnet sich beispielsweise durch die Einbeziehung der Schülerinnen und Schüler bei der Themen- und Aufgabenauswahl und durch Behandlung von Themen, die relevant für das Leben der Schülerinnen und Schüler sind, entsprechend eines unterstützenden, schülerorientierten

Unterrichtsklimas aus (Niemi & Ryan, 2009; Roeser et al., 2000; Wang & Eccles, 2013). Ein unterstützendes, schülerorientiertes Unterrichtsklima ist ebenso für das *Verbundenheitsgefühl mit der Lehrkraft* von Bedeutung. Solch ein Klima wird durch fürsorgliches Lehrerverhalten, emotionale Unterstützung insbesondere bei Lernschwierigkeiten und durch Respekt für und Rücksicht auf interindividuelle Unterschiede zwischen Schülerinnen und Schülern bei Zielsetzungen und Leistungserwartungen begünstigt (vgl. Eccles & Roeser, 2011; Roeser et al., 2000; Urdan & Schoenfelder, 2006).

Es bleibt zu berücksichtigen, dass ein „guter“ Unterricht keineswegs die maximale Ausprägung aller beschriebenen Merkmale umfasst, sondern unterschiedliche Muster „guten“ Unterrichts bestehen, die auch je nach Fach und Bildungsziel variieren können (Helmke, 2010; vgl. Holzberger, Praetorius & Kunter, 2015). Zentral für Effekte der Zusammensetzung ist die Annahme, dass sich in unterschiedlich zusammengesetzten Klassen die Qualität des Unterrichts systematisch unterscheidet.

Bevor detailliert auf theoretische Annahmen und empirische Befunde zur Vermittlung von Effekten der Zusammensetzung durch Merkmale der Unterrichtsqualität und weitere in den Vermittlungsmodellen beschriebene Transmissionswege eingegangen wird, wird nun zunächst die empirische Befundlage zu direkten Effekten der Zusammensetzung auf Schülerleistung und psychosoziale Merkmale zusammenfassend dargestellt.

2.2.4 Empirische Befundlage zu zuwanderungsbezogenen, leistungsbezogenen und sozialen Effekten der Zusammensetzung auf Leistung und psychosoziale Schülermerkmale

In der vorliegenden Arbeit liegt der Fokus auf der zuwanderungsbezogenen Klassenzusammensetzung, deren direkte Effekte auf Leistung und psychosoziale Merkmale nachfolgend dargestellt werden. Zuwanderungsbezogene Zusammensetzung umfasst dabei einerseits der Anteil von Heranwachsenden mit Zuwanderungshintergrund, dessen Effekt unter Kontrolle des individuellen Zuwanderungsstatus als Kompositionseffekt bezeichnet wird. Andererseits wird unter dem Begriff „Zusammensetzung“ auch die zuwanderungsbezogene Heterogenität der Schülergruppen gefasst (vgl. Abschnitt 2.2.2). Ebenso wie bereits auf individueller Schülerebene beschrieben (siehe Kapitel 2.1) sind Merkmale wie Zuwanderungshintergrund, soziale

Herkunft und Schülerleistung miteinander assoziiert, sodass auch auf Klassenebene solche Merkmale der Zusammensetzung nicht unabhängig voneinander sind. Vor diesem Hintergrund werden zunächst auch kurz Effekte der leistungsbezogenen und sozialen Zusammensetzung dargestellt.

2.2.4.1 Effekte der Zusammensetzung auf Schülerleistung

Bei der Untersuchung von Kompositionseffekten zeigt sich weitgehend konsistent für verschiedene Länder und Altersgruppen ein positiver Zusammenhang zwischen der leistungsbezogenen Komposition, abgebildet durch das mittlere Niveau der Vorkenntnisse oder der kognitiven Grundfähigkeiten in einer Klasse oder Schule, und individueller Schülerleistung. Dies bedeutet, dass Schülerinnen und Schüler bessere Leistungsergebnisse erreichen, wenn sie leistungsstärkere Klassen oder Schulen besuchen, unabhängig von ihrem individuellen Hintergrund (vgl. z.B. Baumert et al., 2006; Bellin, 2009; De Fraine et al., 2003; Dumont, Neumann, Maaz, et al., 2013; Gottfried, 2014; Kiss, 2013; Köller et al., 2013; Nikolova, 2011; Opdenakker, Van Damme, De Fraine, Van Landeghem & Onghena, 2002; Opdenakker & Van Damme, 2001; Schümer, 2004; Weinert et al., 1989; Zimmer & Toma, 2000). Einige Studien, die zusätzlich Wechselwirkungen zwischen dem individuellen Leistungsniveau und dem Leistungsniveau in der Klasse untersuchen, zeigen, dass eher leistungsschwache Schülerinnen und Schüler stärker von der Beschulung in einer leistungsstarken Klasse oder Schule profitieren als dass leistungsstarke Heranwachsende durch die Beschulung mit leistungsschwachen Schülerinnen und Schülern negativ beeinflusst sind (vgl. z.B. Kiss, 2013; Zimmer & Toma, 2000). Leistungsschwächere Schülerinnen und Schüler scheinen außerdem in ihrer Leistungsentwicklung besonders von leistungsheterogenen Gruppen zu profitieren (vgl. z.B. Bellin, 2009; Hattie, 2002; Schofield, 2006). Die Aufteilung in ein gegliedertes Schulsystem in Deutschland beeinflusst die Klassenzusammensetzung und die Leistung von Schülerinnen und Schülern bzw. kann mit Aspekten der Zusammensetzung interagieren (siehe auch Abschnitt 2.2.1; vgl. z.B. Baumert, Maaz, Stanat & Watermann, 2009; Baumert et al., 2006; Dumont, Neumann, Nagy, et al., 2013; Köller et al., 2013; Retelsdorf et al., 2012). Die leistungsbasierte Gliederung in verschiedene Bildungsgänge oder Kurssysteme ist dabei häufig auch mit einer sozialen Separierung von Schülerinnen und Schülern verbunden (siehe auch

Abschnitt 2.2.1; vgl. Maaz, Baumert & Trautwein, 2010; Maaz, Trautwein, Lüdtke & Baumert, 2008).

Die soziale Komposition einer Klasse oder Schule zeigt in den meisten Studien einen zusätzlichen, wenn auch geringeren, Effekt auf die individuelle Schülerleistung (vgl. z.B. die Metanalyse von Van Ewijk & Slegers, 2010a; vgl. Ammermueller & Pischke, 2009; Baumert et al., 2006; Caldas & Bankston, 1997; De Fraine et al., 2003; Gottfried, 2014; Lauen & Gaddis, 2013; Opdenakker & Van Damme, 2001; Rumberger & Palardy, 2005; Schofield, 2006; Schümer, 2004). Dies bedeutet, dass Schülerinnen und Schüler unabhängig von ihrem eigenen sozioökonomischen Status höhere Leistungen erreichen, wenn sie Klassen oder Schulen mit einem höheren durchschnittlichen sozioökonomischen Status besuchen. Einige Studien, die zusätzlich Wechselwirkungen zwischen dem individuellen sozialen Hintergrund und dem durchschnittlichen sozioökonomischen Status in der Klasse oder Schule untersuchen, zeigen, dass Effekte der sozialen Zusammensetzung für einzelne Schülerinnen und Schüler unterschiedlicher sozialer Herkunft gleichermaßen gelten (vgl. Driessen, 2002; Rumberger & Palardy 2005; Palardy, 2013).

Für die zugewanderungsbezogene Klassenzusammensetzung ist die Forschungslage insgesamt uneinheitlicher als die zur leistungsbezogenen oder sozialen Komposition. Da diese Merkmale miteinander zusammenhängen, stellt sich insbesondere für die zugewanderungsbezogene Zusammensetzung die Frage, ob auch eigenständige Effekte nach Kontrolle anderer Merkmale der Zusammensetzung zu finden sind. Die empirische Befundlage zur zugewanderungsbezogenen Zusammensetzung wird nachfolgend zuerst bezogen auf internationale Studien und danach auf Studien aus dem deutschen Sprachraum dargestellt.

Metanalysen aus dem anglo-amerikanischen Sprachraum zu Effekten der zugewanderungsbezogenen bzw. ethnischen Klassen- oder Schulkomposition auf Schülerleistung (Mickelson, Bottia & Lambert, 2013; Van Ewijk & Slegers, 2010b) kommen zu dem Schluss, dass geringe, aber statistisch bedeutsame negative Effekte zu finden sind, die auch zum Teil nach Berücksichtigung der leistungsbezogenen und sozialen Komposition bestehen bleiben. In allen Studien der Metaanalysen werden Effekte des allgemeinen Prozentanteils von Schülerinnen und Schülern einer ethnischen Minderheit oder des spezifischen Prozentanteils bezüglich einer bestimmten Minderheitengruppe betrachtet (z.B. *Hispanics*, *African American*). Mickelson et al.

(2013) fassen aufgrund ihrer Analyse von 25 Studien zu Effekten auf Mathematikleistung von Schülerinnen und Schülern zusammen, dass der leicht negative Effekt des Anteils von Schülerinnen und Schülern einer ethnischen Minderheit stärker bei Schülerinnen und Schülern der Sekundarstufe im Vergleich zur Grundschule, stärker bei älteren Schülerinnen und Schülern und stärker bei der Betrachtung von spezifischen Minderheitengruppen anstatt des allgemeinen Prozentanteils von Schülerinnen und Schülern ethnischer Minderheiten ist. Van Ewijk und Slegers (2010b) kommen in ihrer Metaanalyse von 13 Primärstudien (Jahre 1986 – 2006) zu dem Schluss, dass Effekte des Anteils von Afroamerikanern in den USA stärker ausgeprägt sind als Effekte des Anteils von kürzlich Zugewanderten. Außerdem waren die Effekte des Anteils einer Minderheit größer für Schülerinnen und Schüler der entsprechenden Minderheitengruppe als für Schülerinnen und Schüler einer anderen Minderheiten- oder der Mehrheitsgruppe. Es bestanden keine Unterschiede in der Effektgröße zwischen Studien, die Effekte auf Mathematikleistungen analysierten und solchen, die Effekte auf sprachliche Leistungen untersuchten. Die Autoren fanden zudem weder statistisch bedeutsame Effekte des Alters der Schülerinnen und Schüler auf die Analyseergebnisse, noch Unterschiede zwischen Studien, die Effekte der Klassenkomposition, und Studien, die Effekte der Schulkomposition, analysierten. Die Kontrolle des Vorwissens und/oder des sozioökonomischen Status zeigte ebenfalls keinen statistisch bedeutsamen Effekt auf die ermittelten Ergebnisse. Das bedeutet, die Effekte der zuwanderungsbezogenen Komposition waren in den Studien dieser Metaanalyse überwiegend unabhängig von Vorwissen und sozioökonomischem Status der Schülerinnen und Schüler. Van Ewijk und Slegers begründen die fehlenden Belege für erwartungskonforme Unterschiede zwischen Studien unter anderem mit den allgemein geringen Effektgrößen in den 13 Primärstudien sowie mit Unzulänglichkeiten bei der Erfassung des sozioökonomischen Status.

Mittlerweile gibt es neben erstgenannter Metaanalyse auch eine Vielzahl weiterer internationaler Studien, die häufig einen negativen Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund in der Schule oder Klasse und Individuelleistung zeigen (vgl. z.B. Agirdag et al., 2012; Borman et al., 2004; Caldas & Bankston, 1998; Gould, Lavy & Paserman, 2009; Guldmond & Bosker, 2009; Hanushek et al., 2009; Jensen & Würtz Rasmussen, 2011; Karsten et al., 2006; Lee, 2007). Dieser Zusammenhang verliert in internationalen Studien häufig nach Kontrolle des Niveaus in den Vorleistungen und/oder dem durchschnittlichen sozioökonomischen Status die statistische Bedeutung, zum Teil bleibt ein statistisch bedeutsamer Effekt des

Anteils Heranwachsender mit Zuwanderungshintergrund jedoch auch fortbestehen (vgl. z.B. Borman et al., 2004; Caldas & Bankston, 1998; Gould et al., 2009; Hanushek et al., 2009; Karsten et al., 2006; Lee, 2007).

Für die Bedeutung der zuwanderungsbezogenen Zusammensetzung für Heranwachsende mit Zuwanderungshintergrund im Vergleich zu Heranwachsenden ohne Zuwanderungshintergrund ist der Forschungsstand uneinheitlich. Einige Studien finden negative Effekte des Anteils von Heranwachsenden mit Zuwanderungshintergrund vor allem für Schülerinnen und Schüler einer ethnischen Minderheit (z.B. Dronkers & Level, 2007, Hanushek et al., 2009). Andere Studien finden stärkere zuwanderungsbezogene Effekte auf die Leistung von Heranwachsenden ohne Zuwanderungshintergrund (z.B. Jensen & Würtz Rasmussen, 2011) oder keine statistisch bedeutsamen Unterschiede (z.B. Caldas & Bankston, 1998).

Neben diesen Befunden zu leicht negativen Effekten des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund bzw. einer ethnischen Minderheit gibt es teils auch Studien, die einen positiven Effekt des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf individuelle Leistung finden (siehe z.B. Konan et al., 2010; Peetsma, Van der Veen, Koopman & Van Schooten, 2006). Dies scheint vor allem dann der Fall zu sein, wenn Maße der ethnischen oder zuwanderungsbezogenen Heterogenität einbezogen werden. Heterogenitätsmaße beinhalten über den Klassenanteil von Heranwachsenden mit Zuwanderungshintergrund hinausgehende Informationen über die zuwanderungsbezogene Klassenzusammensetzung (vgl. Abschnitt 2.2.2) und sind mit dem Klassenanteil assoziiert: Eine höhere Heterogenität in der Klasse bzw. Schule geht auch mit einem erhöhten Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund bzw. einer ethnischen Minderheit einher. Einige Studien aus den USA und den Niederlanden verweisen auf positive Zusammenhänge zwischen zuwanderungsbezogener bzw. ethnischer Heterogenität und Schülerleistung (vgl. z.B. Benner & Crosnoe, 2011; Denson & Chang, 2009; Gottfredson, Panter, Daye, Allen & Wightman, 2009; Lee & Smith, 1999; Maestri, 2011; Tam & Bassett, 2004; für abweichende Ergebnisse siehe Driessen, 2002). Studien, die Zusammenhänge des Klassenanteils oder eines Heterogenitätsmaßes mit Schülerleistung analysieren, kommen also zu voneinander abweichenden Ergebnissen. Welches der Charakteristika stärker mit Leistung assoziiert ist und ob sie gemeinsam eine genauere Vorhersage der Schülerleistung ermöglichen, ist bislang kaum untersucht worden (vgl. Abschnitt 2.2.5.4).

Studien zu zuwanderungsbezogenen Effekten der Zusammensetzung in Deutschland finden häufig einen negativen Effekt des Anteils von Kindern und Jugendlichen mit Zuwanderungshintergrund auf Schülerleistung, der zumeist nach Kontrolle weiterer Kompositionsmerkmale, wie mittlerem sozioökonomischen Status und mittlerem Vorwissen, nicht mehr statistisch signifikant ist (vgl. z.B. Baumert et al., 2006; Bellin, 2009; Ditton & Krüsken, 2006; Dumont, Neumann, Maaz et al., 2013; Nikolova, 2011; Stanat, Schwippert, et al., 2010; Stanat, 2006b; Walter, 2006). Vereinzelt bleiben bei differenzierter Analyse einzelner Herkunftsgruppen auch nach Kontrolle weiterer Kompositionsmerkmale noch Effekte des Anteils auf die Leistung der Schülerinnen und Schüler bestehen (vgl. Bellin, 2009; Merkens, 2005; Walter & Stanat, 2008). Die wenigen Studien, die die Bedeutung der zuwanderungsbezogenen Komposition für die Leistung von Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund vergleichen, finden keine bedeutsamen Unterschiede (z.B. Stanat, 2006b). Effekte der zuwanderungsbezogenen Heterogenität in der Klasse oder Schule wurden in deutschen Studien bislang nicht untersucht.

2.2.4.2 Effekte der Zusammensetzung auf psychosoziale Schülermerkmale

Analysen von Effekten der Zusammensetzung auf psychosoziale oder motivationale Merkmale umfassen meist Untersuchungen zu leistungsbezogenen Kompositionseffekten auf das akademische Selbstkonzept, auf die in späteren Abschnitten zur Genese von Effekten der Zusammensetzung kurz eingegangen wird (siehe auch Abschnitt 2.2.6; vgl. z.B. Marsh, 1987; Nagengast & Marsh, 2012). Der Annahme von Referenzgruppeneffekten entsprechend zeigt sich, dass Schülerinnen und Schüler, die Schulen mit einem höheren Leistungsniveau besuchen, eine geringere intrinsische Motivation aufweisen als vergleichbare Schülerinnen und Schüler in Schulen mit einem niedrigeren Leistungsniveau (vgl. Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006). Zu Effekten der sozialen Zusammensetzung auf motivationale Schülermerkmale gibt es kaum Forschungsbefunde. Es gibt jedoch Hinweise darauf, dass der durchschnittliche soziale Status in der Klasse nicht bedeutsam mit Lernmotivation in Zusammenhang steht (vgl. Klieme et al., 2010).

Für die zuwanderungsbezogene Zusammensetzung konnte zudem in internationalen Studien gezeigt werden, dass zuwanderungsbezogene bzw. ethnische Homogenität positiv mit der sozioemotionalen Entwicklung (geringere externalisierende

Symptome und stärkere interpersonale Fähigkeiten), Verbundenheitsgefühl mit der Schule, Wahrnehmung eines fairen und wenig diskriminierenden Lernklimas sowie schulbezogenem Optimismus in Zusammenhang steht (Benner & Crosnoe, 2011; Benner et al., 2008; Goldsmith, 2004; Johnson, Crosnoe & Elder, 2001; Seaton & Yip, 2009). Nicht alle genannten Studien berücksichtigen jedoch die Mehrebenenstruktur der Daten und kontrollieren nicht durchgängig weitere Hintergrundmerkmale wie den sozialen Status und das Leistungsniveau der Schülerinnen und Schüler, sodass offen bleibt, ob die Befunde durch diese unberücksichtigten Variablen beeinflusst sind.

Studien aus Deutschland zeigen, dass der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund in positivem Zusammenhang mit der Lernmotivation im Fach Deutsch steht (vgl. Klieme et al., 2010) und dass Schülerinnen und Schüler in einer Schule des Hauptschulbildungsgangs mit einem Anteil von Heranwachsenden mit Zuwanderungshintergrund größer 40 Prozent leicht erhöhte Bildungsaspirationen aufweisen (vgl. Stanat, 2006b). Für die instrumentelle Motivation, das akademische Selbstkonzept und Selbstwirksamkeitsüberzeugungen fanden sich in der Studie von Stanat (2006b) hingegen keine Zusammenhänge mit der zuwanderungsbezogenen Komposition.

Zusammenfassend lässt sich somit festhalten, dass Schülerinnen und Schüler höhere Leistungen erreichen, wenn sie in Klassen oder Schulen mit einer höheren mittleren Leistung und höherem mittleren sozioökonomischen Status lernen. Bezüglich der zuwanderungsbezogenen Klassenzusammensetzung ist die Forschungslage uneinheitlicher. Untersuchungen zu Effekten der Zusammensetzung werden insbesondere in den USA, Großbritannien und den Niederlanden durchgeführt; in Deutschland erfuhren diese Analysen insbesondere im letzten Jahrzehnt stärkere Beachtung (vgl. Bellin, 2009). Internationale Studien zur zuwanderungsbezogenen Komposition, die Effekte des Klassen- oder Schulanteils von Heranwachsenden mit Zuwanderungshintergrund untersuchen, finden meist negative Effekte, die in manchen Studien auch nach Kontrolle der sozialen und leistungsbezogenen Komposition bestehen bleiben. Studien in Deutschland finden meist keinen eigenständigen Effekt des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund nach Kontrolle weiterer Kompositionsmerkmale. In einigen internationalen Studien finden sich jedoch auch positive Effekte der zuwanderungsbezogenen Zusammensetzung. Dies ist insbesondere der Fall wenn es sich um eine zuwanderungsbezogen heterogene Gruppe handelt, was durch entsprechende

Heterogenitätsmaße anstelle des Klassenanteils von Heranwachsenden mit Zuwanderungshintergrund in den Studien abgebildet wird. Die Untersuchung von Effekten der zuwanderungsbezogenen Zusammensetzung und der Vergleich von Effekten des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund und von zuwanderungsbezogener Heterogenität bei Vorhersage von Schülerleistungen stellen also Desiderata der Forschung dar.

Insgesamt wird in den meisten Studien die Schülerleistung als abhängige Variable betrachtet. Es kann davon ausgegangen werden, dass Effekte auf die Leistung größer sind als auf psychosoziale Merkmale (vgl. z.B. Van der Werf et al., 2008). Neben üblicherweise untersuchten leistungsbezogenen Referenzgruppeneffekten kann auch die zuwanderungsbezogene Klassenzusammensetzung mit psychosozialen Merkmalen zusammenhängen. Die wenigen vorhandenen Studien finden mehrheitlich positive Effekte der zuwanderungsbezogenen Homogenität auf psychosoziale Merkmale. Zudem gibt es Hinweise auf positive Effekte des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf psychosoziale Merkmale. Der Vergleich verschiedener Konzeptualisierungen der zuwanderungsbezogenen Zusammensetzung bei Vorhersage psychosozialer Merkmale ist ein bedeutsames Forschungsdesiderat.

Welche theoretischen Annahmen zur Erklärung der hier dargestellten Befunde zu zuwanderungsbezogenen Effekten der Schul- und Klassenzusammensetzung auf Leistung und psychosoziale Merkmale diskutiert werden und welche empirischen Befunde zu vermittelnden Faktoren bislang vorliegen, wird nun in den nächsten beiden Abschnitten (2.2.5 und 2.2.6) detailliert ausgeführt.

2.2.5 Theoretische Annahmen und empirische Befunde zur Genese von Effekten zuwanderungsbezogener Zusammensetzung auf Schülerleistung

In verschiedenen Überblicksaufsätzen und Modellen werden als vermittelnde Prozesse insbesondere das unterrichtliche Angebot und Lehrererwartungen, die normative Peerkultur in der Klasse, Sprachverwendung und Unterstützung der Peers untereinander sowie Vergleichsprozesse zwischen Schülerinnen und Schülern genannt (vgl. z.B. Baumert et al., 2006; Bellin, 2009; Dreeben & Barr, 1988; Dumont, Neumann, Maaz, et al., 2013; Harker & Tymms, 2004; Palardy, 2013; Thrupp, Lauder & Robinson, 2002;

Van Ewijk & Slegers, 2010b; Wilkinson et al., 2000).⁵ Im Folgenden werden nun die Argumente für die verschiedenen miteinander assoziierten Vermittlungsfaktoren von zuwanderungsbezogenen Effekten der Zusammensetzung detaillierter ausgeführt und die jeweilige empirische Befundlage dargestellt. Begonnen wird dabei mit der zentralen Annahme, dass die Unterrichtsqualität Effekte der Zusammensetzung vermittelt. Danach wird kurz auf weitere unter dem Begriff *Peer effects* diskutierte Vermittlungswege eingegangen. Hierbei wird auch auf Annahmen zur zuwanderungsbezogenen Heterogenität der Schülerschaft Bezug genommen, die in klassischen Modellen zu Effekten der Zusammensetzung kaum Berücksichtigung finden.

Der Vielzahl der Studien zu direkten Effekten der Zusammensetzung auf Schülerleistung, die im letzten Abschnitt dargestellt wurde, steht jedoch eine deutlich geringere Anzahl von Studien gegenüber, die die Vermittlung von Effekten der Zusammensetzung untersuchen. Van Ewijk und Slegers (2010b) beschreiben diese Situation in ihrer Metaanalyse folgendermaßen: *„Many causes for the compositional effect of ethnicity have been proposed, although virtually no study quantitatively examines the link between proposed causes and the taking place of a compositional effect.”* (S. 240). Wenn Entstehungs- oder Vermittlungsbedingungen von Effekten der Zusammensetzung untersucht wurden, beziehen sich diese meist auf das Leistungsniveau oder in selteneren Fällen auf den durchschnittlichen sozioökonomischen Status (für einen Überblick siehe Dumont, Neumann, Maaz et al., 2013; Harker & Tymms, 2004; Wilkinson et al., 2000). Der vorliegende Abschnitt stellt überblicksartig den vorhandenen Kenntnisstand zur Genese von zuwanderungsbezogenen Effekten der Zusammensetzung auf Schülerleistung vor und greift auch auf Befunde zur Leistungszusammensetzung und sozialen Zusammensetzung zurück, da diese häufig mit der zuwanderungsbezogenen Zusammensetzung assoziiert sind. Detaillierter werden die relevanten Studien zu Beginn der Teilstudien 1 und 3 beschrieben (vgl. Abschnitt 4.2 und 6.2).

⁵ Eine weitere diskutierte Erklärung dafür, dass sich Kompositionseffekte in Studien nachweisen lassen, insbesondere in den Anfängen der Forschung, ist in möglichen methodischen Schwächen der Studien zu sehen, die zu Phantomeffekten führen (vgl. z.B. Harker & Tymms, 2004; Hauser, 1970; siehe Abschnitt 2.2.7).

2.2.5.1 Unterrichtsqualität und Schülerleistung

Das unterrichtliche Angebot bietet die Möglichkeit, von institutioneller Seite auf die Leistungsentwicklung von Schülerinnen und Schülern einzuwirken. Dies ist mit ein Grund für die besondere Diskussion der Unterrichtsqualität bei Untersuchung der Vermittlung von Effekten der Zusammensetzung (vgl. Abschnitt 2.2.3; z.B. Baumert et al. 2006; Dumont, Neumann, Maaz et al., 2013; Harker & Tymms, 2004; Hattie, 2002; Schofield, 2006). Es wird angenommen, dass ein höherer Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund oder niedrigerem sozioökonomischen Status in der Klasse zu einem Lernumfeld führt, in dem die Qualität des Unterrichts beeinträchtigt wird. Annahmen zu dieser mediierenden Rolle der Unterrichtsqualität beziehen sich neben der Zuordnung von Lehrkräften unterschiedlicher Qualifikation zu Klassen mit unterschiedlicher Zusammensetzung (vgl. Abschnitt 8.2 und 8.4) auf Leistungserwartungen der Lehrkräfte an die Schülerschaft unterschiedlich zusammengesetzter Klassen.

Annahmen zu Lehrererwartungen werden vor allem in der psychologischen Forschung diskutiert und beziehen sich darauf, dass Lehrkräfte von Schülerinnen und Schülern mit Zuwanderungshintergrund oder mit niedrigerem sozioökonomischem Status geringere Leistungen erwarten als von Heranwachsenden ohne Zuwanderungshintergrund oder mit einem sozioökonomisch privilegierten Status (z.B. Alexander und Schofield, 2006; Jussim & Harber, 2005). Diese Lehrererwartungen wirkten sich auf das Verhalten der Lehrkräfte aus, was wiederum die schulischen Leistungen beeinträchtigen kann. Es wird angenommen, dass Lehrkräfte ihre Unterrichts- bzw. Aufgabengestaltung, d.h. Unterrichtsstil und Einsatz anspruchsvollen Materials, für Schülerinnen und Schülern mit Zuwanderungshintergrund an ihren geringeren Erwartungen ausrichten (vgl. Gay, 2000). Dies bedeutet, dass Lehrkräfte in Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund weniger kognitiv aktivierend, langsamer und weniger anspruchsvoll unterrichten als es die Klasse aufgrund ihres tatsächlichen Lernpotenzials verarbeiten könnte (vgl. zusammenfassend Alexander & Schofield, 2006; Eccles & Roeser, 2011; Harker & Tymms, 2004; Hattie, 2002; Thrupp, 1999; Wilkinson et al., 2000). Außerdem könnte erwartungsgesteuertes Verhalten der Lehrkräfte Verhaltensweisen, die das unterstützende Unterrichtsklima in der Klasse prägen, wie Lächeln, unterstützende Reaktionen, physische Nähe, Rückmeldungen durch die Lehrkräfte (wie Lob und Kritik) sowie von den Lehrkräften zur Verfügung gestellte Möglichkeiten zur Beteiligung der Schülerinnen und Schüler beinhalten (vgl. z.B. Babad,

1993; Gay, 2000)⁶. Demnach würden sich Lehrkräfte in Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund oder einem durchschnittlich niedrigeren sozioökonomischen Status weniger schülerorientiert und unterstützend verhalten.

Verschiedene Annahmen zum Zusammenhang zuwanderungsbezogener Zusammensetzung und Unterrichtsqualität beziehen sich neben Lehrererwartungen auf Schwierigkeiten im Unterrichten. Lehrkräfte könnten nicht auf das Unterrichten herkunftsbezogen heterogener Schülergruppen vorbereitet sein, die Einbeziehung der Heterogenität ins Unterrichtsgeschehen kann ihnen Schwierigkeiten bereiten, sie können über wenig Wissen um verschiedene Kulturen und Herkunftssprachen und ihre Relevanz für Lernprozesse und potentielle Lernschwierigkeiten der Schülerinnen und Schüler verfügen und unsicher bei der Förderung von Schülerinnen und Schülern mit Zuwanderungshintergrund sein (vgl. Allemann-Ghionda, Auernheimer, Grabbe & Krämer, 2008; Edelmann, 2008; Gay, 2002, 2010; Vedder et al., 2006). Dies kann die Unterrichtsqualität in Schulklassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund bzw. einer herkunftsbezogen heterogenen Schülerschaft negativ beeinflussen (siehe auch Abschnitt 8.2). Auf der anderen Seite könnte man auch gerade ein besonders unterstützendes Unterrichtsklima von Seiten der Lehrkräfte in solchen Klassen erwarten, beispielsweise da sie das potentielle Leistungsniveau der Schülerinnen und Schüler unterschätzen und die Hilfsbedürftigkeit überschätzen (vgl. Ready & Wright, 2011; Stanat, Weirich & Radmann, 2012), was dazu führen kann, dass sich Lehrkräfte besonders unterstützend und fürsorglich verhalten.

Wie eingangs erwähnt, steht die Vielzahl theoretischer Annahmen in einem Gegensatz zur geringen Anzahl von empirischen Studien zur Vermittlung von Effekten der Zusammensetzung. Die wenigen Studien, die Zusammenhänge von Klassenzusammensetzung und Unterrichtsqualität untersuchen, beschäftigen sich vorrangig mit der leistungsbezogenen und weniger mit der sozialen oder zuwanderungsbezogenen Zusammensetzung.

⁶ Caldas und Bankston (1997) diskutieren weiterführend noch den Umstand, dass geringere Leistungserwartungen der Lehrkräfte bezüglich bestimmter Schülerpopulationen von den Schülerinnen und Schülern wahrgenommen werden, sodass sie letztendlich möglicherweise das Stereotyp geringerer Leistung bedienen (vgl. *stereotype vulnerability*: Steele, 1995; Pygmalion-Effekt: Rosenthal & Jacobson, 1968).

Für die positive Assoziation von Schülerleistung und Unterrichtsmerkmalen, die den bereits in Abschnitt 2.2.3.2 beschriebenen Hauptdimensionen der Unterrichtsqualität – unterstützendes, schülerorientiertes Unterrichtsklima, strukturierte Klassenführung sowie kognitive Aktivierung – zugeordnet werden können, spricht eine Reihe empirischer Befunde (vgl. z.B. Baumert et al., 2010; Klieme et al., 2010; Kunter & Baumert, 2006b; Kunter et al., 2013; Lee & Smith, 1999; Lipowsky et al., 2009; Seidel & Shavelson, 2007; Van Houtte, 2004; Wang et al., 1993; Wubbels & Brekelmans, 2005). Zusätzlich finden sich empirische Belege dafür, dass in Klassen und Schulen, die sich durch ein erhöhtes Leistungsniveau auszeichnen, ein qualitativ hochwertiger Unterricht vorherrscht (vgl. z.B. Baumert, Kunter, et al., 2004; Dreeben & Barr, 1988; Raudenbush, Rowan & Cheong, 1993; Opdenakker & Van Damme, 2001, 2007). Hattie (2002) stellt in seinem Überblicksartikel zwar zusammenfassend dar, dass „guter“ Unterricht wirkungsvoller als die Zusammensetzung der Klasse oder Schule ist, Lehrererwartungen bezüglich der Leistungsfähigkeit ihrer Schülerinnen und Schüler aber mit den Lernergebnissen in Zusammenhang stehen und beispielsweise Lehrkräfte an hierarchiehöheren Schulzweigen mit höherem Leistungsniveau anspruchsvoller, strukturierter und schülerorientierter unterrichten. Einige Studien geben zudem Hinweise darauf, dass Effekte des Leistungsniveaus auf individuelle Schülerleistung teils durch Merkmale des Unterrichts vermittelt werden (vgl. Carbonaro & Gamoran, 2002; Klieme et al., 2008; Opdenakker & Van Damme 2013; Opdenakker et al., 2002; Rindermann, 2007; Weinert et al., 1989; für abweichende Ergebnisse siehe: De Fraine et al., 2003). Diese Studien sprechen also dafür, dass Heranwachsende in leistungsstarken Klassen bzw. Schulen ein vergleichsweise hochwertiges Unterrichtsangebot erhalten und sich entsprechend besonders positiv in ihrer Leistung entwickeln. Es konnte auch gezeigt werden, dass Leistungsunterschiede zwischen Schülerinnen und Schülern verschiedener Schulzweige durch unterschiedliche Unterrichtsqualität und unterschiedliche Leistungserwartungen vonseiten der Lehrkräfte beeinflusst werden (vgl. z.B. Burris, Heubert & Levin, 2006; Van Houtte, 2004).

Schülerinnen und Schüler mit Zuwanderungshintergrund und/oder niedrigerem sozialen Status sind häufiger in Klassen, Schulen und Schularten mit einem niedrigen Leistungsniveau vertreten, sodass die eben dargestellten Befunde insbesondere für diese Schülergruppe bedeutsam sind. Inwiefern zusätzlich zur Leistungszusammensetzung auch Effekte der sozialen und zugewanderungsbezogenen Zusammensetzung durch Unterrichtsqualität vermittelt werden, wird seltener untersucht. In ihrer Zusammenfassung des Forschungsstandes zu Lehrererwartungen an Kinder und

Jugendliche mit Zuwanderungshintergrund kommen Alexander und Schofield (2006) zu dem Schluss, dass vor allem in US-amerikanischen Studien bei Lehrkräften besonders niedrige Leistungserwartungen für Schülerinnen und Schüler mit Zuwanderungshintergrund identifiziert wurden (vgl. z.B. Jussim & Harber, 2005; Tenenbaum & Ruck, 2007) und diese in Zusammenhang mit der Wahrnehmung kultureller Differenz stehen (vgl. z.B. Hauser-Cram, Sirin & Stipek, 2003; Neal, McCray, Webb-Johnson & Bridgest, 2003). Auch im deutschsprachigen Raum gibt es Hinweise auf geringere Lehrererwartungen an Schülerinnen und Schüler mit Zuwanderungshintergrund und/oder niedrigem sozioökonomischem Status (z.B. Lorenz, Gentrup, Kristen, Stanat & Kogan, eingereicht). Eine Metanalyse im US-amerikanischen Raum zeigte zudem, dass Lehrkräfte mit Schülerinnen und Schülern mit Zuwanderungshintergrund bzw. einer ethnischen Minderheit weniger neutrale und positive Kommunikation hatten (z.B. eine Frage an sie richten, Loben) als mit Schülerinnen und Schülern ohne Zuwanderungshintergrund. Des Weiteren gibt es Hinweise verschiedener Studien darauf, dass in Klassen mit einem hohen Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und/oder niedrigerem sozioökonomischem Status ein weniger qualitativ hochwertiger Unterricht stattfindet, der sich insbesondere durch geringere Herausforderung im Sinne von kognitiver Aktivierung auszeichnet (z.B. Palardy, 2015; Schofield, 2006; Stipek, 2004; Van der Veen, 2002).

Für ein unterstützendes, schülerorientiertes Unterrichtsklima als weiterer Qualitätsdimension, die insbesondere für die Lernmotivation relevant sein sollte, scheinen hingegen eher positive Zusammenhänge mit der zuwanderungsbezogenen Heterogenität der Schülerschaft zu bestehen (vgl. Den Brok & Levy, 2005; Levy, Wubbels, Brekelmans & Morganfield, 1997; siehe Abschnitt 2.2.6.1).

Für die soziale Zusammensetzung gibt es zudem Hinweise darauf, dass Effekte des durchschnittlichen sozioökonomischen Status der Schule oder Klasse durch Merkmale der Unterrichtsqualität vermittelt werden (Dumay & Dupriez 2007; Palardy, 2013; Rumberger & Palardy, 2005; Vause, Dupriez & Dumay, 2010; vgl. Schofield 2006). So zeigten sich positive Zusammenhänge zwischen dem durchschnittlichen sozialen Status und Merkmalen der Unterrichtsqualität, deren Kontrolle soziale Kompositionseffekte auf Schülerleistung reduzierte. In der Mehrzahl der genannten Studien wurde jedoch die Schulebene betrachtet, die soziale und zuwanderungsbezogene Zusammensetzung nicht gemeinsam untersucht bzw. weitere leistungsbezogene Merkmale der Zusammensetzung außer Acht gelassen sowie vermittelnde Effekte keiner

direkten empirischen Prüfung unterzogen (für eine ausführlichere Darstellung siehe Teilstudie 1, Abschnitt 4.2). Somit stellt die Untersuchung der vermittelnden Rolle der Unterrichtsqualität für Effekte der sozialen und zuwanderungsbezogenen Klassenzusammensetzung auf Schülerleistung weiterhin ein Forschungsdesiderat dar.

2.2.5.2 Peer effects: Normative Peerkultur und Schülerleistung

Die Vermittlung von Effekten der Zusammensetzung durch eine normative Peerkultur wird vor allem in der Soziologie angenommen (vgl. Baumert et al., 2006; Bellin, 2009). Die Grundannahmen beziehen sich auf die Individualebene: Individuelle Schülerinnen und Schüler mit bestimmten Hintergrundmerkmalen (wie kein Zuwanderungshintergrund, privilegierte soziale Herkunft, hohes Leistungspotenzial) sollen demnach eine hohe Lernmotivation sowie lernförderliche Wertvorstellungen, Überzeugungen und Verhaltensweisen, wie z.B. positive Einstellung gegenüber der Schule und dem Lernen, Optimismus, regelmäßiger Schulbesuch, aufweisen (vgl. Abschnitt 2.1.3.2). Wenn viele solcher Schülerinnen und Schüler in einer Klasse gemeinsam beschult werden, können sie gemeinsam eine normative „Kultur des Erfolgs“ schaffen. Die Kehrseite kann eine Peerkultur darstellen, in der ein gemeinsames Gefühl von Mangel und Selbstzweifel herrscht, geteilte Überzeugungen, wie „Menschen wie ich werden es nie in der Schule zu etwas bringen, auch wenn sie sich anstrengen“, kollektives Stigmatisierungserleben, Gefühl einer Entfremdung von der Schule sowie Akzeptanz abweichenden Verhaltens, wie Störverhalten im Unterricht oder Absentismus (vgl. Agirdag et al., 2012; Barth, Dunlap, Dane, Lochman & Wells, 2004; Baumert et al., 2006; Dreeben & Barr, 1988; Goldsmith, 2011; Harker & Tymms, 2004; Palardy, 2013; Retelsdorf et al., 2012; Rumberger & Palardy, 2005). Dies entspricht auch der von Ogbu (1987) beschriebenen oppositionellen Kultur (vgl. Abschnitt 2.1.3.2). Solche lernförderlichen oder lernhinderlichen Kulturen entstehen dadurch, dass Peers sich in der Interaktion aneinander angleichen. Hier sind Prozesse des Modelllernens und der Verstärkung normkonformen Verhaltens wirksam, die letztendlich zu einer Internalisierung der Gruppenwerte bei den einzelnen Schülerinnen und Schülern führen (vgl. z.B. Barth et al., 2004; Kozaki & Ross, 2011; Ryan, 2000). Ein zusätzlicher Aspekt, der zuwanderungsbezogen heterogene Klassen spezifisch betrifft, ist, dass es in solchen Klassen zu Spannungen zwischen verschiedenen Herkunftskulturen kommen kann, was

einem lernförderlichen Klima entgegenstehe (Hoxby, 2000; zitiert nach Van Ewijk & Slegers, 2010b).

Studien, die eine lernförderliche oder lernhinderliche Peerkultur als Erklärung für Effekte der Zusammensetzung untersuchen, legen ihren Schwerpunkt meist auf leistungsbezogene und soziale Zusammensetzung. Spezifisch für die zuwanderungsbezogene Zusammensetzung liegen kaum Studien vor. So untersuchten Baumert et al. (2006) beispielsweise die normative Kultur an Hauptschulen und bezogen hierfür die Akzeptanz von Leistungsnormen, die Gewaltbereitschaft gegenüber Personen und Gegenständen sowie die Bindung an die Schule (Schulzufriedenheit) ein. Sie konnten zeigen, dass der Anteil gewaltbereiter Schülerinnen und Schüler mit der individuellen Leseleistung negativ assoziiert war. Es bestanden jedoch keine statistisch bedeutsamen Zusammenhänge zwischen den untersuchten Merkmalen der Peerkultur und Merkmalen der Zusammensetzung auf Schulebene. Die betrachteten Merkmale der Peerkultur waren also von der Zusammensetzung der Schülerschaft bezogen auf kognitive Grundfähigkeiten, sozialen Status und Zuwanderungshintergrund weitgehend unabhängig. Auch Barth et al. (2004) konnten nur geringe Zusammenhänge zwischen Merkmalen der Peerkultur definiert über aggressives Verhalten, Lernengagement und Peerbeziehungen mit der sozialen und leistungsbezogenen Schulzusammensetzung finden. Agirdag et al. (2012) konnten hingegen zeigen, dass die normative Peerkultur Effekte des mittleren sozioökonomischen Status der Grundschule auf die Mathematikleistung vermittelt. Als Peerkultur untersuchten die Autoren die geteilte Überzeugung der Schülerschaft, dass Schülerinnen und Schüler wie sie es nie zu etwas in der Schule bringen werden, auch wenn sie sich anstrengen. Unter Einbeziehung dieser Überzeugungen auf Schulebene verschwand der statistisch bedeutsame soziale Kompositionseffekt auf die Leistung. Zu berücksichtigen ist hierbei jedoch, dass die Autoren die leistungsbezogene Zusammensetzung lediglich durch den Anteil von Schülerinnen und Schülern, die eine Klasse wiederholen, als Proxy einbeziehen. Es bleibt somit offen, wie stark das Befundmuster von dem Vorkenntnisniveau in der Schule abhängt und ob ein sozialer Kompositionseffekt auch nach Kontrolle des Leistungsniveaus bestehen bleibt. Eine gemeinsame Analyse der sozialen und zuwanderungsbezogenen Schulzusammensetzung schlossen die Autorinnen und Autoren aufgrund von Multikollinearität aus. Der Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund allein erwies sich als nicht bedeutsam für die individuelle

Mathematikleistung in der Studie, sodass keine darauf bezogenen vermittelnden Effekte untersucht wurden.

In der Studie von Palardy (2013) zur sozialen Schulkompositionseffekten zeigte sich eine vermittelnde Wirkung der Peerkultur auf den High-School-Abschluss und das spätere College-Übergangsverhalten von Schülerinnen und Schülern. Die Peerkultur wurde in der Studie als Schulmittelwert der Schüleranzahl, die die Schule abbrechen, und als Schulmittelwert der Schüleranzahl, die einen Collegeabschluss anstreben, operationalisiert. Schülerinnen und Schüler, die mehr Personen unter ihren Schulfreunden hatten, die die Schule abbrachen, schlossen selbst seltener die High-School ab. Und Schülerinnen und Schüler, die mehr Personen unter ihren Schulfreunden hatten, die einen Collegeabschluss anstrebten, besuchten selbst eher ein zweijähriges Collegeprogramm. Die Effektstärken sind jedoch als klein zu beurteilen. Die Berücksichtigung dieser Merkmale der Peerkultur führte zu einer Reduktion der Effekte der sozialen Schulkomposition auf den High-School-Abschluss und das spätere College-Übergangsverhalten.

Insgesamt können unterschiedliche Konzeptualisierungen von „Peerkultur“ zu unterschiedlichen Befunden der hier referierten Studien beitragen. Es lässt sich jedoch aus den Ergebnissen der verschiedenen Arbeiten zusammenfassen, dass – sofern sich Effekte der Peerkultur auf Schülerleistung überhaupt zeigen – diese gering ausgeprägt sind und meist nicht spezifisch für die zuwanderungsbezogene Zusammensetzung untersucht wurden, sondern eher für Effekte des Leistungsniveaus und des durchschnittlichen sozioökonomischen Status relevant sind. Effekte der normativen Peerkultur werden also als weniger zentral für die Erklärung von Effekten der Zusammensetzung gewertet (vgl. Baumert et al., 2006; Goldsmith, 2011; Dreeben & Barr, 1988; Schofield, 2006).

2.2.5.3 Peer effects: Sprachverwendung und Schülerleistung

Die Sprachverwendung zwischen Peers ist ein zentraler Faktor, der als Transmissionsweg eines Effekts zuwanderungsbezogener Zusammensetzung auf Schülerleistung diskutiert wird. In einer Klasse mit einem hohen Anteil von Heranwachsenden mit Zuwanderungshintergrund sprechen häufig auch viele Schülerinnen und Schüler nicht nur die in der Schule verwendete Instruktionssprache im Elternhaus. Laut Esser (2006) kann der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund sich auf Sprachgewohnheiten und damit einhergehende

Lerngelegenheiten im Klassenraum oder in der Schule auswirken und darüber vermittelt die Schülerleistung beeinflussen. Entsprechend der *language contact hypothesis* (vgl. Peetsma et al., 2006) können Schülerinnen und Schüler mit Zuwanderungshintergrund in der Schule beispielsweise in Pausenzeiten eine andere als die Instruktionssprache miteinander sprechen. Dadurch hätten die Heranwachsenden weniger sprachliche Rollenmodelle der Unterrichtssprache zur Verfügung und würden den korrekten Sprachgebrauch der Unterrichtssprache weniger üben. Zusätzlich beeinflusst diese Klassenzusammensetzung, inwieweit sich die Schülerinnen und Schüler gegenseitig unterstützen können, indem sie sich das Lernmaterial in der Instruktionssprache erklären und bei Aufgaben helfen (Driessen, 2002; Entwisle & Alexander, 1994; Esser, 2006; Peetsma et al., 2006; Van Ewijk & Slegers, 2010b; Walter, 2008b). Dieser beschriebene Transmissionsweg der Sprachverwendung sollte insbesondere für die sprachliche Schülerleistung wirksam sein und darüber vermittelt auch Leistungen in anderen Schulfächern beeinflussen (vgl. z.B. Driessen, 2002; Van Ewijk & Slegers, 2010b; Walter, 2008b). Bei der sprachlichen Klassenzusammensetzung und Effekten auf individuelle Schülerinnen und Schüler mit Zuwanderungshintergrund ist zudem relevant, ob viele verschiedene Herkunftssprachen im Klassenraum vertreten sind oder ein Großteil der Schülerinnen und Schüler mit Zuwanderungshintergrund dieselbe Herkunftssprache spricht. Bei einer sehr heterogenen Sprachzusammensetzung ist zu erwarten, dass die Schülerinnen und Schüler auch untereinander in der Unterrichtssprache kommunizieren. Bei homogener Sprachzusammensetzung wäre zu erwarten, dass die Schülerinnen und Schüler ihre Herkunftssprache verwenden, was wie oben beschrieben negativ die Schülerleistung beeinflusst. Einzelne Autorinnen und Autoren führen jedoch auch an, dass sich Heranwachsende in Klassen mit homogener Sprachzusammensetzung (die nicht der Unterrichtssprache entspricht) auch einfacher gegenseitig unterstützen können, da sie beispielsweise zusätzliche Erklärungen in der Herkunftssprache bieten können oder sprachspezifische Schwierigkeiten leichter verstehen und auflösen können (vgl. Lazear, 1999; Maestri, 2011). Es gibt also sowohl Annahmen zu positiven Effekten als auch zu negativen Effekten einer homogenen Sprachzusammensetzung in Klassen mit Schülerinnen und Schülern mit Zuwanderungshintergrund auf die Schülerleistung.

Empirisch wurden diese Annahmen bislang kaum überprüft. Erste Hinweise geben die Analysen von Stanat, Schwippert et al. (2010) zu Kompositionseffekten auf die individuelle Leseleistung in einem längsschnittlichen Design. Hierbei zeigte sich, dass der

Klassenanteil von Schülerinnen und Schülern, deren primäre Familiensprache nicht Deutsch ist, einen größeren negativen Effekt auf die Leseleistungsentwicklung hat als der Schüleranteil mit mindestens einem im Ausland geborenen Elternteil. Dadurch kann man annehmen, dass eine nicht-deutsche Sprachverwendung zwischen Peers für das Befundmuster relevant ist. Die Effekte der zugewanderungsbezogenen Klassenkomposition konnten letztendlich jedoch durch die soziale und leistungsbezogene Klassenniveau aufgeklärt werden, was gegen die Vermutung der Relevanz der Sprachverwendung zwischen Peers spricht. Die Analysen von Walter (2008b) der PISA-Neuntklässlerstichprobe widersprechen ebenso der Annahme einer Vermittlung von Kompositionseffekten durch die Sprachverwendung zwischen Peers: Für die Leistungen im Lesen, in Naturwissenschaften und Mathematik erwies sich – wenn überhaupt – der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund operationalisiert über das Geburtsland als relevanter für die Individualleistung als ein über die Sprache im Elternhaus operationalisierter Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund. Maestri (2011) fand jedoch für die Niederlande Hinweise auf einen Einfluss der Sprachverwendung zwischen Peers. In ihrer Studie zeigten sich positive Effekte der ethnischen Heterogenität insbesondere für die sprachliche Leistung (im Vergleich zur mathematischen Leistung) von Schülerinnen und Schülern einer ethnischen Minderheit. Dies interpretiert die Autorin dahingehend, dass in sprachlich heterogenen Klassen die Schülerinnen und Schüler stärker dazu neigen, die Unterrichtssprache miteinander zu sprechen (vgl. Lazear, 1999). Der tatsächliche Sprachgebrauch der Schülerinnen und Schüler untereinander wurde in der Studie jedoch nicht direkt erfasst, sodass die Befunde lediglich als erste Hinweise gesehen werden sollten.

Zusammenfassend lässt sich festhalten, dass es bislang wenige Studien gibt, die die Relevanz der Sprachverwendung zwischen Peers für Effekte der Zusammensetzung untersuchen. Vorhandene Studien sprechen überwiegend gegen eine Bedeutung der Sprachverwendung. Weitere Untersuchungen, die direkt die Sprachgewohnheiten und die sprachliche Zusammensetzung der Klasse mit einbeziehen, stehen noch aus.

2.2.5.4 Peer effects: Lernförderlicher Austausch in zugewanderungsbezogenen heterogenen Klassen und Schülerleistung

Die meisten theoretischen Annahmen beziehen sich auf den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und nehmen meist negative Effekte auf die Leistungsentwicklung an. Parallel dazu gibt es jedoch theoretische Annahmen zu positiven Effekten der zugewanderungsbezogenen Heterogenität in der Klasse auf die Schülerleistung. Demnach würde es in Klassen, die bezüglich der Herkunftskulturen der Schülerinnen und Schüler sehr heterogen sind, von Vorteil sein, dass Menschen mit unterschiedlichem Hintergrundwissen und verschiedenen Perspektiven aufeinandertreffen. Dies sei förderlich für die kognitive Entwicklung der Schülerinnen und Schüler (vgl. Benner & Crosnoe, 2011; Gottfredson et al., 2009; Gurin, Dey, Gurin & Hurtado, 2003; Gurin, Nagda & Lopez, 2004; Konan et al., 2010; Maestri, 2011; Tam & Bassett, 2004). Laut Piagets (1977) Konzept des Disäquilibrium bewirke die Konfrontation mit neuer Information, die nicht zu bestehenden Schemata einer Person passt, einen unbehaglichen Zustand, einen kognitiven Konflikt. Die intendierte Abwendung dieses Zustands ruft durch Prozesse der Assimilation und Akkommodation der neuen Wissensinhalte somit eine kognitive Weiterentwicklung, einen Lerngewinn hervor. Hierbei wären positive Effekte der herkunftsbezogenen Heterogenität insbesondere für Problemlösefähigkeiten und das Verständnis komplexerer Sachverhalte zu erwarten, die sich beispielsweise in Gruppendiskussionen entwickeln. Solch eine Annahme der positiven Wirkung von ethnisch heterogenen Gruppen wird in der Arbeits- und Organisationspsychologie unter dem Stichwort „*information/decision making approach*“ für die Problemlösefähigkeiten heterogener Arbeitsteams diskutiert (vgl. Meyer, in press).

Die hier dargestellten Annahmen werden bislang nicht direkt anhand von Informationen über Austauschprozesse zwischen Schülerinnen und Schüler empirisch überprüft. Vielmehr werden Studien, die positive Zusammenhänge zwischen ethnischer Heterogenität und Schülerleistung finden, im Sinne der Annahme interpretiert, dass Schülerinnen und Schüler in zugewanderungsbezogenen heterogeneren Klassen höhere Leistungen erzielen, da die Auseinandersetzung mit unterschiedlichen Perspektiven und der Austausch mit Menschen verschiedenen kulturellen Hintergrunds die kognitive Entwicklung fördere (siehe Abschnitt 2.2.4.1; vgl. z.B. Benner & Crosnoe, 2011; Tam & Bassett, 2004). Wie in den vorherigen Abschnitten jedoch dargestellt, beziehen sich

theoretische Annahmen und empirische Befunde meist auf einen negativen Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund als Charakteristik zuwanderungsbezogener Zusammensetzung und Schülerleistung. Da zuwanderungsbezogen heterogene Klassen auch einen höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund aufweisen, sind die in den letzten Abschnitten diskutierten Faktoren weniger lernförderlicher Bedingungen (vgl. Abschnitt 2.2.5.1 – 2.2.5.3) auch für heterogene Klassen anzunehmen. Dieses Argument würde also – entgegen der in diesem Abschnitt formulierten Annahme – einen negativen Zusammenhang zwischen Heterogenität und Leistung erwarten lassen. In Klassen mit vergleichbarem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund sowie weiteren Merkmalen, wie unter anderem Unterrichtsqualität und Peerkultur, besteht hingegen die Möglichkeit, die in dem Abschnitt dargestellte Annahme positiver Effekte der Heterogenität zu prüfen. Bei Konstanthaltung der weiteren Klassenmerkmale könnten sich also positive Effekte zeigen. Diese Annahme wurde bislang jedoch kaum empirisch geprüft. Die oben zitierten Studien, die positive Effekte zuwanderungsbezogener Heterogenität auf Leistung zeigen (vgl. z.B. Benner & Crosnoe 2011; Tam & Bassett, 2004), kontrollierten auch für weitere Merkmale der Schule bzw. Klasse, wie soziale oder leistungsbezogene Zusammensetzung. Effekte der Heterogenität unter Berücksichtigung verschiedener Kovariaten wurden jedoch nicht berücksichtigt.

Zusammenfassend lässt sich festhalten, dass für zuwanderungsbezogene Effekte der Zusammensetzung auf Schülerleistung, sofern sie nicht vollständig auf die leistungsbezogene und soziale Zusammensetzung der Schülerschaft zurückführbar sind, Hinweise bestehen, dass sie durch Merkmale der Unterrichtsqualität vermittelt werden. Der normativen Peerkultur und Sprachverwendung zwischen Peers kommt bei der Vermittlung eher eine untergeordnete Rolle zu. Einige Studien sprechen auch für positive Effekte von zuwanderungsbezogener Heterogenität als eine weitere Charakteristik zuwanderungsbezogener Zusammensetzung. Die zugrundeliegende Annahme des lernförderlichen Austauschs in heterogenen Gruppen wurde bislang nicht direkt untersucht. Eine Analyse des Effekts der zuwanderungsbezogenen Heterogenität auf Leistung unter Konstanthaltung verschiedener anderer Klassenmerkmale steht ebenfalls noch aus.

Im nächsten Abschnitt werden nun Annahmen und Befunde zu Effekten der Klassenzusammensetzung auf psychosoziale Merkmale dargestellt.

2.2.6 Theoretische Annahmen und empirische Befunde zur Genese von Effekten zuwanderungsbezogener Zusammensetzung auf motivationale Schülermerkmale

Die Schule und die Klasse bieten nicht nur fachliche Lerngelegenheiten, sondern auch Raum für soziale Interaktionen zwischen Individuen und ihre psychosoziale Entwicklung. Die im vorangegangenen Abschnitt beschriebenen Vermittlungswege können demnach mit teils anderer Schwerpunktsetzung auch für motivationale Schülermerkmale bedeutsam sein. Effekte der Zusammensetzung auf motivationale Merkmale rückten insbesondere in den letzten Jahrzehnten durch Arbeiten zur *Self-Determination-Theory* (Deci & Ryan, 1985, 2000; siehe Abschnitt 2.1.3.1) und zur *Stage-Environment-Fit-Theory* (vgl. Eccles et al., 1993), die sich unter anderem mit der Rolle des Schulkontexts für die abnehmende schulbezogene Motivation im Jugendalter auseinandersetzt, stärker in das wissenschaftliche Interesse.

Im folgenden Abschnitt wird der Fokus auf zuwanderungsbezogene Effekte der Zusammensetzung auf die Lernmotivation als zentralem Bildungsziel gelenkt. Theoretische Annahmen und entsprechende empirische Befunde beziehen sich hier auf die Unterrichtsqualität, Verbundenheitsgefühl der Schülerinnen und Schüler, normative Peerkultur sowie Vergleichsprozesse zwischen Schülerinnen und Schülern. Eine detailliertere Beschreibung relevanter Studien zur Ableitung der Forschungsfragen der vorliegenden Arbeit findet sich auch innerhalb der Teilstudien 2 und 3 (Abschnitte 5.2 und 6.2).

2.2.6.1 Unterrichtsqualität und Schülermotivation

Modellannahmen zu den drei Hauptdimensionen der Unterrichtsqualität (vgl. Abschnitt 2.2.3.2) sprechen insbesondere dem unterstützenden, schülerorientierten Unterrichtsklima lernmotivationsförderliche Auswirkungen zu. Dieses Unterrichtsmerkmal bietet zudem Möglichkeiten zur Befriedigung der Grundbedürfnisse Kompetenzerleben, Autonomieerleben und Verbundenheitsgefühl, die mit gesteigerter Lernmotivation einhergehen sollten (vgl. Deci & Ryan, 2000; Abschnitt 2.1.3.1 und 2.2.3.2). Wie bereits in Abschnitt 2.2.5.1 zur Vermittlung von Effekten der

Zusammensetzung auf Leistung ausgeführt, wird in Klassen mit höherem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und/oder niedrigem mittleren sozioökonomischem Status häufig eine geringere Unterrichtsqualität angenommen. Für das unterstützende, schülerorientierte Unterrichtsklima gibt es jedoch auch Annahmen, dass Lehrkräfte das potenzielle Leistungsniveau der Schülerinnen und Schüler mit Zuwanderungshintergrund unterschätzen und deren Hilfsbedürftigkeit überschätzen (vgl. Ready & Wright, 2011; Stanat, Weirich, et al., 2012). Dies kann zu einem besonders fürsorglichen und unterstützenden Lehrerverhalten führen (vgl. Abschnitt 2.2.5.1), das letztendlich motivationsförderlich ist. Somit bietet ein unterstützendes, schülerorientiertes Unterrichtsklima eine mögliche Erklärung für die unter Abschnitt 2.2.4.2 berichteten direkten positiven Effekte zuwanderungsbezogener Zusammensetzung auf motivationale Merkmale.

Empirisch zeigt sich, dass ungeachtet der Klassenzusammensetzung insbesondere ein unterstützendes, fürsorgliches Lehrerverhalten theoriekonform mit erhöhter Lernmotivation der Schülerinnen und Schüler einhergeht (Danielsen, Wiium, Wilhelmsen & Wold 2010; Goodenow, 1993; Kunter et al. 2013; Opendakker, Maulana & den Brok, 2012; Opendakker & Minnaert, 2011; Roeser et al., 2000; Wang & Eccles, 2013).

Für das unterstützende, schülerorientierte Unterrichtsklima geben einzelne empirische Studien zudem Hinweise darauf, dass Lehrkräfte in zuwanderungsbezogenen heterogenen Klassen oder Klassen mit einem erhöhten Anteil von Heranwachsenden mit Zuwanderungshintergrund freundlicher und hilfsbereiter mit den Schülerinnen und Schülern umgehen (vgl. Levy et al., 1997; Den Brok & Levy, 2005). Befunde auf Grundlage von Daten der DESI-Studie zeigten, dass ein positiver Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern nicht-deutscher Herkunftssprache und dem Zuwachs der Lernmotivation im Fach Deutsch im Verlaufe eines Schuljahres bestand (Klieme et al., 2010). Ein unterstützendes Lehrerverhalten war positiv mit Lernmotivationszuwachs assoziiert und der Anteil von Schülerinnen und Schülern nicht-deutscher Herkunftssprache stand in positivem Zusammenhang mit unterstützendem Lehrerverhalten. Inwiefern Effekte der zuwanderungsbezogenen Zusammensetzung auf die Lernmotivation durch unterstützendes Lehrerverhalten vermittelt wurde, wurde jedoch keiner direkten statistischen Prüfung unterzogen, da die Autorinnen und Autoren in ihrem Beitrag insbesondere Prozesse auf der Schul- anstatt der Klassenebene fokussierten (vgl. auch Abschnitt 2.2.1). Die statistische Prüfung der Vermittlung

positiver Effekte der zuwanderungsbezogenen Zusammensetzung auf motivationale Schülermerkmale durch ein motivationsförderliches Unterrichtsklima stellt somit ein weiteres Desiderat der Forschung dar.

2.2.6.2 Peer effects: Verbundenheitsgefühl mit den Peers und Schülermotivation

Das Verbundenheitsgefühl mit den Peers in der Klasse wird als motivationsförderliche Kontextbedingung angesehen, die auch mit der zuwanderungsbezogenen Klassenzusammensetzung in Zusammenhang stehen kann (vgl. *need for relatedness* bei Deci & Ryan, 2000; Niemiec & Ryan, 2009; siehe Abschnitt 2.1.3.1).

Allgemein zeichnen sich Klassen, in denen sich die Schülerinnen und Schüler verbunden mit und zugehörig zu ihren Peers fühlen, durch gegenseitige Unterstützung und Ermutigung der Schülerinnen und Schüler untereinander, positive Interaktionen zwischen den Peers, Freundschaften, wenig Ausschluss von Schülerinnen und Schülern aus der Gemeinschaft sowie ein Gefühl von gegenseitiger Wertschätzung aus (vgl. Goodenow, 1993; Kumar & Maehr, 2010; Urdan & Schoenfelder, 2006; Wang & Eccles, 2013). Solche Klassen bieten einen sicheren Rahmen für das Lernen der Schülerinnen und Schüler, in dem sie sich beispielsweise aktiv am Unterricht beteiligen und Fragen stellen, ohne Angst zu haben, von anderen dafür ausgelacht oder ausgegrenzt zu werden. Gerade für Schülerinnen und Schüler mit Zuwanderungshintergrund kann das Verbundenheitsgefühl mit Klassenkameradinnen und -kameraden von besonderer Bedeutung sein, da sie soziale und kulturelle Hürden überwinden müssen, um sich als Teil der Schulkultur wahrzunehmen (vgl. Crosnoe, Johnson & Elder, 2004; Den Brok & Levy, 2005; Eccles & Roeser, 2011; Goodenow, 1993; Kumar & Maehr, 2010).

Eine Grundlage, auf der Menschen Nähe zueinander aufbauen, Freundschaften schließen und sich verbunden miteinander fühlen, ist laut der *social identity theory* (Tajfel & Turner, 1986) und des *similarity attraction paradigm* (Byrne, 1971) das Ausmaß an wahrgenommener Ähnlichkeit mit der anderen Person, wozu auch der allgemeine Zuwanderungshintergrund oder die landesspezifische kulturelle Herkunft gehören kann (vgl. auch Blau, 1974). Auf Grundlage von Ähnlichkeit zueinander schließen sich Schülerinnen und Schüler zu Gruppen zusammen (*ingroup*) und grenzen sich von anderen ab (*outgroup*). Zuwanderungshintergrund oder ethnische Herkunft können dabei als saliente Merkmale der Identität gelten (vgl. Johnson et al., 2001). Die soziale und

kulturelle Gruppenzugehörigkeit kann gerade in einer pluralistischen Gesellschaft ein Merkmal der Ähnlichkeitswahrnehmung sein (vgl. Kumar & Maehr, 2010; für eine Kritik siehe O'Connor, 1999). Somit kann die Zusammensetzung einer Klasse direkte Konsequenzen für das Verbundenheitsgefühl in der Klasse mit sich bringen: Die zum Beispiel von Benner und Crosnoe (2011) beschriebene *Belongingness Perspective*, die auf Baumeister und Leary (1995) zurückgeht (vgl. Deci & Ryan, 2000), sagt ein weniger starkes Zusammengehörigkeitsgefühl in zuwanderungsbezogen heterogenen Klassen vorher.

Bezogen auf die Lernmotivation könnte dies theoretisch dazu beitragen, dass Schülerinnen und Schüler in heterogenen Klassen allgemein weniger lernmotiviert sind. Für individuelle Schülerinnen und Schüler würde dies konkret bedeuten, dass sie sich stärker mit den Peers verbunden fühlen und stärker lernmotiviert sind, wenn mehr Schülerinnen und Schüler ihres jeweiligen ethnischen Hintergrunds in der Klasse sind. Da zwischen Heranwachsenden mit Zuwanderungshintergrund allgemein trotz unterschiedlicher ethnischer Herkunft eine stärkere wahrgenommene Ähnlichkeit beispielsweise durch eine Zuwanderungsgeschichte oder Mehrsprachigkeit innerhalb der Familie bestehen kann, sind analoge Prozesse des Verbundenheitsgefühls in geringerer Stärke auch für die globale Unterscheidung zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund anzunehmen (vgl. insgesamt Benner et al., 2008; Eitle & Eitle, 2004; Goldsmith, 2004; Johnson et al., 2001; Kumar & Maehr, 2010; Maestri, 2011). In zuwanderungsbezogen besonders heterogenen Klassen sind sehr viele Herkunftskulturen vertreten, wodurch auch eher ein Kräftegleichgewicht herrschen kann, das die Ausgrenzung und Diskriminierung von Schülerinnen und Schüler weniger wahrscheinlich macht (Juvonen, Nishina & Graham, 2006). Dies könnte einen positiven Zusammenhang zwischen Heterogenität und Verbundenheit und somit auch Heterogenität und Lernmotivation erwarten lassen.

Der Zusammenhang zwischen zuwanderungsbezogener Klassenzusammensetzung und dem Verbundenheitsgefühl mit den Peers wurde bislang relativ selten direkt untersucht (vgl. z.B. Johnson et al., 2001). Verschiedene Studien konnten allgemein zeigen, dass ein Verbundenheitsgefühl und eine unterstützende Atmosphäre zwischen den Schülerinnen und Schülern ungeachtet der Klassenzusammensetzung lernmotivationsförderlich ist (vgl. Goodenow, 1993; Wang & Eccles, 2013). Für die zuwanderungsbezogene Schülerzusammensetzung zeigen sich Hinweise darauf, dass

Schülerinnen und Schüler in weniger heterogenen Schulen oder Klassen das Schulklima als fairer und förderlicher für interethnischen Austausch wahrnahmen (Benner et al., 2008), dass sie sich stärker zu Schule zugehörig fühlten (Johnson et al., 2001), dass Schülerinnen und Schüler einer ethnischen Minderheit weniger Diskriminierung empfinden (Seaton & Yip, 2009) und dass es an solchen Schulen stärkere gegenseitige Unterstützung und Akzeptanz für Schülerinnen und Schüler einer ethnischen Minderheit gab (Postmes & Branscombe, 2002). Vereinzelt finden sich auch Studien, die keinen Zusammenhang zwischen zuwanderungsbezogener Zusammensetzung und Verbundenheitsgefühl nachweisen konnten (Van Houtte & Stevens, 2009) oder positive Zusammenhänge zwischen zuwanderungsbezogener Heterogenität und Gefühlen von Sicherheit, Verbundenheit und Akzeptanz bei Schülerinnen und Schülern mit Zuwanderungshintergrund feststellen konnten (Juvonen et al., 2006). Zudem bieten heterogenere Klassen und Schulen mehr Gelegenheit für interethnische Freundschaften (Van Houtte & Stevens, 2009; Vervoort, Scholte & Scheepers, 2011). Zusammenhänge zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und dem Verbundenheitsgefühl mit den Peers wurden bislang nicht untersucht. Hinsichtlich der zuwanderungsbezogenen Heterogenität hingegen sprechen empirische Befunde eher für ein stärkeres Verbundenheitsgefühl in zuwanderungsbezogenen homogenen Klassen. Die Zusammenhänge zwischen verschiedenen Charakteristika zuwanderungsbezogener Zusammensetzung mit dem Verbundenheitsgefühl zwischen Peers und der Lernmotivation wurden bislang nicht untersucht.

2.2.6.3 Peer effects: Normative Peerkultur und Schülermotivation

Wie in Abschnitt 2.2.3.1 beschrieben, ist eine weitere Annahme für die Vermittlung von Effekten der Klassenzusammensetzung die normative Wirksamkeit der Peerkultur. Dieses Argument wird häufig zur Erklärung von Effekten auf Schülerleistung hinzugezogen (Abschnitt 2.2.5.2; vgl. z.B. Dumont, Neumann, Maaz, et al., 2013; Harker & Tymms, 2004; Thrupp et al., 2002), betrifft an sich jedoch zunächst einmal die Lernmotivation und das Lernverhalten der Schülerinnen und Schüler, die nachfolgend deren Leistung beeinflussen. Es ist davon auszugehen, dass Schülerinnen und Schüler vor allem Normen und Werte in der Klasse übernehmen, wenn sie sich mit ihren Peers verbunden fühlen. Schülerinnen und Schüler gleichen sich in ihrer Lernmotivation sowie Normen und Werten durch Prozesse des Modelllernens und Verstärkung normkonformen

Verhaltens aneinander an (vgl. z.B. Eccles & Roeser, 2011; Kozaki & Ross, 2011; Ryan, 2000; Urdan & Schoenfelder, 2006). Für die zugewanderungsbezogene und soziale Klassenzusammensetzung wird, wie in Abschnitt 2.2.5.2 ausgeführt, häufig angenommen, dass Schülerinnen und Schüler unter anderem geringere Lernmotivation und weniger lernförderliche Einstellungen gegenüber der Schule haben. In einer Klasse mit vielen Schülerinnen und Schülern mit Zuwanderungshintergrund würden sich Peers in der Interaktion aneinander angleichen und gemeinsam eine eher lernhinderliche Peerkultur entwickeln (vgl. Agirdag et al., 2012; Baumert et al., 2006; Goldsmith, 2011; Palardy, 2013). Dies würde jedoch im Gegensatz zu dem in Abschnitt 2.1.3.2 berichteten häufigen Befund stehen, dass auf Individualebene entweder keine Unterschiede zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund oder eine erhöhte Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund bestehen.

Empirisch gibt beispielsweise die Studie von Frenzel, Pekrun und Goetz (2007) Hinweise auf das Vorliegen einer normativen Peerkultur mit geteilten Werten und Überzeugungen sowie deren Verbindung zu motivationalen Merkmalen allgemein. In dieser Studie zum wahrgenommenen Lernumfeld und Emotionen im Mathematikunterricht zeigte sich, dass solche Schülerinnen und Schüler mehr Lernfreude und weniger Angst im Mathematikunterricht hatten, die der Meinung waren, dass viele ihrer Peers Mathematik mochten. Barth et al. (2004) konnten zudem zeigen, dass Schülerinnen und Schüler im Ausmaß ihres Lernengagements, ihres aggressiven Verhaltens und ihrer positiven Beziehungen mit den Peers mit der durchschnittlichen Ausprägung dieser Merkmale in ihrer Klasse übereinstimmten. So vertraten zum Beispiel Schülerinnen und Schüler stärker akademische Werte, wenn sie in Klassen lernten, die sich durch eine hohe Akzeptanz akademischer Werte auszeichnete. Die Akzeptanz akademischer Werte von Schülerinnen und Schülern sowie deren gemittelte Ausprägung auf Klassenebene wurden jedoch über Lehrereinschätzungen gewonnen, sodass dieses Befundmuster beispielsweise auch durch Urteilstendenzen beeinflusst sein könnte.

Für die zugewanderungsbezogene Heterogenität geben die Ergebnisse von Benner et al. (2008) Hinweise darauf, dass Schülerinnen und Schüler an weniger heterogenen Schulen stärker soziale Normen wahrnehmen, die die Anstrengung für Schulaufgaben und Leistung befürworten. Die Folgen dieser Normen konnten empirisch nicht eindeutig geklärt werden. Auf der einen Seite stand das Ausmaß dieser sozialen Normen in positivem Zusammenhang mit schulbezogenem Lernengagement aus Schülersicht, auf

der anderen Seite in negativem Zusammenhang mit Lernengagement aus Sicht der Lehrkraft, sodass die Mediationsanalysen zu uneinheitlichen Ergebnissen führten. Bivariat zeigten sich keine bedeutsamen Zusammenhänge zwischen zuwanderungsbezogener Heterogenität und Lernengagement.

Zusammenfassend lässt sich festhalten, dass es Hinweise für die Bedeutsamkeit der normativen Peerkultur für die Lernmotivation von Schülerinnen und Schülern gibt. Die normative Peerkultur wurde kaum in Zusammenhang mit zuwanderungsbezogener Zusammensetzung untersucht.

2.2.6.4 Peer effects: Vergleichsprozesse zwischen Peers und Schülermotivation

Schülerinnen und Schüler vergleichen sich mit anderen Schülerinnen und Schülern ihrer Klasse, ihrer Schule oder desselben Bildungsgangs, um Informationen über ihre eigenen Fähigkeiten zu erlangen. Diese Information nutzen sie zur Bildung ihrer eigenen Fähigkeitsselbsteinschätzung bzw. des akademischen Selbstkonzepts (vgl. *Social Comparison Theory* von Festinger, 1954). Das allgemeine akademische Selbstkonzept in einer Domäne und die situationsspezifischere Selbstwirksamkeitserwartung einer Person gelten als motivationale Merkmale von Schülerinnen und Schülern entsprechend der Erwartungskomponente von Erwartungs-mal-Wert-Modellen der Motivation (z.B. Eccles, 2009). Demnach kann eine positive Fähigkeitsselbsteinschätzung eine höhere Lernfreude und Lernverhalten sowie letztendliche positive Fähigkeitsentwicklung fördern.

Bezogen auf Merkmale der Klassenzusammensetzung, scheint zunächst einmal das Fähigkeitsniveau der Mitschülerinnen und Mitschüler in der Klasse für die Fähigkeitsselbsteinschätzung und nachfolgende Lernmotivation, Lernverhalten und Leistungsentwicklung relevant: Der von Herbert W. Marsh (1987) formulierte *Big-Fish-Little-Pond-Effect* besagt, dass eine Schülerin bzw. ein Schüler in einer Klasse mit einem höheren durchschnittlichen Leistungsniveau ihre/seine eigene Leistung geringer einschätzt als wenn sie/er in einer Klasse mit einem durchschnittlich niedrigeren Leistungsniveau beschult würde. Das Fähigkeitsniveau der Referenzgruppe ist also für die eigene Fähigkeitsselbsteinschätzung relevant, wobei oft der eben beschriebene kontrastierende Effekt angenommen wird (vgl. z.B. Marsh, 1987; Nagengast & Marsh, 2012). Gleichzeitig bieten sich aber auch in Klassen mit einem durchschnittlich höheren Leistungsniveau mehr Möglichkeiten, sich mit leistungsstarken Schülerinnen und Schülern zu vergleichen und sich ihnen anzugleichen (vgl. *upward social comparison*

z.B. bei Collins, 1996; Dijkstra, Kuyper, van der Werf, Buunk & van der Zee, 2008). Diese Vergleiche seien informativer für die eigene Leistungsentwicklung, da sie Auskunft darüber geben könnten, wie man sich selbst verbessern kann. Somit könnten sie u.U. nicht nur zu einer niedrigeren Fähigkeitsselbsteinschätzung im Sinne von eben beschriebenen Kontrasteffekten führen, sondern sich auch in einem Assimilationseffekt positiv auf die eigene Fähigkeitseinschätzung und nachfolgende Lernmotivation auswirken (Collins, 1996). Bezüglich der schulischen Lernsituation wird dieser sogenannte *Basking-in-reflected-glory-effect* (Cialdini et al., 1976), also die persönliche Selbstaufwertung durch das Zugehörigkeitsgefühl zu einer positiv bewerteten Gruppe, zumeist auf die Situation von Eliteschulen oder höheren Bildungsgängen im deutschen Schulsystem bezogen (vgl. Baumert et al., 2006; Chmielewski, Dumont & Trautwein, 2013; Köller, 2004; Marsh, Kong & Hau, 2000; Trautwein et al., 2006). Demnach sollten beispielsweise Schülerinnen und Schüler des Gymnasiums eine erhöhte Fähigkeitsselbsteinschätzung und erhöhte Lernmotivation aufweisen, da sie sich zum ranghöchsten Bildungsgang des Schulsystems zugehörig fühlen. Die Kehrseite wären Wahrnehmungen von Schülerinnen und Schülern an Hauptschulen, dass sie in der rangniedrigsten Schulart beschult werden, was zu einem Stigmatisierungserleben und geringerer Fähigkeitsselbsteinschätzung sowie Lernmotivation führen kann (vgl. Baumert et al., 2006; Knigge, 2009). Für Vergleichsprozesse sollten also insgesamt vor allem die leistungsbezogene Klassen- oder Schulzusammensetzung bzw. der Bildungsgang in gegliederten Schulsystemen relevant sein. Für Heranwachsende mit Zuwanderungshintergrund lässt sich ergänzen, dass Festinger (1954) allgemein in seiner *Social Comparison Theory* postuliert hat, dass Personen sich vor allem mit anderen Personen vergleichen, die ihnen ähnlich sind. Ein Merkmal der Ähnlichkeit kann die Zugehörigkeit zu derselben sozialen Gruppe bzw. *ingroup* im Sinne der oben erwähnten *Social Identity Theory* (Tajfel & Turner, 1986) darstellen. Demnach könnte es für Schülerinnen und Schüler mit Zuwanderungshintergrund hilfreich sein, andere Schülerinnen und Schüler mit Zuwanderungshintergrund oder desselben ethnischen Hintergrunds in der Klasse als Vergleichsperson hinzuziehen zu können. Insgesamt sind viele parallel verlaufende Vergleichsprozesse mit verschiedenen Referenzgruppen und aus unterschiedlichen Motiven möglich (vgl. z.B. Dijkstra et al., 2008; Köller, 2004).

Studien, die sich mit Vergleichsprozessen zwischen Peers als Mediator von Effekten der Zusammensetzung auf motivationale Schülermerkmale auseinandersetzen,

erfragen meist nicht direkt, mit wem sich Schülerinnen und Schüler vergleichen. Vielmehr wird indirekt durch Zusammenhänge zwischen Leistungskomposition und individuellem Fähigkeitsselbstkonzept als motivationalem Merkmal auf zugrundeliegende Vergleichsprozesse geschlossen. Sehr konsistent lassen sich in verschiedenen Ländern und Altersgruppen Referenzgruppeneffekte im Sinne des Big-Fish-Little-Pond-Effekts nachweisen (vgl. z.B. Baumert et al., 2006; Marsh et al., 2012; Nagengast & Marsh, 2012). Schülerinnen und Schüler in leistungstärkeren Klassen oder Schulen weisen also ein niedrigeres akademisches Selbstkonzept auf als vergleichbare Schülerinnen und Schüler in einer leistungsschwächeren Klasse oder Schule. Außerdem lassen sich im deutschen gegliederten Schulsystem Assimilationseffekte, d.h. Reputations- und Stigmatisierungseffekten nachweisen (vgl. z.B. Baumert et al., 2006; Chmielewski et al., 2013; Knigge, 2009; Köller, 2004; Schümer, 2004; Trautwein et al., 2006); in ihrer Stärke sind sie meist jedoch geringer als Referenzgruppeneffekte auf Klassenebene (z.B. Baumert et al., 2006; Trautwein et al., 2006). Zusätzlich finden sich Hinweise darauf, dass Schülerinnen und Schüler dazu neigen, sich mit leistungstärkeren Schülerinnen und Schülern zu vergleichen (*upward social comparison*; Collins, 1996; Dijkstra et al., 2008). Zusammenhänge mit dem Selbstkonzept der Schülerinnen und Schüler sind dabei in verschiedenen Studien uneinheitlich; es lassen sich Belege sowohl für ein negativeres als auch positiveres Selbstkonzept bei Aufwärtsvergleichen finden (Collins, 1996; Dijkstra et al., 2008; Köller, 2004).

Bezogen auf Schülerinnen und Schüler mit Zuwanderungshintergrund deuten einige Studien darauf hin, dass diese aufgrund wahrgenommener Ähnlichkeit ihre Leistung eher mit anderen Schülerinnen und Schülern mit Zuwanderungshintergrund vergleichen (vgl. Dijkstra et al., 2008; Meisel & Blumberg, 1990). Des Weiteren gibt es erste Hinweise dafür, dass der Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund positiv mit dem akademischen Selbstkonzept von Schülerinnen und Schülern in Zusammenhang steht, was mit stereotypen Wahrnehmungen eines geringeren Leistungsniveaus von Schülerinnen und Schülern mit Zuwanderungshintergrund assoziiert ist (vgl. Darstellung von Dijkstra et al., 2008). Inwiefern solch differenzierte Vergleichsprozesse bedeutsam für motivationale Merkmale von Schülerinnen und Schüler mit Zuwanderungshintergrund sind, stellt ein Desiderat der Forschung dar.

2.2.7 Methodische Überlegungen zur Analyse von Effekten der Klassenzusammensetzung

Analysen zu den Effekten von Aggregatsmerkmalen galten insbesondere in den Anfangsjahren der Forschung teilweise als wenig vertrauenswürdig. Dies betraf insbesondere Analysen von Kompositionseffekten, also die Effektschätzung eines Aggregatmerkmals unter Kontrolle des entsprechenden Individualmerkmals (vgl. Abschnitt 2.2.2). Befunde von Kompositionsanalysen seien lediglich das Resultat fehlspezifizierter Modelle. Über in den letzten Abschnitten dargestellte Vermittlungsmöglichkeiten und Ursachen bräuchte man nicht zu diskutieren, da Kompositionseffekte vor allem Artefakte darstellen (vgl. z.B. Harker & Tymms, 2004; Hauser, 1970; Van Ewijk & Slegers, 2010b). Mittlerweile wird davon ausgegangen, dass es sich bei Kompositionseffekten nicht nur um Artefakte handeln muss (Thrupp et al., 2002), was nicht zuletzt der technischen Weiterentwicklung von differenzierteren Analysemodellen zu verdanken ist (vgl. Meece & Schaefer, 2010). Nachfolgend wird nun kurz allgemein auf Analysemodelle für Kompositionseffekte am Beispiel der Klassenkomposition sowie auf in der Literatur diskutierte Herausforderungen bei der Untersuchung von Kompositionseffekten eingegangen.

Die Analyse von Kompositionseffekten schließt per se die Berücksichtigung von Daten auf verschiedenen Ebenen (z.B. Individuum, Schulklasse, Schule etc.) ein. Somit ist die Anwendung von Mehrebenenanalysen (*multilevel modeling*, *hierarchical linear models*) erforderlich. Mehrebenenanalysen ermöglichen es, die geschachtelte Datenstruktur und damit verbundene Abhängigkeiten (z.B. stärkere Ähnlichkeit von Schülerinnen und Schülern innerhalb einer Klasse im Vergleich zu Schülerinnen und Schülern einer anderen Klasse) bei der Bestimmung von Standardfehlern zu berücksichtigen und erlauben die simultane Verwendung von Variablen auf der jeweils entsprechenden Ebene sowie die Abbildung von Beziehungen unter ihnen (vgl. z.B. Ditton, 1998; Raudenbush & Bryk, 2002; Snijders & Bosker, 2012). Es werden dabei unter anderem Random-Intercept-Modelle und Random-Slope-Modelle unterschieden. In Random-Intercept-Modellen werden verschiedene ebenenspezifische Regressionsgleichungen formuliert, bei denen zum Beispiel auf Individualebene (L1, Effekte innerhalb von Klassen) der Wert eines Schülers auf der Kriteriumsvariable (z.B. sein Lesetestscore) durch den Klassenmittelwert der Kriteriumsvariable (Lesetestscore) und Merkmale des Individuums (z.B. individueller sozioökonomischer Status [SES]) erklärt werden. Auf Klassenebene (L2, Effekte zwischen Klassen) stellt der

Klassenmittelwert (des Lesetest scores) die abhängige Variable in der Regressionsgleichung dar, die durch den Mittelwert in der Kriteriumsvariable über alle Klassen sowie unabhängige Variablen – zum Beispiel den durchschnittlichen SES der Klasse – vorhergesagt werden kann. In diesem Modell wird also davon ausgegangen, dass systematische Mittelwertsunterschiede auf der abhängigen Variable zwischen Klassen bestehen, die es zu erklären gilt. Abbildung 2.7 zeigt schematisch die im Rahmen von Mehrebenenmodellen beschriebenen Effekte am Beispiel von Effekten des SES auf den Lesetest score. In der Abbildung repräsentieren die drei grauen Linien die Regressionsgeraden für drei verschiedene Klassen. Die Klassen in Abbildung 2.7a unterscheiden sich nur in ihrem durchschnittlichen SES und in dem durchschnittlichen Lesetest score; die Regressionsgeraden verlaufen parallel zueinander. Die unterbrochene Linie zeigt die Regressionsgerade für Effekte zwischen Klassen.

Random-Slope-Modelle nehmen zusätzlich zu Mittelwertsunterschieden an, dass sich Klassen auch in der Beziehung bzw. dem Regressionsgewicht (*slope*) zwischen Prädiktor- und Kriteriumsvariable unterscheiden können. So könnte beispielsweise der Zusammenhang zwischen Lesetest score und SES oder zwischen Zuwanderungshintergrund und Motivation in manchen Klassen stärker als in anderen ausgeprägt sein (siehe Abbildung 2.7b). Sollen beispielweise Aussagen über Effekte der Zusammensetzung für bestimmte Schülergruppen getroffen werden, wie „Der Zusammenhang zwischen dem SES-Niveau in der Klasse und dem Lesetest score von Heranwachsenden mit niedrigem SES ist stärker ausgeprägt als der Zusammenhang mit dem Lesetest score von Heranwachsenden mit hohem SES“, sind Cross-Level-Interaktionen (Slope-as-outcome-Modell) zu analysieren. Hierbei darf das Regressionsgewicht für den SES auf Individualebene zwischen verschiedenen Klassen variieren (*random slope*). Dieses Regressionsgewicht stellt dann auf Klassenebene eine weitere abhängige Variable dar, die durch Merkmale der Klasse, wie zum Beispiel das SES-Niveau in der Klasse, vorhergesagt werden kann (vgl. Aguinis, Gottfredson & Culpepper, 2013; Marsh et al., 2009).

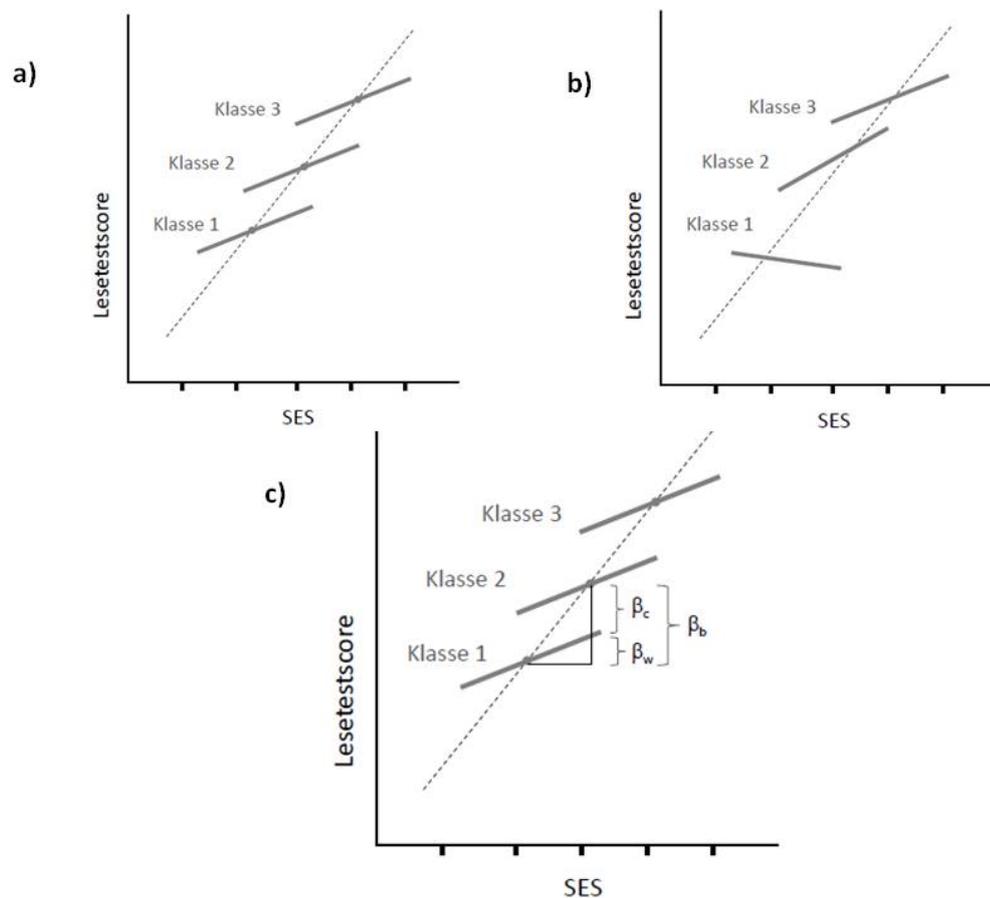


Abbildung 2.7: Schematische Darstellung von Effekten innerhalb und zwischen Klassen (entnommen und leicht angepasst aus Lüdtke, Robitzsch, Trautwein & Kunter, 2009, S. 129 und Raudenbush & Bryk, 2002, S. 140)

Wie in Abschnitt 2.2.2 eingeführt, beschreibt ein Kompositionseffekt die Effektschätzung eines Aggregatmerkmals unter Kontrolle des entsprechenden Individualmerkmals (beispielsweise individueller SES und Klassenmittelwert des sozioökonomischen Status zur Vorhersage von Schülerleistung). Um Kompositionseffekte zu untersuchen wird ein oben beschriebenes Random-Intercept-Modell verwendet. In so einem Modell können zunächst ebenenspezifische, voneinander unabhängige Zusammenhänge analysiert werden. In dem Analysebeispiel zu Lesetestscores und SES besteht auf Individualebene ein positiver Effekt; individuelle Schülerinnen und Schüler mit höherem SES erreichten einen höheren Lesetestscore (*within effect*). Das Regressionsgewicht (β_w) für den SES auf Individualebene gibt den erwarteten Unterschied im Lesescore von zwei Schülerinnen und Schülern an, die sich um eine Einheit des SES unterscheiden. Auf Klassenebene gibt es ebenfalls einen positiven Effekt, Klassen mit einem höheren SES-Niveau haben auch einen höheren mittleren Testscore (*between effect*). Das Regressionsgewicht (β_b) des mittleren SES gibt

den erwarteten Unterschied zwischen zwei Klassen an, die sich in ihrem mittleren SES um eine Einheit unterscheiden. Ein Kompositionseffekt zeigt nun, ob der mittlere SES der Klasse einen Effekt auf die Leseleistung hat, wenn man dafür kontrolliert, dass individuelle Schülerinnen oder Schüler mit einem höheren SES einen höheren Lesetestscore erreichen. Er zeigt also den erwarteten Unterschied im Lesetestwert von zwei Schülerinnen und Schülern mit demselben SES, die Klassen besuchen, deren mittlerer SES sich um eine Einheit unterscheidet. Ein Kompositionseffekt liegt vor bei einer statistisch bedeutsamen Differenz zwischen *between effect* und *within effect* ($\beta_c = \beta_b - \beta_w$; siehe Abbildung 2.7c; vgl. z.B. Ditton, 1998; Enders & Tofighi, 2007; Lüdtke et al., 2008; Lüdtke et al., 2009; Raudenbush & Bryk, 2002)⁷. Analog lässt sich ein zuwanderungsbezogener Kompositionseffekt als der Effekt des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund unter Kontrolle des individuellen Zuwanderungshintergrunds beschreiben.

Zentrale Kritikpunkte an Kompositionsanalysen, die auch unter Stichworten wie „*phantom effect*“ oder „*contextual fallacy*“ bekannt wurden, ist eine mögliche Unreliabilität oder starke Messfehlerbehaftetheit von Individualmerkmalen, die statistisch bedeutsame Kompositionseffekte hervorrufen können (Harker & Tymms, 2004; Hauser, 1970; vgl. Lüdtke et al. 2011; Lüdtke et al., 2008; Lüdtke, Robitzsch & Köller, 2002). Werden diese messfehlerbehafteten Individualmerkmale aggregiert, um ein Klassenmerkmal zu bilden, kann ein statistisch bedeutsamer Kompositionseffekt entstehen, da die Effekte des Klassenmerkmals in ungenügender Weise für Effekte des Individualmerkmals kontrolliert werden (Aggregationsbias; z.B. Harker & Tymms, 2004; Lüdtke et al. 2011; Marsh et al., 2012; vgl. Baumert et al., 2006). Eine weitere mögliche Fehlerquelle auf Klassenebene besteht in nur einer kleinen Anzahl von Angaben der Schülerinnen und Schülern, die aggregiert werden, um ein Klassenmerkmal abzubilden (*sampling error*; Lüdtke et al., 2011; Lüdtke et al., 2008).

Insbesondere in querschnittlich angelegten Studien besteht zudem das Simultanitäts- bzw. Reziprozitätsproblem (vgl. Baumert et al., 2006; Duncan, Magnuson & Ludwig, 2004; Duncan & Raudenbush, 1999), das dadurch entsteht, dass Schülerinnen

⁷ Diese beispielhafte Beschreibung bezieht sich auf eine Zentrierung des individuellen SES am Klassenmittelwert. Wird der individuelle SES am Gesamtmittelwert aller Schülerinnen und Schüler der Analysestichprobe zentriert, repräsentiert das Regressionsgewicht des daraus gebildeten Klassenmittelwerts des sozioökonomischen Status direkt den Kompositionseffekt, also den Effekt des Aggregatmerkmals unter Kontrolle des Individualmerkmals.

und Schüler nicht nur von ihrem Klassenumfeld beeinflusst werden, sondern ihr Klassenumfeld auch aktiv mitgestalten. Vor allem in Querschnittstudien ist dadurch die Aussagekraft von Kompositionsanalysen eingeschränkt. Gerade im Falle der Leistungsvorhersage durch Kompositionsmerkmale sind Maße des Vorkenntnisniveaus der Klasse und Individuen in einem längsschnittlichen Design vonnöten; eine Vernachlässigung kann zur Überschätzung von Kompositionseffekten führen. Zwar konnten Baumert et al. (2006) zeigen, dass in Querschnittsdesigns Maße der kognitiven Grundfähigkeiten als Proxy für Vorkenntnisse verwendet werden können. Es bleibt jedoch zu berücksichtigen, dass auch die Entwicklung der kognitiven Grundfähigkeiten durch die Schule und Schülerinnen und Schüler beeinflusst wird (vgl. z.B. Köller & Baumert, 2002; Becker et al., 2012) und somit ein ähnliches Reziprozitätsproblem vorliegen kann. Durch Verwendung eines mit der abhängigen Variable zeitgleich erhobenen Maßes der kognitiven Fähigkeiten könnte es auch zu einer Überkontrolle auf individueller Ebene und damit zu einer Unterschätzung von Kompositionseffekten kommen. Im allgemeinen Fall wird die mangelnde Berücksichtigung zentraler erklärender Variablen als Endogenitätsproblem bezeichnet (vgl. Duncan et al., 2004; Duncan & Raudenbush, 1999; Hauser, 1970; vgl. Abschnitt 8.4). Bleiben zentrale Individualmerkmale unberücksichtigt, die mit der abhängigen Variable und dem Klassenkontext in Zusammenhang stehen, kann dies zur Überschätzung von Kompositionseffekten führen. Besonders problematisch ist es, wenn bei der Interpretation von Kompositionseffekten auf Erklärungen zurückgegriffen wird, die in Analysemodellen nicht direkt berücksichtigt werden, wie es Hauser (1970) mit seinem *consexual effect* illustrierte. Ebenso kann die mangelnde Berücksichtigung weiterer Klassen- oder Kontextmerkmale zu Überschätzungen von Kompositionseffekten führen (vgl. z.B. Baumert et al. 2006; Harker & Tymms, 2004; Legewie, 2012; Rumberger & Palardy, 2005), beispielsweise wenn bei Analysen zur Leistungskomposition keine Informationen über Bildungsgänge in gegliederten Systemen einbezogen werden (siehe allgemein auch *omitted-variable bias*).

Bei Analysen von Effekten der Zusammensetzung sind auch Fragen der Operationalisierung des Konstrukts beispielsweise als Mittelwert oder Maß der Heterogenität zentral. Im Zusammenhang mit der zuwanderungsbezogenen Klassenzusammensetzung wurde bereits in Abschnitt 2.2.2.2 auf unterschiedliche Formen der Operationalisierung eingegangen, die auch in den anschließenden Abschnitten der Darstellung des theoretischen Hintergrunds und empirischen Forschungsstands

berücksichtigt wurden. In Teilstudie 3 (siehe Kapitel 6) wird ebenfalls ausführlich das Thema der Operationalisierung zuwanderungsbezogener Zusammensetzung behandelt.

2.3 Zusammenfassung

Das schulische Lernumfeld ist zentral für die Leistungs- und psychosoziale Entwicklung der Schülerinnen und Schüler und leistet in Ergänzung zur in Abschnitt 2.1 dargestellten individuellen Ebene einen Beitrag zur Erklärung und Vorhersage der Schulleistung und Lernmotivation. Der Lernkontext besteht dabei aus verschiedenen Ebenen, wie dem Bildungssystem eines Landes, dem Bildungsgang innerhalb gegliederter Systeme, der konkreten Schule und der Schulklasse. Letzterer kommt als proximales Umfeld die stärkste Bedeutung in der Entwicklung zu: Schülerinnen und Schüler verbringen einen Großteil ihrer Zeit im Klassenraum, in dem sie mit der Lehrkraft sowie ihren Mitschülerinnen und Mitschülern interagieren. Auf der Ebene der Schulklasse ist die Klassenzusammensetzung ein entwicklungsrelevantes Merkmal. Die Klassenzusammensetzung umfasst allgemein verschiedene Charakteristika wie die Heterogenität einer Schülergruppe und die durchschnittliche Merkmalsausprägung in der Gruppe, deren Effekt in Kompositionsanalysen untersucht wird. Klassenkompositionseffekte beschreiben allgemein Effekte eines Aggregatmerkmals auf Klassenebene unter gleichzeitiger Berücksichtigung des entsprechenden Merkmals auf Individualebene. Konsistent über verschiedene Fächer, Altersgruppen und Länder zeigt sich ein positiver leistungsbezogener Kompositionseffekt auf individuelle Schülerleistung. Dies bedeutet, dass Schülerinnen und Schüler – ungeachtet ihres eigenen Vorkenntnisniveaus oder Lernpotentials – höhere Leistungen erreichen, wenn sie in leistungsstärkeren Klassen lernen. Ein hohes mittleres Leistungsniveau wirkt sich zudem eher negativ auf das akademische Selbstkonzept und die Lernmotivation individueller Schülerinnen und Schüler aus. Des Weiteren zeigen Studien häufig einen positiven Effekt der sozialen Komposition auf individuelle Schülerleistung. Dieser ist zu einem Großteil mit Effekten der Leistungskomposition konfundiert, jedoch weist die soziale Komposition meist einen kleinen eigenständigen Effekt auf. Schülerinnen und Schüler mit Zuwanderungshintergrund besuchen häufiger Schulklassen mit einem niedrigen Leistungsniveau und niedrigerem mittleren sozialen Status, sodass die dargestellten Befunde insbesondere für diese Schülergruppe relevant sind.

Die empirische Befundlage bezüglich Effekten zuwanderungsbezogener Zusammensetzung ist uneinheitlich. Unter zuwanderungsbezogener Zusammensetzung wird einerseits der in Kompositionsanalysen untersuchte Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und andererseits die zuwanderungsbezogene Heterogenität als Klassenmerkmal gefasst. Zuwanderungsbezogene Kompositionseffekte beschreiben den Effekt des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund unter Kontrolle des individuellen Zuwanderungshintergrunds. Insgesamt ist über die Leistungs- und soziale Komposition hinaus eher von keinem eigenständigen oder einem leicht negativen zuwanderungsbezogenen Kompositionseffekt auszugehen. Vereinzelt zeigen sich jedoch auch positive Effekte zuwanderungsbezogener Zusammensetzung auf Schülerleistung. Dies ist vor allem in Studien zu Effekten zuwanderungsbezogener Heterogenität der Fall. Eine erhöhte zuwanderungsbezogene Heterogenität geht mit einem erhöhten Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund einher, sodass teils konkurrierende Effekte in verschiedenen Studien zu zuwanderungsbezogener Zusammensetzung auf Schülerleistung vorliegen. Bezogen auf psychosoziale Merkmale sind empirische Befunde ebenfalls uneinheitlich. Es gibt Hinweise darauf, dass zuwanderungsbezogene Heterogenität eher negativ mit der Lernmotivation zusammenhängt. Für den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund finden sich vereinzelt auch positive Effekte auf die Lernmotivation. Effekte der beiden Charakteristika zuwanderungsbezogener Zusammensetzung wurden kaum gemeinsam untersucht.

Des Weiteren gibt es bislang nur vereinzelt empirische Studien, die die Vermittlung von Effekten der Klassenzusammensetzung allgemein und insbesondere der zuwanderungsbezogenen Zusammensetzung auf Schülerleistung untersuchen. Insbesondere die Unterrichtsqualität wird in theoretischen Annahmen als Mediator diskutiert. Empirisch zeigten sich bislang erste Hinweise darauf, dass Merkmale der Unterrichtsqualität Effekte der Zusammensetzung partiell vermitteln. Für die zuwanderungsbezogene Zusammensetzung gibt es Hinweise, dass der Unterricht in Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund weniger kognitiv aktivierend gestaltet wird. Für ergänzende theoretische Annahmen zur Vermittlung von Effekten der zuwanderungsbezogenen Zusammensetzung auf Schülerleistung, wie einer normativen Peerkultur, der Sprachverwendung zwischen Peers und dem lernförderlichen Austausch, gibt es bislang wenige Untersuchungen. Aufgrund der bisherigen Befundlage ist von einer geringeren

Bedeutung dieser Mediatoren im Vergleich zur Unterrichtsqualität auszugehen. Studien beziehen sich dabei meist auf Effekte des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund und kaum auf zuwanderungsbezogene Heterogenität.

Für die Vermittlung von Effekten der zuwanderungsbezogenen Klassenzusammensetzung auf psychosoziale Schülermerkmale, wie der Lernmotivation, lassen sich bislang ebenso nur erste Hinweise finden. Einige Studien zeigten, dass ein durch die Lehrkraft beeinflusstes, unterstützendes, motivationsförderliches Unterrichtsklima in Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und in zuwanderungsbezogen heterogenen Klassen vorliegt, was positive Effekte zwischen Zusammensetzung und Lernmotivation erwarten lässt. Andererseits gibt es Hinweise auf ein weniger starkes Verbundenheitsgefühl in zuwanderungsbezogen heterogenen Klassen, was eine geringere Lernmotivation in heterogenen Klassen erwarten lässt. Für den Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und dem Verbundenheitsgefühl mit den Peers liegen bislang keine empirischen Studien vor. Zu weiteren theoretisch angenommenen Mediatoren von Effekten zuwanderungsbezogener Zusammensetzung, wie der normativen Peerkultur und Vergleichsprozessen zwischen Peers, liegen bislang ebenfalls kaum Studien vor. Vor dem Hintergrund der bisherigen Befundlage scheint ihnen eine geringere Bedeutung zuzukommen.

3

Ziele und Fragestellungen der vorliegenden Arbeit

3. Ziele und Fragestellungen der vorliegenden Arbeit

Untersuchungsgegenstand der vorliegenden Arbeit ist die Leistungs- und psychosoziale Entwicklung von Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund. Hierbei wird insbesondere die Rolle des proximalen Umfelds der Klassenzusammensetzung für die akademische Entwicklung Heranwachsender und die Vermittlung von Effekten der Zusammensetzung fokussiert. Die durchschnittlich geringeren erreichten Kompetenzniveaus von Schülerinnen und Schülern mit Zuwanderungshintergrund im deutschen Bildungssystem wurden vielfach dokumentiert (vgl. Kapitel 1). Zur Erklärung dieser zuwanderungsbezogenen Disparitäten werden bislang vor allem individuelle und innerfamiliäre Determinanten hinzugezogen (vgl. Kapitel 2.1), die zusätzliche Bedeutung des Lernumfelds für die akademische Entwicklung der Schülerinnen und Schüler rückte in den letzten Jahrzehnten jedoch immer stärker in das Forschungsinteresse (vgl. Kapitel 2.2). Ob die zuwanderungsbezogene Klassenzusammensetzung in Zusammenhang mit der Leistung und psychosozialen Schülermerkmalen steht, ist nicht eindeutig geklärt. Dies betrifft zum einen die Untersuchung von verschiedenen konzeptuellen Bestandteilen der zuwanderungsbezogenen Klassenzusammensetzung, also dem Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der zuwanderungsbezogenen Heterogenität im Klassenzimmer. Zum anderen betrifft dies die Frage, ob von leistungsbezogener und sozialer Zusammensetzung unabhängig zuwanderungsbezogene Effekte der Zusammensetzung auf Leistung und psychosoziale Schülermerkmale zu finden sind und ob diese durch Merkmale der Unterrichtsqualität vermittelt werden. Die Auseinandersetzung mit diesen Forschungslücken ist wichtig, da zuwanderungsbezogene Zusammensetzung innerhalb des Bildungssystems und unterrichtlicher Umgang mit ihr zunehmend von Bedeutung ist. Es wird davon ausgegangen, dass der Umgang mit ethnischer oder zuwanderungsbezogener Zusammensetzung und Heterogenität in der Gesellschaft und folglich innerhalb des Bildungssystems z.B. durch Arbeitsmigration und veränderte Geburtenraten eine der zentralen Herausforderungen der kommenden Jahrzehnte darstellt (vgl. z.B. Autorengruppe Bildungsberichterstattung, 2014). Die Institution Schule bietet einen Raum, der gesellschaftliche soziale Integration durch den zwischenmenschlichen Kontakt in der Schule und Klasse ermöglichen kann (vgl. Blau, 1977). Es ist also diesbezüglich von zentraler Bedeutung, Informationen darüber zu gewinnen, wie alle Schülerinnen und

Schüler ungeachtet ihrer ethnischen oder sozialen Herkunft gefördert werden können, welche Auswirkungen der zuwanderungsbezogenen Klassenzusammensetzung zu erwarten sind und wie ihnen ggf. begegnet werden kann.

Vor diesem Hintergrund setzt sich die vorliegende Arbeit das Ziel, Effekte der zuwanderungsbezogenen Schulklassenzusammensetzung im Sinne eines breiten Verständnisses von Bildungserträgen sowohl auf verschiedene Leistungsmaße als auch auf motivationale Schülermerkmale zu untersuchen. Besonderes Augenmerk soll dabei einerseits auf die Operationalisierung der zuwanderungsbezogenen Klassenzusammensetzung und andererseits auf mögliche Transmissionswege der Effekte durch die Unterrichtsqualität als Mediator (vgl. Abschnitt 2.2.5.1 und 2.2.6.1) gelegt werden. Grundlage der Untersuchung bilden repräsentative Stichproben der Grundschule und Sekundarstufe I verschiedener Large-Scale-Assessment-Studien, die teils längsschnittlich angelegt sind. Abbildung 3.1 stellt das der vorliegenden Arbeit zugrundeliegende Arbeitsmodell dar. Es wurde zur Systematisierung der untersuchten Fragestellungen für diese Arbeit entwickelt und bildet ausschließlich die in dieser Arbeit untersuchten Fragestellungen ab. Es hat somit nicht den Anspruch eines vollständig umfassenden Erklärungsmodells. Die Grundlage hierfür bildeten insbesondere das Vermittlungsmodell für Kontexteffekte von Baumert et al. (2006; vgl. Abschnitt 2.2.3.1) und das Modell zu Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung von Klieme und Rakoczy (2008; vgl. Abschnitt 2.2.3.2). Nachfolgend werden nun die in den drei Teilstudien verfolgten Fragestellungen vorgestellt.

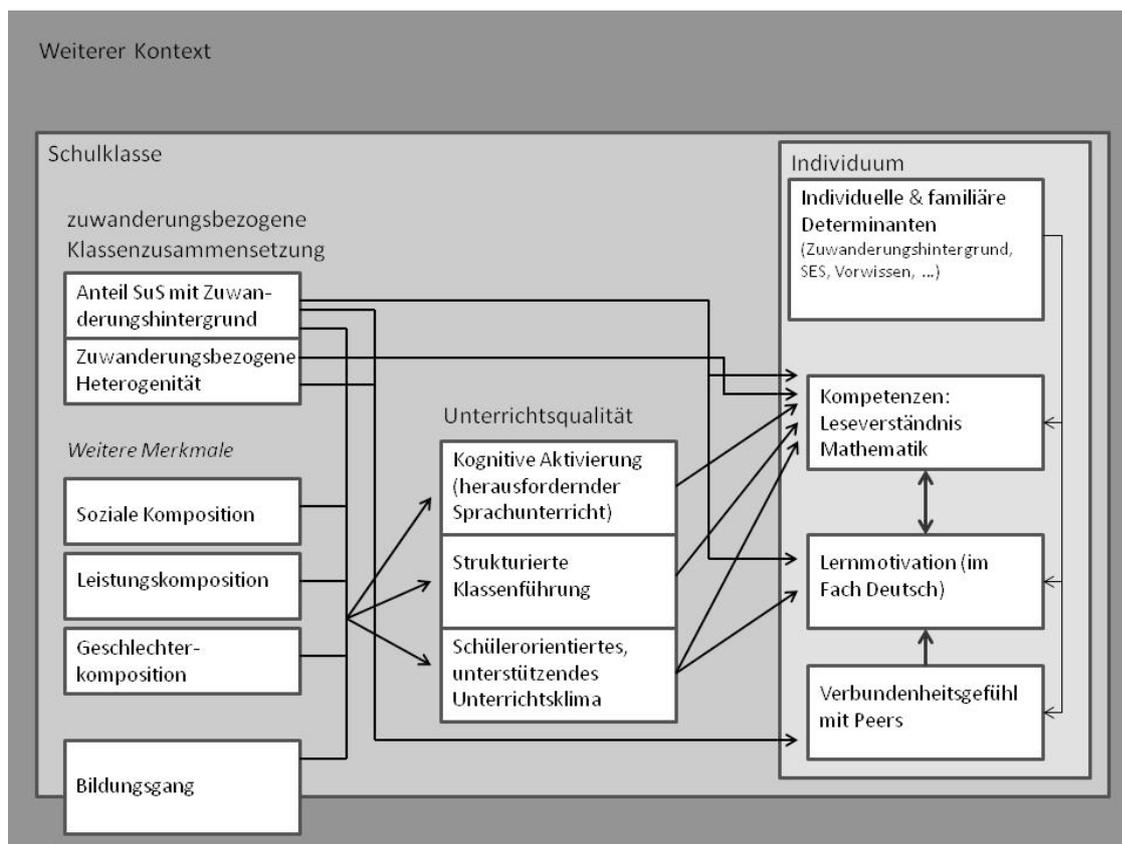


Abbildung 3.1: Arbeitsmodell der vorliegenden Arbeit

Teilstudie 1: Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality

Die erste Teilstudie setzt sich zunächst das Ziel, das häufig in Studien in Deutschland gefundene Ergebnismuster zu replizieren, dass bei Analyse von Effekten des Klassenanteils von Heranwachsenden mit Zuwanderungshintergrund und der sozialen Zusammensetzung letztere für Individuelleistung bedeutsamer ist (vgl. Abschnitt 2.2.4). Dazu wird auf Daten der längsschnittlich angelegten DESI-Studie mit ihrer bundesweit repräsentativen Stichprobe in der Sekundarstufe zurückgegriffen („Deutsch Englisch Schülerleistungen International“; DESI-Konsortium, 2008). Angenommen wird ein leicht negativer Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und dem individuellen Leseverständnis am Ende von Klasse 9 unter Kontrolle der individuellen Ausgangsleistung zu Beginn der Klasse 9. Dieser Zusammenhang sollte jedoch bei Kontrolle der sozialen Zusammensetzung und weiterer Merkmale der Klassenzusammensetzung sowie Schulmerkmale seine statistische Bedeutung verlieren. In Einklang mit dem aktuellen Forschungsstand sollte sich ein positiver sozialer Kompositionseffekt auf individuelles Leseverständnis zeigen.

In einem zweiten Schritt verfolgt diese Teilstudie das Ziel, über den aktuellen Kenntnisstand hinausgehend die mediiierende Rolle der Unterrichtsqualität in einem Vermittlungsmodell auf Klassenebenen zu prüfen. Erwartet wird hierbei entsprechend der in den Abschnitten 2.2.3.1 und 2.2.5.1 vorgestellten Annahmen, dass Effekte der sozialen Zusammensetzung auf die Individualleistung durch Merkmale der Unterrichtsqualität partiell vermittelt werden. Es wird also die Annahme überprüft, dass geringere Leistung individueller Schülerinnen und Schüler in Klassen mit einer weniger sozial privilegierten Komposition in Zusammenhang steht mit einer geringeren Unterrichtsqualität in diesen Klassen. Zur Operationalisierung der Unterrichtsqualität wird auf etablierte Modelle zurückgegriffen, die zwischen kognitiver Aktivierung, strukturierter Klassenführung und schülerorientiertem, unterstützendem Unterrichtsklima unterscheiden (siehe Abschnitt 2.2.3.2). Abbildung 3.2 zeigt den Fokus der in Teilstudie 1 geplanten Untersuchung.

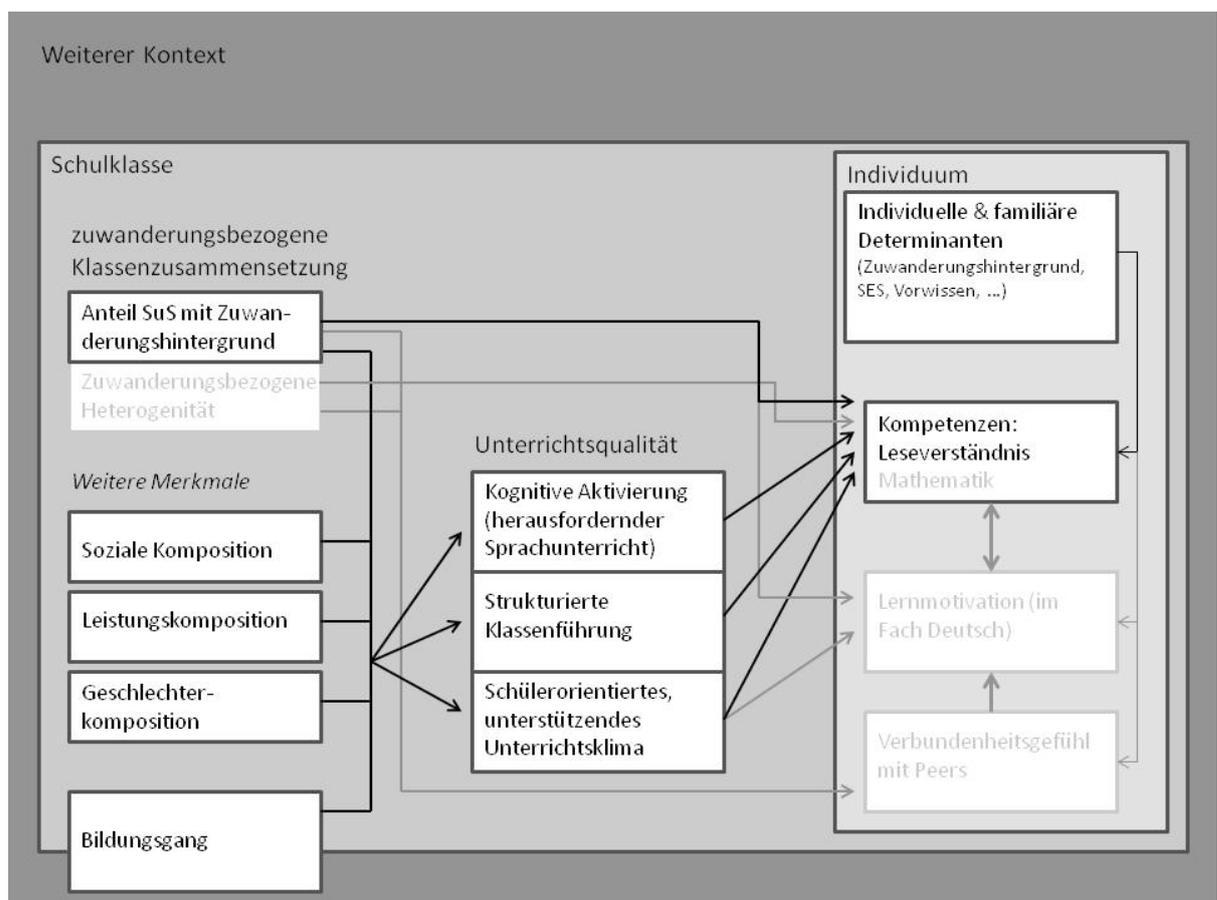


Abbildung 3.2: Überblick über Teilstudie 1

Teilstudie 2: Classroom composition and language minority students' motivation in language lessons

Die zweite Teilstudie setzt sich das Ziel, den Kenntnisstand zu Effekten der zuwanderungsbezogenen Klassenzusammensetzung zu erweitern und anstelle von Leistungsmaßen die Lernmotivation von Schülerinnen und Schülern im Verlauf eines Schuljahres in Abhängigkeit von der Klassenzusammensetzung zu untersuchen (vgl. Abschnitt 2.2.4). Erneut wird hierbei auf Daten von Schülerinnen und Schülern der DESI-Studie (DESI-Konsortium, 2008) zurückgegriffen, die eine Stichprobe von neunten Klassen zu zwei Messzeitpunkten umfassen. Des Weiteren zielt die Studie darauf ab, die Annahme zu prüfen, dass ein von der Lehrkraft gestaltetes, schülerorientiertes, unterstützendes Unterrichtsklima Effekte der Klassenzusammensetzung auf Lernmotivation vermittelt (vgl. Abschnitt 2.2.6.1). Erwartet wird ein positiver Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und individueller Lernmotivation, der partiell durch ein besonders schülerorientiertes, unterstützendes Unterrichtsklima in solchen Klassen vermittelt wird. Weiterführend wird der Frage nachgegangen, ob sich der Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und Lernmotivation für Schülerinnen und Schüler mit und ohne Zuwanderungshintergrund unterscheidet. In Anlehnung an Annahmen zur Bedeutung der Verbundenheit mit den Peers, die sich zum Teil auch auf den Zuwanderungshintergrund der Mitschülerinnen und Mitschüler beziehen (vgl. Abschnitt 2.2.6.2), wird hier vorhergesagt, dass Schülerinnen und Schüler mit Zuwanderungshintergrund eine höhere Lernmotivation aufweisen, wenn sie in Klassen mit einem höheren Anteil von Heranwachsenden mit Zuwanderungshintergrund unterrichtet werden. Analog sollten Schülerinnen und Schüler ohne Zuwanderungshintergrund eine höhere Lernmotivation aufweisen, wenn sie in Klassen mit einem höheren Anteil von Heranwachsenden ohne Zuwanderungshintergrund unterrichtet werden. Dies würde, teils im Widerspruch zur weiter oben formulierten Annahme, einen u-förmigen Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Lernmotivation erwarten lassen. Abbildung 3.3 gibt einen Überblick über die Schwerpunkte von Teilstudie 2. Da in dieser Studie keine direkten Informationen über das Verbundenheitsgefühl der Schülerinnen und Schüler mit den Peers vorliegen, hat letztgenannter Fragenkomplex lediglich explorativen Charakter und ist nicht in der Abbildung hervorgehoben.

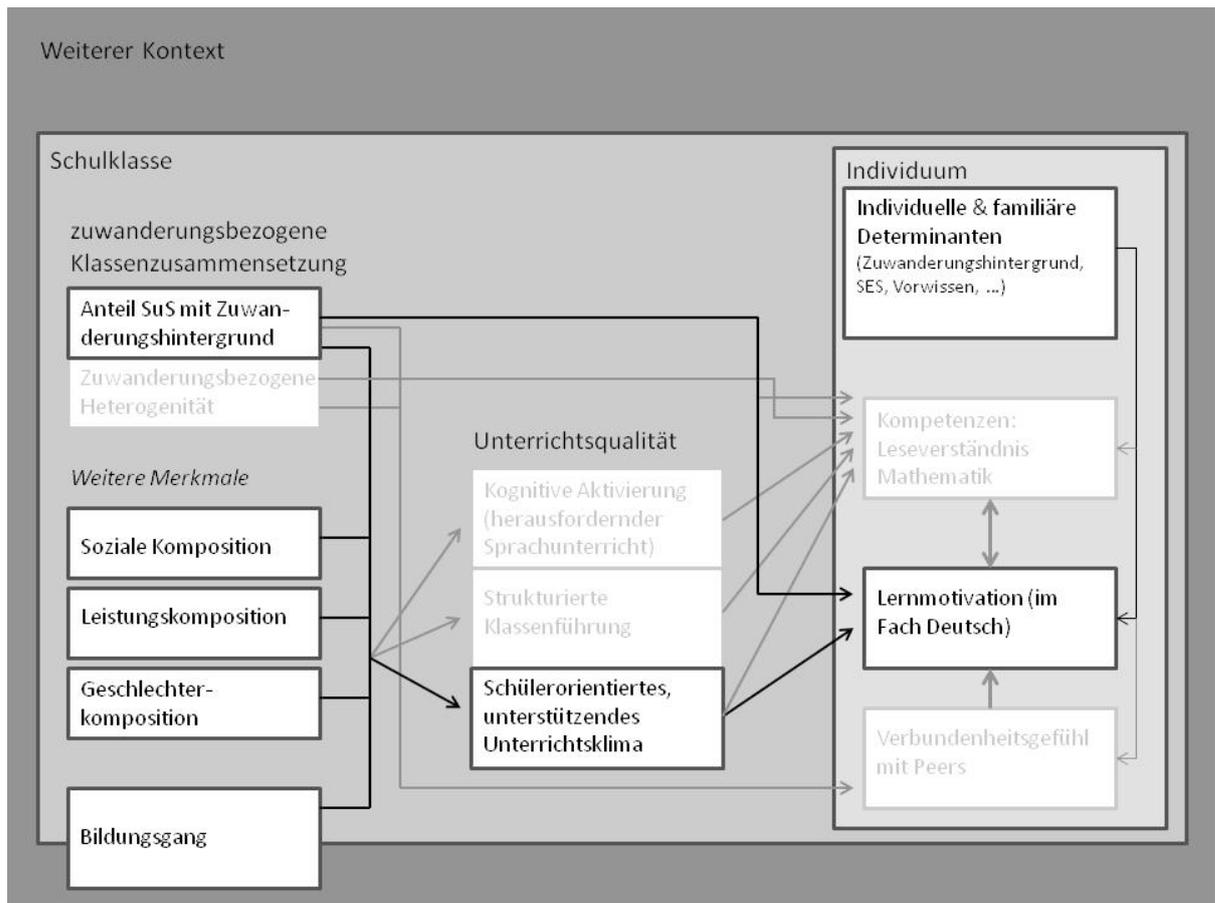


Abbildung 3.3: Überblick über Teilstudie 2

Teilstudie 3: Ethnic composition and heterogeneity in the classroom–Their measurement and relationship with student outcomes

Die dritte Teilstudie setzt sich differenziert mit der Operationalisierung von zuwanderungsbezogener Heterogenität als Charakteristik der zuwanderungsbezogenen Zusammensetzung auseinander. Ziel ist es, Maße zur Operationalisierung zuwanderungsbezogener Heterogenität unter Rückgriff auf Heterogenitätsindizes verschiedener Disziplinen zu identifizieren und ihre Zusammenhänge mit Schülerleistungen sowie dem Verbundenheitsgefühl mit den Peers in Anschluss an Teilstudie 2 zu untersuchen (vgl. Abschnitt 2.2.2, 2.2.4 und 2.2.6.2). Darüber hinaus soll der Frage nachgegangen werden, ob die zuwanderungsbezogene Heterogenität zusätzlich zum überwiegend in der Bildungsforschung verwendeten Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund als Merkmal der zuwanderungsbezogenen Zusammensetzung sowie zusätzlich zu weiteren Merkmalen der Klassenzusammensetzung mit Bildungsergebnissen der Schülerinnen und Schüler

zusammenhängt. Für die Analyse wurde auf Daten des querschnittlich angelegten IQB-Ländervergleichs 2011 (Stanat, Pant, Böhme & Richter, 2012) mit seiner repräsentativen Stichprobe von Viertklässlerinnen und Viertklässlern zurückgegriffen. Angenommen wird ein leicht negativer Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und individueller Schülerleistung (vgl. Abschnitt 2.2.4). Dieser Zusammenhang sollte vor allem durch weitere Merkmale der Klassenzusammensetzung sowie vermittelnde Prozesse im Klassenumfeld bedingt sein (vgl. Abschnitte 2.2.4 und 2.2.5). Analog wird auch für zuwanderungsbezogene Heterogenität angenommen, dass sie als alleiniger Prädiktor betrachtet in negativem Zusammenhang mit individueller Leistung steht. Zuwanderungsbezogen heterogene Klassen haben auch einen höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund. Demnach sind in solchen Klassen ebenfalls weniger lernförderliche Rahmenbedingungen und Prozesse zu erwarten. Bei Analyse des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund und zuwanderungsbezogener Heterogenität als gemeinsame Prädiktoren sowie bei Kontrolle weiterer Merkmale der Klassenzusammensetzung können hingegen theoretisch angenommene positive Zusammenhänge zwischen Heterogenität und Schülerleistung (vgl. Abschnitt 2.2.5.4) aufgedeckt werden. Diese Erwartung beruht auf der Annahme, dass bei Kontrolle des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund sowie weiterer Merkmale der Zusammensetzung auch tendenziell Prozesse im Klassen- und Schulumfeld konstant sind. Dies bietet die Möglichkeit, von diesen Rahmenbedingungen unabhängig der Frage nachzugehen, ob Heterogenität positiv mit Schülerleistung zusammenhängt.

Bezogen auf das Verbundenheitsgefühl mit den Peers wird für den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund ein u-förmiger Zusammenhang und für die zuwanderungsbezogene Heterogenität ein negativer Zusammenhang angenommen (vgl. Abschnitt 2.2.6.2, siehe auch Teilstudie 2 oben). Bei gemeinsamer Analyse von zuwanderungsbezogener Heterogenität und dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund als Prädiktoren sowie bei Kontrolle von weiteren Merkmalen der Zusammensetzung, sollte insbesondere die zuwanderungsbezogene Heterogenität für das Verbundenheitsgefühl mit den Peers von Bedeutung sein. Dies gründet auf der Annahme einer größeren Bedeutung eines herkunftsspezifischen Hintergrunds, wie er in Heterogenitätsmaßen abgebildet wird, im Vergleich zur allgemeinen Unterscheidung zwischen Heranwachsenden mit und ohne

Zuwanderungshintergrund für das Verbundenheitsgefühl (vgl. Abschnitt 2.2.6).

Abbildung 3.4 gibt einen Überblick über die Schwerpunkte von Teilstudie 3. Die anschließenden drei Kapitel enthalten die Teilstudien.

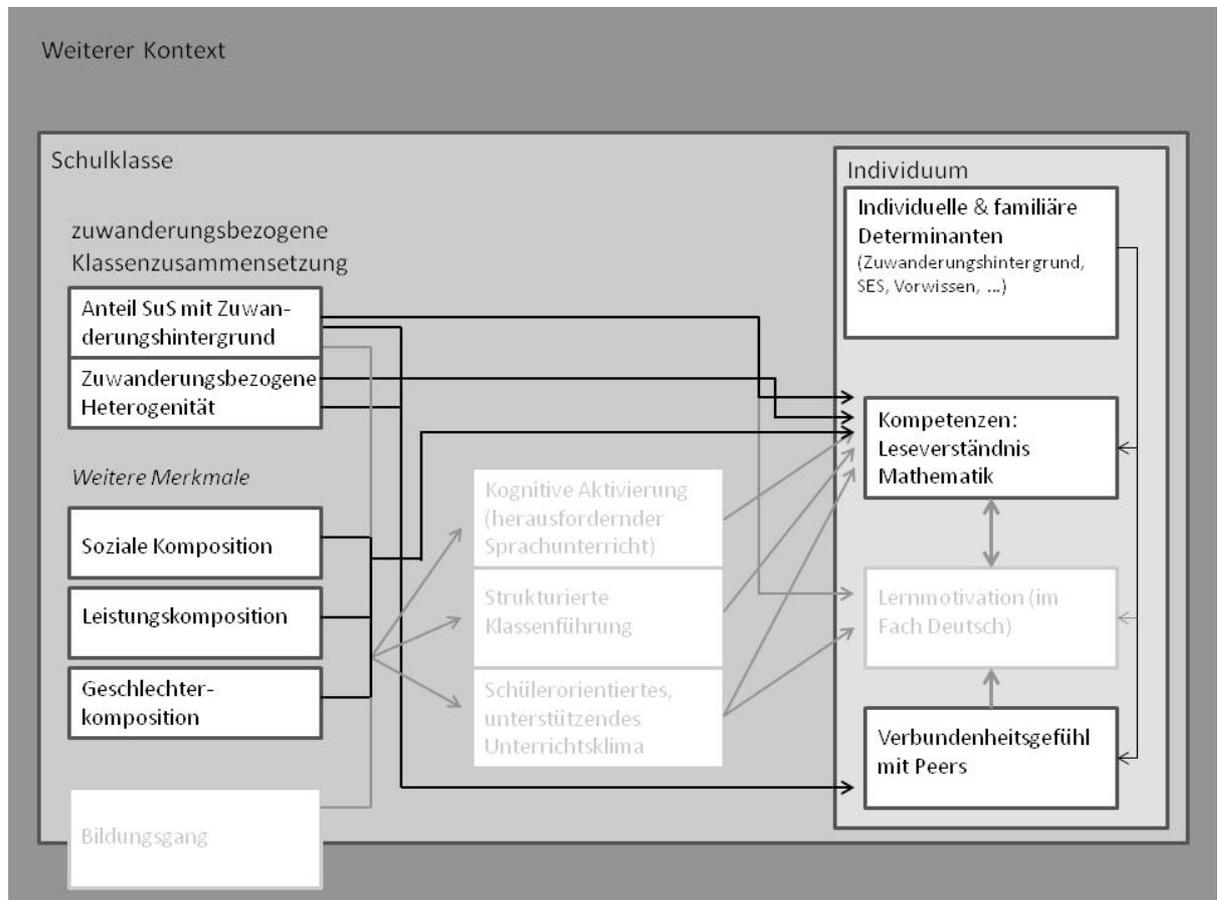


Abbildung 3.4: Überblick über Teilstudie 3

4

Teilstudie 1:

Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality

Der nachfolgende Zeitschriftenbeitrag ist veröffentlicht in der Zeitschrift „Learning and Instruction“.

© 2014 Elsevier Ltd. All rights reserved. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <http://creativecommons.org/licenses/by-nc-nd/4.0/>
<http://dx.doi.org/10.1016/j.learninstruc.2014.01.007>

Er ist zu zitieren als:

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction*, 32, 63–72. doi: 10.1016/j.learninstruc.2014.01.007

4. Teilstudie 1: Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality

Abstract

In this study, we investigated effects of socioeconomic and language minority classroom composition on students' reading achievement and explored the mediating role of central features of instructional quality, namely focus on language, student-oriented climate, and structured classroom management. Analyses were based on data collected from 352 German ninth-grade classrooms across two measurement points (t_1 , t_2) in a multilevel framework. We found socioeconomic composition to be related to individual reading achievement at t_2 after accounting for the corresponding baseline assessment at t_1 . However, the proportion of German language learners had no additional effect on reading achievement. Our results also suggest that the effect of the socioeconomic composition on achievement may be mediated partially by the teacher's focus on language during instruction. We conclude that more attention has to be paid to providing equal opportunities to all students in language classes.

Keywords: compositional effects, socioeconomic status, language minority students, instructional quality, mediation

4.1 Introduction

Being able to read and understand texts is paramount to academic achievement and essential for preparing children and adolescents for their educational and occupational future. Findings from various international large-scale assessment studies indicate, however, that students from families with low socioeconomic status (SES) and students with a language minority background reach lower levels of reading achievement (e.g., OECD, 2010). Currently, many education systems around the world are attempting to find the best way to serve their heterogeneous bodies of students, as schools and classrooms face high concentrations of low SES and language minority students. Although the classroom composition can affect learning (e.g., Van Ewijk & Slegers, 2010a, 2010b), little is known about the factors underlying such compositional effects.

A crucial factor might be instructional quality. Although the acquisition of reading comprehension skills also occurs outside the school, it is a major focus of classroom instruction. Instruction is influenced by the curriculum prescribed by a state or school district, by programs at individual schools, by the characteristics of the individual teachers as well as by the characteristics of the students in the class. Literature suggests that instruction depends on classroom characteristics (Helmke, 2010) and disadvantaged students attend classes where instructional quality is lower, yet this discussion is mostly based on theoretical arguments and qualitative research (see Hattie, 2002). Little is known about the effects of instructional quality on the development of reading comprehension in classrooms with varying compositions of SES and language minority students. Such knowledge, however, is important to understanding possible mechanisms causing classroom composition to affect learning.

In the present study, we aim to shed light on the relationship among classroom composition, instructional quality, and reading achievement in the course of one school year. We present a first attempt to systematically investigate the assumption of the mediating role of instructional quality for classroom composition effects in a large sample controlling for relevant background characteristics. In the literature review, we first describe the theories and present findings suggesting that characteristics of students in a class may be relevant to learning (Section 4.1.1). These aspects lead to our research question on the relationship between classroom composition characteristics and reading achievement. Second, we explore the role of instructional quality for achievement development in classrooms that differ in terms of their average SES and proportion of

language minority students (Section 4.1.2), referring to our research questions on the mediating role of instructional quality for classroom composition effects.

4.1.1 Classroom composition and achievement

Empirical findings indicate that students' learning outcomes are related to the composition of their learning environment. They show that students tend to reach higher levels of achievement when they are learning in schools or classrooms with a high average SES even after the average prior achievement level of the school or classroom is controlled (see meta-analysis by Van Ewijk & Slegers, 2010a).

Research on how achievement is affected by the proportion of language minority students in a class has provided less conclusive results. A number of studies found that the proportion of language minority students has no additional relevance for achievement after controlling for the average SES of students in a class and prior achievement level (e.g., Stanat, Schwippert, et al., 2010). In contrast, some studies report that a higher proportion of language minority students in a school or class may even lead to higher achievement controlling for SES (e.g., Benner & Crosnoe, 2011) or controlling for SES and prior achievement (e.g., Peetsma et al., 2006), as students in such learning environments encounter and have to work through contradictions and discrepancies in everyday life, which helps them to expand their intellectual capacities (Benner & Crosnoe, 2011).

Although effects of school and classroom composition on achievement are - at least for SES and mean prior achievement - quite well documented in different countries, few studies have investigated empirically why students perform differently in classrooms with varying compositions of SES and language minority students (Van Ewijk & Slegers, 2010a). Several interrelated factors have been suggested to make the SES or language minority composition of the learning environment relevant to learning: (1) peer processes, (2) school resources, and (3) instructional quality (see Thrupp, 1999; Thrupp et al., 2002; for a methodological perspective by Harker & Tymms, 2004, and Hauser, 1970, see Section 4.5.1).

First, peers with a high SES may positively affect a student's individual achievement because these students may share values, beliefs, and behaviors associated with learning and achievement (e.g., positive attitudes towards school, optimism, and regular school attendance). In classes with a large number of such students, a positive

learning culture fostering motivation to learn as well as achievement may be likely to emerge (see Agirdag et al., 2012; Goldsmith, 2011; Palardy, 2013). Likewise, students may be better able to support each other by explaining learning materials. In contrast, a high concentration of language minority peers is assumed to have negative effects on a student's language-related achievement because there are fewer opportunities to learn the language of instruction from interacting with classmates (see Peetsma et al., 2006; Van Ewijk & Slegers, 2010b).

Second, students with a high SES whose native language is the language of instruction often attend schools with better resources and favorable organizational and structural features such as, class size, teacher qualifications, and programs that encourage learning (Hattie, 2002; Rumberger & Palardy, 2005; Wilkinson et al., 2000). Parents with a high SES whose native language is the language of instruction may deliberately choose schools which are thought to offer better learning conditions due to the resources they offer.

Third, classes with many high SES students whose native language is the language of instruction are often provided with better learning opportunities, for example, more challenging tasks and a more student-oriented climate (see Section 4.1.2.2). This hypothesis, suggesting that instructional quality serves as a possible mediator for the relationship between classroom composition and student achievement, is explored in the present paper. After describing the construct of instructional quality in the next section, we explain this argument in more detail.

4.1.2 The role of instruction for student achievement in classrooms with varying compositions of SES and language minority students

4.1.2.1 Instructional quality and its relationship with student achievement

It has been suggested that three basic dimensions of instructional quality (as opposed to instructional quantity) related to student learning are student-oriented climate, efficient classroom management, and cognitively challenging learning opportunities (Baumert et al., 2010; Klieme et al., 2006; Pianta & Hamre, 2009). These dimensions have been investigated mostly in mathematics lessons but are assumed to be relevant in other school subjects as well. This is most obviously the case for student-oriented climate and structured classroom management. Cognitively challenging learning opportunities should be more domain-specific in nature but also play an important role in most subjects.

For language instruction, such as reading instruction, an indicator of cognitively challenging learning opportunities is teachers' focus on students' language learning (see Klieme et al., 2008).

A focus on language during instruction refers to the targeted attention a teacher pays to language-related aspects like writing or speaking grammatically correct. A teacher who places a special focus on language during instruction is believed to set clear and high demands and to offer challenging learning opportunities (Klieme et al., 2008). Klieme et al. (2008) showed that a focus on language was associated with higher gains in reading comprehension from the beginning to the end of ninth grade. Following Seidel and Shavelson's meta-analysis (2007), focus on language can be classified as an indicator of domain-specific components of instruction as it specially refers to the content domain of language instruction. In classes with a student-oriented climate, teachers offer support for individual learners and focus on making assignments comprehensible to students. These characteristics are thought to influence students' interest and involvement during lessons and to promote learning (Weinert et al., 1989; see "goal-setting and orientation" in Seidel & Shavelson, 2007). Structured classroom management refers to teacher behaviors such as pointing out important aspects of a lesson or making sure that students pay attention throughout the entire lesson. Well-structured classroom management allows students to spend more time on a task and on working with the learning material (Kunter & Baumert, 2006b; Weinert et al., 1989; see "organization of learning" in Seidel & Shavelson, 2007).

There is some inconsistency in the relative importance of these dimensions of instructional quality across studies investigating the effects on student outcomes. For instance, Seidel and Shavelson's (2007) meta-analysis found the largest effects on learning outcomes for domain-specific components of teaching (that we deem represented by focus on language in the present study) and only very small effects for more distal components, such as the organization of learning. In contrast, in Wang et al.'s (1993) review, classroom management as well as student and teacher social interactions - which are part of a student-oriented climate - were among the five most influential categories. Despite this inconsistency, it can be assumed that focus on language, student-oriented climate, and structured classroom management are main features of instructional quality fostering student achievement in general independently of student characteristics.

4.1.2.2 Relationship among classroom composition, instruction, and student achievement

Instructional quality not only influences achievement but may also be related to characteristics of classroom composition, such as average SES of students, the proportion of language minority students or average prior achievement. More specifically, aspects of instructional quality may mediate the relationship between the average SES or the proportion of language minority students in a class and individual student achievement. Harker and Tymms (2004) describe this perspective as “the school’s provision of education is moulded by the nature of the student body in ways that influence the achievement of individual pupils within it” (p. 180). Prominent proponents of this perspective state that more attention should be paid to how instructional quality in the classroom influences differences in outcomes between classes (Hattie, 2002; Wilkinson et al., 2000).

Theoretical explanations why instructional quality may be lower in classrooms with a low average SES or a high proportion of language minority students independent of the prior achievement level in the classroom include a number of aspects that potentially altogether affect classroom instruction: First, teachers of socially disadvantaged and/or language minority students tend to have lower achievement expectations for those students (see Jussim & Harber, 2005; Rumberger & Palardy, 2005; Tenenbaum & Ruck, 2007). Lowered expectations shape teachers’ behavior, which in turn may cause lower student achievement. This behavior includes offering fewer challenging learning opportunities and teaching at a slower pace (see Babad, 1993; Gay, 2000). Teachers may also have fewer positive interactions with these students (e.g., praise, affirmation) associated with a less supportive student climate. This argument commonly stems from research in the US that assumes and finds, among others, stereotype-based behavior of European American teachers who treat Afro-American or Latino students differently than European American students with whom they share the same ethnic background (Casteel, 1998; for a meta-analysis see Tenenbaum & Ruck, 2007).

Second, students and teachers jointly shape the learning environment in which instructional processes take place, and student characteristics may therefore be important. For instance, students’ work ethic or disciplinary climate affects the possibilities of how instruction can be delivered (Barth et al., 2004). This argument is in line with assumptions that peers in low SES classrooms share fewer values, beliefs, and behaviors

associated with learning and achievement (Agirdag et al., 2012; Goldsmith, 2011; see Section 4.1.1).

Third, low SES and language minority students often attend lower tracks which adhere to less demanding curricula, and teachers in these tracks might be less qualified to deliver high quality instruction (Hattie, 2002; Rumberger & Palardy, 2005; Wilkinson et al., 2000).

The main question is whether the theoretically assumed mediation of compositional effects through instructional quality can be empirically confirmed. Thus far, only few studies have analyzed the relationship among SES or language minority classroom composition, instructional quality, and student achievement. Most studies have focused on higher level school composition. For instance, Raudenbush et al. (1998) explored social and ethnic inequality in access to resources for mathematics learning in eighth grade. They found that, at the school level, students whose parents had lower education or an ethnic background other than European American were concentrated in schools providing less access to favorable disciplinary climates, schools that offer algebra, teachers trained in mathematics, and classrooms that emphasize reasoning. However, the authors did not analyze the relationship with students' achievement and thus a possible mediation model in their study.

Peetsma et al. (2006) referred to a Dutch study conducted by Van der Veen (2002) which found evidence for the relationship between SES and language minority composition on the one hand and challenging instruction on the other hand. In schools with a high proportion of immigrant children or students with less educated parents, teachers focused more on basic skills and less on higher order skills, such as employing reading strategies. However, the authors did not provide information on the achievement composition of the schools.

A recent study in the US found an effect of the average school SES on students' high school graduation and college enrollment (Palardy, 2013). The effect was partially mediated by school practices that emphasize academics (i.e., academic climate, teacher quality and teaching climate, and social engagement), controlling for a large number of student and school background variables. However, classroom instruction and effects on students' achievement were not analyzed in this study.

At the classroom level, Vause, Dupriez, and Dumay (2010) found that students gained higher scores in a mathematics test at the end of grade six when they were in classrooms with a high average level of parental education and a large number of books.

This result remained even after controlling for prior achievement at individual level at the beginning of the school year. Aspects of instructional quality played a partially mediating role, especially classroom management and teacher-led course rhythm (i.e., a teacher who pays attention to each student's time-on-task). The compositional effect lost its strength when aspects of instructional quality were taken into account, indicating that these factors mediate effects of classroom composition related to SES. However, the authors did not control for the achievement composition of the classrooms.

Evidence for a possible mediating effect of instructional quality also stems from a study investigating the joint and net effects of classroom composition and class processes in eighth-grade mathematics classrooms in Belgium, England, the Netherlands, and the US (Dumay & Dupriez, 2007). Multilevel analyses revealed that different aspects of classroom composition - proportion of students who speak the test language at home, mean educational resources, parents' education, and students' educational expectations – were directly related to students' mathematics test scores. Additionally, these classroom characteristics jointly affected test scores together with classroom processes – classroom climate, school climate, perception of safety in school, curriculum covered, and emphasis on homework. The proportion of between-class variance associated with these joint effects varied between eleven and twenty-three percent in the four countries. These results suggest that classroom composition effects may be partly indirect, thus supporting the mediation hypothesis. However, a mediation model was not directly tested and as students' prior achievement was not taken into account, it remains unclear if these effects are mainly due to differences in achievement composition of the classrooms under study.

In sum, the available research evidence suggests that the SES composition of a classroom or school affects achievement independently of students' prior achievement, whereas the effects of the proportion of language minority students are less clear. There is a number of theoretical arguments why both SES and minority student composition may together be relevant for individual achievement and classrooms with a low average SES and high proportions minority students constitute overall disadvantageous learning environments. A few studies have been conducted on the relationship among classroom composition characteristics, aspects of instructional quality, and achievement. Their results suggest that these relationships may be in line with a mediation hypothesis. However, none of these investigations has explored the relative effects of SES and

language minority classroom composition controlling for prior achievement and the extent to which these effects are mediated by aspects of instructional quality.

4.2 The present study

The present study explored classroom composition effects with respect to average SES and proportion of language minority students on students' reading achievement. The following two research questions were addressed:

1. Are the SES and language minority classroom composition related to students' reading achievement when relevant background characteristics are controlled? We expected students in classes with a high average SES to show higher achievement than students in classes with a low average SES (Hypothesis 1a). Moreover, in line with other studies on effects of the overall proportion of language minority students, we expected the proportion of language minority students to have no additional independent effect on reading achievement beyond the effect of the SES composition (Hypothesis 1b).
2. Are classroom composition effects mediated by characteristics of instructional quality? We hypothesized that the expected positive relationship between average SES and students' reading achievement could be explained by characteristics of instructional quality. In particular, we expected that instruction in classes with a high average SES would have a stronger focus on language, a more student-oriented climate, and a more effectively structured classroom management (see Section 4.1.2). These features, in turn, should be positively related to reading achievement. We thus hypothesized that focus on language during instruction (Hypothesis 2a), student-oriented climate (Hypothesis 2b), and structured classroom management (Hypothesis 2c) would partially mediate the expected SES composition effect.

4.3 Method

4.3.1 Participants and procedure

Our analyses were based on data from the Assessment of Student Achievements in German and English as a Foreign Language (DESI) study which investigated the

development of language competencies of ninth-graders in a large, nationally representative sample of 220 randomly selected schools in Germany (DESI-Konsortium, 2008). In each school, two classes were randomly selected and all students in these classes participated in language assessments at two measurement points: at the beginning and at the end of school year 2003-2004. The DESI study was carried out by a consortium of researchers headed by the German Institute for International Educational Research (DIPF).

The sample of the present study included students in the eleven western states of Germany (8047 students in 352 classrooms and 181 schools; average number of students per classroom $M = 23$). The analyses were limited to the western states due to the very small proportion of language minority students in schools in eastern states (see Federal Statistical Office Germany, 2012). The mean student age was $M = 14.86$ years ($SD = 0.74$) at the first measurement point and 52% of the students were female. Of the students, 22% indicated that their native language was not German or another language in addition to German. We further refer to this group as “German language learners” [GLLs] (classroom level distribution: $M = 23.70\%$, $SD = 18.02\%$, Min. = 0%, Max. = 88.20%). The average classroom SES measured by the HISEI with a theoretical range from 16 to 90 (see Section 4.3.2) was $M = 50.16$ ($SD = 8.14$, Min. = 31.41, Max. = 70.19). The correlation between average SES and proportion of GLLs was $r = -.50$, indicating that these composition characteristics are related to each other to some degree but far from being completely collinear.

In the German education system, students are assigned to different school tracks after completing primary education. In 14 German states, secondary school starts after four years, in two states after six years of primary education (KMK, 2012). “Hauptschule” is the lower track, also referred to as vocational track, providing a basic general education and designed to prepare students for occupations that require apprenticeships. “Realschule” is the intermediate track, providing students with a more extensive general education. “Gymnasium” is the academic track, leading to the degree necessary for university entrance. “Gesamtschule” basically combines all three tracks in a comprehensive school type (see KMK, 2013a). The different school tracks differ in their students’ achievement level at secondary school entry, but also in their curricula and organizational structures. Therefore, school tracks must be controlled for in analyses that aim at estimating net effects of classroom composition and instructional processes.

Of the classes, 38% were located in a vocational track school, 29% in an intermediate track school, 25% in an academic track school, and 8% in a comprehensive school. There were only 27 classes of comprehensive schools in the sample, and comprehensive and lower track schools have no strict enrollment requirements, leading to student populations rather similar to each other. Therefore, we created the compound category “schools with a vocational track” from these two school types. Additional analyses differentiating between these two school types revealed that the effects are comparable to the results in the present study.

4.3.2 Measures

At the end of ninth grade, students evaluated the instructional quality of their German lessons on a questionnaire. These student reports reflect instructional quality in a relative sense from the point of view of students in a classroom. The questionnaire included 12 aspects composed of several items which were rated on a 4-point Likert scale. The 12 aspects represented three features of instructional quality, namely focus on language, student-oriented climate, and structured classroom management. Table 4.1 gives an overview of these aspects. The majority of the aspects were judged by all students in each class, however for reasons of practicality 4 of the 12 aspects were randomly assigned to approximately 50% of the students in each class. The ICC2 indicates the reliability of an aggregated construct at the classroom level. It is estimated by applying the Spearman-Brown formula to the intraclass correlation (ICC) measuring the agreement of student ratings in a classroom (Marsh et al., 2012). The ICC2 values of the aspects indicate acceptable levels of reliability (Lüdtke et al., 2009).

Students’ reading achievement in German was assessed at two measurement points with a standardized test designed by the DESI research group to cover the grade 9 curriculum of language instruction. The complete test at t_1 and t_2 comprised the same four literary and four factual texts and a total of 41 items. The items were presented mainly as multiple choice and short answer questions. At both time points, each student received a sample of tasks in one out of four booklets (matrix design). The booklets were rotated so that no student received a task more than once. The items covered various competencies, such as identifying details, making inferences, using prior knowledge, or creating mental models (Willenberg, 2007). A generalized Rasch model was used to estimate student achievement scores on a common scale. Because of the booklet rotation design used in

the assessment, scores were generated using the plausible values (PV) technique (Adams, Wu, & Carstensen, 2007). Five PVs were generated for each student at each measurement point, scaled to have a mean score of 500 and a standard deviation of 100 in the German student population at t_2 (current sample distribution t_1 : $M = 502.90$, $SD = 88.35$, t_2 : $M = 504.21$, $SD = 95.42$). EAP/PV reliabilities were .85 (t_1) and .86 (t_2).

The individual student characteristics included in the analyses are language background, SES, gender, and prior achievement in reading (reading score at t_1). Students reported their native language on the questionnaire. We compared students with only German native language to “German language learners” [GLLs], that is, students who indicated that their native language was not German or another language in addition to German (dummy coding; German = 0, GLL = 1). We assessed the SES of a student’s family by the “International Socio-Economic Index of Occupational Status” (HISEI; Ganzeboom, De Graaf, Treiman, & De Leeuw, 1992). The HISEI represents the highest occupational status of the student’s mother or father, based on a classification of parents’ occupation which the parents reported on a questionnaire. Additional information on parents’ occupation was also reported by the students on their questionnaire. We used student information only when parents’ information was missing (see Section 4.3.3.1). We assessed student gender in the tracking form. Information was available for all students (dummy coding; male = 0, female = 1).

Classroom level characteristics included composition characteristics of the classroom and school track. The individual level student characteristics were aggregated to form proportion of GLLs, average SES, proportion of females, and average reading score at t_1 . For school tracks, we created two dummy variables “academic track” and “school with a vocational track”, with “intermediate track” as the reference category.

Table 4.1 Overview of Aspects and Features of Instructional Quality: Number of Items, Reliabilities and Item Wording

Aspect	N (items)	Cron- bach’s alpha	ICC2	Item wording example
<i>Feature: focus on language</i>				
Importance of language-related competencies ^a	10	.88	.67	How important is it for your teacher that...

				...you can write grammatically correctly?
Importance of oral competencies ^a	3	.67	.67	...your pronunciation is clear and correct?
<i>Feature: student-oriented climate</i>				
Support	3	.86	.83	If I need help, I get it from my German teacher.
Student orientation	6	.88	.85	If somebody has a good idea, my German teacher acknowledges it.
Focus on student interests	2	.80	.84	My German teacher makes the lessons interesting to me.
Comprehensibility	3	.78	.83	The assignments in the German classes are clear and comprehensible to me.
Classroom climate during Instruction	3	.85	.79	I get along well with my German teacher.
Positive culture of making mistakes ^a	14	.89	.83	I learn something new when my German teacher corrects my mistakes.
<i>Feature: structured classroom management</i>				
Organization of instruction	3	.79	.79	My German teacher points out the important aspects of a lesson.
Instrumental motivation	3	.84	.79	My German teacher emphasizes that good German skills are important for many professions.
Classroom management	2	.78	.86	My German teacher ensures the students pay attention throughout the entire lesson.
Organization of group work ^a	3	.74	.84	Group work results are shared with the class at the end of the German lesson.

Note. ICC2 = reliability of classroom mean ratings.

^a These aspects were randomly assigned to approximately 50% of the students in a class, all other aspects were assigned to each student in a class.

4.3.3 Data analysis

We used structural equation modeling to analyze the data (see Bovaird, 2007). Four random intercept multilevel structural equation models were estimated to explore (1) the relationship between the classroom composition characteristics and students' reading achievement after controlling for prior achievement and the other covariates, and (2) the mediating role of instructional quality for compositional effects on individual achievement. The software Mplus (Version 6.1; Muthén & Muthén, 1998-2010) was used for all analyses. The analyses were conducted with sample weights at both levels adjusting for different selection probabilities of students and classrooms to obtain representative results (see analysis strategy described by Nagengast & Marsh, 2012).

Metric background variables at the student level (SES, reading score) were standardized, which implies centering at their grand mean. Categorical variables (GLL, female) were neither centered nor standardized. Consequently, the estimates of background variables aggregated at the classroom level are interpretable as compositional effects (Raudenbush & Bryk, 2002). The proportion variables at the classroom level (GLLs, females) were divided by 10 to indicate steps of 10%. For example, an estimated regression coefficient of 0.5 for the proportion of GLLs in the class would imply that a 10% increase in GLLs is associated with an increase in the dependent variable by 0.5 standard deviations of the student distribution. With an average of about 23 students per class, 10% corresponded to approximately two students.

The three features of instructional quality were analyzed as latent variables at the classroom level with their aspects as indicator variables (see Table 4.1). For instance, the latent variable "focus on language" is defined by the two indicators "importance of language-related competencies" and "importance of oral competencies". We used a doubly latent approach with cross-level measurement invariance (e.g., Lüdtke et al., 2011; Marsh et al., 2012). This approach comprises latent measurement models at both levels and latent aggregation for the classroom level constructs. An important advantage of a model with these features is that it corrects possible measurement and sampling errors associated with designs in which variables measured at the individual level are used to operationalize a construct at the classroom level. Our focus was on classroom level effects of instructional quality as they represent the shared perception of instructional characteristics of students in each classroom (e.g., Lüdtke et al., 2009). However, it is important to model additional relationships at the student level to control for relationships between individual background variables and students' perception of instructional quality.

Confirmatory factor analyses showed a satisfactory fit of the measurement models for the features of instructional quality (student-oriented climate: $\chi^2 = 166.484$, $df = 23$, $p < .05$, RMSEA = .028, CFI = .988, SRMR_{within} = .017, SRMR_{between} = .022; structured classroom management: $\chi^2 = 43.751$, $df = 7$, $p < .05$, RMSEA = .026, CFI = .985, SRMR_{within} = .017, SRMR_{between} = .048; focus on language: the model is saturated so that $df = 0$ and the fit is perfect). A model with three factors of instructional quality demonstrated a slightly better fit than a model with only one overall factor (*3 factor model*: $\chi^2 = 928.642$, $df = 112$, $p < .05$, RMSEA = .031, CFI = .961, SRMR_{within} = .033, SRMR_{between} = .077; *1 factor model*: $\chi^2 = 1405.208$, $df = 119$, $p < .05$, RMSEA = .037, CFI = .939, SRMR_{within} = .041, SRMR_{between} = .102; model fit comparison⁸: delta CFI = .022, see Cheung & Rensvold, 2002). Therefore, it seemed justified to separately analyze the three different features of instructional quality. See Figure 4.1 for a schematic depiction of the mediation model using the feature focus on language as an example.

⁸ Due to our large sample size the χ^2 statistic is significant in each model. The χ^2 -difference test is also affected by sample size. We used delta CFI to compare the models because it is less affected by sample size.

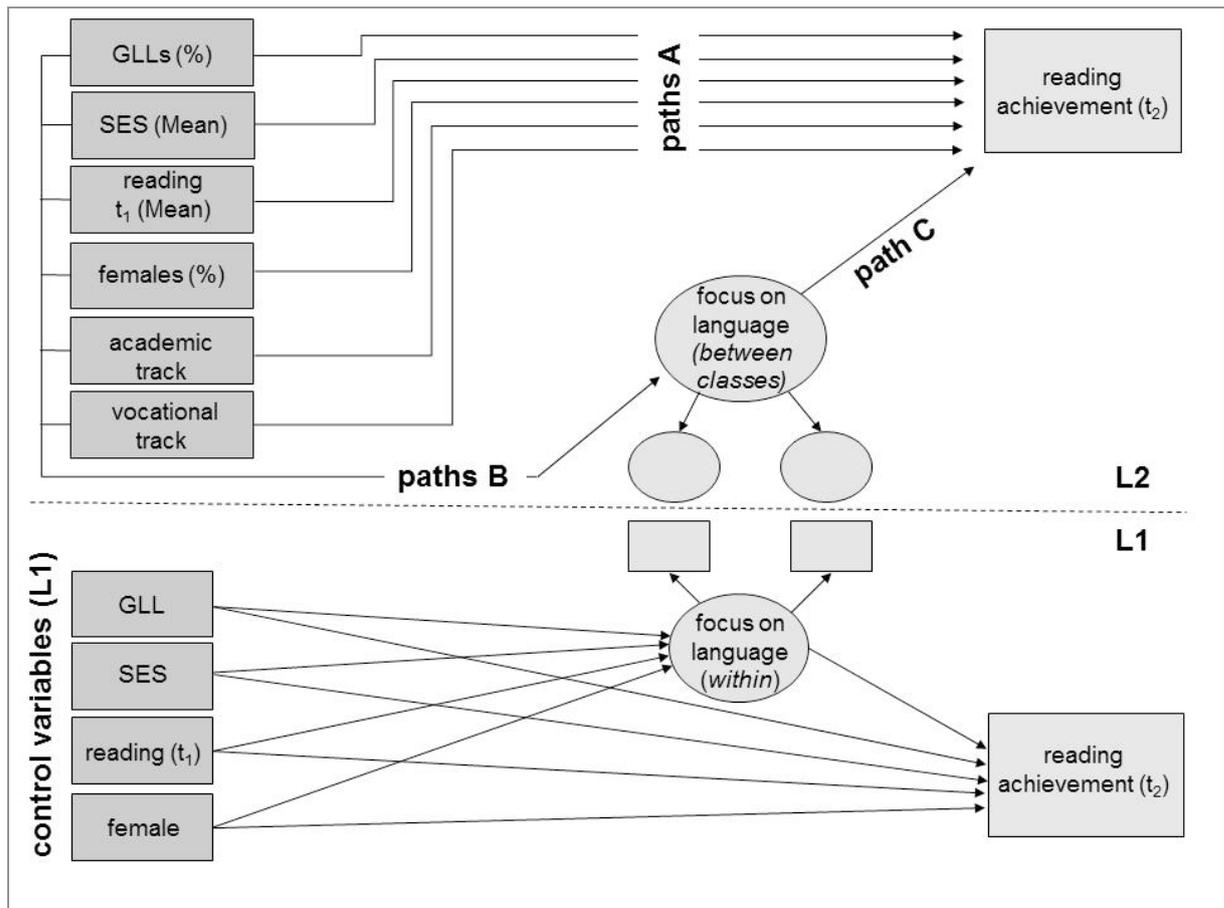


Figure 4.1: Schematic depiction of the mediation model using the feature focus on language as an example.

Note. GLLs = German language learners, SES = socioeconomic status, L2 = classroom level, L1 = student level. Paths A and B = compositional effects, Path C = classroom level (i.e., between) effect.

4.3.3.1 Treatment of missing data

We adopted several measures to deal with missing values. Plausible values for reading achievement were generated for all students, including those with missing data on test item information. The SES scores were based on information in the parent questionnaire which 62% completed. The DESI study authors replaced missing SES scores by estimated values using multiple imputation (e.g., Rubin, 1987) on the basis of additional information from the student questionnaire. For 92% of the students, SES scores could be generated. In our reanalysis, we used these imputed scores. Students without any information on their SES and/or GLL status (9%) were excluded from the analyses. In doing so, we used 91% of the original sample. To deal with item nonresponse on the instructional quality variables, which were not imputed in the original data set, we used the full information maximum likelihood (FIML) estimator implemented in Mplus.

This estimator applies a model-based approach to missing data (see Enders, 2010), using all information available from the model variables to estimate the model parameters.

4.4 Results

4.4.1 Relationship between classroom composition and reading achievement

In the first part of our analyses, we investigated whether the SES and language minority classroom composition were related to students' reading achievement (see model 1 in Table 4.2; a model which included only paths A of Fig. 4.1). For results of the student level variables and fit indices see the Appendix A.1. The results indicated that the SES classroom composition was significantly related to individual reading achievement even after controlling for prior achievement level in reading and the general achievement level associated with school tracks.⁹ Students in classrooms with a one standard deviation higher average SES showed half a standard deviation better reading achievement.

No significant compositional effect of the proportion of GLLs in a class on reading achievement emerged. Analyses not presented in Table 2 revealed a negative compositional effect of the proportion of GLLs on reading achievement when included as the sole predictor at the classroom level ($\beta = -.10, p < .001$). This effect, however, disappeared when the SES composition was also taken into account. Thus, the proportion of language minority students had no additional independent effect on reading achievement beyond the effects of the SES composition.

4.4.2 Mediation of classroom composition effects by instructional quality

In the second part of our analyses, we investigated whether classroom composition effects are mediated by characteristics of instructional quality (see models 2 - 4 in Table 4.2). The pattern of compositional effects on reading achievement in models 2-4 was comparable to that of model 1. Again, the SES composition was the most powerful predictor of individual achievement (paths A in Fig. 4.1).

In a next step, we focused on the relationship between classroom composition and instructional quality (paths B in Fig. 4.1). The results indicated that the three aspects of instructional quality were associated with classroom composition characteristics. In classes with a higher average SES, students reported a strong focus on language learning.

⁹ The predictors at the student and classroom levels were not highly interrelated, thus multicollinearity was not a concern (Variance inflation factor [VIF] for SES: level 1 = 1.16, level 2 = 4.78, $\beta_{\text{SES mean on GLLs\%}} = -.184$; VIF for GLL: level 1 = 1.10, level 2 = 1.53).

In classes with a higher proportion of GLLs, students also reported a stronger focus on language, and—in addition—a more student-oriented climate and structured classroom management. However, these effect sizes were small.

In a next step, we focused on the relationship between instructional quality and reading achievement (path C in Fig. 4.1). The results of model 2 indicated a positive effect of the class's overall perception of focus on language during instruction on the individuals' achievement. Students in classes with a stronger focus on language showed higher reading achievement after one school year. However, the results of model 3, which included student-oriented climate, and those in model 4, which included structured classroom management, indicated that these two aspects of instructional quality were not related to reading achievement after one school year. Thus, these aspects do not seem to mediate compositional effects on reading achievement.

The average SES of students in a class was associated with focus on language, which in turn was associated with achievement. This suggests that the SES composition effect could partially be mediated by focus on language. To test the mediation we computed the indirect effect and applied the delta method to determine the corresponding standard error (see Nagengast & Marsh, 2012; Raykov & Marcoulides, 2004). As indicated by the associated delta test, the indirect effect of average SES on reading achievement (not presented in Table 4.2) was, in fact, marginally significant ($\beta = .14$, $SE = .08$, $p = .086$). This suggests that a small part of the SES composition effect—that is, the result that students showed higher achievement in classes with a higher average SES after correction for individual SES and the other covariates—may be due to differences in the focus on language during instruction. In classes with a higher SES composition, there seemed to be a stronger focus on language, which in turn was associated with higher achievement. As the proportion of GLLs was not directly related to reading achievement (see Section 4.4.1) we did not investigate mediating effects for this composition characteristic.

Table 4.2. Results of Multilevel Structural Equation Models with the Reading Achievement at Measurement Point Two as Outcome Variable and Features of Instructional Quality as Mediating Variables (Classroom Level Results)

	Model 1 (without instruction)		Model 2 (focus on language)		Model 3 (student-oriented climate)		Model 4 (structured classroom management)	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
A) Reading achievement (t_2) on composition variables								
GLLs (%)	-.02	(.02)	-.04*	(.02)	-.03	(.02)	-.03	(.02)
SES (Mean)	.51***	(.12)	.42***	(.12)	.51***	(.12)	.51***	(.12)
Reading achievement t_1 (Mean)	-.04	(.09)	-.07	(.09)	-.04	(.09)	-.05	(.09)
Female (%)	.02	(.02)	.02	(.02)	.02	(.02)	.02	(.02)
Academic track	.08	(.10)	.11	(.10)	.08	(.10)	.08	(.10)
Vocational track	-.21*	(.09)	-.26**	(.09)	-.23*	(.09)	-.23*	(.09)
B) Instruction on composition variables								
GLLs (%)	-	.09**	(.03)	.06*	(.02)	.09**	(.03)	
SES (Mean)	-	.45**	(.17)	.16	(.16)	.12	(.15)	
Reading achievement t_1 (Mean)	-	.12	(.12)	.01	(.10)	.08	(.12)	
Female (%)	-	.00	(.02)	-.04	(.03)	-.04	(.03)	
Academic track	-	-.15	(.14)	-.03	(.14)	-.13	(.14)	
Vocational track	-	.24	(.13)	.36**	(.13)	.34*	(.15)	
C) Reading achievement (t_2) on instruction (Level 2)								
Instruction	-	.30*	(.12)	.08	(.06)	.12	(.08)	
R^2 (Level 2 reading achievement t_2)	.48		.51		.48		.48	
R^2 (Level 1 reading achievement t_2)	.60		.61		.61		.61	

Note. * $p < .05$, ** $p < .01$, *** $p < .001$, Level 1 = student level, Level 2 = classroom level, GLLs = German language learners, SES = socioeconomic status. Regression coefficients were standardized by the total variance of student level outcome variable (see Marsh et al., 2012).

4.5 Discussion

In this study, we aimed to shed light on the relationship among two aspects of classroom composition, instructional quality, and reading achievement. We investigated whether SES and language minority classroom composition are related to students' reading achievement. In line with Hypothesis 1a, we found that students in classes with a higher average SES reached higher achievement levels, replicating results of former studies (e.g., Van Ewijk & Slegers, 2010a). That is, independent of their individual background characteristics, students in classes with a higher average SES reached higher test scores after one school year. As predicted by Hypothesis 1b, the proportion of language minority students had no effect on students' reading achievement beyond class-average SES and prior achievement (see Stanat, Schwippert, et al., 2010).

Next, we explored whether the classroom composition effects are mediated by characteristics of instructional quality. Regarding Hypothesis 2a on the mediating role of focus on language during instruction, the findings suggest that in classrooms with a lower average SES, teachers tend to focus less on challenging language instruction (see also Dumay & Dupriez, 2007; Van der Veen 2002, cited by Peetsma et al., 2006). These students seem to receive instruction that fosters their reading development less than students in classes with a high average SES. After taking focus on language into account, the direct SES composition effect on achievement declined by about one fifth. Our large sample size enabled us to reveal this reduction of the direct SES composition effect when including focus on language during instruction. However, the indirect SES effect on achievement was rather small and only marginally significant. Even though the instructional effect and the explained variance were rather small, the findings can inform future research. In studies on teaching and learning, small effects are common, due to the large number of factors potentially influencing student achievement (Lanahan, McGrath, McLaughlin, Burian-Fitzgerald, & Salganik, 2005; Seidel & Shavelson, 2007). However, the only marginally significant mediating effect suggests that this characteristic of instructional quality operationalized in its present form was no powerful mediator of composition effects. It is up to future research to investigate whether a more specific conceptualization of a focus on language during instruction shows more decisive results.

Furthermore, we found no support for the hypothesis that student-oriented climate (Hypothesis 2b) or structured classroom management (Hypothesis 2c) mediates the SES composition effect. The low association between the two instructional features and reading achievement contradict some studies indicating a positive relationship between

these characteristics and achievement in mathematics (e.g., Kunter & Baumert, 2006b; see Wang et al., 1993). However, after taking differences between school tracks into account, Kunter and Baumert (2006b) did not find an effect of classroom management. In conclusion, findings about the effects of classroom management on achievement are inconclusive. Furthermore, our study focused on reading achievement, an outcome that is less strongly related to the curriculum as mathematics achievement, making strong effects less likely. Student-oriented climate may be more relevant for psychosocial outcomes, such as learning motivation (see Klieme et al, 2010).

4.5.1 Limitations and future research

The present study has some limitations which should be addressed in future research. First, instructional quality was based solely on student ratings which are valid indicators for some instructional features (De Jong & Westerhof, 2001; Kunter & Baumert, 2006a; Wagner et al., 2013). Student ratings are superior to those of teachers because they include a large number of raters and they provide more reliable information than external observations as students share a common experience in the classroom based on a large number of observations. Despite the advantages of student ratings, a common frame of reference across classrooms cannot be taken for granted. It is also possible that students in classrooms with varying compositions of SES and language minority students systematically differ in their expectations about what instruction should be like. As all data sources go along with specific strengths and weaknesses (Turner & Meyer, 2000), future studies should include classroom observations and external ratings (Praetorius, Lenske, & Helmke, 2012) to gain more comprehensive insights into classroom instruction and its mediating role for compositional effects.

Second, the students in our sample were already attending their respective classes for up to four years at measurement point one (grade 9), providing a long period of time for classroom composition to affect learning. Therefore, additional effects of classroom composition on reading achievement at the end of grade 9 were expected to be small. Moreover, reading comprehension skills improve only marginally during one school year in this age group, making large effects less likely. That is, our study may provide only a snapshot of processes that may have been more pronounced or even different in kind during the preceding school years. Future studies should explore composition and

mediating effects over a longer period and start for example with the beginning of secondary school when classes are composed.

Third, achievement development is not the only educational goal. Future studies should examine the effects of classroom composition and instructional quality on diverse goals of schooling in the cognitive, affective, and social domains (e.g., interest in reading or reading motivation) to be able to gain further insight into the relationship between composition, instruction, and educational outcomes.

Fourth, more data is needed to disentangle the relationships among school tracks, prior achievement, composition characteristics, and instructional quality. For this purpose, future studies should analyze student achievement in several classrooms that are taught by the same teacher but vary in their composition. Such a strategy could not be pursued in the present study because there were no such cases in our sample.

Finally, beyond the restrictions of our study, compositional analyses in general are criticized because models are allegedly inadequately specified (Harker & Tymms, 2004; Hauser, 1970). More precisely, these authors argue that spurious effects arise if level-1 variables were assessed with low reliability and validity or covariates at level 1 were omitted. However, our variables analyzed at level 1 constitute reliable and valid measures related to students' reading achievement that are theoretically sound and commonly used in international large-scale assessment studies (OECD, 2010). In addition, Thrupp et al. (2002) concluded from their review that compositional effects exist (see also Van Ewijk & Slegers, 2010a), even though one can never be completely sure about an omitted variable bias.

4.5.2 Conclusion

The findings of the present study confirm once again that it is not the proportion of language minority students but rather the SES classroom composition that matters for the academic achievement of students. This may suggest that SES desegregation could be advantageous for students' achievement development in low SES areas. However, it may be possible to counteract segregation effects if factors that reduce compositional effects (i.e., instructional quality) are identified. Our findings suggest that focus on challenging language instruction might affect compositional effects on reading achievement. Thus, characteristics of instructional quality that are closely linked to language achievement seem to be more relevant for reducing compositional effects on reading than

characteristics that are more loosely connected to language achievement (i.e., classroom management). In our analyses we were able to gain insight into these mechanisms based on a large, representative sample of complete classrooms investigated at two measurement points. The analyses represent a first systematic investigation of the assumption that students in classrooms with a high number of disadvantaged students attain poorer achievement because they receive instruction of lower quality.

In particular, we investigated effects in language instruction, a domain that is less researched than mathematics and science instruction. The findings suggest that more attention has to be paid to providing equal opportunities to all students in classrooms with varying SES compositions in language classes. Yet, it is necessary to gain more insights into which specific activities are most relevant to foster cognitively challenging learning opportunities—represented in this study by teachers' focus on language in language instruction. The present study suggests that, from a student perspective, teachers in classrooms with varying SES compositions may differ in the emphasis of correct language use during language-related activities, such as writing or speaking, in the standards the teacher sets in instruction, and the level of challenging instruction. However, these results do not provide information on teachers' activities in their classroom. Therefore, more research is needed about how these general characteristics translate into action, that is, for example, frequent instruction of grammar, ongoing student feedback and correction of language mistakes, or explicit explanation of the importance of correct language use in education.

In conclusion, our findings can add to the body of knowledge on compositional effects and in particular to the research on the mediating role of instructional quality. These are necessary steps to gain insight into how instruction can be organized to provide equal opportunity to all students and thereby allow common educational goals to be met.

Teilstudie 2:

Classroom composition and language minority students' motivation in language lessons

Der nachfolgende Zeitschriftenbeitrag ist veröffentlicht in der Zeitschrift „Journal of Educational Psychology“.

© 2015 American Psychological Association

<http://dx.doi.org/10.1037/edu0000035>

"This article may not exactly replicate the authoritative document published in the APA journal. It is not the copy of record."

Er ist zu zitieren als:

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., & Stanat, P. (2015). Classroom Composition and Language Minority Students' Motivation in Language Lessons. *Journal of Educational Psychology*. 107(4), 1171-1185. doi: 10.1037/edu0000035

5. Teilstudie 2: Classroom composition and language minority students' motivation in language lessons

Abstract

In this study, we investigated effects of the proportion of language minority students in classrooms on the development of students' intrinsic motivation in language lessons and the mediating role of instructional climate (e.g., teacher support, focus on student interests). In addition, we explored the interaction between the proportion of language minority students and individual language background affecting students' motivation. Analyses were based on data collected in 352 German 9th-grade classrooms across 2 measurement points in a multilevel framework. We found that students in classrooms with a higher proportion of language minority students were more intrinsically motivated to learn. However, we found no statistical evidence for the prediction that the compositional effect should be mediated by the instructional climate in the classroom. Cross-level interaction analyses revealed that only language minority students were increasingly motivated in classrooms with a higher proportion of language minority students. Our findings add to the growing evidence indicating that the nature and quality of the social environment are important determinants of school-related motivation.

Keywords: classroom composition, motivation, language minority students, instructional climate, multilevel analysis

5.1 Introduction

International large-scale assessment studies commonly report an achievement gap between language minority students, whose native language is not the language of instruction, and language majority students (e.g., OECD, 2010). These students—as well as students from families with low socioeconomic status (SES)—tend to be disadvantaged in terms of educational success. In addition, they are more likely to attend schools and classrooms with higher concentrations of disadvantaged students. High proportions of these students in a school or classroom are reported to negatively affect the achievement development (Van Ewijk & Slegers, 2010a, 2010b). Yet, less is known about composition effects on school-related motivation.

Motivation is one of the main forces driving learning. We understand motivation to learn as malleable construct as opposed to a rather stable trait that some students have and others do not (see Van der Werf et al., 2008). Students reporting a higher level of motivation to learn are more likely to engage in learning activities, to reach higher achievement levels, to keep track in school, and to show greater life satisfaction as well as mental health (OECD, 2004; Roeser et al., 2000). Thus, the development of motivation is an important educational goal of schooling in addition to achievement development (Van der Werf et al., 2008).

Motivation of individuals develops in relation to others in their social context (e.g., Eccles et al., 1993; Eccles & Roeser, 2011; Urdan & Schoenfelder, 2006). Therefore, institutional, instructional and interpersonal experiences in every day school life may foster or hinder increased motivation to learn. Context characteristics that meet students' need for competence, autonomy, and relatedness are prone to foster intrinsic motivation (Niemi & Ryan, 2009, see Eccles & Roeser, 2011). Two main structure and process characteristics of the context are assumed to be the composition of the student body and the instructional climate in a classroom—besides institutional characteristics such as school size (see Lee, Smerdon, Alfeld-Liro, & Brown, 2000) or resources and school tracks. Classroom composition is assumed to affect instructional climate (see Hattie, 2002) which, in turn, has been shown to influence students' motivation (Urden & Schoenfelder, 2006). In addition, classroom composition may be related to students' feeling of belonging to their peers (see Goodenow, 1993) which should affect students' motivation differentially depending on their individual background characteristics. Yet, little is known about the relationship between language minority student classroom

composition on the one hand and instructional climate and student motivation on the other (see Den Brok & Levy, 2005).

The aim of the present study is to shed light on (1) the effect of the proportion of language minority students in a classroom on the development of students' motivation and (2) the mediating role of instructional climate. Furthermore, we explore if (3) the proportion of language minority students has different effects on motivation for minority students than for majority students. In our analyses of data from a short-term longitudinal study, we focus on intrinsic motivation in German lessons of secondary school students in Germany. In the following literature review, we first present theories and findings on minority students' school-related motivation. Second, we explore the role of language minority classroom composition and instructional climate for motivational development.

5.1.1 Minority Students' Motivation to Learn

Minority students often show poorer school achievement than majority students (e.g., OECD, 2010). From the perspective of self-determination theory, this may be related to reduced future motivation to learn and lower interest (Deci & Ryan, 2000) with reciprocal relationships between motivation and achievement over time (e.g., Marsh et al., 2005). In addition to an assumed reduced future motivation of students due to their poorer school achievement in the past, theories that address minority students in particular suggest that they differ from majority students in their level of motivation to learn. More specifically, on the one hand Ogbu's cultural ecological theory suggested that some minority students may be especially amotivated (Ogbu, 1987, 2004; see Conchas, 2001). Originally addressing motivation of Afro-American students in the US, the theory describes minority students as being assigned a subordinate status, being devaluated in school and feeling vulnerable. As a consequence, these students become alienated from school, reject educational values, and are amotivated to learn. A recent review on schooling, cultural diversity, and student motivation also reports decreased motivation of some minority students due to blocked opportunities and limited options (Kumar & Maehr, 2010).

On the other hand, the immigrant optimism hypothesis suggests that some minority students' level of motivation to learn may be higher than that of majority students (Kao & Tienda, 1995; see Barnett, Sonnert, & Sadler, 2012; Siahhan, Lee, & Kalist, 2014). Ethnic minority students whose parents or who themselves immigrated to

the receiving country are expected to show higher motivation than majority students. It is assumed that immigrants often aim at offering their children a better future and therefore have higher achievement goals which may translate into students' heightened motivation to learn. In the US, especially positive attitudes towards education can be found in particular for Asian American students (see Chen & Stevenson, 1995) but also for other groups (see Fuligni, 2001; Fuligni et al., 2005). Yet, the extent to which these assumptions can be generalized across minority groups and countries of residence is an open question.

In an analysis of the German sample of the OECD-led Programme for International Student Assessment (PISA), Walter and Taskinen (2008) found only weak evidence for the hypothesis of immigrant optimism for students' motivation in science classes. Students who immigrated with their parents to Germany showed only heightened motivation in families from the former Soviet Union whereas students whose parents were born in Turkey expressed lowered motivation compared to non-immigrant students. There were no significant differences between non-immigrant students and immigrant students from other countries. Stanat and Christensen (2006), in contrast, found higher levels of interest in and enjoyment of mathematics for immigrant students in the majority of OECD countries, including for instance the US and Germany (see also Stanat, Segeritz, et al., 2010).

In summary, most studies reviewed show either no difference between minority and majority students or slightly higher levels of motivation for minority students. Yet, most studies pertain to subjects such as mathematics or science. Motivation in language classes may be different due to the potential difficulties in language classes for students whose native language is not the language of instruction: Language classes focus on reading and production of larger text units while in mathematics or science classes language is a means to impart content knowledge in a specific domain.

5.1.2 School and Classroom Context

5.1.2.1 Language Minority Classroom Composition and Student Motivation

In schools and classrooms students develop their individual interests (see Renninger, 2000) as well as interact with each other and shape each other's attitudes and behavior (Eccles et al., 1993; Eccles & Roeser, 2011; Urda & Schoenfelder, 2006). One can assume that a multitude of interrelated processes among peers in a classroom occur

that potentially influence a student's level of motivation. Of the three characteristics that are prone to foster increased motivation—competence, autonomy, and relatedness—(Niemic & Ryan, 2009), especially the relatedness or belongingness to the peers may be associated with the composition of the student body: Emotional *belonging to the peers* plays a pivotal role in adolescence when the significance of peer relations and acceptance increases (see Goodenow, 1993; Kumar & Maehr, 2010; Wang & Eccles, 2013). It is especially important for language minority students as they must cope with ethnic, socioeconomic, and sociolinguistic differences between their family environment and school majority environment (see Den Brok & Levy, 2005; Eccles & Roeser, 2011). A classroom in which students feel emotionally related to their peers is characterized by positive interactions among students, mutual encouragement and assistance, as well as the absence of exclusion from the peers (see Goodenow, 1993). Students in such classrooms are more likely to feel secure, more motivated to learn and to engage in learning activities (see Wang & Eccles, 2013).

One important characteristic of peer composition related to the feeling of belongingness is the proportion of minority students. Especially for minority students, minority versus majority group membership may act as a lens through which individuals in a culturally pluralistic society view each other (see Johnson et al., 2001; Kumar & Maehr, 2010). In their recent review Kumar and Maehr (2010) argue that minority students often feel rejected from peers of the majority group and therefore show less motivation to learn. By implication, a classroom with a high proportion of minority students should strengthen minority students' feeling of relatedness and facilitate their motivation to learn. These assumptions are in line with the so-called belongingness perspective postulating social relatedness as a fundamental need related to adjustment and well-being (see Baumeister & Leary, 1995; Deci & Ryan, 2000). In general, these assumptions mostly refer to the proportion of students with the same specific ethnic background as a source of belongingness, yet, they also apply to majority versus minority group membership in society: According to social identity theory (Tajfel & Turner, 1986) and similarity attraction paradigm (Byrne, 1971) group membership and feeling of belonging may be based on similarity between students. Minority students—even with varying ethnic backgrounds—may be more similar to each other than to majority students for instance in terms of multilingual language experiences and immigration history within the family. Empirical findings showed that minority students in primary school were more socio-emotionally adjusted in schools with more peers of the same ethnic

background (Benner & Crosnoe, 2011). Furthermore, an analysis of US Caucasian, African American, and Hispanic students in grade 8 distinguished between the broader categories of majority schools (over 50 % of Caucasian students) and minority schools (less than 50 % of Caucasian students) (Goldsmith, 2004). The findings suggested that African American and Hispanic students were more looking forward to instruction and believed that the topics they learnt will be useful in the future when they attended segregated minority schools.

In conclusion, the belongingness hypothesis suggests that minority students are more motivated in classrooms with high proportions of minority students and majority students to be more motivated in classrooms with high proportions of majority students. In addition to the belonging to the peers, the instructional climate stimulated by the teacher may be related to the classroom composition and to students' motivation. In a first step we now briefly review theories and present findings on characteristics of instructional climate that are prone to foster motivation. In a second step we discuss possible relationships between classroom composition, instructional climate, and students' motivation.

5.1.2.2 Instructional Climate and Student Motivation

According to the self-determination theory, an instructional climate stimulated by the teacher that fosters a feeling of competence, autonomy, and relatedness or belongingness can be expected to promote students' increased motivation to learn (Deci & Ryan, 2000; Niemiec & Ryan, 2009).

Instructional climate that meets students' need for *competence* is characterized by optimally challenging learning activities that allow students to test and expand their academic abilities. Teachers who foster students' feeling of competence provide their students with appropriate tools and feedback promoting success and feelings of efficacy, and they are clear about their expectations (Eccles & Roeser, 2011; Niemiec & Ryan, 2009; Urdan & Schoenfelder, 2006; Wang & Eccles, 2013). For instance, a multilevel analysis of secondary school students in the US found that, at the school level, students' mathematics-related motivational beliefs (i.e., self-concept and task value) and school engagement were higher when teachers provided clear expectations as well as consistent and contingent responses to student questions (Wang & Eccles, 2013).

Instructional climate that meets students' need for *autonomy* includes material that

relates to students' interests, provides students opportunities for participation, allows for student input, and minimizes the salience of evaluative pressure and sense of coercion (Niemiec & Ryan, 2009; Roeser et al., 2000; Urdan & Schoenfelder, 2006; Wang & Eccles, 2013). Again, Wang and Eccles (2013) found students in schools whose curriculum included meaningful topics that reflected students' personal goals and interests to show more positive motivational beliefs and higher levels of school engagement.

Instructional climate that meets students' need for *belongingness to the teacher and the classroom* is characterized by caring teachers who provide a secure and caring learning environment and emotional support and encouragement during the learning process, especially when difficulties are encountered (Eccles & Roeser, 2011; Goodenow, 1993; Opdenakker et al., 2012; Roeser et al., 2000). For instance, a multilevel analysis of secondary school students in Norway found that, at the class level, students' perceived support provided by the teachers was positively related to students' academic initiative (Danielsen et al., 2010).

5.1.2.3 Instructional Climate in Classrooms with Varying Proportions of Language Minority Students: Mediating Effects on Motivation

Research suggests that the instructional and, hence, the motivational climate stimulated by the teacher may vary as a function of the school or classroom composition (see Hattie, 2002; Wilkinson et al., 2000). It is often assumed that schools and classrooms with higher concentrations of disadvantaged students (i.e., low SES and minority students) are characterized by an instructional climate which hampers a positive motivational development and improvement of students' abilities. Underlying assumptions are based on theories and findings showing teachers' lower achievement expectations towards disadvantaged students (see Ready & Wright, 2011) that may result in offering fewer challenging learning opportunities and fewer positive interactions with these students (e.g., praise, affirmation) at the individual student level (for a meta-analysis and review see Den Brok & Levy, 2005; Tenenbaum & Ruck, 2007). These characteristics are associated with a less supportive student climate. Furthermore, disadvantaged students often attend schools with fewer resources and less favorable organizational and structural features (i.e., class size, programs that encourage learning), and their teachers tend to be less qualified to deliver high quality or motivating instruction

(Hattie, 2002; Wilkinson et al., 2000). This may lead to the assumption of a negative relationship between the proportion of language minority students in a classroom and the instructional climate in the class which in turn negatively affects students' level of motivation. However, this assumption of a less motivating climate in classrooms with high proportions of minority students lacks empirical evidence so far.

In contrary, research rather suggests that in classrooms with high proportions of language minority students, there may be an instructional climate fostering motivation. Teachers in such classrooms may become more aware of student backgrounds and their possible educational disadvantages, as some teachers tend to underestimate the achievement level and overestimate the need for additional help of minority students (see Ready & Wright, 2011; Stanat, Weirich, et al., 2012). As a consequence, they may try harder to tailor their instruction to students' needs and interests and act more supportive and caring resulting in an overall more motivating climate. One study providing partial evidence for this idea analyzed high school students' perceptions of interpersonal teacher behavior in 38 classes (Levy et al., 1997). The findings suggest that teachers behave differently in class depending on its cultural composition. In classrooms with a large ethnic heterogeneity (defined by place of birth and primary language spoken at home), teachers seemed to behave more cooperatively (acting friendly, assisting students). In those classrooms, teachers also seemed to structure and guide the learning process more instead of giving responsibility to students (for a review on teacher-student relationships in multicultural classes see Den Brok & Levy, 2005). However, the relationship between teacher behavior and student outcomes, such as motivation, was not analyzed. In a study on US ninth-graders, students perceived the school climate to be fairer and more directed towards academics and interracial understanding if they attended schools with a high proportion of peers with the same ethnic background (Benner et al., 2008). Students also reported to engage more in learning in schools with a high proportion of same ethnic background students. In addition, school diversity indirectly affected student-reported school engagement through perceived academic climate and school climate (belongingness, fairness, and interracial climate). Analyses of the German DESI study investigating ninth-graders' motivation in German lessons also showed effects pointing in this direction (Klieme et al., 2010). The study focused on the impact of school-level quality variables on motivational gain scores in a three-level design controlling for other composition variables and supportive teacher behavior at the classroom level. The analyses revealed positive relationships between the proportion of language minority

students in a classroom and gains in motivation. In addition, a supportive teacher behavior was positively related to gains in motivation. Yet, whether the effect of the proportion of minority students on gains in motivation is mediated by supportive teacher behavior was not explored.

In summary, literature suggests that students in classrooms with higher proportions of minority students may be more motivated to learn. A large proportion of minority students in the classroom may be associated with a more motivating instructional climate affecting all students in a classroom. In addition, students' increased motivation may be fostered by high levels of belonging to the peers. As minority students' feeling of belongingness to the peers should be stronger in classrooms with high proportions of minority students, high proportions of minority students should positively affect minority students' level of motivation compared to majority students' level of motivation.

5.2 The Present Study

The present study explored classroom composition effects with respect to language¹⁰ minority composition on students' motivation in language lessons in the German secondary school system. We examined effects of classroom and individual student background characteristics including their motivation at the beginning of the school year on students' motivation at the end of the school year. We additionally took into account the instructional climate during that school year, based on student reports at the end of the school year (see Figure 5.1). Our research questions were as follows:

- 1) (a) Do language minority students differ from language majority students in their motivation to learn in language lessons after one school year after controlling for individual and compositional background characteristics? (b) Is the proportion of language minority students in a classroom related to students' motivation to learn in language lessons after controlling for individual and compositional background characteristics?

¹⁰ Theories exploring minority students' motivation to learn focus especially on ethnic minority students defined by country of birth. The present paper focuses on students' language minority background because it is an important process characteristic of ethnic minority families which is expected to mediate the relationship between ethnic or immigrant background and school success (e.g., Baumert et al., 2003).

- 2) Is the possible classroom composition effect on motivation mediated by the instructional climate in the classroom?
- 3) Has the proportion of language minority students different effects on motivation for minority students than for majority students?

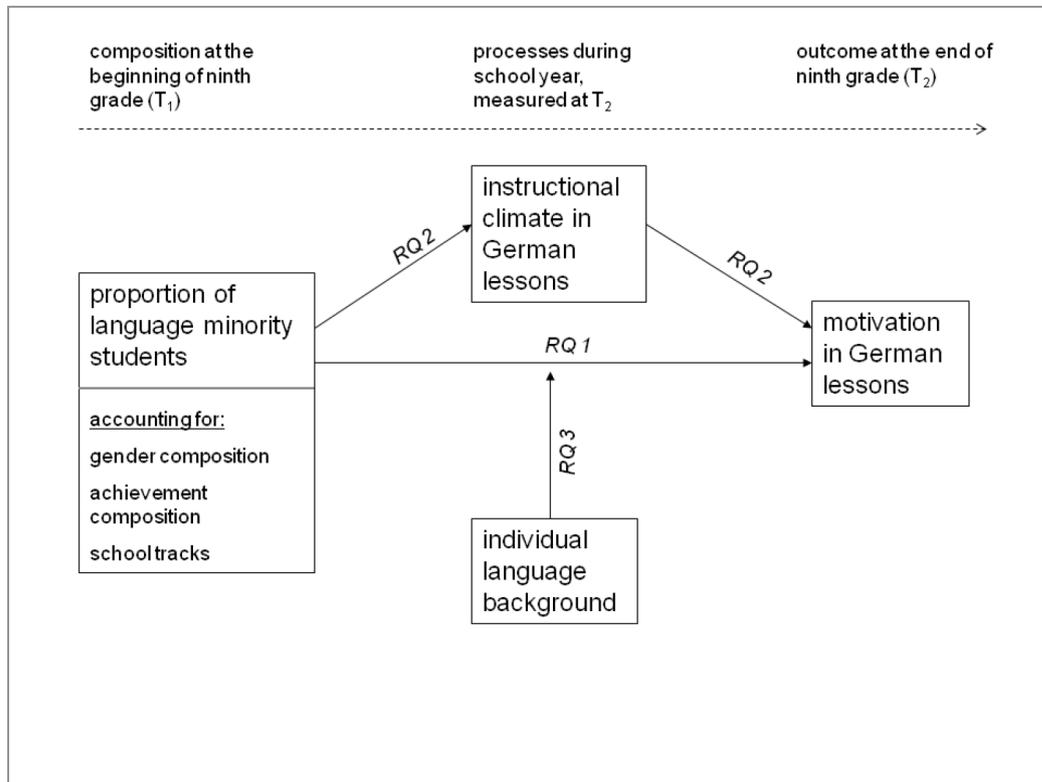


Figure 5.1: Hypothesized model of classroom composition effects on students' motivation including Research Questions (RQ) 1 to 3.

In line with results of former studies, we expected no difference in the level of motivation between individual minority and majority students or a slightly lower level of motivation of minority students (Research Question 1a). Research so far revealed either no difference or slightly higher levels of motivation in math and science learning for minority students. As our study investigates motivation in language learning—which might be more challenging for students whose native language is not the language of instruction—we rather expected no difference or a slightly lower level of motivation of minority students. Moreover, we expected the proportion of minority students to be positively related to students' level of motivation (Research Question 1b). This positive relationship may be mediated by the instructional climate in the classroom. In classrooms with high proportions of minority students, the level of motivating instructional climate may be higher, affecting all students' motivation in that classroom (Research Question 2).

However, in line with the belongingness hypothesis and to some extent competing with the hypotheses mentioned before, the proportion of language minority students may have different effects on the level of motivation for minority students than for majority students. According to this hypothesis, minority students should be increasingly motivated in classrooms with high proportions of minority students and majority students should be increasingly motivated in classrooms with high proportions of majority students (Research Question 3). Following the assumptions of the belongingness hypothesis related to a comparison between majority and minority background students, the proportion of minority students should be related to the average level of motivation at the classroom level in a u-shaped manner.

In all analyses we took relevant background characteristics at the individual student level and at the classroom level into account. At the student level we controlled for students' achievement at T_1 as lower achieving students tend to have lower motivational beliefs, with reciprocal relationships between motivational beliefs and achievement over time (e.g., Guay, Marsh, & Boivin, 2003; Marsh et al., 2005). Additional characteristics include students' gender and SES. Girls tend to be more motivated to learn than boys and students from families with a high SES tend to be more motivated than low SES students (e.g., OECD, 2010).

At the classroom level we controlled for the achievement composition in the classroom and school track which may be related to students' motivation through social comparisons among peers (for big-fish-little-pond effect see Marsh, 1987; Nagengast & Marsh, 2012; for motivational beliefs related to tracking see Chmielewski et al., 2013; Trautwein et al., 2006). Furthermore, we took the gender and SES composition into account.

5.3 Methods

5.3.1 Participants and Procedure

Our analyses are based on data from a nationally representative sample of secondary school students in Germany who participated in the DESI study ("Assessment of Student Achievements in German and English as a Foreign Language [DESI] study");

DESI-Konsortium, 2008)¹¹. The study was carried out by a consortium of researchers headed by the German Institute for International Educational Research (DIPF).

Our reanalyses of this data set were based on 7,650 students attending 352 classrooms in 181 schools (average number of students per classroom $M = 22$) who participated at both measurement points (T_1 , T_2), at the beginning and at the end of ninth grade. The analyses were limited to the western states of Germany due to the very small proportion of language minority students in the eastern states (see Federal Statistical Office of Germany, 2012).

The German education system assigns students to different school tracks after primary education. In 14 German federal states, secondary school starts after four years, in two states after six years of primary education (KMK, 2013a). Depending on the federal state, ninth grade is the fifth respectively the third year of secondary school. These tracks differ in the level of education the students typically receive—from the vocational track “Hauptschule” providing a basic general education and the intermediate track “Realschule” providing students with a more extensive general education to the academic track “Gymnasium” leading to the degree necessary for university entrance. “Gesamtschule” combines all three tracks in a comprehensive school type that exists parallel to the separate track schools. The vocational track and the comprehensive school type have no strict enrollment requirements and have student populations that are quite similar to each other. In our analyses, we combined them into a single category referred to as “school with a vocational track”. The three school tracks not only differ in their students’ achievement level at secondary school entry, but also in their curricula and organizational structures. Therefore, school tracks must be controlled for in analyses that aim at estimating net effects of classroom composition and instructional processes.

Our key independent variable was the proportion of language minority students in classrooms. In general, the German education system—like other countries in the world—is facing an increasing population of students with immigrant background that often does not only speak the language of instruction (in our case German) within their families. Ordinarily, largest language minority groups in Germany represent Turkish, Russian and other Slavic languages. While students of Turkish origin commonly face

¹¹ We used the same dataset as Klieme et al. (2010) cited within the theoretical background section. Our study differs from that study in various aspects: Klieme and colleagues focused on effects of school-level processes and consequently analyzed a three-level model without examining mediation effects. We took Klieme and colleagues findings as a starting point and focused especially on the language minority classroom composition. We additionally explored mediation effects and interaction effects between the proportion of language minority students and individual minority background.

pronounced disadvantages in the educational system, students from the former Soviet Union do relatively better (see Segeritz et al., 2010). However, the various language minority groups are on average more similar than different in terms of socioeconomic status, educational success and prestige in society. Language minority students are on average socioeconomically and educationally disadvantaged in comparison to language majority students (see Federal Statistical Office of Germany, 2012; Kristen et al., 2011; OECD, 2006). In the sample of the present study the largest language minority groups were Russian (21.4% of the language minority students [LMiSs] in the sample), Turkish (19.4% of LMiSs), and Polish (10.8% of LMiSs)¹².

For a general sample description see Table 5.1. For information on the distribution of LMiSs across classrooms see Figure 5.2.

Table 5.1 Descriptive Sample Statistics of Demographic Variables

Variable	Level	
	Classroom (Level 2)	Individual (Level 1)
	<i>N</i> = 352	<i>N</i> = 7,650
LMiSs (%)		
Mean	23.70	
<i>SD</i>	18.02	
Range	0.00–88.20	
%		21.90
SES		
Mean	48.29	
<i>SD</i>	8.00	
Range	27.00–70.59	16.00–90.00

¹² These three largest LMiS groups did not differ in their intrinsic motivation in German lessons at T₁ [$F(2, 384) = 2.74, p = .066, \omega^2 = .01; M_{tur} = 2.65, SD_{tur} = 0.72; M_{rus} = 2.71, SD_{ru} = 0.68; M_{pol} = 2.48, SD_{pol} = 0.63$]. And there was only a tendency of Turkish students showing a slightly higher level of motivation at T₂ [$F(2, 724) = 3.28, p = .038, \omega^2 = .01; M_{tur} = 2.65, SD_{tur} = 0.67; M_{rus} = 2.50, SD_{ru} = 0.72; M_{pol} = 2.55, SD_{pol} = 0.73$]. However there were slight differences between students from different LMiS groups on other covariates: Students with Turkish background showed lower German achievement scores than students with Polish or Russian language background at T₁ ($F(2, 807) = 64.07, p < .001, \omega^2 = .13; M_{tur} = -1.01, SD_{tur} = 0.90; M_{rus} = -0.27, SD_{ru} = 0.95; M_{pol} = -0.19, SD_{pol} = 1.07$) and lower SES than students with Polish language background ($F(2, 800) = 6.96, p = .001, \omega^2 = .01; M_{tur} = 37.68, SD_{tur} = 11.58; M_{rus} = 39.34, SD_{ru} = 12.72; M_{pol} = 42.21, SD_{pol} = 13.20$). We decided to not to differentiate between single language groups in our main analyses for several reasons: We do not have detailed information on the specific language background for some of the students in the sample (7.00 %) and there is a large number of small language groups in the sample (up to N=64 different language groups, depending on the categorization of the groups). At the classroom level, there would have been a lot of classes with very low proportions of different native languages. This makes effect estimates rather imprecise and does not allow conclusions about specific student groups.

Female (%)

Mean	50.45	
SD	17.05	
Range	7.40–100.00	
%		51.90

School tracks, %

AT	25
IT	29
VT	46

Age (years)

Mean	14.85
SD	0.74

Note. LMISs= language minority students; SES= socioeconomic status; AT = academic track (“Gymnasium”), IT= intermediate track (“Realschule”), VT= school with a vocational track, i.e. combined category of vocational track (“Hauptschule”) and comprehensive schools (“Gesamtschule”). For operationalization of these variables see Measures section. Sample description based on weighted cases (see Data Analysis section).

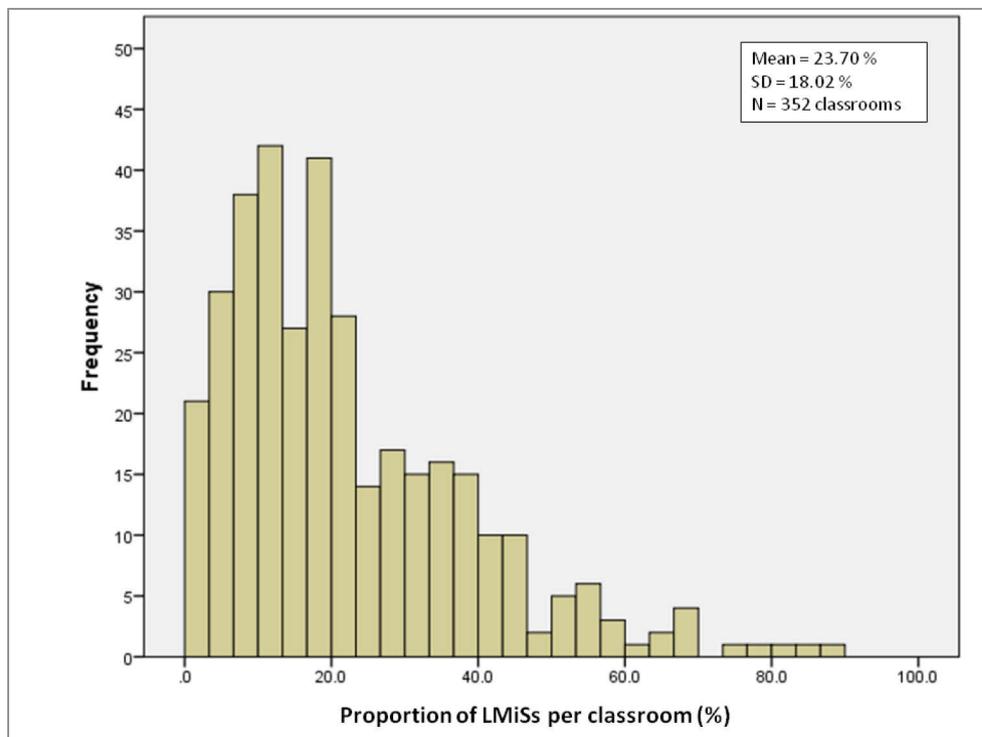


Figure 5.2: Distribution of language minority students (LMISs) across the 352 sampled classrooms.

5.3.2 Measures

German achievement tests were administered at T_1 , student questionnaires at T_1 and T_2 as well as parent questionnaires at T_2 .

Outcome variable: intrinsic motivation in German lessons (T_2). Students rated their intrinsic motivation in German lessons in the student questionnaire at T_1 (control variable; see below) and T_2 (outcome variable). The scales captured intrinsic aspects of motivation referring to enjoyment and interest in German lessons and were adapted from existing scales on motivation in other school subjects (e.g., PISA-Konsortium Deutschland, 2006). The scale at T_2 was composed of three items rated on a 4-point Likert scale (1 “fully disagree” to 4 “fully agree”) by all students. One item example is “German is fun.” (manifest $M = 2.55$, $SD = 0.76$; Cronbach’s Alpha = .68). These items represented the latent construct of motivation at T_2 in our analyses.

Student level background variables. *Language minority [LMiS] status.* Information on the language minority status was gathered with the student questionnaire (T_1) as well. Students who reported a) only another language than German or b) another language *and* German as a native language were categorized as language minority students [LMiSs] (dummy coding; German = 0, LMiS = 1). Of the LMiSs, 65.66% reported only another language than German and 34.34% another language and German as their native language.

Gender. Student gender was assessed in the tracking form completed by the school administration. Information was available for all students (dummy coding; male = 0, female = 1).

German achievement at T_1 . Tests in German language achievement were administered to all students in the sample at the beginning of ninth grade. The test covered the domains reading comprehension, writing, and language awareness, that is, grammatically correct and stylistically appropriate language use. It was developed by experts in German language instruction and designed to cover the curriculum of grade 9 German language instruction (DESI-Konsortium, 2008). A generalized Rasch model was used to estimate student achievement scores on a common scale. The scores were generated as weighted likelihood estimates (WLE) which are based on maximum likelihood estimation considering additional individual test information (Warm, 1989).

The score distribution was set to a mean of zero and a standard deviation of one in the calibration model.

Intrinsic motivation in German lessons (T₁). In our analyses we controlled for intrinsic motivation at T₁. This scale was the same as at T₂ (see above), yet, it was included in only one of two questionnaire versions which were randomly assigned to the students in each classroom (see for the rationale of planned missing data designs, Graham, Taylor, Olchowski, & Cumsille, 2006). Consequently, scale values were missing by design for about 50% of the students in each classroom (manifest $M = 2.57$, $SD = 0.75$, Cronbach's Alpha = .67). Motivation at T₁ is treated as a latent construct with the three manifest items as indicators.

Socioeconomic status (SES). The SES of a student's family was represented by the "International Socio-Economic Index of Occupational Status" (ISEI; Ganzeboom et al., 1992). This index has a theoretical range from 16 to 90. The HISEI represents the highest occupational status of the student's mother or father, based on a classification of parents' occupation reported in the parent questionnaire (T₂).

Classroom level background variables. Classroom level background variables included composition characteristics of the classroom and school track. The aforementioned individual background variables were aggregated to form proportion of LMiSs, proportion of females, average German achievement at T₁, and average SES. For school tracks, we created two dummy variables "academic track" and "school with a vocational track", with "intermediate track" as the reference category.

Mediator: instructional climate. Students evaluated the instructional climate of their German lessons at T₂. The questionnaire included six scales composed of several items which were rated on a 4-point Likert scale. Table 5.2 gives an overview of these scales. They represent aspects of instructional climate that are deemed to meet students' needs for autonomy, competence, and relatedness—and, hence, to foster their intrinsic motivation (see section "instructional climate and student motivation")—including supportive teacher behavior and a focus on student interests. Five scales were administered to all students in each class, yet, for reasons of practicality the longest scale ("positive culture of making mistakes") was randomly assigned to approximately 50% of the students in each class. The intraclass correlation 2 (ICC2) indicates the reliability of an aggregated construct at the classroom level. It is estimated by applying the Spearman-

Brown formula to the intraclass correlation (ICC1), measuring the agreement of student ratings in a classroom (Marsh et al., 2012). The ICC2 values of the instructional climate measures reveal acceptable levels of reliability (Lüdtke et al., 2009). Together, these six scales represented the latent construct “instructional climate” in our analyses. Multilevel confirmatory factor analyses of the six scales revealed that the correlation at the classroom level between the six constructs ranged between $r = .88$ and $r = .98$, justifying a parsimonious model with one overall latent construct. High values of this construct indicated a more motivating instructional climate.

Table 5.2 Overview of Scales of Instructional Climate: Number of Items (N), Reliabilities, and Item Example

Scale	<i>N</i> (items)	Cronbach's alpha	ICC1	ICC2	Item example
Support	4	.86	.19	.83	If I need help, I get it from my German teacher.
Student orientation	6	.88	.21	.85	If somebody has a good idea, my German teacher acknowledges it.
Focus on student interests	2	.80	.21	.84	My German teacher makes the lessons interesting to me.
Comprehensibility	3	.78	.19	.83	The assignments in the German classes are clear and comprehensible to me.
Classroom climate during instruction	3	.86	.15	.79	I get along well with my German teacher.
Positive culture of making mistakes ^a	13	.89	.18	.82	I learn something new when my German teacher corrects my mistakes.

Note. ICC1 = intraclass correlation, ICC2 = reliability of classroom mean ratings.

^a This scale was randomly assigned to approximately 50% of the students in each classroom; all other scales were assigned to each student in a classroom.

5.3.3 Data Analysis

We used structural equation modeling to analyze the data (see Bovaird, 2007). Two random intercept multilevel structural equation models were estimated to explore (1) the relationship between the proportion of language minority students and students' motivation at T_2 after controlling for individual background characteristics including prior motivation as well as potentially confounding classroom composition characteristics at T_1 , and (2) the mediating role of the instructional climate for the composition effects on individual motivation. The software Mplus (Version 6.1; Muthén & Muthén, 1998-2010) was used for all analyses. The analyses were conducted with sample weights at both levels adjusting for different selection probabilities of students and classrooms to obtain representative results (see analysis strategy described by Nagengast & Marsh, 2012).

Metric background variables at the student level (achievement, SES, manifest items of the latent motivation construct at T_1) were standardized, which implies centering at their grand mean. Categorical variables (LMiS, female) were neither centered nor standardized. Consequently, the estimates of background variables aggregated at the classroom level are interpretable as compositional effects (Raudenbush & Bryk, 2002). The proportion variables at the classroom level (LMiSs, females) were divided by 10 to indicate steps of 10%. For example, an estimated regression coefficient of 0.5 for the proportion of LMiSs in the class would imply that a 10% increase in LMiSs is associated with an increase in the dependent variable by 0.5 standard deviations. With an average of about 22 students per class, 10% corresponded to approximately 2 students. We additionally included the quadratic term of the proportion of LMiSs in a classroom, in order to explore a potential nonlinear relationship between the proportion of LMiSs and motivation (see Hypothesis 3b). Prior to calculating the quadratic term, we centered the proportion of LMiSs at its mean to counteract multicollinearity. All regression coefficients in the results table were standardized using the total variance (within + between) of the outcome variable (see Marsh et al., 2009).

For addressing research question 2 we used a doubly latent approach with cross-level measurement invariance for the construct of instructional climate (see Table 5.2). This approach comprises latent measurement models at both levels and latent aggregation for the classroom level construct (e.g., Lüdtke et al., 2011; Marsh et al., 2012). An important advantage of a model with these features is that it corrects possible measurement and sampling errors associated with designs in which variables measured at the individual level are used to operationalize a construct at the classroom level. This study focused on classroom-level effects of instructional climate as they represent the shared perception of instructional characteristics of students in each classroom (e.g., Lüdtke et al., 2009). However, we modeled additional relationships at the student level to control for variables potentially confounded with students' perception of instructional climate.

In regard to research question 3, we estimated a latent cross-level interaction effect (i.e., slope-as-outcome model). In this analysis model, we used the same background variables as in the first model but LMiS was group-mean centered and the regression coefficient of motivation T_2 on LMiS was allowed to vary across classrooms (i.e., random slope). The random slope was regressed on the proportion of language minority students at the classroom level (see Aguiní et al., 2013; Marsh et al., 2009). In a

last step, we computed a model including all background variables, instructional climate, and the cross-level interaction effect to explore if a possible differential effect of the proportion of LMISs on motivation of minority students compared to majority students is related to differences in the instructional climate.

We applied several measures to deal with missing values. WLE achievement scores were generated for all students, including those with missing data on single test items. Students who answered no test item (4.7%) were excluded from the analyses. The SES scores were based on information from the parent questionnaire which were completed by 62% of the participants. The DESI study authors (see Hartig, Jude, & Wagner, 2008) replaced missing SES scores by estimated values using multiple imputation (e.g., Rubin, 1987) on the basis of additional information from the student questionnaire. SES scores could be generated for 92% of the students. In our reanalysis, we used these imputed scores. Students without any information on their SES, achievement score, and/or LMIS status (9%) were excluded from the analyses. Overall, we were able to include 91% of the original sample in our analyses. To deal with item nonresponse on the instructional climate variables and motivation at T₁ and T₂ in the planned missing data design, we used the full information maximum likelihood (FIML) estimator implemented in Mplus. This estimator applies a model-based approach to missing data (see Enders, 2010), using all information available from the model variables to estimate the model parameters.

5.4 Results

5.4.1 Preliminary Analyses

Measurement model fit. Prior to the analyses of the research questions, we tested the fit of the measurement models of our latent constructs motivation at T₂, motivation at T₁, and instructional climate in three separate analyses.

The first measurement model for motivation at T₂ specifying parallel measurement models at the individual and at the classroom level and cross-level measurement invariance showed acceptable model fit ($\chi^2 = 6.814$, $df = 2$, $p < .05$, RMSEA = .018, CFI = .997, SRMR_{within} = .006, SRMR_{between} = .040). Such an unconditioned model without any predictors provides information on the amount of variance at both levels necessary to compute the intraclass correlation (ICC1). The ICC1 estimates the proportion of the total

variance due to differences between classrooms. If there was no variation in motivation between classrooms, multilevel analyses would not be feasible. In our study the proportion of variance between classrooms was 14%. As is commonly the case for non-cognitive variables, this proportion is lower than the variation typically found for achievement between classrooms (see Trautwein et al., 2006).

In the second model, we tested the fit of the motivation model at T_2 together with the student level control variable motivation at T_1 . We constrained the loadings of indicator variables to be equal across measurement points to ensure that the same construct was measured across time, and we assumed additional correlated uniqueness for the same items measured at T_1 and T_2 . The model fit was acceptable ($\chi^2 = 72.679$, $df = 9$, $p < .05$, RMSEA = .030, CFI = .987, SRMR_{within} = .018, SRMR_{between} = .031). The correlation between motivation at T_1 and T_2 was $r = .73$.

In the third model, we explored the fit for instructional climate. The fit for this doubly latent measurement model with cross-level invariance was acceptable as well ($\chi^2 = 155.936$, $df = 23$, $p < .05$, RMSEA = .028, CFI = .988, SRMR_{within} = .017, SRMR_{between} = .024). The proportion of variance at the classroom level (ICC1) was 26% and the reliability of this aggregated construct was ICC2 = .88.

Measurement invariance across student populations. In additional multi-group analyses at the student level we compared the fit and structure of the measurement models for language minority students and language majority students to test for measurement invariance across the groups (Byrne, 2008; see Segeritz & Pant, 2013). The analyses revealed no systematically different answer pattern on the scales (i.e., scalar measurement invariance) depending on students' language background for the models of intrinsic motivation in German lessons at two measurement points and the model of instructional climate.

Single predictor analyses and multicollinearity. We tested separate multilevel models for our background variables to gain insight into their relationship to motivation at T_2 controlling for individual motivation at T_1 (see Table A2.1 in the appendix). All variables were significantly related to motivation at T_2 except the SES.

We also tested for multicollinearity among background variables and found the average SES at the classroom level to be highly to moderately correlated with the average achievement at the classroom level ($r = .85$) and with the proportion of LMiSs ($r = -.48$). It was also strongly related to school tracks (average SES in the academic track: $M = 58.43$, $SD = 4.54$; average SES in the intermediate track: $M = 48.01$, $SD = 3.09$, average

SES in schools with a vocational track: $M = 41.92$, $SD = 4.43$; $F(2, 349) = 509.621$, $p < .001$, $\omega^2 = .74$). As SES was not related to the outcome variable and socioeconomic background information was already represented through its strong relation to achievement and school tracks we chose a more parsimonious model without SES as control variable¹³.

Our independent variable (LMiSs %) was far from being collinear to the control variables (correlation at the classroom level to achievement: $r = -.50$, and to proportion of females: $r = -.14$; mean differences across tracks: average LMiSs% in the academic track: $M = 15.16$, $SD = 11.34$, average LMiSs% in the intermediate track: $M = 19.30$, $SD = 14.29$, average LMiSs% in schools with a vocational track: $M = 32.33$, $SD = 19.28$; $F(2, 349) = 41.651$, $p < .001$, $\omega^2 = .19$). As expected, average German achievement at T₁ and school tracks were strongly related (average achievement in the academic track: $M = 1.36$, $SD = 0.39$; average achievement in the intermediate track: $M = 0.33$, $SD = 0.39$, average achievement in schools with a vocational track: $M = -0.91$, $SD = 0.61$; $F(2, 349) = 715.202$, $p < .001$, $\omega^2 = .80$).

5.4.2 Students' Motivation in Classrooms with Varying Proportions of Language Minority Students

The first set of our analyses investigated whether the language minority classroom composition¹⁴ was related to students' level of motivation in German lessons (see Model 1 in Table 5.3).

At the individual level there was no significant difference between LMiSs and students whose native language is only German in their level of motivation at T₂ controlling for motivation at T₁ (see Table A2.2 in the appendix for results of the student level variables and fit indices). Regarding the control variables at the classroom level (see Model 1 in Table 5.3) the analyses showed that students' level of motivation in German lessons was lower in classrooms with a higher average achievement level at T₁ controlling for individual achievement and higher in classes within the academic school

¹³ Including SES in the analysis models did not change the overall pattern of the findings reported in the results section.

¹⁴ Excluding students who reported another language and German as a native language from the analyses only comparing students with German and students with non-German native language did not change the overall pattern of the findings reported in the results section.

track. The language minority classroom composition was significantly related to the individual level of motivation after controlling for gender composition, prior German achievement level, the general achievement level associated with school tracks and characteristics at the student level including motivation at T_1 . An increase of the proportion of language minority students from 0% to 10%—corresponding to about two students—was associated with an increase in motivation by .05 standard deviations ($\beta = .05$, $SE = .02$, $p = .014$). The quadratic term of the proportion of minority students was not significant ($\beta = -.01$, $SE = .01$, $p = .106$), indicating a linear relationship between the proportion of minority students and motivation. In the end, the effect of the proportion of language minority students is a combined effect of the linear and quadratic term of LMiSs (%), with a tendency to depend on the specific percentage point assumed (i.e., effect for 10% of LMiSs: $0.05*1 - 0.01*1^2 = 0.04$; increase from 0% to 10% of LMiSs corresponded to a motivational increase by $0.04 - 0.00 = 0.04$ standard deviations and an increase from 10% to 20% of LMiSs corresponded to a motivational increase by $0.06 - 0.04 = 0.02$ standard deviations). As the quadratic term of the proportion of minority students was not significantly related to the outcome variable, we decided to remove it from the following models.

Table 5.3 Results of Multilevel Structural Equation Models Predicting Motivation in German Lessons at T₂ (Classroom Level Results)

	Model 1		Model 2		Model 3		Model 4	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
Motivation (T ₂) on composition variables								
LMiSs (%)	.05*	(.02)	.02	(.01)	.03	(.02)	.02	(.01)
LMiSs (%) quadratic term	-.01	(.01)	–		–		–	
Female (%)	-.03	(.02)	-.02	(.01)	-.03	(.02)	-.02	(.01)
German achievement T ₁ (Mean)	-.19**	(.07)	-.15**	(.05)	-.17**	(.07)	-.15**	(.05)
Academic track	.26**	(.09)	.17**	(.06)	.24**	(.08)	.17**	(.06)
Vocational track	-.04	(.08)	-.15**	(.06)	-.02	(.08)	-.15*	(.06)
Instructional climate on composition variables								
LMiSs (%)	–		.04	(.02)	–		.03	(.02)
LMiSs (%) quadratic term	–		–		–		–	
Female (%)	–		-.03	(.02)	–		-.03	(.02)
German achievement T ₁ (Mean)	–		-.01	(.07)	–		-.01	(.07)
Academic track	–		.08	(.10)	–		.08	(.10)
Vocational track	–		.22*	(.11)	–		.22*	(.11)
Motivation (T ₂) on instructional climate								
Instructional climate	–		.26**	(.02)	–		.26**	(.02)
Cross-level interactions								
LMiS x LMiSs (%)	–		–		.10**	(.02)	.07**	(.02)
R ² Level 2 (motivation T ₂)	.27		.77		– ^a		– ^a	
R ² Level 1 (motivation T ₂)	.53		.62		–		–	
N (Level 2)	352		352		352		352	

Note. Level 1 = student level; Level 2 = classroom level; LMiSs = Language Minority Students. Regression coefficients were standardized using the total variance (within + between) of the outcome variable. Note that interaction analyses (Model 3 and Model 4) change the meaning of main effects (i.e., LMiS, LMiSs %) representing an estimate of the regression in case that the other variables equal to zero (Hox, 2010).^a

Explained variance is not reported in Mplus for random slope models, since in these models the variance of the outcome variable varies as a function of the predictor variables.

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

5.4.3 Mediation of the Language Minority Classroom Composition Effect by Instructional Climate

In a next step, we tested if the positive relationship between proportion of LMiSs in a classroom and level of motivation in German lessons after one school year was mediated by a more motivating instructional climate from the students' perspective (see Model 2 in Table 5.3). The results show that the shared perception of the instructional climate in the classroom was positively related to the level of motivation after one school year ($\beta = .26, SE = .02, p < .001$) and that the effects of most classroom composition variables on motivation slightly decreased. However, the change was rather small. We computed the indirect effect and applied the delta method to determine the corresponding standard error in order to test the mediation effect (see Nagengast & Marsh, 2012; Raykov & Marcoulides, 2004). The indirect effect of the proportion of LMiSs on motivation (not presented in Table 3) was not significant ($\beta = .01, SE = .01, p = .067$). This suggests that the positive effect of the proportion of language minority students on motivation after one school year was not mediated by the instructional climate.

5.4.4 Differential Effects for Language Minority and Language Majority Students

We next explored the differential effect of the proportion of LMiSs in a classroom on the level of motivation for LMiSs compared to students with only German as their native language. In a first step, we estimated Model 1 without instructional climate and the quadratic term of the proportion of minority students (Table 5.3) and computed an additional latent interaction between the effects of individual language background on motivation in classrooms with varying proportions of LMiSs. The interaction effect was significant ($\beta = .10, SE = .02, p < .05$; see Model 3 in Table 5.3). The proportion of LMiSs in a classroom explained 63% of the total variability in the effect of individual language status on motivation across classrooms (Aguinis et al., 2013). Figure 5.3 illustrates the results of these interaction analyses controlling for other background variables at both levels. Values of background variables were assumed to be zero in this case, representing average scores of continuous variables and the category "male" and "intermediate track" of dummy variables. The y-axis shows values of motivation at T_2 one standard deviation above and below the mean standardized by the total variance. The x-axis indicates the proportion of LMiSs in a classroom ranging from 0 to 90 (the range in

our sample was 0 to 88.2%, see sample description). While the motivation at T_2 of all students increased in classrooms with higher proportions of LMiSs, the effect was only statistically significant for LMiSs. Being a LMiS in a classroom with a low proportion of LMiSs (below about 25%) was associated with below-average motivation. For instance, a LMiS in a classroom with 10% of LMiSs had a standardized motivational score of -0.18 . Being a LMiS in a classroom with a higher proportion of LMiSs (above about 25%), in contrast, was associated with above-average motivation: for example, a LMiS in a classroom with 60% of LMiSs had a standardized motivational score of 0.45 . An additional analysis including instructional climate (see Model 4) suggested that this differential effect was not due to differences in instructional climate. In supplemental analyses presented in the appendix (see Table A.3 and Figure A.1) we controlled for further teacher characteristics and class size. These analyses confirmed the differential effect as well.

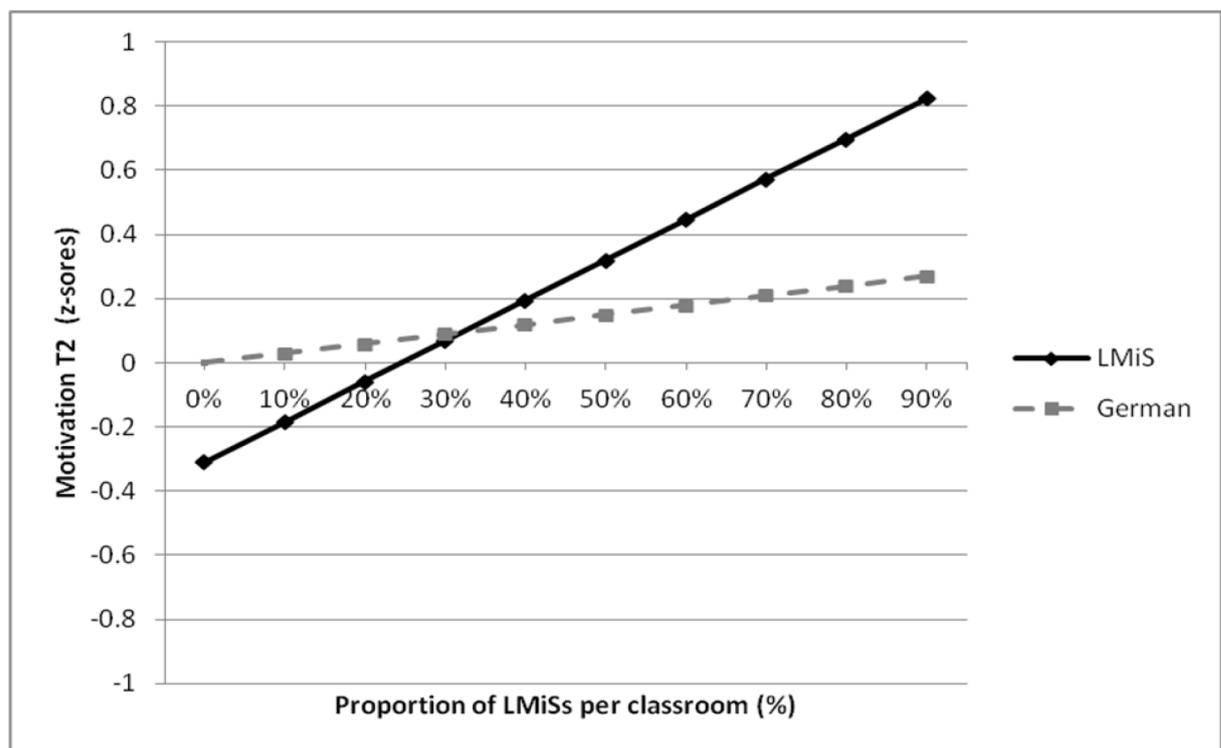


Figure 5.3: Differential effect of the proportion of language minority students (LMiSs) on the level of motivation for LMiSs and students with only German native language (Model 3 in Table 5.3).

5.5 Discussion

The aim of the present study was to shed light on the relationship between the proportion of language minority students in a classroom and students' level of motivation in language lessons over the course of one school year. Our findings suggest that students in classrooms with a higher proportion of minority students were increasingly motivated to learn in German lessons. This effect was only significant for language minority students. We found no statistical evidence that the composition effect was mediated by the instructional climate in the classroom.

The analyses for our first research question revealed a slightly positive relationship between the proportion of language minority students and the level of intrinsic motivation (Hypothesis 1b). The regression coefficient represented an increase of the proportion of language minority students by 10 percentage points and showed overall small effects. However, this means, that increases by larger proportions of language minority students are associated with larger effect sizes. Our analyses for research question 3 revealed that this association was only significant for language minority students. We conclude that it expresses relevant effects on students' level of motivation in real classroom settings. At the same time, one has to consider that the proportion of language minority students can be simultaneously related to various student outcomes, such as motivation and achievement, in competing ways (for effects on German reading achievement based on the same dataset see Rjosk et al., 2014). As to individual differences in the level of motivation (Hypothesis 1a), we did not find general significant differences between language minority students and majority students at the student level. Thus, our findings did not support the assumption of a heightened intrinsic motivation in language lessons of minority students that would predict the immigrant optimism hypothesis (Kao & Tienda, 1995). Heightened intrinsic motivation of minority students was reported particularly for mathematics (Stanat & Christensen, 2006). Overall, learning the language of instruction in language lessons may have a slightly different meaning for language majority students and language minority students, even though analyses of measurement invariance showed that they are psychometrically comparable (see Segeritz & Pant, 2013): Learning in language lessons may represent special challenges for language minority students related to possible lower expectancies of success. Simultaneously, extrinsic utility value (i.e., usefulness of language abilities in reaching future goals) and costs (i.e., expended effort and time) of learning the language of instruction might vary between language minority students and language majority

students (see Dörnyei, 1998; Eccles & Wigfield, 2002). This suggests that analyses of motivational differences indeed should be subject-specific and should furthermore include a wide array of motivational beliefs such as attainment value, utility value, and costs (Eccles & Wigfield, 2002; see also Fuligni et al., 2005).

As to the classroom level background characteristics, the findings suggest that attending a higher achieving classroom is negatively related to students' motivation after holding constant the school track. This finding is in line with the big-fish-little-pond effect or contrast effect describing originally the phenomenon that students tend to feel less competent in classrooms with a large number of high achieving peers (Nagengast & Marsh, 2012; Marsh, 1987). Recent research successfully extended this concept to explain effects on other motivation-related variables, such as interest in a school subject (see Trautwein et al., 2006). Furthermore, attending the academic track was positively related to students' motivation. This may suggest a basking in reflected glory effect or assimilation effect describing positive attitudes towards school achievement and motivation of high track students related to placement within the educational system (Cialdini et al., 1976; see Chmielewski et al., 2013; Trautwein et al., 2006).

The analyses for our second research question revealed that the instructional climate in the classroom was related to the level of motivation at the end of the school year after accounting for motivation at the beginning of that school year. This supports previous studies on the relevance of the instructional context for motivation (e.g., Eccles & Roeser, 2011; Roeser et al., 2000; Opdenakker et al., 2012; Wang & Eccles, 2013). However, the instructional climate did not vary as a function of the proportion of language minority students, and thus it did not mediate the effect of the proportion of minority students on intrinsic motivation in language lessons.

The analyses for our third research question revealed a differential relationship between the proportion of language minority students and the level of motivation for language minority students compared to language majority students. The positive effect of a high proportion of language minority students was only significant for language minority students themselves. These results are partly in line with the belongingness hypothesis (Baumeister & Leary, 1995). Language minority students were increasingly motivated in language lessons in classrooms with more other students with a language minority background. In principle, the belongingness hypothesis would also predict higher motivation levels of language majority students in classrooms with higher proportions of majority students, yet this was not the case in our study. Our findings are

rather in line with the assumption that majority/minority group membership and emotional belongingness to peers may be especially important for minority students (see Den Brok & Levy, 2005; Eccles & Roeser, 2011; Kumar & Maehr, 2010). However, we have no information on students' perceived belonging to peers. We assume that the differential effect is related to the peers in the classroom as our additional analyses suggested that it is not associated with competing contextual characteristics that potentially explain motivational differences between classrooms. More precisely, we took into account perceived instructional climate, class size, and the classroom teacher's years of teaching experience and teacher training for language lessons as a proxy for teaching quality. In the end, the explanation of the differential effect remains an open question for future research.

In summary, our study using short-term longitudinal data from a large sample of classrooms provided optimal opportunities to gain first insights into the relationships between classroom composition, instructional climate, and students' level of motivation. Most studies so far investigated the more distal context characteristic of school composition or did not investigate actual entire classrooms (Urda & Schoenfelder, 2006). Our findings thus add to growing evidence that the nature and quality of the social environment are important in promoting motivation in school (see Wang & Eccles, 2013). When asked about motivation of language minority students in language lessons, we now would clearly say "it depends". With an increasing proportion of minority students in the classroom the motivational situation may change.

5.5.1 Limitations and Future Research

The present study has some limitations which should be addressed in future research. First, we investigated students in grade nine who attended their classrooms already for two to four years (depending on the German federal state), and we followed them for a period of only one school year. We would assume that even if relationships explored over a short duration are modest, their interconnections over an entire educational history have the potential to produce important cumulative differences among students (see Johnson et al., 2001). Based on our data however, we cannot make inferences about relationships across longer time periods. Future studies should explore students' motivation over a longer period and start, for example, with the beginning of secondary school when classes are composed.

Second, we were not able to test the belongingness perspective directly as we have no information on students' perception of their emotional relatedness or attachment to the peers in their classroom. We interpreted the stronger association between the proportion of language minority students and motivation for language minority students compared to language majority students in light of the belongingness hypothesis. Future studies should test this assumption more directly.

Third, we did not explore the classroom composition due to proportions of specific language groups. Our study showed that the distinction between language minority and majority students is a viable categorization that seems to affect students' motivation in language lessons. Yet, we did not consider the degree of homogeneity in language backgrounds of the students in a classroom—which should be even stronger related to a feeling of belonging and student motivation than a broader distinction between majority and minority students in terms of possible perceived ingroup and outgroup affiliations (see Tajfel & Turner, 1986). Investigating differential effects for separate language groups as well as using measures such as diversity indices (see e.g., Benner et al., 2008) could be a useful addition.

Fourth, we explored compositional effects on only one student outcome—intrinsic motivation in language lessons. Future research should include further motivational outcomes such as utility value and costs which might differ for language minority and majority students (see Fuligni 2001; Fuligni et al., 2005).

Finally, beyond the restrictions of our study, compositional analyses in general are sometimes criticized because models are inadequately specified (Harker & Tymms, 2004; Hauser, 1970). More precisely, some authors argue that spurious effects may arise if level-1 variables were assessed with low reliability and validity or covariates at level 1 were omitted. However, our variables analyzed at level 1 constitute reliable and valid measures related to students' motivation that are theoretically sound and commonly used in international large-scale assessment studies (OECD, 2010).

5.5.2 Conclusion

In conclusion, the findings indicate that the classroom context related to students' language background is relevant for students' motivational development. That is, it seems a viable addition to include context characteristics in future studies that intend to explain students' motivational development. The results of the present study confirm once again

that the instructional climate, which is assumed to be stimulated by the teacher and created in relationship between students and teachers, affects students' level of motivation in language lessons across one school year. Such findings call for research on measures designed to assist teachers in promoting an instructional climate that is likely to facilitate increased motivation.

Yet, we found no statistical evidence that the instructional climate varies as a function of language minority classroom composition. Rather the proportion of language minority students affected students differentially. Language minority students reported a heightened motivation to learn in language lessons only in classrooms with high proportions of language minority students. This points to the environment's potential in terms of heightened motivation to learn especially for minority students. Further research should explore the mechanisms underlying this pattern. This might help us to support language minority students' motivation and, in this way, to improve language minority students' performance in the end.

6

Teilstudie 3:

Ethnic composition and heterogeneity in the classroom—Their measurement and relationship with student outcomes

Draft version 07.10.2015. This paper has not been peer reviewed. Please do not copy or cite without author's permission.

Rjosk, C., Richter, D., Lüdtke, O., & Eccles, J. S. (submitted). Ethnic composition and heterogeneity in the classroom—Their measurement and relationship with student outcomes. *Journal of Educational Psychology*.

6. Teilstudie 3: Ethnic composition and heterogeneity in the classroom—Their measurement and relationship with student outcomes

Abstract

This study explores various measures of the ethnic makeup in a classroom and their relationship with student outcomes. We examine whether measures of ethnic diversity are related to achievement (mathematics, reading) and feeling of belonging with one's peers over and above commonly investigated composition characteristics. Multilevel analyses were based on data from a representative sample of 18,762 elementary school students in 903 classrooms. The proportion of minority students and diversity measures showed negative associations with student outcomes in separate models. Including diversity measures and the proportion of minority students, diversity of minority students mostly lost its significance. However, students showed higher levels of mathematics achievement in more diverse classrooms. The various measures of diversity led to comparable results.

Keywords: ethnic composition, diversity, multilevel analysis, academic achievement, classroom

6.1 Introduction

As societies become more diverse in terms of ethnic background, the composition of the student body within educational systems diversifies as well. Various theoretical accounts assume a negative relationship between the proportion of ethnic minority students and student achievement based on school resources and mediated by instructional quality, language spoken with peers, and learning culture (Driessen, 2002; Goldsmith, 2011; Raudenbush et al., 1998; Stipek, 2004).

Research exploring the relationship between the ethnic makeup of schools or classrooms and student achievement shows mixed results: The proportion of ethnic minority students in a school or classroom often has no or slightly negative predictive effects on student achievement (Mickelson et al., 2013; Van Ewijk & Slegers, 2010b). For ethnic heterogeneity, some studies report that a higher proportion of ethnically heterogeneous students may lead to higher achievement. The proportion of minority students and heterogeneity represent two different characteristics of the ethnic makeup but are associated with each other. Several authors assume that students in ethnically heterogeneous learning environments encounter and have to work through contradictions and discrepancies in everyday life and therefore may be able to expand their intellectual capacities (e.g., Benner & Crosnoe, 2011; Gurin et al., 2003; Peetsma et al., 2006; Tam & Bassett, 2004).

Most studies have dealt only with a broad distinction between ethnic minority and majority without addressing and measuring ethnic heterogeneity or diversity. The present study compares various measures of ethnic composition and heterogeneity used in different disciplines with the goal of better understanding the relationship between the ethnic makeup of classrooms and student outcomes. Our aim is to investigate whether the measures of ethnic diversity are related to student achievement and psychosocial outcomes over and above commonly investigated characteristics of classroom composition.

6.2 Relationship between Characteristics of the Student Body and Individual Student Outcomes

Students differ in their educational success and level of achievement outcomes. This variability is associated with individual background characteristics, such as cognitive

abilities, prior knowledge, and the socioeconomic background of their families and associated home learning environment (e.g., OECD, 2010). In addition to these individual and family characteristics, the composition of the student body matters for individual outcomes. For instance, students tend to show higher achievement in classrooms that are characterized by a high average prior achievement level and a high average socioeconomic status (SES) of the student body (Van Ewijk & Slegers, 2010a; but see Hauser, 1970). However, research is inconclusive on whether the ethnic composition is related to student outcomes (e. g. Driessen, 2002), independent of the average prior achievement and average SES of the classroom.

The present paper focuses on the relationship between the ethnic makeup of classrooms and students' achievement as well as psychosocial outcomes. More precisely, the term ethnic makeup may pertain to two characteristics that represent different strands of theory and research: (1) the proportion of ethnic minority students is the measure of ethnic composition commonly used in educational research and (2) ethnic heterogeneity measured by various indices analogous to the concept of diversity operationalized in a large number of different disciplines.

6.2.1 Definitions: Ethnic Composition and Heterogeneity

Educational research that addresses questions of the ethnic makeup of classrooms commonly operationalizes the ethnic composition by calculating the *proportion of ethnic minority students* in a classroom or school. For instance, international meta-analyses on ethnic composition and student achievement with about 38 primary studies consistently distinguish between ethnic minority and majority students (Mickelson et al., 2013; Van Ewijk & Slegers, 2010b). This approach is sometimes referred to as “simplistic majority-minority approach” (Budescu & Budescu, 2012), as it only draws a superficial picture of the actual ethnic classroom composition. Related to the ethnic composition is the idea of heterogeneity: It is often implicitly assumed that a high proportion of ethnic minority students represents a heterogeneous student body. However, the majority-minority-distinction does not provide much information on heterogeneity.

The concept of *heterogeneity* or *diversity* plays a key role in a large number of disciplines, such as ecology (e.g., McCann, 2000), economics (e.g., Hall & Tideman, 1967), organizational psychology (e.g., Hoppe, Fujishiro, & Heaney, 2014; Meyer, in press), communication (e.g., Dimmick & McDonald, 2001), and geography (e.g., Les & Maher, 1998). Their operationalizations of diversity can be used in educational research

as well. Yet, they are less common in this field (as an example see Benner & Crosnoe, 2011). In the current study, diversity is understood as “the distribution of population elements along a continuum of homogeneity to heterogeneity with respect to one or more variables” (Liebersohn, 1969 p. 851 cited in Budescu & Budescu, 2012; cf. Teachman, 1980). This concept is also referred to as “variety diversity” describing differences in group compositions according to a categorical variable (see Harrison & Klein 2007 for a classification of diversity concepts). Existing operationalizations capture the information of diversity with a measure that represents a single, dual, or threefold concept (see Junge, 1994; McDonald & Dimmick, 2003; Stirling, 2007). That is, the measure includes one or more of the following pieces of information: number of categories, distribution of elements across categories, and a numerical distance measure that expresses how similar the various categories are to each other. These pieces of information are combined using relative frequencies (e.g., Simpson’s *D*) or logarithms of those frequencies (e.g., Shannon’s *H*). The present study applies single and dual concept measures that address only the number of categories (i.e., number of ethnicities present in a classroom) or additionally the distribution of elements across categories (i.e., how many students of each ethnicity are there in a classroom, for instance included in the measures Simpson’s *D* and Shannon’s *H*)¹⁵. A detailed overview of the most common diversity measures can be found in Table A3.1 in the appendix¹⁶.

These measures and their comparison have received relatively little attention in educational research thus far. Reviews in other fields showed that some measures are more sensitive to the number of categories (e. g., Junge’s *H*) or to changes in the largest proportion of categories (e. g., Simpson’s *D*) than other measures, and that the measures are close in agreement when using them to quantify diversity (McDonald & Dimmick, 2003). McDonald and Dimmick (2003) concluded that the measures Simpson’s *D* and Shannon’s *H* are most appropriate if one is interested in a measure that is simultaneously sensitive to the number of categories and the maximum proportion of categories. A recent review by Budescu and Budescu (2012) in the educational field that explored two

¹⁵ The present study does not include threefold concepts of diversity measures because we do not focus on numerical distance measures expressing similarity or dissimilarity between countries of birth or ethnicities and because there are no appropriate distance measures available.

¹⁶ The present study covers diversity as a characteristic at the classroom level and does not include operationalizations such as the proportion of students with the same background as an individual student at the student level (see Benner & Crosnoe, 2011; Hoppe et al., 2014).

diversity measures (Simpson's *D* and Shannon's *H*) found that ranking schools according to these two indices did not lead to differences in ranking. However, Shannon's *H* was related slightly stronger to school-level achievement than Simpson's *D*. Overall, there is a dearth of research examining the applicability of different diversity measures in the educational field and how they are related to student outcomes. Therefore, the present study starts by describing a variety of different diversity measures within the preliminary analyses section (see section "Comparison and selection of diversity measures") before including these measures into the analysis models.

6.2.2 Proportion of Ethnic Minority Students and Individual Student Outcomes

Theories and empirical research on the proportion of ethnic minority students typically address effects on student achievement and assume negative relationships between these two characteristics. Research does not suggest that the proportion of minority students per se is related to student outcomes, rather that there are mediating processes and associated aspects that induce these relationships. Interrelated factors that explain why the proportion of ethnic minority students—independent of the prior achievement and socioeconomic composition as well as individual background characteristics—may be negatively related to student achievement outcomes are: (a) school resources, (b) instructional quality, (c) minority language usage, and (d) learning culture.

First, ethnic minority students often have less access to schools with good resources and favorable organizational and structural features such as class size, teacher qualifications, and programs that encourage learning. They are more likely to attend residential neighborhood schools with poor resources in the segregated areas they live in (Betts, Rueben, & Danenberg, 2000; see Palardy, 2013).

Second, classes with high proportions of ethnic minority students may encounter less beneficial learning opportunities in terms of instructional quality, for example, less challenging tasks and a less student-oriented climate. This is based on the assumption that teachers show lower achievement expectations towards ethnic minority students (Ready & Wright, 2011), which, in turn, may cause them to offer fewer challenging learning opportunities and to engage in less positive interactions with these students (for a meta-analysis and review see Den Brok & Levy, 2005; Tenenbaum & Ruck, 2007). In addition,

teachers in segregated neighborhood schools with poor resources that ethnic minority students frequent tend to be less qualified to deliver high quality instruction. These inequalities arise, for example, because the most highly qualified teachers gradually shift to less-disadvantaged schools within an area as they typically have first right of transfer when vacancies appear (Betts et al., 2000).

Third, classrooms with high proportions of ethnic minority students also tend to have high proportions of students who do not speak the language of instruction at home. As a consequence, these students may be less able to support each other by explaining learning materials in the language of instruction. Furthermore, they may not speak the language of instruction with each other in situations such as school breaks, which results in fewer learning opportunities and may negatively affect students' language-related achievement (see Driessen, 2002; Entwisle & Alexander, 1994; Peetsma et al., 2006; Van Ewijk & Slegers, 2010b).

Fourth, ethnic minority students may share values, beliefs, and behaviors associated less with learning and achievement (e.g., negative attitudes towards school, pessimism, and irregular school attendance). Originally focusing on motivation of Afro-American students in the US, Ogbu's cultural ecological theory (1987) assumes that minority students are assigned a subordinate status, devaluated in school and feel vulnerable. As a consequence, these students may come to feel alienated from school, to reject educational values, and to be less motivated to learn (see Kumar & Maehr, 2010; Ogbu, 2004). In classes with a large number of such students, peers transmit these values and beliefs through interacting with each other. Thus, a less beneficial learning culture negatively affecting motivation to learn and achievement may emerge (see Agirdag et al., 2012; Goldsmith, 2011; Palardy, 2013).

The aforementioned factors would predict a negative relationship between the proportion of minority students and student achievement, but a few theories also assume a positive relationship for some psychosocial outcomes, such as feelings of belonging with one's classmates and learning motivation. According to the self-determination theory, social relatedness or belongingness is one of the three basic needs whose fulfillment is assumed to foster motivation to learn (Deci & Ryan, 2000; Niemiec & Ryan, 2009). Especially for minority students, minority versus majority group membership may act as a lens through which individuals in a culturally pluralistic society view each other and on which they build their sense of belonging (see belongingness perspective by Baumeister & Leary, 1995; Johnson et al., 2001; Kumar & Maehr, 2010). This assumption mostly

refers to students with the same specific ethnic background as a source of belongingness, yet, it may also apply to majority versus minority group membership in society. According to social identity theory (Tajfel & Turner, 1986) and the similarity attraction paradigm (Byrne, 1971), group membership and feeling of belonging are based on similarity between students. Minority students may view themselves to be more similar to each other than to majority students for instance in terms of multilingual experiences and immigration history within the family. In their review, Kumar and Maehr (2010) argue that minority students often feel rejected by peers from the majority group in society and therefore show less motivation to learn. By implication, a classroom with a high proportion of minority students should strengthen minority students' feeling of belonging with their peers and facilitate their motivation to learn.

International meta-analyses (Mickelson et al., 2013; Van Ewijk & Slegers, 2010b) commonly find a substantial but small negative effect of the proportion of minority students predicting student achievement. This effect varies in size depending on the minority groups explored, the students' age, the control variables included, and the operationalization of the constructs. Although more research on mediating processes is clearly needed, some findings support the hypotheses of ethnic inequalities in access to schools with high resources (Raudenbush et al., 1998), of lower instructional quality in classrooms with high proportions of ethnic minority students (Palardy, 2015; Stipek, 2004), and a less favorable learning culture in schools with high proportions of ethnic minority students and average low SES (Agirdag et al., 2012; Goldsmith, 2011). Furthermore, a few studies suggest that minority students look forward to instruction more and believe that the topics they learnt will be more useful in the future if they attend schools with high proportions of minority students (Goldsmith, 2004). Furthermore, language minority students are more motivated to learn in language lessons in classrooms with higher proportions of language minority students (Rjosk, Richter, Hochweber, Lütke, & Stanat, 2015).

6.2.3 Ethnic Diversity and Individual Student Outcomes

Theories and empirical research on ethnic diversity focus not only on students' achievement development but also on psychosocial outcomes such as students' feeling of belonging with their peers.

According to Piaget's (1977) concept of disequilibrium, ethnic diversity should have positive effects on students' cognitive development. Being faced with new

information that does not fit into one's schemas—for instance through exposure to multiple perspectives from people with varying ethnic backgrounds—induces a state of unpleasantness. This state drives the learning process through assimilation and accommodation of new ideas and thus fosters cognitive development. As a consequence, several authors assume positive effects of ethnic diversity in the classroom on students' achievement (Benner & Crosnoe 2011; Gottfredson et al., 2009; Gurin et al., 2003; Tam & Bassett, 2004). This line of argument is similar to the information/decision-making perspective taken in organizational psychology to explain positive effects of work team diversity (see Meyer, in press). Studies in educational research found, for instance, that students in ethnically more heterogeneous kindergartens showed higher achievement levels in mathematics and reading, after controlling for the socioeconomic composition, proportion of minority students and other school characteristics (Benner & Crosnoe, 2011). Students from ethnically diverse high schools also showed university GPAs in the first semester that were a one-fourth to one-half point higher than that of students from a non-diverse high school, after controlling for achievement composition and quality of high schools (Tam & Bassett, 2004).

Although positive relationships are assumed between ethnic diversity on the one hand and cognitive development as well as school achievement on the other, a negative association between diversity and students' feeling of belonging with their classmates may exist (Benner & Crosnoe, 2011; Benner et al., 2008). According to the belongingness perspective (see Baumeister & Leary, 1995; Byrne, 1971; Tajfel & Turner, 1986) ethnic diversity should be negatively associated with attachment to the peers. These theories assume that and students' sense of belonging is more strongly related to the specific ethnicity of the peers than to the broader category of minority status. For instance, in a study on US ninth-graders, students perceived the school climate to be fairer and more directed towards academics and interracial understanding if they attended ethnically less diverse schools (Benner et al., 2008). Furthermore, ethnically less diverse schools and classrooms were characterized by stronger attachment to school (Johnson et al., 2001) and lower levels of perceived cultural discrimination (Seaton & Yip, 2009).

In sum, educational research commonly assumes a negative relationship between the proportion of ethnic minority students and student achievement. Empirical findings indicate that the relationships may be as predicted, yet there is not much research on the

underlying mechanisms inducing compositional effects. However, some studies also find a positive relationship between ethnic classroom diversity and students' achievement.

To conclude, the two strands of theories and findings presented so far would lead to contradictory predictions: A high proportion of ethnic minority students in a classroom has been shown to be negatively related to student achievement. Simultaneously, a high proportion of ethnic minority students partly corresponds to a higher ethnic diversity which is assumed to be positively related to school achievement. This raises the question of how the ethnic composition is related to school achievement when one tries to disentangle the effects of the proportion of ethnic minority students and ethnic diversity. Furthermore, we are interested in the question how ethnic diversity and the broader distinction between ethnic majority and ethnic minority students are related to students' feeling of belonging with their classmates as an example of a non-cognitive student outcome.

6.3 The Present Study: Research Questions and Hypotheses

The aim of the present study is to examine the relationship between the ethnic makeup of classrooms and student achievement in mathematics and reading comprehension and students' feeling of belonging with their classmates. The competing theoretical assumptions and empirical findings presented in the last sections form the basis of our study. We examine the relationships between student outcomes and the proportion of minority students as well as ethnic diversity in German elementary school classrooms in a cross-sectional design. The average SES in a classroom and the average prior achievement are classroom level background variables that have been shown to matter for individual achievement outcomes in former studies (e.g., Van Ewijk & Slegers, 2010a). Our analyses include these variables as covariates. We take the average prior achievement into account using the average cognitive abilities in the classroom as a proxy. Our research questions and hypotheses are as follows:

1) Relationship with achievement scores

1a) Is the proportion of ethnic minority students in a classroom related to students' achievement in mathematics and reading comprehension? We predict that the proportion of minority students will be negatively related to achievement outcomes. The underlying assumption is that in classrooms with high proportions of minority students there is a less favorable learning environment characterized by poor school resources, lower instructional quality, non-German language usage with peers, and less favorable learning

culture (see section “Proportion of ethnic minority students and individual student outcomes”).

1b) Is ethnic diversity in the classroom—operationalized by various diversity measures—related to students’ achievement in mathematics and reading comprehension? Contrary to some studies reviewed in the last sections, we predict a negative relationship between ethnic diversity measures and achievement for the same reasons as those as described in Hypothesis 1a.

1c) Do diversity measures explain additional variance in student achievement over and above the proportion of ethnic minority students? We predict that measures of ethnic diversity provide additional information on the classroom composition and therefore will be related to student achievement outcomes over and above the proportion of ethnic minority students in a classroom. We furthermore assume that controlling for the proportion of ethnic minority students in a classroom and other background characteristics also controls for characteristics of the learning environment associated with classroom composition. Holding that constant might offer the opportunity to investigate if diversity has also positive effects on achievement as predicted in the literature. We predict additional positive effects of ethnic diversity.

2) Relationship with feeling of belonging with the peers

2a) Is the proportion of ethnic minority students related to students’ feeling of belonging with their classmates? We assume that the proportion of ethnic minority students reveals different relationships with the feeling of belonging for minority students than for majority students. According to the belongingness hypothesis, minority students should feel more attached in classrooms with high proportions of minority students and majority students should feel more attached in classrooms with high proportions of majority students. That is, we predict an interaction between the proportion of minority students and individual minority status. In line with this assumption, the proportion of minority students should be on average related to the feeling of belonging with one’s peers in a u-shaped manner.

2b) Is ethnic diversity in the classroom—operationalized by various diversity measures—related to students’ feeling of belonging with their classmates? We assume that ethnic diversity is negatively related to the feeling of belonging because a large diversity in a classroom corresponds to a low number of students from the same ethnic background (see section “Ethnic diversity and individual student outcomes”).

2c) Do diversity measures explain additional variance in students' feeling of belonging with the peers over and above the proportion of ethnic minority students? We assume that ethnic diversity is more strongly related to the feeling of belonging than the proportion of minority students as students should build their sense of belonging on the specific ethnicity of the peers in their classroom rather than on the broader category of minority status (see section "Ethnic diversity and individual student outcomes"). When analyzed together as independent variables, we predict that ethnic diversity will still show a significant association with the feeling of belonging while the proportion of minority students in a classroom loses its significance.

6.4 Method

6.4.1 Participants

Our analyses are based on data from a nationally representative sample of elementary school students in Germany who participated in the 2011 National Assessment Study of student achievement in elementary schools [*IQB-Ländervergleich*] (Stanat, Pant, et al., 2012) of the German Institute for Educational Quality Improvement (IQB). The data include 27,081 students of complete 4th grade classrooms in 1,349 randomly selected German public schools (see Richter et al., 2012).

We excluded special-needs schools, schools from the former GDR—because they have very low proportions of ethnic minority students (see Federal Statistical Office Germany, 2012)—and classrooms with a large number of missing values for ethnic background information (see section "Missing data treatment"). Consequently, our analyses were based on 18,762 students attending 903 classrooms in 903 schools (average number of students per classroom $M = 21$). For a sample description see Table 6.1.

Table 6.1. Descriptive Sample Statistics for Demographic Variables

	<i>M (SD)</i>	Min.	Max.
Individual level (L1), <i>N</i> = 18,762:			
Ethnic minority status	37.60 %	–	–
SES	50.23 (16.19)	10	89
Female	49.50 %	–	–
Age in years	10.41 (0.50)	6.83	13.17
classroom level (L2), <i>N</i> = 903:			
Proportion of German background	60.52% (22.08)	0%	100%
Proportion of Turkish background	6.53 % (9.50)	0%	61.11%
Proportion of former USSR background	4.91% (7.63)	0%	55.56%
Proportion of Polish background	2.13% (4.00)	0%	41.67%
Proportion of former Yugoslavia background	2.96% (4.61)	0%	29.41%
Proportion of other background	9.71% (9.09)	0%	52.94%
Proportion of students with missing background information	13.24% (10.67)	0%	48.15%
Proportion of ethnic minority students	39.48% (22.08)	0%	100%
SES	49.64 (8.15)	25.75	78.07

Note. SES= socioeconomic status. For operationalization of SES and students' ethnic background see measures section.

6.4.2 Measures

We use information from standardized achievement tests, student questionnaires and parent questionnaires.

Student level background variables. *Ethnic background of students.* We categorized the ethnic background of students using information from the parent questionnaire. If the parent response was missing, we used information from the student questionnaire (see section “Missing data treatment”). The ethnic background was categorized based on the country where the parents were born. To be categorized as a student with minority status, at least one parent had to be born abroad. For instance, a student with a Turkish background has parents who were both born in Turkey or one parent in Turkey and one parent in Germany. If one parent was born in Turkey and one parent in another country (i.e. not Germany), the student was assigned to the category

“other country”. Within the category “other country”, countries most represented were Iran and Arab countries. In our analyses we distinguish six groups (see Table 6.1) corresponding to the largest groups in this sample (see Stanat, Pant, et al., 2012). If questionnaire information was completely missing, we did not exclude the student—as it would distort the classroom level analyses—but assigned him or her to the category “unidentifiable” which we included in the analyses. For analyses comparing broadly ethnic majority and ethnic minority students, students with a German background (i.e. both parents born in Germany) were categorized as “ethnic majority” and the remaining groups as “ethnic minority”.

Students’ socioeconomic status (SES). As the measure of SES we used the HISEI or highest “International Socio-Economic Index of Occupational Status” (Ganzeboom et al., 1992; Ganzeboom, 2010). This index is a classification of parents’ occupation based on income and education with a score range of 10 (e.g., a kitchen helper) to 89 (e.g., a medical doctor). We used information about the current occupation that the parents provided in the questionnaire in an open answer format.

Prior achievement. To approximate students’ prior achievement within the cross-sectional design, we used the figural subtest of the cognitive abilities test (KFT 4–12+R, Heller & Perleth, 2000; see Baumert et al., 2006 on the validity of such a proxy). The figural subtest consists of 25 items that ask the students to choose one figure out of several possibilities in analogy to a given pair of figures.

Gender. Student gender was recorded in the tracking form completed by the classroom teacher (dummy coding; male = 0, female = 1).

Classroom level independent and background variables. The key classroom level variables of this study were the proportion of ethnic minority students in a classroom and ethnic diversity. Covariates were the classroom level SES and prior achievement.

Proportion of ethnic minority students. We calculated the proportion of ethnic minority students in each classroom in accordance with the individual student level categorizations described above as relative frequency (see below for an example).

Ethnic diversity. We operationalized ethnic classroom diversity calculating various diversity measures (see Table A6.1) based on individual ethnic background information described above.

In the following, we explain how four exemplary measures to operationalize the ethnic makeup of classrooms are computed: the proportion of ethnic minority students and three diversity measures—number of categories, Simpson’s D , and Shannon’s H .

Imagine two fictitious classrooms, classroom A and classroom B. Both classrooms have 20 students. In classroom A there are 6 students with German background, 13 with Turkish background, and 1 with Polish background. In classroom B there are 6 students with German background, 5 with Turkish background, 4 with parents from the former USSR, 4 with parents from the former Yugoslavia, and 1 with another ethnic background. The *proportion of ethnic minority students* in classroom A is 0.7 ($\text{prop}_A = (13+1)/20 = 0.7$) and in classroom B it is also 0.7 ($\text{prop}_B = (5+4+4+1)/20 = 0.7$). The *number of ethnic groups* in classroom A equals 3 ($N_{\text{cat}_A} = \text{German, Turkish, Polish} = 3$) and in classroom B it equals 5 ($N_{\text{cat}_B} = \text{German, Turkish, former USSR, former Yugoslavia, other} = 5$). Simpson’s D represents the probability that two students selected at random from a classroom belong to different ethnicities—thus the greater the value of D , the greater the diversity. Its minimum value is 0 and its maximum is achieved when the distribution across the c ethnic groups in the classroom is uniform (in our study $c = 7$, i.e. $D_{\text{max}} = (7-1)/7 = 0.857$). Shannon’s H involves a logarithmic transformation of probabilities ($H_{\text{min}} = 0$; H_{max} in our study $= \ln(c) = \ln(7) = 1.946$). Simpson’s D and Shannon’s H both use the relative frequencies (p_i) of each ethnic group in the classroom. For classroom A the relative frequencies are $p_{A,\text{German}} = 0.3$, $p_{A,\text{Turkish}} = 0.65$, and $p_{A,\text{Polish}} = 0.05$ and for classroom B they are $p_{B,\text{German}} = 0.3$, $p_{B,\text{Turkish}} = 0.25$, $p_{B,\text{USSR}} = 0.2$, $p_{B,\text{Yugoslavia}} = 0.2$, and $p_{B,\text{other}} = 0.05$. The measure *Simpson’s D* for classroom A equals $D_A = 0.485$ ($D_A = 1 - \sum p_i^2 = 1 - ((0.3 \times 0.3) + (0.65 \times 0.65) + (0.05 \times 0.05)) = 1 - 0.515 = 0.485$) and for classroom B $D_B = 0.765$ ($D_B = 1 - \sum p_i^2 = 1 - ((0.3 \times 0.3) + (0.25 \times 0.25) + (0.2 \times 0.2) + (0.2 \times 0.2) + (0.05 \times 0.05)) = 1 - 0.235 = 0.765$). The measure *Shannon’s H* for classroom A equals $H_A = 0.791$ ($H_A = - \sum p_i \ln(p_i) = - ((0.3 \times \ln(0.3)) + (0.65 \times \ln(0.65)) + (0.05 \times \ln(0.05))) = 0.791$) and for classroom B it equals $H_B = 1.501$ ($H_B = - \sum p_i \ln(p_i) = - ((0.3 \times \ln(0.3)) + (0.25 \times \ln(0.25)) + (0.2 \times \ln(0.2)) + (0.2 \times \ln(0.2)) + (0.05 \times \ln(0.05))) = 1.501$).

In sum, our example shows that, even though the proportion of minority students in classroom A and B are the same, their student body differs in terms of ethnic homogeneity. Classroom B is ethnically more diverse than classroom A, which is

reflected in a larger number of ethnicities in the classroom and higher values of Simpson's D and Shannon's H (see Table 6.2).

Table 6.2 Ethnic makeup of two different classrooms: exemplary computations

	classroom A	classroom B
Proportion of minority students	.70	.70
Number of ethnic groups	3	5
Simpson's D	.485	.765
Shannon's H	.791	1.501

Note. Both classrooms involve 20 students but differ in their distribution of ethnic makeup (classroom A: 6, 13, and 1; classroom B: 6, 5, 4, 4, and 1).

Classroom level socioeconomic status. We aggregated the average SES score for each classroom based on the individual student scores (see above).

Classroom level prior achievement. We aggregated the average cognitive abilities test score for each classroom as a proxy for prior achievement based on the individual student scores (see above).

Outcome variables. Student achievement in reading comprehension and mathematics. Trained test administrators conducted standardized achievement tests in German reading comprehension and mathematics in the classrooms. The tests were designed by a team of experienced teachers and scientists in partnership with the German Institute for Educational Quality Improvement (IQB) to measure these achievement domains in accordance with the German national educational standards. Task units measuring reading comprehension consisted of a literary or factual text and several items mainly presented as multiple choice and short answer questions. Each student received two to four out of 11 task units (booklet design).

The mathematics achievement test covered the five content domains “numbers and operations”, “space and shape”, “patterns and structures”, “measurement”, and “probability” using a variety of tasks, such as simple computations, extracting information from charts, and reflecting shapes (see Winkelmann, Heuvel-Panhuizen, & Robitzsch, 2008).

A generalized Rasch model was used to estimate student achievement scores on a common scale for each achievement domain. Mathematics and reading scores were generated using the plausible values (PV) technique (Adams et al., 2007). Fifteen PVs

were generated for each student in each domain, which were scaled to have a mean score of 500 and a standard deviation of 100 in the German student population (current sample distribution mathematics: $M = 489.05$, $SD = 96.00$, reading: $M = 494.12$, $SD = 90.87$). EAP/PV reliabilities in the calibration model were .91 (mathematics) and .73 (reading).

Feeling of belonging with one's peers. Students rated their feelings of belonging with their classmates in the student questionnaire. The scale is part of a questionnaire measuring emotional and social experiences in school (Rauer & Schuck, 2003). It consists of four items rated on a 4-point Likert scale (1 “fully disagree” to 4 “fully agree”) by all students. Item examples are “My classmates are nice to me.” and “When I am sad, my classmates comfort me.” (manifest $M = 3.3$, $SD = 0.6$, Cronbach’s Alpha = .71).

6.4.3 Data Analysis

As preliminary analyses, we calculated eight diversity measures (see Table A3.1) and analyzed their relationship with each other using Pearson correlations (see section “Comparison and selection of diversity measures”).

We used structural equation modeling to explore our research questions 1 and 2 (see Bovaird, 2007). Several random intercept multilevel structural equation models were estimated to analyze the relationship between (I) the proportion of ethnic minority students in a classroom and student outcomes, (II) ethnic diversity operationalized by various measures and student outcomes, as well as (III) both classroom characteristics and student outcomes. We employed a stepwise model building procedure. We used the software Mplus (Version 6.1; Muthén & Muthén, 1998-2010) for all analyses.

Metric background variables at the student level (SES, proxy for prior achievement) were standardized, which implies centering at their grand mean. Categorical variables (ethnic minority status, female) were neither centered nor standardized. The estimates of background variables aggregated at the classroom level can be interpreted as compositional effects (Raudenbush & Bryk, 2002). Classroom level variables were standardized at the classroom level. We additionally included the quadratic term of the proportion of ethnic minority students in a classroom, in order to explore a potential nonlinear relationship between the proportion and student outcomes. Prior to calculating the quadratic term, we centered the proportion of ethnic minority students at its mean to counteract multicollinearity. All regression coefficients in the result tables were standardized using the total variance (within + between) of the outcome variable.

For addressing research question 2, we additionally used a doubly latent approach with cross-level measurement invariance for the construct of feeling of belonging with one's peers. This approach comprises latent measurement models at both levels and latent aggregation for the classroom level construct (e.g., Lüdtke et al., 2011; Marsh et al., 2012). An important advantage of a model with these features is that it corrects possible measurement and sampling errors associated with designs in which variables measured at the individual level are used to operationalize a construct at the classroom level. The focus of our analyses was at the classroom level. However, we conducted additional cross-level interaction analyses for research question 2a on the association between individual ethnic minority versus majority status and the relationship between the proportion of minority students and feeling of belonging with one's peers.

Missing data treatment. Students in the sampled schools were obliged to participate in the achievement tests. Yet, individual students could be excluded from the study by the school if they met one of the following three criteria: (a) students with permanent physical impairment that made it impossible to participate, (b) severe intellectual or emotional impairment, and (c) students who were less than one year in Germany and could neither speak nor read in German. The response rate of the student questionnaire in the total sample was 87.3%, that is, it was lower than the response rate of achievement tests of 98.3% because participation was not mandatory in some federal states of Germany. The rate of the questionnaire varied between 76% in Hamburg and 98% in Hesse. The response rate for the parents questionnaires was 81.4%.

We excluded special-needs schools ($N = 51$ schools), schools from the former GDR ($N = 398$ schools) and classrooms with a large number of missing values concerning ethnic background information ($> 50\%$; $N = 17$ schools) from the analyses. This led to a sample size of 19,457 students in 908 schools. The remaining sample did not include any missing values on the ethnic background variable because we classified missing information as unidentifiable and kept the student in the data set. Information on student gender was missing for 0.64 % of the sample, on the proxy for prior achievement for 6.91 %, on mathematics achievement for 4.75 %, on reading achievement for 4.77 %, on SES for 29.62 %, and on all four belonging items for 19.27 % of the sample. To deal with item nonresponse, we used the full information maximum likelihood (FIML) estimator implemented in Mplus for all variables except student gender. This estimator applies a model-based approach to missing data (see Enders, 2010), using all information available from the model variables to estimate the model parameters. In doing so, we

were able to use 96.43 % of the intended sample. A total of 3.57 % of the students had either missing gender information or missing values on all estimated variables (SES, proxy for prior achievement, and outcome variables) and was excluded during the analyses.

6.5 Results

6.5.1 Preliminary Analyses

6.5.1.1 Comparison and selection of diversity measures

In a first step we computed the composition and diversity measures (see section “Classroom level independent and background variables” and Table A3.1). Inspection of bivariate Pearson correlations at the classroom level (see Table A3.2 in the appendix) showed that the measures of ethnic diversity are highly correlated with each other and that they are also highly correlated with the commonly used “majority-minority approach”, that is, the proportion of minority students in a classroom (first column in Table A3.2).

To gain further insight into the relationship between the proportion of minority students in a classroom and the diversity measures, we plotted the proportions of ethnic minority students for each classroom and the values for Simpson’s D as an example (see Figure 6.1). The two measures are directly dependent on one another: In a classroom with very low proportions of minority students and consequently very high proportions of majority students, ethnic diversity is lower. In classrooms with less or equal to 50% of minority students, the correlation between the proportion of minority students and Simpson’s D is $r = .99$ ($p < .01$) and in classrooms with a proportion of minority students greater than 50% it is $r = .33$ ($p < .01$). That is, classrooms with high proportions of minority students vary in their ethnic heterogeneity. The correlation pattern for other diversity measures was comparable.

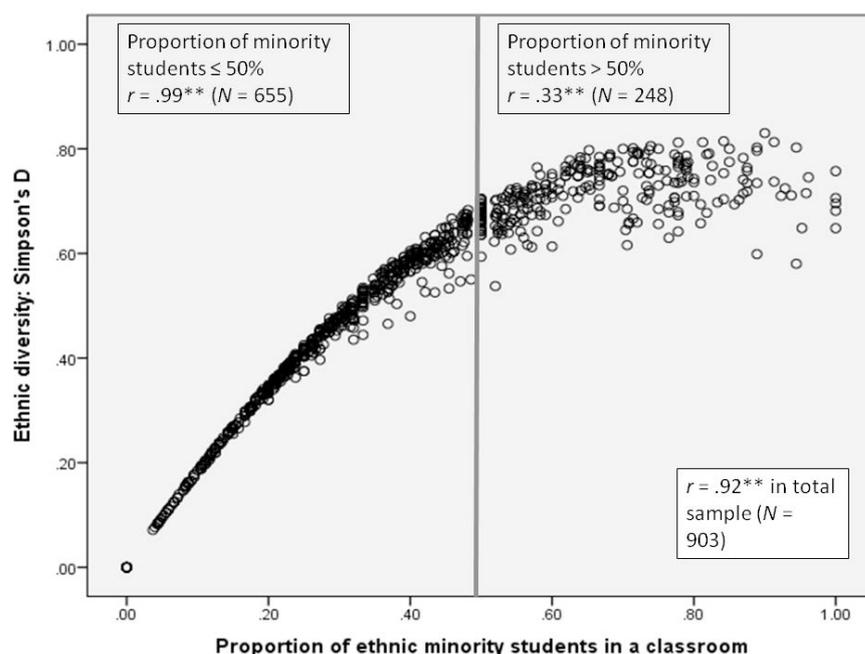


Figure 6.1: Joint distribution of the proportion of ethnic minority students in a classroom and Simpson's D for $N = 903$ classrooms. (Students with only German background, i.e. majority students, were counted as one ethnic category when calculating Simpson's D).

As the high intercorrelations between the proportion of minority students in a classroom and the various measures of diversity may cause problems of multicollinearity and make it difficult to disentangle effects of these two classroom characteristics, we decided to calculate the diversity measures without counting the proportion of majority students—that is German students—as a category. These measures thus depict the diversity of the proportion of ethnic minority students. This approach shows similar results to analyses that are based on a reduced sample of classrooms with a proportion of minority students greater than 50% but simultaneously allows us to use the complete sample with its greater power. The intercorrelations among the optimized measures as well as their bivariate correlation with the independent and outcome variables can be found in Table A3.3 in the appendix. These analyses indicate that the correlation patterns for the various measures of ethnic diversity are very similar to each other. For reasons of parsimony, we only show the main analyses for three exemplary measures of diversity in the result section of this paper (for results of further analyses with the remaining measures see appendix Table A3.7 to A3.9).

6.5.1.2 Measurement model fit and variance of outcome variables between classrooms

For the feeling of belonging with one's classmates as an outcome, we first explored the fit of the doubly latent model with cross-level measurement invariance which showed acceptable model fit ($\chi^2 = 379.680$, $df = 7$, $p < .05$, RMSEA = .058, CFI = .951, SRMR_{within} = .037, SRMR_{between} = .041). Such an unconditioned model without any predictors provides information on the amount of variance at both levels necessary to compute the intraclass correlation (ICC1). The ICC1 estimates the proportion of the total variance due to differences between classrooms. If there was no variation in motivation between classrooms, multilevel analyses would not be meaningful. In our study the proportion of variance between classrooms was 9%. As is commonly the case for non-cognitive variables, this proportion is lower than the variation typically found for achievement between classrooms (see Trautwein et al., 2006). This was also true for the current analyses, for mathematics achievement the proportion of variance between classrooms was 22% and for reading achievement 20%.

6.5.2 Relationship between the Proportion of Ethnic Minority Students, Ethnic Diversity and Individual Student Outcomes

6.5.2.1 Proportion of ethnic minority students in a classroom and student achievement

The first set of our multilevel analyses explored the relationship between measures of the ethnic makeup of classrooms and students' individual achievement in mathematics and reading comprehension. The result pattern of the classroom level variables is shown in Table 6.3 for mathematics and in Table 6.4 for reading comprehension as an outcome (for individual student level results see Table A3.4 and A3.5 in the appendix). In a first step (Model M.1 in Table 6.3 and Model R.1 in Table 6.4, see research question 1a), we investigated the association between the proportion of ethnic minority students and achievement controlling for individual students' ethnic background, gender, proxy for prior achievement, and SES. The results show that—independent of individual background characteristics—a student in a classroom with a one standard deviation higher proportion of ethnic minority students reached mathematics scores that were 0.15 standard deviations lower on average than a student in a class with a low proportion of ethnic minority students (Model M.1 in Table 6.3). Furthermore, the significant regression weight of the quadratic term of the proportion of ethnic minority students in model M.1b indicates a slightly u-shaped relationship with a tendency of lower

mathematics scores in classrooms with very low or very high proportions of ethnic minority students. The same result pattern emerged for reading comprehension as an outcome (see Model R.1 and R.1b in Table 6.4).

6.5.2.2 Ethnic diversity in a classroom and student achievement

The next models (M.2 to M.4 in Table 6.3 and R.2 to R.4 in Table 6.4) show results for selected diversity measures each as single predictor at the classroom level (see research question 1b). The first line of the result tables indicates the respective diversity measure used (Ncat = number of ethnicities in the classroom, SimD = Simpson's D , ShaH = Shannon's H ; see Table A3.7 to A3.9 in the appendix for analyses with further diversity measures). Their coefficients are slightly smaller in size than the coefficient for the proportion of minority students in the classroom, but they also show a negative relationship with individual mathematics and reading achievement. That is, in ethnically more diverse classrooms, students reach slightly lower levels of mathematics and reading achievement scores.

Table 6.3. Results of Multilevel Structural Equation Models Predicting Mathematics achievement (Classroom Level Results; for Student Level Results see Table A3.4)

	Model M.1	Model M.1b	Model M.2 (Ncat)	Model M.3 (SimD)	Model M.4 (ShaH)	Model M.5 (Ncat)	Model M.6 (SimD)	Model M.7 (ShaH)	Model M.8 (Ncat)	Model M.9 (SimD)	Model M.10 (ShaH)
minority students %	-.17**	-.16**	-	-		-.19**	-.18**	-.18**	-.15**	-.14**	-.14**
minority students %: quadratic term	-	-.04**	-	-	-	-.03	-.03	-.03	-.00	-.00	-.00
diversity measure	-	-	-.08**	-.05**	-.06**	.04*	.04*	.05*	.04*	.03*	.03*
prior achievement (mean)	-	-	-	-	-	-	-	-	.07**	.07**	.07**
SES (mean)	-	-	-	-	-	-	-	-	.06**	.06**	.06**
R ² L2	0.25	0.26	0.05	0.02	0.03	0.27	0.27	0.27	0.39	0.39	0.39
R ² L1	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.34	0.34	0.34

Note. * $p < .05$, ** $p < .01$, Ncat = number of categories (i. e. ethnicities), SimD = Simpson's D , ShaH = Shannon's H , SES = socioeconomic status, L1 = student level, L2 = classroom level; line "diversity measure" shows coefficients of the respective measure indicated in the first line. Covariates at the student level: ethnic background, SES, proxy for prior achievement, and gender. Regression coefficients were standardized by the total variance (within + between) of the outcome variable.

Table 6.4. Results of Multilevel Structural Equation Models Predicting German Reading Achievement (Classroom Level Results; for Student Level Results see Table A3.5)

	Model R.1	Model R.1b	Model R.2 (Ncat)	Model R.3 (SimD)	Model R.4 (ShaH)	Model R.5 (Ncat)	Model R.6 (SimD)	Model R.7 (ShaH)	Model R.8 (Ncat)	Model R.9 (SimD)	Model R.10 (ShaH)
minority students %	-.15**	-.14**	-	-	-	-.13**	-.13**	-.13**	-.10**	-.10**	-.10**
minority students %: quadratic term	-	-.04**	-	-	-	-.04*	-.04*	-.04*	-.01	-.01	-.01
diversity measure	-	-	-.08**	-.06**	-.07**	-.01	-.00	-.01	.00	.00	.00
prior achievement (mean)	-	-	-	-	-	-	-	-	.05**	.05**	.05**
SES (mean)	-	-	-	-	-	-	-	-	.07**	.07**	.07**
R ² L2	0.21	0.23	0.07	0.03	0.05	0.23	0.23	0.23	0.30	0.30	0.30
R ² L1	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.25	0.25	0.25

Note. * $p < .05$, ** $p < .01$, Ncat = number of categories (i. e. ethnicities), SimD = Simpson's D , ShaH = Shannon's H , SES = socioeconomic status, L1 = student level, L2 = classroom level; line "diversity measure" shows coefficients of the respective measure indicated in the first line. Covariates at the student level: ethnic background, SES, proxy for prior achievement, and gender. Regression coefficients were standardized by the total variance (within + between) of the outcome variable.

6.5.2.3 Relationship between the proportion of ethnic minority students, ethnic diversity, and student achievement

When the proportion of ethnic minority students in a classroom and a diversity measure are analyzed simultaneously as predictors at the classroom level (Models M.5 to M.7 in Table 6.3 and Models R.5 to R.7 in Table 6.4; see research question 1c) and also controlling for the socioeconomic composition and prior achievement level (Models M.8 to M.10 in Table 6.3 and Models R.8 to R.10 in Table 6.4), the coefficients for the proportion of ethnic minority students remain negative and significant for mathematics achievement and reading comprehension as an outcome. However, the associations between ethnic diversity in the classroom and student achievement are different from the single predictor models: After taking into account the proportion of minority students and the classroom composition with regard to prior achievement and SES, respectively, ethnic diversity was positively related to individual mathematics achievement (Models M.5 to M.10 in Table 6.3)¹⁷. The identical analyses using reading achievement as an outcome (Models R.5 to R.10 in Table 6.4) showed no significant effect of ethnic diversity. In all analyses, the proportion of explained variance did not differ much between models using various measures of ethnic diversity.

6.5.2.4 Relationship between the proportion of ethnic minority students, ethnic diversity, and students feeling of belonging with their peers

Analogous to the models presented so far, Table 6.5 shows the results for analyses exploring students' feeling of belonging with their peers as an outcome. The models using the proportion of ethnic minority students or a diversity measure as single predictor at the classroom level controlling for individual student characteristics showed negative relationships with students' feeling of belonging with their peers (Models B.1 to B.4; see research questions 2a and 2b). The strongest predictor was the proportion of ethnic minority students in a classroom. Its non-significant quadratic term showed that the relationship between the proportion of ethnic minority students and feeling of belonging was linear (Model B.1b).

¹⁷ The predictors at the classroom level were not highly interrelated, thus multicollinearity was not a concern (Variance inflation factor [VIF] for proportion of minority students = 1.53, for Simpson's D = 1.38, for average prior achievement = 1.32, and for average SES = 1.27). For correlation tables see appendix Table A3.3.

Table 6.5. Results of Multilevel Structural Equation Models Predicting Feeling of Belonging with one’s Peers (Classroom Level Results; for Student Level Results see Table A3.6)

	Model B.1	Model B.1b	Model B.2 (Ncat)	Model B.3 (SimD)	Model B.4 (ShaH)	Model B.5 (Ncat)	Model B.6 (SimD)	Model B.7 (ShaH)	Model B.8 (Ncat)	Model B.9 (SimD)	Model B.10 (ShaH)
minority students %	-.09**	-.08**	-	-	-	-.10**	-.10**	-.10**	-.07**	-.07**	-.08**
minority students %: quadratic term	-	-.01	-	-	-	-.00	.00	.00	.00	.01	.01
diversity measure	-	-	-.04**	-.02	-.03*	.02	.02	.02	.01	.02	.02
prior achievement (mean)	-	-	-	-	-	-	-	-	.08**	.08**	.08**
SES (mean)	-	-	-	-	-	-	-	-	-.02	-.02	-.02
R ² L2	0.11	0.12	0.03	0.01	0.01	0.12	0.12	0.12	0.19	0.19	0.19
R ² L1	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03

Note. * $p < .05$, ** $p < .01$, Ncat = number of categories (i. e. ethnicities), SimD = Simpson’s D , ShaH = Shannon’s H , SES = socioeconomic status, L1 = student level, L2 = classroom level; line “diversity measure” shows coefficients of the respective measure indicated in the first line. Covariates at the student level: ethnic background, SES, proxy for prior achievement, and gender. Regression coefficients were standardized by the total variance (within + between) of the outcome variable.

Additional cross-level interaction analyses not presented in Table 6.5 revealed that individual ethnic minority students felt a stronger sense of belonging with their peers in classrooms with a higher proportion of ethnic minority students and ethnic majority students a stronger sense of belonging in classrooms with a higher proportion of ethnic majority students (interaction term: $\beta = .23$, $SE = .04$, $p < .01$). In Models B.2 to B.4, the regression coefficients of the diversity measures were rather small and reached statistical significance only in some models. However, these measures represent diversity within the group of ethnic minority students. Models using diversity measures including German students as one category (not presented in Table 6.5) partly showed stronger associations (for Ncat: $\beta = -.04$, $SE = .01$, $p < .01$; for SimD: $\beta = -.07$, $SE = .01$, $p < .01$; for ShaH: $\beta = -.07$, $SE = .01$, $p < .01$). When analyzed simultaneously as predictors at the classroom level (Models B.5 to B.7; see research question 2c) and also including classroom level covariates (Models B.8 to B.10), the coefficients of the proportion of ethnic minority students remained negative and significant and ethnic diversity within the group of minority students was not significantly related to student's feeling of belonging with their peers.

6.6 Discussion

The present study investigated the relationship between various measures of ethnic composition and heterogeneity or diversity used in different disciplines on the one hand and achievement and psychosocial student outcomes on the other hand. The aim was to explore whether measures of ethnic diversity are related to student outcomes over and above commonly investigated characteristics of classroom composition.

In a first step we therefore collected detailed information on possible diversity measures from research conducted in disciplines such as communication, geography, and biology. Our preliminary analyses comparing these measures operationalizing the ethnic makeup of classrooms led to the conclusion that they are highly intercorrelated (see McDonald & Dimmick, 2003). That is why we considered a selection of diversity measures within our main analyses. Overall, using different diversity measures as independent variable predicting student outcomes led to comparable results. Furthermore, the proportion of minority students—which is mostly used in educational research to describe the ethnic makeup—was highly correlated with diversity measures of the complete student population in a classroom. Thus, even if diversity and proportion of minority students are not the same from a content perspective, they should lead to

comparable associations with student outcomes. In our study, we adapted the diversity measures to represent diversity only among minority students.

Our main analyses showed that students had lower achievement scores in classrooms with a higher proportion of ethnic minority students (see Hypothesis 1a). These findings are in line with international research (Mickelson et al., 2013; Van Ewijk & Slegers, 2010b). Possible mediating factors—which were not the focus of the present study—that may induce such findings are instructional quality, motivational processes among peers, non-German language spoken with peers, and school resources (Agirdag et al., 2012; Palardy, 2015; Raudenbush et al., 1998; Stipek, 2004; Van Ewijk & Slegers, 2010b). Similarly, students had lower achievement scores in ethnically more diverse classrooms. This result is in line with our hypothesis (see Hypothesis 1b) but contradicts studies showing an advantage for academic achievement development in ethnically diverse classrooms (e.g., Benner & Crosnoe, 2011). However, when the proportion of ethnic minority students and a diversity measure were analyzed as joined predictors, controlling for average SES and average prior achievement in the classroom, we found different patterns of results: After accounting for differences in the proportion of minority students, ethnic diversity in the classroom was not significantly related to reading achievement, but did show a positive association with students' level of mathematics achievement. This is partly in line with assumptions of advantaged achievement development in ethnically diverse classrooms (Benner & Crosnoe, 2011; Gottfredson et al., 2009; Gurin et al., 2003; Tam & Bassett, 2004). Thus, a positive effect for mathematics achievement became visible only after controlling for the proportion of minority students and the level of socioeconomic status and prior achievement at the classroom level. We assume that controlling for these classroom characteristics also controlled for the less favorable learning environment and cumulated disadvantages in classrooms that are associated with a higher proportion of ethnic minority students.

The question arises why this pattern emerged only for mathematics achievement as an outcome and not for reading comprehension. If the pattern was only a consequence of methodological issues such as multicollinearity the positive effect of diversity should have emerged for both achievement scores. The models have identical independent variables and additional analyses showed no worrying variance inflation and intercorrelations. It is possible that the assumed benefits of ethnic classroom diversity develop more easily for mathematics achievement because it is more tied to instruction than reading comprehension. Reading is more tied to out-of-school-activities and

individual family background (Baumert & Schümer, 2001). Alternatively, it could be the case that a positive effect of ethnic diversity emerges for a large number of achievement outcomes but did not for reading comprehension because reading was tested in German and ethnic minority students might speak not German at home. In addition we can assume that in ethnically diverse classrooms there is a larger number of different language backgrounds present which might be related to different kinds of student difficulties and also different cultural background knowledge needed to understand texts. In such classrooms it might be more difficult for the teacher to react to all students' needs during language instruction, exacerbating positive diversity effects.

Our multilevel analyses taking students' feeling of belonging with one's peers as an outcome revealed that students felt less related to their classmates in classrooms with a high proportion of ethnic minority students and in ethnically more diverse classrooms. Additional analyses showed that individual minority students felt a stronger sense of belonging in classrooms with higher proportions of minority students and the opposite was true for individual ethnic majority students. This finding was in line with the belongingness perspective predicting higher sense of belonging in more homogeneous groups (Baumeister & Leary, 1995; Benner & Crosnoe, 2011; Byrne, 1971; Tajfel & Turner, 1986). Interestingly, the diversity measures were not more strongly related to the feeling of belonging than the proportion of ethnic minority students was. Studies analyzing ethnic composition and belonging commonly argue that students build their belonging based on the specific ethnic background (see Benner & Crosnoe, 2011) rather than on the broader distinction between ethnic minority and majority. Our finding did not support this argument related to diversity among minority students.

6.6.1 Limitations and Future Research

The present study has five important limitations. First, we analyzed data from a cross-sectional design which renders it impossible to make statements about the origins and further development of the classroom effects. One important background characteristic of classrooms that determines future student achievement is the average prior achievement in a classroom. We included it as a covariate using cognitive ability scores as a proxy. These test scores were collected at the same time point as the outcome variables and may therefore lead to a bias underestimating compositional effects (see Duncan et al., 2004). At the same time, it is possible that our analyses overestimated classroom level effects as the proportion of minority students in a classroom is also

confounded, for instance, with less favorable residential environments and segregated areas in large cities (for a methodological discussion of compositional effects see Harker & Tymms, 2004; Hauser, 1970). Effects of the ethnic composition usually are smaller and often lack statistical significance when controlling for prior achievement level and SES level in German studies (Dumont, Neumann, Maaz, & Trautwein, 2013). Future research thus should favor longitudinal designs and include a large range of context characteristics.

Second, we were lacking background information on the country of origin for some students and created the category “unidentifiable” to include them into the analyses in order to get a full picture of the complete classroom. This may have led to distorted estimations of classroom level effects as they are treated as one category besides other ethnic categories. Future research should gain this kind of background information for instance from school reports available for every student to avoid missing data.

Third, we only analyzed relationships between the ethnic makeup of classrooms and students’ feeling of belonging with one’s classmates as well as their achievement in mathematics and reading. Including a variety of achievement measures that are more likely to be related to the benefits of exposure to diversity, such as creative thinking and problem solving (see Gurin et al., 2003), could be a useful addition in future studies.

Fourth, we do not know how the students in our study perceived ethnic diversity in their classroom and if their perception mattered for their attachment to classmates and group formation (for examples of perceived team diversity in organizational psychology see Shemla, Meyer, Greer, & Jehn, 2014). Future research should involve students’ point of view to a larger degree.

Fifth, there may be further student background characteristics jointly constituting diversity in addition to the ethnic background. Recent developments in organizational psychology investigating the alignment of multiple diversity attributes and creating a hypothetical dividing line between homogeneous groups (“faultline”; see Thatcher & Patel, 2012) could be a model for future educational research as well.

Finally, our study pictures only one aspect of ethnic diversity—that is, the distribution of students in a classroom according to their families’ country of birth expressed in a number that quantifies the degree of diversity. We are aware that ethnicity and ethnic identity go far beyond these numbers and would like to encourage more quantitative and qualitative research in this domain based on a diversity of methods. The aim of the present study was to explore measures of the ethnic makeup of classrooms in large-scale assessment frameworks. Thus, it provides a basis for future research that is

concerned with recommendations on school and classroom composition and ways to address it.

6.6.2 Conclusion

In conclusion, the findings indicate that the ethnic makeup of classrooms matters for individual achievement and psychosocial outcomes. The ethnic diversity measures collected for this educational research study ended up being closely intertwined. Thus, using one measure or the other should lead to comparable results. The general proportion of ethnic minority students showed the strongest relation with student outcomes but ethnic diversity revealed differential result patterns for some outcomes. While the proportion of ethnic minority students in a classroom was negatively associated with individual student outcomes, ethnic diversity was positively related to mathematics achievement controlling for relevant classroom background characteristics associated with less favorable learning environments. Future research in the field of education consequently should not ignore diversity measures completely. Depending on the research question, subgroup and school subject diversity measures can give us more insight into how the ethnic makeup of the classroom is related to student outcomes.

7

Zusammenfassung zentraler Befunde der Teilstudien

7. Zusammenfassung zentraler Befunde der Teilstudien

In der *ersten Teilstudie* wurden Effekte der zuwanderungsbezogenen und sozialen Klassenzusammensetzung auf individuelles Leseverständnis sowie die mediiierende Rolle der Unterrichtsqualität anhand von Daten von Schülerinnen und Schülern der 9. Klasse in einem längsschnittlichen Design untersucht. Als Maß der zuwanderungsbezogenen Klassenzusammensetzung wurde der Prozentanteil von Schülerinnen und Schülern, die nicht nur Deutsch als Muttersprache angaben, verwendet. Erwartungskonform stand der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund in leicht negativem Zusammenhang mit individueller Leseleistung am Ende von Klassenstufe 9, unter Kontrolle der Individuelleistung zu Beginn der Klassenstufe 9 und weiterer leistungsassoziierter Schülermerkmale auf individueller Ebene. Bei zusätzlicher Berücksichtigung der sozialen Zusammensetzung anhand des durchschnittlichen sozioökonomischen Status und bei Kontrolle weiterer Klassenmerkmale verschwand der statistisch bedeutsame Effekt der zuwanderungsbezogenen Zusammensetzung. Die soziale Zusammensetzung erwies sich also für die individuelle Leseleistung bedeutsamer als die zuwanderungsbezogene Zusammensetzung, wobei diese beiden Klassenmerkmale miteinander assoziiert waren. Zusätzlich zeigten sich bedeutsame Leistungsunterschiede zwischen Bildungsgängen mit Hauptschulzweig und Realschulen. Weitere leistungsbezogene Kontrollvariablen auf Klassenebene standen nicht in statistisch bedeutsamem Zusammenhang mit individueller Leseleistung am Ende der Klassenstufe 9, wobei Effekte der Leistungszusammensetzung und des Bildungsgangs miteinander konfundiert sind.

Merkmale der Deutschunterrichtsqualität wurden unter Verwendung latenter Messmodelle als aggregierte Klassenwerte von Schülereinschätzungen einbezogen. Es wurden die theoretisch abgeleiteten drei Hauptqualitätsdimensionen „strukturierte Klassenführung“, „schülerorientiertes, unterstützendes Unterrichtsklima“ und „kognitive Aktivierung“ unterschieden. Die kognitive Aktivierung, die in ihrer Ausgestaltung stärker abhängig vom Unterrichtsfach ist, wurde für den Deutschunterricht über die Einschätzung eines herausfordernden Sprachunterrichts abgebildet. Die Untersuchung von Zusammenhängen zwischen Merkmalen der Klassenzusammensetzung, Unterrichtsqualität und Leseleistung am Ende der Klassenstufe 9 zeigte leicht positivere Beurteilungen der strukturierten Klassenführung und des schülerorientierten, unterstützenden Unterrichtsklimas in Klassen mit einem höheren Anteil von

Heranwachsenden mit Zuwanderungshintergrund. Diese beiden Unterrichtsqualitätsmerkmale wurden zusätzlich von Schülerinnen und Schülern eines Bildungsgangs mit Hauptschulzweig positiver beurteilt. Strukturierte Klassenführung und schülerorientiertes, unterstützendes Unterrichtsklima standen jedoch nicht mit der Leseleistung am Ende von Klassenstufe 9 in bedeutsamem Zusammenhang. Demnach kam diesen beiden Merkmalen der Unterrichtsqualität keine Bedeutung bei der Vermittlung der betrachteten Kompositionseffekten zu. Ein aus Schülersicht herausfordernder Sprachunterricht stand jedoch in positivem Zusammenhang mit der Leseleistung am Ende von Klassenstufe 9. Zudem wurde der Sprachunterricht in Klassen mit einem durchschnittlich höheren sozioökonomischen Status in stärkerem Maße als herausfordernd eingeschätzt. Der indirekte Effekt der Vermittlung von Zusammenhängen zwischen der sozialen Zusammensetzung und individueller Leseleistung am Ende von Klassenstufe 9 durch das Ausmaß herausfordernden Sprachunterrichts wies eine geringe Effektstärke auf, war jedoch marginal statistisch bedeutsam. Diese Teilstudie bietet somit Hinweise darauf, dass Effekte der sozialen Klassenzusammensetzung auf Schülerleistungen zum Teil durch Merkmale der Unterrichtsqualität vermittelt werden.

In der *zweiten Teilstudie* bestand das Ziel darin, Zusammenhänge zwischen der zuwanderungsbezogenen Klassenzusammensetzung und der intrinsischen Lernmotivation der Schülerinnen und Schüler im Deutschunterricht sowie die mediierende Rolle eines schülerorientierten, unterstützenden Unterrichtsklimas zu untersuchen. Weiterführend wurde der Frage nachgegangen, ob differentielle Zusammenhänge zwischen der zuwanderungsbezogenen Klassenzusammensetzung und der Lernmotivation von Schülerinnen und Schülern mit Zuwanderungshintergrund im Vergleich zu Heranwachsenden ohne Zuwanderungshintergrund bestehen. Die Analysen basierten auf denselben Daten wie Teilstudie eins. Erneut wurde als Maß der zuwanderungsbezogenen Klassenzusammensetzung der Prozentanteil von Schülerinnen und Schülern, die nicht nur Deutsch als Muttersprache angaben, verwendet. Die Untersuchung des Zusammenhangs von Merkmalen der Klassenkomposition und Lernmotivation am Ende von Klassenstufe 9 zeigte, dass unter Kontrolle individueller Hintergrundmerkmale und leistungsbezogener Merkmale auf Klassenebene ein leicht positiver Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Lernmotivation im Fach Deutsch besteht. Die Befunde zu Kovariaten auf Klassenebene zeigten, dass Schülerinnen und Schüler einer leistungsstärkeren Klasse eine niedrigere Lernmotivation

aufwiesen und dass Schülerinnen und Schüler des Gymnasiums im Vergleich zu Schülerinnen und Schülern der Realschule eine erhöhte Lernmotivation aufwiesen. Der sozioökonomische Status der Schülerinnen und Schüler sowie die soziale Zusammensetzung wurden in Teilstudie zwei nicht in Analysemodellen berücksichtigt, da Voranalysen keine bedeutsamen systematischen Zusammenhänge mit der Lernmotivation (am Ende von Klassenstufe 9 über die Lernmotivation zu Beginn von Klassenstufe 9 hinaus) zeigten.

Zur Abbildung eines schülerorientierten, unterstützenden Unterrichtsklimas wurden latent aggregierte Klassenwerte von Schülereinschätzungen verwendet. Theoriekonform stand das Unterrichtsklima in positivem Zusammenhang mit der Lernmotivation am Ende von Klassenstufe 9. Merkmale der Klasse waren jedoch unabhängig vom Unterrichtsklima. Lediglich an Bildungsgängen mit Hauptschulzweig war das Unterrichtsklima aus Schülersicht in stärkerem Maße schülerorientiert und unterstützend als in der Realschule. Es zeigten sich somit keine Hinweise darauf, dass Effekte der Klassenzusammensetzung auf die Lernmotivation durch das Ausmaß eines schülerorientierten, unterstützenden Unterrichtsklimas vermittelt wurden.

Weiterführende Analysen zeigten, dass der ermittelte positive Zusammenhang zwischen dem Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Lernmotivation nur für Heranwachsende mit Zuwanderungshintergrund statistisch bedeutsam war. Während ihre Lernmotivation in Klassen mit einem geringen Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund unterdurchschnittlich ausgeprägt war, war sie überdurchschnittlich ausgeprägt in Klassen mit einem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund ab ungefähr 25 Prozent. Die Lernmotivation im Deutschunterricht von Schülerinnen und Schülern ohne Zuwanderungshintergrund hing dagegen nicht zusammen mit dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund. Diese positive Assoziation spezifisch für Heranwachsende mit Zuwanderungshintergrund konnte nicht durch das Unterrichtsklima erklärt werden. In Zusatzanalysen zeigte sich zudem, dass die positive Assoziation nicht durch weitere Lehrer- und Klassenmerkmale (Klassengröße, Lehrerfahrung, Lehrerausbildung) erklärbar war. Es wurde angenommen, dass der positive Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Lernmotivation für Jugendliche mit Zuwanderungshintergrund direkt durch Merkmale der Schülergruppe, wie einem Verbundenheitsgefühl zu den Peers, beeinflusst sein könnte.

In der *dritten Teilstudie* wurden verschiedene Varianten der Messung zuwanderungsbezogener Klassenzusammensetzung und deren Prädiktion von Mathematikleistung und Leseverständnis und des Verbundenheitsgefühls mit den Peers untersucht. Die Analysen basieren auf Daten von Kindern der vierten Klassenstufe in einem querschnittlichen Design. Charakteristika der zuwanderungsbezogenen Klassenzusammensetzung wurden über mehrere Maße abgebildet, welche einerseits den Anteil der Schülerinnen und Schüler mit Zuwanderungshintergrund und andererseits die zuwanderungsbezogene Heterogenität in der Klasse repräsentieren. Grundlage zur Kategorisierung eines Zuwanderungshintergrunds bildete das Geburtsland der Eltern. Vergleichbar zu Teilstudie eins und zwei wurde der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund betrachtet. Darüber hinaus wurden insgesamt acht verschiedene Heterogenitätsmaße, die in verschiedenen Disziplinen Verwendung finden, zur Operationalisierung zuwanderungsbezogener Heterogenität verwendet.

Die Analysen zur Schülerleistung als Kriterium verweisen auf einen leicht negativen Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Kompetenz in Mathematik sowie im Leseverständnis. Zudem bestehen Hinweise auf nicht-lineare Zusammenhänge. Geringe Zusammenhänge bleiben auch nach Kontrolle der sozialen Zusammensetzung sowie der Zusammensetzung bezüglich kognitiver Grundfähigkeiten als Proxy für das Vorwissen der Kinder bestehen. Die zuwanderungsbezogenen Heterogenitätsmaße standen als alleinige Prädiktoren auf Klassenebene ebenfalls in leicht negativem Zusammenhang mit der Schülerleistung. Die Effektgrößen des Zusammenhangs lagen bei den verschiedenen Heterogenitätsmaßen in vergleichbarer Größenordnung. Das Leseverständnis war unabhängig von der zuwanderungsbezogenen Heterogenität, wenn simultan der Anteil von Kindern mit Zuwanderungshintergrund als Prädiktor einbezogen wurde sowie bei Kontrolle der weiteren Merkmale der Zusammensetzung. Die Analysen zur mathematischen Kompetenz zeigten jedoch leicht positive Zusammenhänge zwischen zuwanderungsbezogener Heterogenität und Schülerleistung, die in Einklang mit theoretischen Annahmen einer lernförderlichen Wirkung zuwanderungsbezogen heterogener Klassen stehen.

Analoge Untersuchungen zum Zusammenhang von zuwanderungsbezogener Zusammensetzung und dem Verbundenheitsgefühl mit den Peers zeigten leicht negative Effekte. Entgegen der Erwartung erwies sich der Anteil von Kindern mit

Zuwanderungshintergrund als bedeutsamer für das Verbundenheitsgefühl der Kinder als die zuwanderungsbezogene Heterogenität. Zusätzliche Cross-Level-Interaktionsanalysen verwiesen darauf, dass solche Schülerinnen und Schüler mit Zuwanderungshintergrund ein stärkeres Verbundenheitsgefühl zu den Peers äußerten, die Klassen mit einem höheren Anteil von Kindern mit Zuwanderungshintergrund besuchten. Ebenso äußerten diejenigen Schülerinnen und Schüler ohne Zuwanderungshintergrund ein stärkeres Verbundenheitsgefühl, die in Klassen mit einem höheren Anteil von Kindern ohne Zuwanderungshintergrund beschult wurden.

Teilstudie drei zeigt also eine allgemein stärkere Prädiktionskraft des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund als Maß für die zuwanderungsbezogene Zusammensetzung. Bei Kontrolle weiterer Merkmale der Zusammensetzung wurden für Maße der zuwanderungsbezogenen Heterogenität geringe positive Effekte auf Mathematikleistung gefunden.

Gesamtdiskussion und Ausblick

8. Gesamtdiskussion und Ausblick

8.1 Die Bedeutung zuwanderungsbezogener Klassenzusammensetzung für Schülerleistung und psychosoziale Merkmale: Zu erwartende Richtung und Größe von Effekten

Internationale und nationale Forschung zeigt, dass insbesondere die Bedeutung der zuwanderungsbezogenen Klassenzusammensetzung für die akademische Entwicklung von Schülerinnen und Schülern uneindeutig ist (vgl. Abschnitt 2.2.4). Unklar ist, ob sich Effekte der zuwanderungsbezogenen Zusammensetzung auch unter Kontrolle weiterer Merkmale der Klasse nachweisen lassen, ob die zuwanderungsbezogene Zusammensetzung in negativem Zusammenhang mit Individualleistung steht, welche Maße zur Abbildung verschiedener Charakteristika zuwanderungsbezogener Zusammensetzung herangezogen werden sollten und welche Effekte neben der Leistung als Kriterium auch auf psychosoziale Schülermerkmale zu erwarten sind.

In den nachfolgenden Teilabschnitten wird nun die zu erwartende Richtung und Größe von zuwanderungsbezogenen Effekten der Zusammensetzung auf Grundlage der Befunde der vorliegenden Arbeit sowie anderer Studien diskutiert. Dabei wird auf Effekte des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund und von Heterogenitätsmaßen als unabhängige Variablen zur Beschreibung verschiedener Charakteristika der Zusammensetzung eingegangen. Ebenso werden Effekte auf Leistungsmaße und auf psychosoziale Merkmale diskutiert.

8.1.1 Effekte des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf Leistungsmaße

In der empirischen Forschung zur zuwanderungsbezogenen Zusammensetzung werden Effekte des Anteils von Heranwachsenden mit Zuwanderungshintergrund als Kompositionseffekte auf individuelle Schülerleistungen in Mathematik oder im Leseverständnis analysiert (vgl. Abschnitt 2.2.4). Die Befunde der ersten Teilstudie stellen eine Replikation des Forschungsstandes der Mehrzahl von Analysen aus dem deutschsprachigen Raum dar: Der Prozentanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund steht in leicht negativem Zusammenhang mit individueller Leseleistung; unter Kontrolle der sozialen und leistungsbezogenen Komposition verliert die zuwanderungsbezogene Komposition jedoch ihre statistische Bedeutung (vgl. z.B.

Baumert et al., 2006; Dumont, Neumann, Maaz et al., 2013; Stanat, Schwippert, et al., 2010; Stanat, 2006b). Die soziale Zusammensetzung erweist sich also als bedeutsamer für die Individualleistung als die zuwanderungsbezogene Komposition. Die Koeffizienten liegen für den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund unter Kontrolle weiterer Kompositionsmerkmale und des Bildungsgangs bei $\beta = -.02$ und für den mittleren SES bei $\beta = .51$. Der Effekt des mittleren SES ist groß, aber – wie zu erwarten – geringer als Effekte auf Individualebene (beispielsweise Effekt des individuellen Leseverständnisses zum ersten Messzeitpunkt: $\beta = .76$)¹⁸. Im Vergleich mit anderen Kompositionsanalysen nationaler Studien zeigt sich, dass die Leistungskomposition und der Bildungsgang die größten Effekte auf Individualleistung hatten, was in Teilstudie eins nicht der Fall war (vgl. z.B. Baumert et al., 2006; Köller et al., 2013)¹⁹. Mögliche Ursachen können in der Abhängigkeit des verwendeten Leseleistungstests vom Bildungsgang sowie einer allgemeinen Konfundierung von Kontextmerkmalen liegen (siehe unten). Im Vergleich zu anderen Studien im deutschen Sprachraum liegt also ein erhöhter Effekt der sozialen Komposition vor, der sich jedoch in einer in internationalen Studien vorfindbaren Größenordnung bewegt (vgl. Van Ewijk & Slegers, 2010a).

In Teilstudie drei wurde ebenfalls der Effekt des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf die Leistung im Leseverständnis und in Mathematik bestimmt. Den Befunden der ersten Teilstudie und vorherrschenden Ergebnismustern nationaler Studien entgegen, zeigten sich in Teilstudie drei leicht negative Zusammenhänge zwischen der zuwanderungsbezogenen Komposition und Individualleistung auch nach Kontrolle der sozialen und leistungsbezogenen Komposition. Die unter Verwendung der Individual- und Klassenebenenvarianz der

¹⁸ In Teilstudie eins wurden alle Koeffizienten unter Verwendung der Individualebenenvarianz der abhängigen Variablen standardisiert.

¹⁹ In Teilstudie eins zeigten sich in bivariaten Analysen auf Klassenebene erwartungskonforme Zusammenhänge der Leseleistung zum zweiten Messzeitpunkt mit dem mittleren Leistungsniveau zum ersten Messzeitpunkt ($r = .59$) und den Bildungsgängen ($r_{\text{Gym}} = .53$, $r_{\text{Haupts}} = -.58$). Die bivariate Korrelation zwischen sozialer Komposition und Leseleistung zum zweiten Messzeitpunkt ist vergleichsweise hoch ($r = .68$).

Zusatzanalysen der Kompositionseffekte von Teilstudie eins getrennt für Bildungsgänge verweisen darauf, dass erwartungsgemäß die Varianz der Leistungskomposition zum ersten Messzeitpunkt und der sozialen Komposition in einzelnen Bildungsgängen geringer ist, die Befunde zur zuwanderungsbezogenen und sozialen Komposition jedoch dem allgemeinen Befundmuster von Teilstudie eins entsprechen. Lediglich in der Realschule verliert der Effekt der sozialen Komposition seine statistische Bedeutung, was vermutlich mit der sehr geringen Zwischenklassenvarianz der sozialen Komposition in Verbindung steht.

abhängigen Variablen standardisierten Koeffizienten zur Prädiktion des Leseverständnisses und der Mathematikleistung lagen für den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund bei $\beta_{\text{Lesen}} = -.10$ und $\beta_{\text{Mathe}} = -.13$, für den mittleren SES bei $\beta_{\text{Lesen}} = .08$ und $\beta_{\text{Mathe}} = .06$ und für den Proxy für Vorwissen bei $\beta_{\text{Lesen}} = .05$ und $\beta_{\text{Mathe}} = .08$ (Vergleichswert auf Individualebene z.B. Proxy für Vorwissen $\beta_{\text{Lesen}} = .39$ und $\beta_{\text{Mathe}} = .48$)²⁰. Für das abweichende Befundmuster bestehen mehrere Erklärungsmöglichkeiten, die Gegenstand weiterführender Forschung sein sollten. Zum einen besteht die Möglichkeit, dass zuwanderungsbezogene Kompositionseffekte in der Grundschule (Teilstudie drei) stärker ausgeprägt sind als in der Sekundarstufe (Teilstudie eins). In der Literatur vertretene Annahmen zur Altersabhängigkeit von Kompositionseffekten beziehen sich häufig darauf, dass für jüngere Schülerinnen und Schüler erwachsene Personen, wie Lehrkräfte, einflussreicher sind und für Jugendliche die Peergruppe stärkere Bedeutung zeigt. Demnach könnte man stärkere vermittelnde Effekte der Unterrichtsqualität im Grundschulalter erwarten. Eine vermittelnde Bedeutung von Peerprozessen, wie einer normativen Peerkultur, sollte stärker in der Sekundarstufe wirksam sein. Der empirische Forschungsstand zur Altersabhängigkeit von Kompositionseffekten ist jedoch sehr uneinheitlich (vgl. die Befunde von drei Metaanalysen: Mickelson et al., 2013; Van Ewijk & Sleegers, 2010a, 2010b). Eine zusätzliche Erklärungsmöglichkeit wäre, dass der Grundschulbesuch stärker an den direkten Wohnort gebunden ist als der Besuch einer weiterführenden Schule (vgl. Abschnitt 2.2.1). Dadurch können etwaige Kontexteffekte der Nachbarschaft durch Selektion stärker mit Kompositionseffekten in der Grundschule konfundiert sein. Dies könnte dazu führen, dass sich auch nach Kontrolle weiterer Kompositionsmerkmale ein Zusammenhang zwischen zuwanderungsbezogener Komposition und Individualleistung zeigt.

Der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund als Charakteristik zuwanderungsbezogener Zusammensetzung hat somit, wenn überhaupt, einen kleinen negativen Effekt auf die Leistung. Dass Effekte einzelner Klassenkompositionsmerkmale im Vergleich zu Merkmalen auf Individualebene weniger stark sind, ist dabei grundsätzlich nicht verwunderlich. So ermittelten Metanalysen im

²⁰ Die hier berichteten Koeffizienten entsprechen nicht den in Teilstudie drei dargestellten. In den Modellen von Teilstudie drei wurde zusätzlich ein quadratischer Term des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund einbezogen, der hier aus Gründen der Vergleichbarkeit verschiedener Teilstudienergebnisse nicht berücksichtigt wurde.

internationalen Bereich beispielsweise Kompositionseffekte in der Größenordnung von -.26 bis .07 (Mickelson et al., 2013; Van Ewijk & Slegers, 2010b; vgl. Abschnitt 2.2.4.1). Hierbei ist zu berücksichtigen, dass auch kleine Effekte, die sich zu einem bestimmten Zeitpunkt zeigen, aufgrund sich potentiell kumulierender Effekte im Laufe einer Schullaufbahn einflussreich für die akademische Entwicklung von Schülerinnen und Schülern sind (vgl. Diskussion bei Johnson et al., 2001; Van Ewijk & Slegers, 2010b). Die der vorliegenden Arbeit zugrundeliegenden Daten lassen jedoch keine Schlüsse über vorhergehende oder längerfristige Effekte der Klassenzusammensetzung zu.

8.1.2 Effekte der zuwanderungsbezogenen Heterogenität auf Leistungsmaße

Bei Verwendung verschiedener Indizes zur Abbildung zuwanderungsbezogener Heterogenität als Merkmal der Zusammensetzung zeigten sich ebenfalls leicht negative Zusammenhänge zwischen zuwanderungsbezogener Heterogenität und Schülerleistung. Die für die Heterogenitätsmaße ermittelten Effektgrößen (Prädiktion von Mathematikleistung: $\beta_{\text{Ncat}} = -.08$ bis $\beta_{\text{GleaD}} = -.00$; Prädiktion von Leseleistung: $\beta_{\text{Ncat}} = -.08$ bis $\beta_{\text{var}} = -.02$) waren geringer ausgeprägt als die des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund (Prädiktion von Mathematikleistung: $\beta = -.17$, Prädiktion von Leseleistung: $\beta = -.15$). Die einzelnen Heterogenitätsmaße unterschieden sich dabei in ihrer Effektgröße nicht stark voneinander (vgl. auch McDonnald & Dimmick, 2003)²¹. Auch wenn Heterogenitätsmaße und der allgemeine Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund inhaltlich voneinander unterscheidbare Merkmale der zuwanderungsbezogenen Zusammensetzung abbilden, erklären beide mehrheitlich denselben Varianzanteil zwischen Klassen, wobei der Prozentanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund stärkere Zusammenhänge mit der Leistung aufweist. Zusätzlich zum Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund zeigten sich leicht positive Zusammenhänge zwischen zuwanderungsbezogener Heterogenität und Mathematikleistung. Die Größe der Effekte unter Kontrolle weiterer

²¹ Die Prüfung von Unterschieden zwischen Regressionskoeffizienten der Heterogenitätsmaße in den verschiedenen Modellen anhand von z-Tests unter Anwendung einer Bonferroni-Korrektur zeigte lediglich einen bedeutsamen Unterschied zwischen dem Koeffizienten der Kategorienanzahl (Ncat) und Gleasons *D* (GleaD) als alleinige Prädiktoren auf Klassenebene bei Vorhersage der Mathematikleistung (Modelle M.2 und M.15 in Teilstudie drei). Alle weiteren Koeffizienten verschiedener Heterogenitätsmaße und Modelle waren statistisch nicht voneinander verschieden.

Kompositionsmerkmale ist erneut als gering einzustufen ($\beta = .04$ bis $.02$). Sie entsprechen beispielsweise in etwa Effekten von verschiedenen Unterrichtsmerkmalen auf Schülerleistungen (z.B. Schaffung eines sozialen Lernklimas; Verdeutlichung von Lernzielen), die in der Metanalyse von Seidel und Shavelson (2007) berichtet wurden. Dieser Befund entspricht Ergebnissen anderer Studien zu positiven Zusammenhängen zwischen ethnischer Heterogenität und Leistungsmaßen, die von verschiedenen Autorinnen und Autoren als Anzeichen für einen leistungsförderlichen Austausch zwischen Peers interpretiert werden (z.B. Benner & Crosnoe 2011; Maestri, 2011; Tam & Bassett, 2004; siehe Abschnitt 2.2.5.4). Die vorliegende Arbeit zeigt erstmals auch für das deutsche Bildungssystem, dass ethnische Heterogenität positiv mit Schülerleistung assoziiert sein kann. Für die Leseleistung zeigte die zugewanderungsbezogene Heterogenität über den Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund hinaus keine bedeutsamen Effekte. Inwiefern der Grad an Sprachabhängigkeit der als Kriterium betrachteten Schülerleistung eine Rolle für das in Teilstudie drei gefundene differentielle Ergebnismuster für Mathematik und Leseverständnis spielt (vgl. Abschnitt 6.6), sollte Gegenstand weiterführender Forschung sein.

Die in der vorliegenden Arbeit vertretene breitere Perspektive auf zugewanderungsbezogene Klassenzusammensetzung entspricht der Forderung von Benner und Crosnoe in ihrem Aufsatz von 2011: „*Taken as a whole, [...], as a field, we need to adopt a more nuanced view of school racial/ethnic composition and its relation to children’s developmental competencies.*“ (S. 642). Gleichzeitig stellt sie lediglich einen ersten Schritt dar, der weitere Fragen eröffnet. Zum Beispiel sollte weiterführend untersucht werden, ob eine „kritische Anzahl“ an Schülerinnen und Schülern desselben ethnischen Hintergrunds und eine zusätzliche Heterogenität der anderen Schülerinnen und Schüler der Klasse sich positiv auf die Schülerleistung auswirkt (vgl. Benner & Crosnoe, 2011; Garces & Jayakumar, 2014) und wie sich Heterogenität anhand multipler Schülermerkmale parallel abbilden lässt (vgl. Meyer, in press; siehe auch Abschnitt 8.5.2).

8.1.3 Effekte der zugewanderungsbezogenen Klassenzusammensetzung auf psychosoziale Schülermerkmale

Merkmale der Klassenzusammensetzung können nicht nur für die Leistung von Schülerinnen und Schülern relevant sein, sondern auch für psychosoziale Merkmale (z.B.

Benner et al., 2008; Eccles et al., 1993; Klieme et al., 2010; vgl. Abschnitt 2.2.6). In der zweiten Teilstudie zeigte sich ein positiver Effekt des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf die Lernmotivation im Fach Deutsch von Heranwachsenden mit Zuwanderungshintergrund. Der Annahme in Teilstudie zwei, dass dieses Ergebnismuster mit dem Zugehörigkeitsgefühl zu den Peers als einem von drei motivationsförderlichen Grundbedürfnissen (Baumeister & Leary, 1995; Deci & Ryan, 1985, 2000) in Verbindung steht, wurde in Teilstudie drei nachgegangen. Es wurde deutlich, dass sich Schülerinnen und Schüler mit Zuwanderungshintergrund erwartungskonform stärker zu den Mitschülerinnen und Mitschülern zugehörig fühlten, wenn ihre Klasse einen höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund aufwies. Umgekehrt fühlten sich Heranwachsende ohne Zuwanderungshintergrund stärker zu ihren Mitschülerinnen und Mitschülern zugehörig, wenn ihre Klasse einen höheren Anteil von Heranwachsenden ohne Zuwanderungshintergrund aufwies. Die zuwanderungsbezogene Heterogenität hing indes kaum mit dem Zugehörigkeitsgefühl der Schülerinnen und Schüler zusammen. Lediglich für zwei von acht Heterogenitätsmaßen als alleiniger Prädiktor auf Klassenebene zeigten sich statistisch bedeutsame negative Effekte, die als klein zu beurteilen sind ($\beta_{\text{Ncat}} = -.04$, $\beta_{\text{ShaH}} = -.03$)²². Bei simultaner Berücksichtigung des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund und einem Heterogenitätsmaß zeigte sich lediglich ein bedeutsamer Effekt des Anteils. Die Befunde geben also erste Hinweise darauf, dass das Vorhandensein eines Zuwanderungshintergrunds bedeutsamer für das Zugehörigkeitsgefühl ist als die Homogenität in der Klasse bezogen auf ein konkretes Herkunftsland. Diese Befundlage entspricht nicht theoretischen Annahmen der besonderen Salienz eines herkunftslandspezifischen Hintergrunds als Nähe stiftendes Schülermerkmal (vgl. Byrne, 1971; Tajfel & Turner, 1986; siehe Abschnitt 2.2.6.2). Die Zugehörigkeit zur Gruppe „ohne Zuwanderungshintergrund“ oder „mit Zuwanderungshintergrund“ scheint also stärker zur Gruppenbildung von „ingroup“ und „outgroup“ innerhalb der Klasse beizutragen. Vorstellbar wäre zum Beispiel, dass Schülerinnen und Schüler mit Zuwanderungshintergrund aufgrund Mehrsprachigkeit und Migrationsgeschichte innerhalb der Familie eine stärkere Ähnlichkeit zu anderen Heranwachsenden mit Zuwanderungshintergrund als zu Heranwachsenden ohne

²² Bei Berechnung der Heterogenitätsmaße einschließlich der Gruppe „deutsch“ zeigten sich etwas größere Effekte, die zwischen $\beta_{\text{ShaH}} = -.07$ und $\beta_{\text{Ncat}} = -.04$ lagen (siehe Abschnitt 6.5.2.4).

Zuwanderungshintergrund wahrnehmen. Dies beeinflusst das Gefühl der Zugehörigkeit zu einer Gruppe und die Abgrenzung von anderen Gruppen. Neben einer notwendigen Replikation der Befunde in zukünftiger Forschung wären auch Studien wünschenswert, die die kulturelle Identität der Schülerinnen und Schüler (vgl. Edele, Stanat, Radmann & Segeritz, 2013) als potentiellen Moderator der Zusammenhangsstärke zwischen individuellem Zuwanderungshintergrund und zugewanderungsbezogener Zusammensetzung einbeziehen.

8.1.4 Zuwanderungsbezogene Klassenzusammensetzung und die Bedeutung von zusätzlichen Merkmalen der Klassenzusammensetzung und des weiteren Kontexts

Bei der Beurteilung der Bedeutung zugewanderungsbezogener Klassenzusammensetzung ist, wie bereits erwähnt, zu berücksichtigen, dass diese eines unter mehreren potentiell bedeutsamen Merkmalen des Lernumfelds darstellt (vgl. Abschnitt 2.2.1). Die soziale und die Leistungskomposition sowie der Bildungsgang sind allgemein bedeutsam mit individueller Leistung und psychosozialen Merkmalen assoziiert (vgl. Abschnitt 2.2.4). In den Teilstudien eins und drei waren bivariat ebenfalls bedeutsame Zusammenhänge mit der Schülerleistung nachweisbar. In den Mehrebenenregressionsmodellen zeigten sich jedoch nicht immer statistisch bedeutsame Effekte aller Klassen- bzw. Schulmerkmale. Das scheint vor allem an einer hohen Konfundierung von sozialer und leistungsbezogener Zusammensetzung sowie des Bildungsgangs im deutschen Schulsystem zu liegen²³. Die Verwendung konfundierter erklärender Variablen kann zu Suppressionseffekten führen. Die Merkmale „unterdrücken“ also ihre Effekte gegenseitig bzw. es werden durch Bindung von Varianzen Zusammenhänge zwischen einzelnen Variablen und dem Kriterium vergrößert (z.B. Eid, Gollwitzer & Schmitt, 2010). So zeigten schon Baumert et al. (2006) anhand von PISA-Daten, dass Effekte leistungsbezogener und sozialer Kontextmerkmale auf Schülerleistung zu großem Anteil miteinander konfundiert sind. Insgesamt ist von differentiellen Lern- und Entwicklungsmilieus auszugehen, die mehr oder weniger förderlich für die akademische Entwicklung sein können. In der Schulpraxis zeichnen

²³ Beispielsweise lagen die bivariaten Korrelationen auf Klassenebene in Teilstudie eins zwischen Lesekompetenz zu T₁ und SES bei $r = .59$, zwischen Lesekompetenz zu T₁ und Bildungsgängen: $r_{\text{Gym}} = .73$, $r_{\text{HS}} = -.77$ und zwischen SES und Bildungsgängen: $r_{\text{Gym}} = .78$, $r_{\text{HS}} = -.70$.

sich somit häufig Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und zuwanderungsbezogen heterogenere Klassen auch durch einen durchschnittlich niedrigeren sozioökonomischen Status sowie ein niedrigeres Vorkenntnisniveau aus. Die statistische Trennung von Effekten einzelner Kontextmerkmale kann vor diesem Hintergrund als künstlich für die schulische Praxis erscheinen, dient jedoch der Überprüfung differentieller, theoretisch abgeleiteter Annahmen zur Bedeutung der zuwanderungsbezogenen und sozialen Klassenzusammensetzung über Merkmale der Leistungskomposition und des Bildungsgangs hinaus (vgl. Abschnitt 2.2.5 und 2.2.6). Durch die Einbeziehung des Bildungsgangs (Teilstudien eins und zwei) und des mittleren Vorleistungsniveaus der Klasse (Teilstudien eins und zwei, in Teilstudie drei Proxy für Vorleistungsniveau) sollten also leistungsrelevante Kontextmerkmale kontrolliert werden.

Der in Teilstudie zwei ermittelte kleine, jedoch statistisch bedeutsame Effekt der Leistungskomposition auf die Lernmotivation entsprechend eines Big-Fish-Little-Pond-Effekts sowie der gleichzeitig bedeutsame Unterschied zwischen Schülerinnen und Schülern des Gymnasiums und der Realschule im Sinne eines Assimilationseffekts entspricht nicht dem üblichen Befundmuster zum Vergleich von Assimilations- und Kontrasteffekten im deutschen Bildungssystem (vgl. Baumert et al., 2006; Chmielewski et al., 2013; Köller, 2004; Trautwein et al., 2006). Häufig zeigt sich, dass Kontrasteffekte bedeutsamer für das Selbstkonzept und Interesse sind und assimilative Effekte des Bildungsgangs darüber hinaus ihre statistische Relevanz verlieren. Dieser Befund ist insbesondere für das mathematische Selbstkonzept bestätigt und zeigt sich vor allem in Stichproben mit zwischen Schulen vorgenommener Leistungsgruppierung im Vergleich zu kursspezifischer Leistungsgruppierung (vgl. Chmielewski et al., 2013; Trautwein et al., 2006). Die stärkeren Effekte des Bildungsgangs auf die Lernmotivation am Ende von Klasse 9 unter Kontrolle der Motivation zu Beginn von Klasse 9 in Teilstudie zwei könnten zum Beispiel auch darauf hindeuten, dass nicht so stark direkte Leistungsvergleiche zwischen Schülerinnen und Schülern die Lernmotivation beeinflussen, sondern eher bildungsgangspezifische Merkmale wie zum Beispiel curriculare Vorgaben. Das Ergebnismuster von Teilstudie zwei entspricht zudem Befunden von Köller (2004, Studie 1) zum sprachlichen Selbstkonzept im Fach Englisch.

Zusätzlich ist bei der Interpretation der Ergebnisse zu berücksichtigen, dass diese auch von weiteren Kontextmerkmalen, wie städtischen oder ländlichen Schuleinzugsgebieten und der Zusammensetzung des Wohnumfelds, mit bedingt werden

können und somit möglicherweise keine reinen Schul- und Klasseneffekte darstellen (vgl. Abschnitt 2.2.1 und Abschnitt 8.4). Des Weiteren kann die Bedeutung von Effekten des Wohnumfelds und der Schul- und Klassenzusammensetzung auch zwischen Bundesländern, beispielsweise aufgrund unterschiedlicher Bevölkerungs- und schulischer Angebotsstrukturen, variieren (vgl. Abschnitt 2.2.1; z.B. Baumert et al., 2006).

Zusammenfassend lässt sich festhalten, dass die Ergebnisse der vorliegenden Arbeit zu Effekten der zuwanderungsbezogenen Klassenzusammensetzung insgesamt leicht negative Effekte des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund auf Schülerleistung zeigen, wobei diese Effekte meist vollständig auf die soziale und leistungsbezogene Komposition zurückführbar sind. Es liegen also meist keine eigenständigen zuwanderungsbezogenen Effekte der Zusammensetzung vor. Darüber hinaus zeigten sich auch leicht positive Effekte der zuwanderungsbezogenen Heterogenität für mathematische Leistung sowie positive Effekte des Klassenanteils von Schülerinnen und Schülern mit Zuwanderungshintergrund für psychosoziale Merkmale von Heranwachsenden mit Zuwanderungshintergrund. Diese positiven Effekte geben Hinweise auf mögliche förderliche Faktoren von Charakteristika der zuwanderungsbezogenen Zusammensetzung für Schülerinnen und Schüler mit Zuwanderungshintergrund. Es lässt sich also durch die vorliegenden Analysen die Annahme erhärten, dass geringere Schülerleistungen in Klassen mit höherem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund vor allem auf insgesamt schwierige Lernmilieus zurückzuführen sind und parallel für Heranwachsende mit Zuwanderungshintergrund auch positive Effekte vor allem im psychosozialen Bereich zu erwarten sind. Effekte der zuwanderungsbezogenen Zusammensetzung scheinen letztendlich mannigfaltig zu sein und können sowohl förderliche wie auch hinderliche Bedingungen für die akademische Entwicklung bieten.

8.2 Unterrichtsqualität als Mediator von Effekten der Klassenzusammensetzung

Die Untersuchung der Transmissionswege von Effekten der Klassenzusammensetzung dient dem Ziel, die Vermittlung nachzuvollziehen und so Informationen zur Ableitung von Maßnahmen zur Verringerung dieser Effekte zu

erhalten (vgl. z.B. Schofield, 2006; Walter & Stanat, 2008). Differenzierte Maßnahmen sind notwendig, da eine zentral gesteuerte Zuteilung von Schülerinnen und Schülern zu Schulen und Klassen, die einer segregierten Klassenzusammensetzung entgegenwirkt, für die schulische Situation im deutschen Bildungssystem eine kaum realistisch durchführbare Möglichkeit darstellt (vgl. Abschnitt 8.5.1).

In der vorliegenden Arbeit wurde die Unterrichtsqualität als Mediator untersucht, da sie zum einen von theoretischer Seite einen weithin angenommenen, jedoch kaum empirisch untersuchten Transmissionsweg darstellt (vgl. Abschnitt 2.2.5). Zum anderen wird erwartet, dass die Unterrichtsqualität einen manipulierbaren Mechanismus darstellt, um Effekte der Zusammensetzung zu beeinflussen (z.B. Baumert et al., 2006; Dumont, Neumann, Maaz, et al., 2013; Harker & Tymms, 2004; Schofield, 2006). Teilstudie eins gab Hinweise darauf, dass Effekte der sozialen Klassenzusammensetzung auf die Leseleistung unter Kontrolle weiterer Merkmale der Zusammensetzung partiell durch einen kognitiv aktivierenden, herausfordernden Sprachunterricht mediiert werden. Die Größe des indirekten Effekts ist als gering einzustufen und der Effekt war lediglich marginal signifikant ($\beta = .14$, $SE = .08$, $p = .086$). Dennoch ist die Effektgröße vor dem Hintergrund allgemein zu erwartender geringer Effektgrößen von Unterrichtsmerkmalen auf Schülerleistung (vgl. die Metanalyse von Seidel und Shavelson, 2007; siehe auch Abschnitt 8.1.2) als erster Hinweis auf eine vermittelnde Bedeutung der kognitiven Aktivierung zu bewerten. Die Befunde stehen also in Einklang mit vorheriger Forschung, die überwiegend auf Schulebene Hinweise auf Zusammenhänge zwischen Zusammensetzung, Unterrichtsmerkmalen und Schülerleistungen aufzeigte, allerdings meist ohne vermittelnde Effekte explizit zu analysieren (Dumay & Dupriez, 2007; Palardy, 2015, 2013; Rumberger & Palardy, 2005; Stipek, 2004; Vause et al., 2010; vgl. Schofield 2006). Die weiteren in Teilstudie eins betrachteten Unterrichtsqualitätsmerkmale, strukturierte Klassenführung und schülerorientiertes, unterstützendes Unterrichtsklima, hingen nicht mit der Leseleistung zusammen und vermittelten nicht Kompositionseffekte auf Leistung. Es ist zu erwarten, dass diese Merkmale stärker mit psychosozialen Schülermerkmalen wie der Lernmotivation in Zusammenhang stehen. Es zeigte sich theoriekonform in Teilstudie zwei, dass das Unterrichtsklima in positivem Zusammenhang mit der Lernmotivation stand. Es war jedoch nicht mit der zuwanderungsbezogenen Klassenzusammensetzung assoziiert. Das

Unterrichtsklima vermittelte also nicht die zuwanderungsbezogenen Klassenkompositionseffekte auf die Lernmotivation im Fach Deutsch.

Der Annahme der Vermittlung von Effekten der Klassenzusammensetzung durch Unterrichtsqualitätsmerkmale, die in der Arbeit getroffen wurde, stehen weitere alternative Annahmen zum Verhältnis von Zusammensetzung der Schülerschaft und Unterricht gegenüber. Es wird von verschiedenen Forschenden angenommen, dass unterschiedliche Bedürfnisse der Schülerinnen und Schüler in unterschiedlich zusammengesetzten Klassen bestehen und somit dieselben Unterrichtsmerkmale unterschiedliche Effekte hervorrufen können (vgl. Brophy, 1986; Muijs, Campbell, Kyriakides & Robinson, 2005; Opdenakker & Minnaert, 2011; Palardy, 2008; Vause et al., 2010). Es wird hier also von einem Moderationseffekt anstelle eines Mediationseffekts ausgegangen. Annahme ist dabei, dass individuelle Schülerinnen und Schüler mit Zuwanderungshintergrund und niedrigem sozioökonomischen Status sowie Klassen mit einem hohen Anteil solcher Schülerinnen und Schüler einen anders gearteten, z.B. stärker strukturierten und an positivem Feedback orientierten Unterricht benötigen. In den genannten Studien zeigt sich jedoch zumeist, dass Haupteffekte des Unterrichts auf Schülerleistung bedeutsamer waren als Interaktionseffekte, die Schülermerkmale oder Merkmale der Zusammensetzung einbezogen. Unterschiedliche Schülerpopulationen benötigen demnach nicht anderen Unterricht, sondern leistungsschwächere Schülerinnen und Schüler (unabhängig von weiteren Hintergrundmerkmalen) bedürfen ggf. zusätzlicher Unterstützung.

Annahmen der Vermittlung durch Merkmale der Unterrichtsqualität beziehen sich darauf, dass Lehrkräfte auf Basis der Klassenzusammensetzung differentielle Erwartungen an die Schülerinnen und Schüler haben und sich unterschiedlich verhalten (vgl. Abschnitt 2.2.5.1). Die Verbindung von Klassenzusammensetzung und Unterricht entsteht also durch eine Reaktion der Lehrkraft auf diese Klasse und sie würde in einer anderen Klasse auf andere Art und Weise unterrichten. Alternative Erklärungsansätze für eine Verbindung von Klassenzusammensetzung und Unterrichtsmerkmalen, die davon abgegrenzt werden sollten, beziehen sich darauf, dass Lehrkräfte mit geringerer Qualifikation systematisch Klassen mit einer bestimmten Zusammensetzung unterrichten. Dies kann durch systematische Zuordnungsmechanismen auf Klassen- und Schulebene sowie unterschiedliche Ressourcenausstattungen von Schulen beeinflusst sein (vgl. Abschnitt 8.4).

In der vorliegenden Arbeit wurden lediglich in Zusatzanalysen zu Teilstudie zwei weitere Lehrermerkmale einbezogen, die potentiell den Zusammenhang zwischen Klassenzusammensetzung und Unterrichtsqualität erklären könnten. Es zeigten sich keine statistisch bedeutsamen Hinweise auf Zusammenhänge zwischen dem Klassenanteil von Heranwachsenden mit Zuwanderungshintergrund und der Lehrqualifikation, operationalisiert über die Lehrerfahrung in Jahren und das Studienfach (vgl. Abbildung A2.1). Dies entspricht dem Befund von Stipek (2004), die entgegen ihrer Annahmen keinen statistisch bedeutsamen negativen Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Lehrqualifikation (Lehrerfahrung in Jahren und Ausbildungsabschlussniveau) nachweisen konnte. Die Autorin begründet diese nicht-erwartungskonforme Befundlage mit einer möglichen Varianzeinschränkung innerhalb der Stichprobe, die 155 Schulen in drei Staaten der USA umfasst. Zusätzlich ist zu berücksichtigen, dass das Ausbildungsniveau und die Anzahl an Lehrjahren möglicherweise eher distale Schätzungen für Lehrqualifikationen darstellen und allgemein die in Reviews zusammengefasste Befundlage von Zusammenhängen zwischen dem Ausbildungsniveau und der Lehrerfahrung einerseits und Schülerleistungszuwachs andererseits alles andere als eindeutig ist (Hanushek, 1997; Wayne & Youngs, 2003). Lehrkräfte in der Studie von Stipek (2004) äußerten jedoch geringere Erwartungen an Unterstützungsmöglichkeiten innerhalb der Familie in Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund, was wiederum mit einem weniger anspruchsvollen und weniger kognitiv aktivierenden Unterricht in solchen Klassen assoziiert war. Aus anderen Untersuchungen bestehen ebenso Hinweise darauf, dass Lehrkräfte geringere Leistungserwartungen an Schülerinnen und Schülern mit Zuwanderungshintergrund besitzen (vgl. z.B. Alexander & Schofield, 2006; Lorenz et al., eingereicht; Tenenbaum & Ruck, 2007). Die angeführten Studien stützen somit größtenteils die Annahme einer vermittelnden Bedeutung der Unterrichtsqualität.

Weitere Untersuchungen, die Unterrichtsqualität, Klassenzusammensetzung sowie Erwartungen der Lehrkräfte und Zuordnungsmechanismen innerhalb der Schule berücksichtigen, sind notwendig, um differenzierte Informationen zu Ansatzmöglichkeiten zur Verringerung von unterrichtsbezogenen Effekten der Zusammensetzung zu erhalten (vgl. Abschnitt 8.5.1). Vor dem Hintergrund der oben diskutierten geringen Effektgrößen ist zu erwarten, dass durch Veränderung von

Unterrichtsprozessen Effekte der Klassenzusammensetzung nicht völlig verschwinden. Das Einwirken auf Unterrichtsprozesse stellt aber einen Ansatzpunkt unter mehreren dar.

8.3 Die Bedeutung der Klassenzusammensetzung für das Bestehen zuwanderungsbezogener Disparitäten

Wie einleitend in dieser Arbeit angeführt, werden zur Erklärung zuwanderungsbezogener Disparitäten in Schülerleistungen sowie Unterschieden in psychosozialen Merkmalen vornehmlich Merkmale der Schülerinnen und Schüler und ihrer Familien hinzugezogen (vgl. Kapitel 1 und 2.1). Erklärungen, die ausschließlich individuelle und innerfamiliäre Determinanten einbeziehen, vermögen jedoch nicht, zuwanderungsbezogene Disparitäten zufriedenstellend aufzuklären (vgl. Abschnitt 2.1.4). Verschiedene Autorinnen führen an, dass auch Merkmale des schulischen Kontexts als potentielle Erklärungsfaktoren für die Entstehung und Aufrechterhaltung von Disparitäten mit einbezogen werden sollten (Bellin, 2009; Diefenbach, 2010; Stanat, 2006a). Vor diesem Hintergrund gibt die vorliegende Arbeit Hinweise darauf, dass die Klassenzusammensetzung und Merkmale der Unterrichtsqualität systematisch zwischen Klassen variieren und in einigen Klassen als vorteilhafter für die akademische Entwicklung der Schülerinnen und Schüler einzustufen sind als in anderen Klassen. Für die Schülerleistungen wurden keine differentiellen Effekte der Klassenzusammensetzung für Schülerinnen und Schüler mit Zuwanderungshintergrund im Vergleich zu Schülerinnen und Schülern ohne Zuwanderungshintergrund aufgedeckt²⁴. In Klassen mit einem niedrigen sozioökonomischen Niveau sind demnach von Heranwachsenden mit und ohne Zuwanderungshintergrund gleichermaßen geringere Leistungen zu erwarten; die Zusammensetzung scheint keine stärkere Bedeutung für Schülerinnen und Schüler mit Zuwanderungshintergrund zu haben. Wohl aber werden Schülerinnen und Schüler mit Zuwanderungshintergrund häufiger in Klassen mit einem höheren Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund beschult und besuchen auch

²⁴ Beispielsweise bestanden keine statistisch bedeutsamen Cross-Level-Interaktionseffekte zwischen zuwanderungsbezogener bzw. sozialer Klassenzusammensetzung und dem individuellen Zuwanderungshintergrund bzw. sozioökonomischem Status bei Vorhersage von Schülerleistungen in Teilstudie eins (Interaktion zwischen Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und individuellem Zuwanderungshintergrund: $\beta = -.04$, $SE = .02$, $p = .146$, Interaktion zwischen sozioökonomischem Klassenniveau und individuellem Zuwanderungshintergrund: $\beta = .10$, $SE = .09$, $p = .277$, Interaktion zwischen sozioökonomischem Klassenniveau und individuellem sozioökonomischem Status: $\beta = .01$, $SE = .04$, $p = .757$). Diese Informationen wurden in der Teilstudie nicht berichtet.

häufiger Klassen, die sich durch ein niedriges sozioökonomisches und vorleistungsbezogenes Niveau auszeichnen, als Schülerinnen und Schüler ohne Zuwanderungshintergrund²⁵. Dies bedeutet, dass Schülerinnen und Schüler mit Zuwanderungshintergrund häufiger unter Bedingungen lernen, die als ungünstiger für die Leistungsentwicklung einzustufen sind. Es ist davon auszugehen, dass dieser differentielle schulische Lernkontext auch zu zuwanderungsbezogenen Disparitäten in Schülerleistungen beiträgt. In Teilstudien zwei und drei wurden auch positive Zusammenhänge zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und leistungsförderlichen psychosozialen Aspekten wie der Lernmotivation und dem Zugehörigkeitsgefühl zu Peers für Schülerinnen und Schüler mit Zuwanderungshintergrund ermittelt. Die negativen leistungsbezogenen, sozialen und teilweise auch zuwanderungsbezogenen Klassenkompositionseffekte auf die Leistungsentwicklung scheinen allerdings demgegenüber ein stärkeres Gewicht zu haben. Dies betrifft ebenso den leicht positiven Zusammenhang zwischen zuwanderungsbezogener Heterogenität innerhalb von Klassen und Mathematikleistung (Teilstudie drei), dessen Effektgröße ebenfalls geringer ist als die anderer Merkmale der Zusammensetzung.

Die Bedeutung des schulischen Lernumfelds für zuwanderungsbezogene Disparitäten wirft Fragen der Bildungsgerechtigkeit oder Chancengerechtigkeit auf. Chancengerechtigkeit lässt sich beschreiben als „... die faire Chance zur freien Teilhabe an der Gesellschaft, die auch gewährleistet wird durch eine gerechte Institution Schule, in der Schülerinnen und Schüler aufgrund ihrer sozialen und natürlichen Merkmale keine zusätzlichen Nachteile erfahren, sowie durch eine Förderung der Befähigung aller und durch eine wechselseitige Anerkennung der an Schule beteiligten Personen.“ (Berkemeyer et al., 2014, S. 15). Chancen auf gesellschaftliche Teilhabe sind mit der akademischen Entwicklung der Personen verbunden. Alle Menschen sollen so unterstützt werden, dass sie möglichst hohe akademische Kompetenzen erwerben, möglichst hohe Bildungsabschlüsse erzielen und psychosozial angepasst sind. Die wechselseitige Anerkennung der an Schule beteiligten Personen wird dabei als respektvolles Miteinander beschrieben, das durch ein Gefühl von Angenommensein die Ausbildung von

²⁵ Ergebnisse von zusätzlichen Varianzanalysen zum Vergleich des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund in Klassen unterschiedlichen sozioökonomischen und vorleistungsbezogenen Niveaus in Teilstudie eins zeigen statistisch bedeutsame Unterschiede. Siehe auch Abbildung A4.1 und Abbildung A4.2 im Anhang.

Selbstvertrauen fördert, was dienlich für die Teilhabe an der Gesellschaft ist (Honneth, 2011). Andere Autorinnen und Autoren heben hervor, dass allen Schülerinnen und Schülern gleiche Möglichkeiten zur gesellschaftlichen Teilhabe bereitgestellt werden sollen und dass sie zur Teilhabe gleichermaßen befähigt werden sollen (vgl. z.B. Hopf, 2011). Dies bedeutet, dass die Institution Schule insgesamt qualitativ hochwertige Lerngelegenheiten bereitstellen sollte, um auf gerechte Weise Möglichkeiten zur gesellschaftlichen Teilhabe zu bieten. Die „Förderung der Befähigung aller“ zielt darauf ab, dass die Schule darüber hinaus gerade durch differentielle Unterstützungsbemühungen bestrebt sein sollte, durch familiäre Hintergründe bedingte ungleiche Startbedingungen im schulischen Kompetenzerwerb mit der Zeit auszugleichen und zu bewirken, dass Schülerinnen und Schüler unabhängig von ihrer gesellschaftlichen Stellung dieselben Erfolgsaussichten haben (vgl. Berkemeyer et al., 2014; Giesinger, 2007; Hopf, 2011; Stanat, 2006a). Bezogen auf diese Beschreibung von Bildungsgerechtigkeit verweisen die Ergebnisse zu Effekten der Klassenzusammensetzung auf Schülerleistung sowie die Hinweise auf eine vermittelnde Rolle der Unterrichtsqualität der vorliegenden Arbeit auf bestehende Bildungsungerechtigkeit: Es liegt eine ungleiche Verteilung von schulischen Lerngelegenheiten vor, die auch zum Teil eine differentielle Bereitstellung von Unterrichtsqualität beinhaltet. Dies stellt also eine nachteiligere Bildungssituation für Schülerinnen und Schüler mit Zuwanderungshintergrund dar. Hierbei ist zu berücksichtigen, dass Bildungsbenachteiligungen sowohl auf individueller und familiärer Ebene (vgl. Kapitel 2.1) als auch auf Klassen- und Schulebene bestehen und Maßnahmen zur Verringerung von Bildungsungerechtigkeiten an verschiedenen Ebenen ansetzen müssen (vgl. auch Diskussion bei Rumberger & Palardy, 2005). Die Befunde der vorliegenden Arbeit verweisen darauf, dass eine Ursachenzuschreibung von zugewanderungsbezogenen Disparitäten allein auf der Individualebene zu kurz greift (vgl. Bellin, 2009).

8.4 Methodische Bewertung und Grenzen der Arbeit

Die vorliegende Arbeit ermöglicht durch ihre Anlage über den bisherigen empirischen Kenntnisstand hinausgehende, vertiefte Einblicke in Effekte der Klassenzusammensetzung sowie deren Vermittlung. Sie bezieht repräsentative Stichproben verschiedener Klassenstufen im deutschen Bildungssystem ein, zu denen teils Informationen zu mehreren Messzeitpunkten vorliegen. Dies bietet eine solide

Grundlage zur Einschätzung der Bedeutsamkeit von Effekten der Zusammensetzung der Schülerschaft auf die akademische Entwicklung von Kindern und Jugendlichen. Zusätzlich wurden – im Gegensatz zu weiter verbreiteten Analysen auf Schulebene – in der vorliegenden Arbeit die Zusammensetzung in kompletten Schulklassen betrachtet sowie Merkmale der Unterrichtsqualität in der Klasse direkt in die Analysen mit einbezogen. Die über Schülereinschätzungen gewonnenen Informationen zur Unterrichtsqualität wurden in latenten Mehrebenenmodellen modelliert, die es ermöglichen Mess- und Stichprobenfehler in angemessener Weise zu berücksichtigen (Doubly-latent-Approach mit Cross-level-Messinvarianz). Zudem wurde weitgehend die Messinvarianz der verwendeten Konstrukte für Schülerinnen und Schüler mit und ohne Zuwanderungshintergrund sichergestellt. Die Analyse einer Vielzahl von Merkmalen der Zusammensetzung, die Einbeziehung verschiedener Maße zur Abbildung verschiedener Charakteristika zuwanderungsbezogener Zusammensetzung sowie die Analyse von Effekten nicht nur auf verschiedene Schülerleistungen, sondern auch auf psychosoziale Merkmale, ermöglichten eine vertiefte Auseinandersetzung mit Effekten der Klassenzusammensetzung.

Die Nutzung von Daten aus groß angelegten Schulleistungsstudien, die diese aufgeführten Analysevorteile ermöglichten, sind auch mit Grenzen der vorliegenden Arbeit verbunden, die teilweise bereits in den vorangegangenen Abschnitten angesprochen wurden. So ermöglichen diese Studien zwar die Einbeziehung einer großen Anzahl an Schülerinnen und Schülern bzw. Klassen und Lehrkräften, stützen sich dabei jedoch überwiegend auf mit Fragebögen gewonnene Selbstberichtsdaten. Für die vorliegende Arbeit wäre insbesondere eine zusätzliche Durchführung von Unterrichtsbeobachtungen nützlich, um eine stärkere Vergleichbarkeit der Unterrichtsqualität zwischen Klassen zu gewährleisten (zur Diskussion der Messung von Unterrichtsqualität siehe Abschnitt 4.5.1). Unterrichtsbeobachtungen durch externe Beurteilerinnen und Beurteiler würden auch eine umfassendere Einschätzung des Ausmaßes der kognitiven Aktivierung im Deutschunterricht ermöglichen, die in Teilstudie eins eher abstrakt über Schülereinschätzungen der Wichtigkeit verschiedener sprachlicher Aspekte für ihre Lehrkraft abgebildet wurde (siehe Abschnitt 4.3.2). So konnten beispielsweise Taylor, Pearson, Peterson und Rodriguez (2003) anhand von externen Beobachtungen des Leseunterrichts an Grundschulen der ersten bis fünften Klasse aufzeigen, dass Schülerinnen und Schüler in Klassen, deren Lehrkraft kognitive Prozesse höherer Ordnung durch ihre Fragen und Aufgaben anregte, den größten

Leselernzuwachs im Verlauf eines Schuljahres hatten²⁶. Aufgaben, die zur kognitiven Aktivierung anregen, beinhalteten beispielsweise die Betonung der Interpretation von Charakteren, Textzusammenfassungen oder Vorhersagen zur geschichtlichen Handlung vor oder während des Lesens. Das Ausmaß an kognitiver Aktivierung wurde anhand der Aufgabenhäufigkeit bestimmt, die dem Beurteilerrating entnommen wurde. Solche Aspekte könnten unter anderem in einer externen Beobachtung des Deutschunterrichts zur Operationalisierung der Unterrichtsqualität in zukünftigen Studien herangezogen werden (vgl. z.B. Videobeobachtungen im Fremdsprachenunterricht im Rahmen der DESI-Studie, Helmke et al., 2008). Insbesondere im Sprachunterricht ist eine tiefere Auseinandersetzung mit der Operationalisierung kognitiv aktivierender Unterrichtsprozesse vonnöten, bislang stützen sich ausführliche Beschreibungen der Unterrichtsqualität vorzugsweise auf den Mathematikunterricht (siehe z.B. die COACTIV-Studie: Kunter et al., 2011).

Eine weitere mit groß angelegten Schulleistungsstudien verbundene Situation ist, dass als Leistungsmaße überwiegend Leseverstehen und Mathematikleistungen erhoben werden, da sie zentrale Schülerkompetenzbereiche darstellen, deren Testverfahren zudem ökonomisch administriert und ausgewertet werden können. Für die auf Sekundäranalysen beruhende vorliegende Arbeit war eine Einbeziehung anderer Leistungsmaße deshalb nicht möglich. Wünschenswert wäre einerseits für Teilstudie drei auch die Erhebung von Leistungsmaßen, wie komplexe Problemlösefähigkeiten oder Kreativität, die der Annahme eines lernförderlichen Austauschs entsprechend stärkere Zusammenhänge mit zuwanderungsbezogener Heterogenität erwarten lassen würden (Gurin et al., 2003; vgl. Abschnitt 2.2.5.4 und Abschnitt 6.6.1). Andererseits wären für Teilstudie eins Maße wünschenswert, für die eine größere Leistungsentwicklung innerhalb des neunten Schuljahres zu erwarten wäre, beispielsweise im literarischen Textverstehen (vgl. Roick, Meier, Henschel & Frederking, 2014).

Eine weitere Schwierigkeit der vorliegenden Arbeit liegt in der Trennung von Effekten der Klassenzusammensetzung, die innerhalb der konkreten Klasse wirksam sind (z.B. durch Lehrererwartungen beeinflusste Unterrichtsqualität oder Peerprozesse) und Effekten weiterer Merkmale, die mit der Zuordnung von Schülerinnen und Schülern zu

²⁶ Der zitierte Beitrag analysiert den Zusammenhang zwischen verschiedenen Unterrichtsmerkmalen und Leistungszuwachs der Schülerinnen und Schüler allgemein. Zusammenhänge mit der Klassenzusammensetzung wurden nicht untersucht.

Schulen und Klassen assoziiert sind (siehe in Abschnitt 2.2.7 beschriebenes Endogenitätsproblem).

Systematische Zuordnungen von Schülerinnen und Schülern zu Schulen entstehen beispielsweise durch Schulwahl in Wohnortnähe und Nachbarschaftszusammensetzung, durch systematische Zusammenhänge zwischen Merkmalen des Elternhauses bzw. der Schülerinnen und Schüler und der Schulwahl sowie durch gegliederte Bildungssysteme. Diese Selektion hat zur Folge, dass bereits zum Zeitpunkt der Zusammenstellung der Klassen unterschiedliche Lernvoraussetzungen zwischen Schulen und Klassen bestehen. Dies kann zu Überschätzungen von Effekten der Zusammensetzung führen, wenn entsprechende Merkmale nicht ausreichend kontrolliert werden, da diese Lernvoraussetzungen auch mit der als Kriterium betrachteten Schülerleistung in Zusammenhang stehen. Im Falle von gegliederten Schulsystemen, wie dem deutschen, sollten also die Bildungsgänge in Analysen mit einbezogen und Eingangsfähigkeiten der Schülerinnen und Schüler anhand von Längsschnittstudien kontrolliert werden (siehe Vorgehen in Teilstudie eins und zwei). In Querschnittsstudien wird die Approximation der Vorkenntnisse durch kognitive Grundfähigkeiten empfohlen (Baumert et al., 2006), was in der querschnittlich angelegten Teilstudie drei der vorliegenden Arbeit angewendet wurde. Es bleibt dabei jedoch zu berücksichtigen, dass die Verwendung von Maßen der kognitiven Grundfähigkeit dennoch einerseits zu Überschätzungen von Effekten führen kann, da kognitive Grundfähigkeiten den Leistungsstand von Schülerinnen und Schülern weniger gut vorhersagen als domänenspezifische Vorkenntnisse (vgl. Diskussion bei Stanat, 2006b). Andererseits kann die Verwendung von Maßen der kognitiven Grundfähigkeit auch zu Unterschätzungen von Effekten der Zusammensetzung führen, da die Entwicklung der kognitiven Grundfähigkeiten selbst durch die Schule beeinflusst wird und zudem Phänomene wie *stereotype threat* differentiell für verschiedene Schülergruppen auf die Entwicklung der kognitiven Fähigkeiten einwirken können (vgl. Becker et al., 2012; Köller et al., 2013; Köller & Baumert, 2002; Schofield, 2006). Zudem ist damit zu rechnen, dass bei gleichzeitiger Erhebung von kognitiven Fähigkeiten und Kriteriumsmaßen deren Messfehler korrelieren, was zu einer zu starken Kontrolle der Vorleistung und Unterschätzung von Kompositionseffekten führen kann (Duncan & Raudenbush, 1999). Vor allem für Teilstudie drei ist der Einfluss weiterer Variablen bei Schätzung der Effekte der Zusammensetzung nur bedingt abschätzbar. Es lässt sich somit allgemein nicht ausschließen, dass die in der vorliegenden Arbeit ermittelten Effekte sich

zumindest zu einem Teil aus unberücksichtigten, systematischen Unterschieden zwischen Schulen ergeben.

Die Frage nach unberücksichtigten Drittvariablen ist auch für den Zusammenhang von Klassenzusammensetzung und Unterricht (vgl. Abschnitt 8.2) relevant. Einige Studien konnten zeigen, dass Schülerinnen und Schüler mit Zuwanderungshintergrund und/oder niedrigem sozioökonomischem Status häufiger an Schulen und auch Schularten mit geringerer Ressourcenausstattung vertreten sind, die sich durch weniger günstige organisationale und strukturelle Merkmale, wie höhere Schul- und Klassengröße, weniger qualifizierte Lehrkräfte und weniger zusätzliche Förderprogramme, auszeichnen (Baumert & Schümer, 2001; Hanushek et al., 2009; Raudenbush et al., 1998; Roscigno, 2000). Lehrkräften mit geringer Qualifizierung könnte es aufgrund fehlender professioneller Kompetenzen schwerer fallen, einen qualitativ hochwertigen Unterricht anzubieten, was zu einem negativen Zusammenhang zwischen dem Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund und der Unterrichtsqualität führen sollte (z.B. Eccles & Roeser, 2011). Effekte der Klassenzusammensetzung auf Unterrichtsqualität könnten also nicht, wie in der vorliegenden Arbeit diskutiert, auf differentielle Lehrererwartungen und Verhalten zurückführbar sein (siehe Abbildung 8.1 a), sondern auf die geringere Qualifikation von Lehrkräften an Schulen, die Schülerinnen und Schüler mit Zuwanderungshintergrund und/oder niedrigem sozioökonomischem Status besuchen (siehe Abbildung 8.1 b). Die Verbindung von Klassenzusammensetzung und Unterrichtsqualität kann des Weiteren durch eine systematische Zuordnung von Lehrkräften zu Klassen innerhalb von Schulen entstehen (vgl. Abschnitt 8.2). Dabei wären verschiedene Szenarien vorstellbar, etwa wenn als kompetent beurteilte Lehrkräfte von der Schulleitung zum Beispiel besonders leistungsschwache oder leistungsstarke Klassen oder gerade Klassen mit einem hohen oder niedrigen Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund zugewiesen bekommen. Auch durch Einflussnahme von Eltern und Lehrkräften kann es dazu kommen, dass bezogen auf Hintergrundmerkmale ungünstig zusammengesetzte Klassen vor allem von unerfahrenen Lehrkräften unterrichtet werden. Empirisch bestehen einige Hinweise auf die systematische Zuordnung innerhalb und zwischen Schulen (Burns & Mason, 2002; Kalorides & Loeb, 2013; Lauen & Gaddis, 2013; Palardy, 2015). Um solche Effekte besser verstehen zu können, wäre es vonnöten, die Zusammensetzung mehrerer Klassen derselben Schule zu betrachten (vgl. Legewie, 2012) und ggf. weitere Informationen über die Lehrkräfte und Zuteilungsprozesse zu besitzen (vgl. Abschnitt 8.2). In der

vorliegenden Arbeit lag jedoch nur eine Klasse pro Schule (Teilstudie drei) bzw. ein bis zwei Klassen pro Schule (Teilstudie eins und zwei) vor, was eine zufriedenstellende Trennung von Schul- und Klasseneffekten nicht ermöglicht.

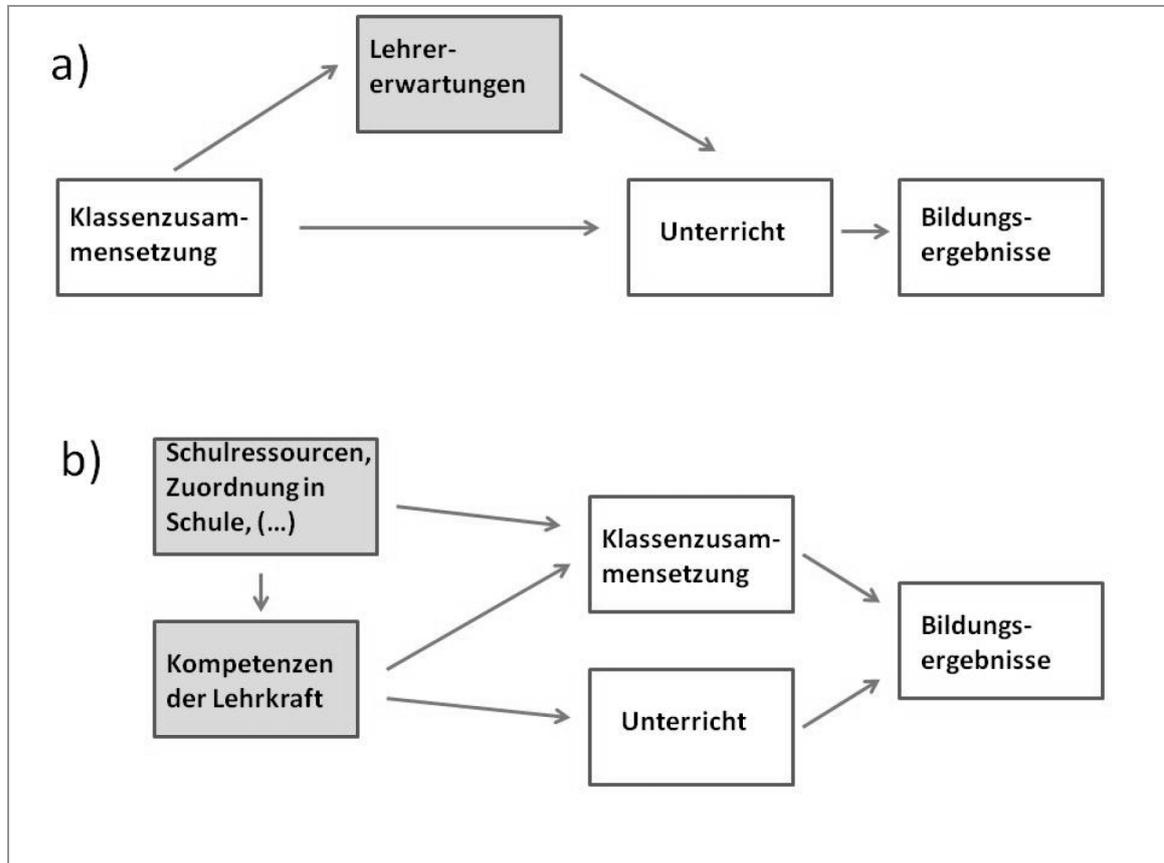


Abbildung 8.1: Beispielhafte schematische Darstellung möglicher Zusammenhänge zwischen Klassenzusammensetzung, Unterrichtsqualität und Bildungsergebnissen der Schülerinnen und Schüler

Auch die Analyse von mehreren Klassen, die von derselben Lehrkraft unterrichtet werden, sich jedoch in der Klassenzusammensetzung unterscheiden, würde zusätzlich belastbarere Ergebnisse zum Zusammenhang von Zusammensetzung, Unterrichtsqualität und Leistungsentwicklung liefern (vgl. Abschnitt 8.2 und Abschnitt 4.5.1). Allgemein böte die Nutzung von Zufallsexperimenten, wie sie beispielsweise von Duncan et al. (2004; vgl. Duncan & Raudenbush, 1999) oder Legewie (2012) beschrieben werden²⁷,

²⁷ Als Beispiele für Experimente zur Untersuchung von Kontexteffekten werden das Gautreau-Programm in den USA genannt, in dem verpflichtend Familien innerhalb des Raums Chicago umgesiedelt wurden und dies nachfolgend genutzt wurde um die Verbindung des Wohnumfelds und der schulischen Entwicklung der Kinder zu analysieren (Rosenbaum, 1991). Ein weiteres Beispiel stellt das Tennessee STAR Experiment dar, bei dem Schülerinnen und Schüler zufällig zu unterschiedlich großen Klassen zugeteilt

eine gute Möglichkeit belastbare, weniger verzerrte Effekte der Zusammensetzung zu schätzen – ein Umstand, der in zukünftiger Forschung stärker angestrebt werden sollte.

8.5 Implikationen für die pädagogische Praxis und zukünftige Forschung

8.5.1 Implikationen für die pädagogische Praxis

Die Befunde der vorliegenden Arbeit zur Bedeutung der Klassenzusammensetzung für die akademische Entwicklung von Schülerinnen und Schülern sowie Hinweise auf die vermittelnde Bedeutung der Unterrichtsqualität verweisen darauf, dass die Klassenzusammensetzung auch in der pädagogischen Praxis eine Rolle spielt. Zudem wurde argumentiert, dass die Befundlage für einen Beitrag von Effekten der Zusammensetzung zu zuwanderungsbezogenen Disparitäten spricht, was Bildungsgerechtigkeit entgegensteht (vgl. Abschnitt 8.3).

Diese Befunde werfen die Frage auf, wie in der pädagogischen Praxis mit segregierten Klassenzusammensetzungen umgegangen werden kann bzw. was getan werden kann, um negative Effekte entsprechender Zusammensetzungen auf die akademische Entwicklung abzuschwächen. In der Literatur diskutierte Reaktionsmöglichkeiten umfassen unter anderem (a) Desegregationsbemühungen durch gezielte Umstrukturierung der Schülerschaft von Schulen, (b) Veränderungen rechtlicher Vorgaben zum Beispiel zur Schulwahl, (c) Veränderung der Ressourcenallokation und (d) Maßnahmen zur Veränderung der Lehrerbildung und des Unterrichts.

(a) Programme der Desegregation durch extern gesteuerte Umverteilung von Schülerinnen und Schülern sind insbesondere aus den USA bekannt. Im Rahmen des „*busing*“ wurden vor allem in den 1970er und 1980er Jahren Schülerinnen und Schüler zu Schulen außerhalb ihres direkten Wohnorts transportiert, um vor allem die wohnortbedingte ethnische Segregation an Schulen zu verringern (z.B. Orfield, 2014). Einige Autorinnen und Autoren sprechen sich für solche gezielten Maßnahmen als einzig wirksame Möglichkeit zur Reaktion vor allem auf sozioökonomische Segregation aus (z.B. Kahlenberg, 2001 zitiert nach Rumberger & Palardy, 2005). Empirisch bestehen Hinweise darauf, dass Desegregation auf die Leistungen von afroamerikanischen

wurden. Diese Klassen wurden nachfolgend genutzt, um Kompositionseffekte auf Schulleistungen zu schätzen (Boozer & Cacciola, 2001).

Schülerinnen und Schüler einen kleinen positiven Effekt haben kann, ohne die Leistungen der anderen Schülerinnen und Schüler zu beeinflussen (Schofield, 1995; vgl. Schofield, 2006). Dennoch werden Maßnahmen wie das *busing* von mehreren Autoren durchaus kritisch gesehen (z.B. Schofield, 2006; Thrupp, 1995; Wilkinson et al., 2000). Abgesehen von praktischer Realisierbarkeit, ethischer Bedenken gegen zwangsweise Beschulungen an einem anderen Ort und finanziellen Kosten, wird die Wirksamkeit im Vergleich zu anderen Maßnahmen, wie dem Einwirken auf die Unterrichtsgestaltung, als geringer eingeschätzt (z.B. Thrupp, 1995; Wilkinson et al., 2000). Zudem ist zu bedenken, dass Studien zur reinen Desegregation keine Rückschlüsse über zugrundeliegende Prozesse, die etwaige Veränderungen in der Schülerleistung bewirkten, zulassen (Schofield, 2006).

(b) Eben benannte, zentral gesteuerte Umstrukturierungsmaßnahmen setzen meist Veränderungen rechtlicher Vorgaben zur Schulstruktur voraus. Die in den USA vorgenommenen Maßnahmen gehen zentral auf die gerichtlichen Beschlüsse im Rahmen von *Brown v. Board of Education* zurück, die in den 1950er Jahren festlegten, dass ethnische Trennungen an Schulen gegen den Gleichheitsgrundsatz der Vereinigten Staaten von Amerika sprechen (vgl. z.B. Orfield, 2014). Bereits seit den 1990er Jahren ist jedoch eine steigende Resegregation an US-amerikanischen Schulen zu beobachten (vgl. z.B. Lauen & Gaddis, 2013; Orfield, 2014)²⁸. Im europäischen Raum werden beispielsweise rechtliche Vorgaben zur Aufhebung fester Schulbezirke als eine Maßnahme diskutiert, um Beziehungen zwischen segregierten Wohnumfeldern und segregierten Schul- und Klassenzusammensetzungen abzuschwächen (vgl. z.B. Sachverständigenrat deutscher Stiftungen für Integration und Migration, 2012; Makles & Schneider, 2012; van Ackeren, 2006). Veränderungen der Vorgaben zur Schulwahl in England, Frankreich und den Niederlanden lassen jedoch vermuten, dass eine Erhöhung der Schulwahlmöglichkeiten eher eine stärkere Segregation mit sich bringt. Da bildungsnahe Eltern sich stärker mit der Schulwahl auseinandersetzen, Informationen zu Schulen leichter beschaffen können, den Ruf der Schule und Leistungsergebnisse stärker bei ihrer Wahl berücksichtigen und ggf. längere Schulwege eher in Kauf nehmen als bildungsferne Eltern, ist mit einer zunehmenden Entmischung der Schülerschaft zu

²⁸Die Resegregation steht in Verbindung mit veränderten Verfassungsinterpretationen. Das Merkmal Zuwanderungshintergrund bzw. ethnischer Hintergrund darf z.B. seit 2007 nicht mehr als Kriterium für oder gegen eine Schul- oder Universitätsaufnahme von Personen hinzugezogen werden, wodurch eine gesteuerte Desegregation erschwert wird (*Parents Involved v. Seattle School District No. 1*; vgl. Lauen & Gaddis, 2013; Orfield, 2014; Van Ewijk & Slegers, 2010b).

rechnen (van Ackeren, 2006). Aktuelles Beispiel für schulstrukturelle Veränderungen ist die in den letzten Jahren durchgeführte Berliner Schulreform, in der unter anderem die Wohnortnähe als Auswahlkriterium der weiterführenden Schule aufgehoben wurde (Maaz, Baumert, Neumann, Becker & Dumont, 2013). Für das Auswahlverfahren stark nachgefragter Schulen wurde eingeführt, dass nach Aufnahme von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf 60 Prozent der Plätze über Kriterien der Schule (z.B. Eingangstests, Notendurchschnitt) vergeben werden. Durch diesen angeregten Wettbewerb zwischen Schulen um leistungsstarke Schülerinnen und Schüler ist eine Entmischung der Schülerschaft zu erwarten. Um diesem Umstand zu begegnen wurde festgelegt, dass die Vergabe von 30 Prozent der Plätze an diesen Schulen über ein Losverfahren entschieden wird. Die Schulwahl beruht aber zunächst auf dem Elternwillen, sodass hier bereits systematische Unterschiede im Wahlverhalten von Eltern verschiedener sozialer Herkunft anzunehmen sind. Eine Untersuchung der Schulzusammensetzungen vor und nach der Berliner Schulstrukturereform steht in diesem Rahmen noch aus.

(c) Eine weitere Möglichkeit zum Umgang mit schulischer Segregation und der dadurch entstehenden Benachteiligung von Schülerinnen und Schülern ist die Veränderung der Ressourcenausstattung von Schulen, d.h. die zusätzliche finanzielle bzw. materielle und personelle Unterstützung von Schulen mit weniger lernförderlicher Zusammensetzung der Schülerschaft. Wie im theoretischen Rahmen der vorliegenden Arbeit zusammengefasst, sind Schülerinnen und Schüler mit Zuwanderungshintergrund und/oder niedrigem sozioökonomischen Status häufiger an Schulen vertreten, die sich durch eine geringere Ressourcenausstattung und dadurch weniger lernförderliche Bedingungen auszeichnen (vgl. Abschnitt 2.2.1). Daraus lässt sich eine Erhöhung von Ressourcen als Maßnahme an solchen Schulen ableiten (vgl. z.B. Bellin, 2009; Guldemond & Bosker, 2009; van Ewijk & Slegers, 2010b), auch wenn ihre Wirksamkeit teils infrage gestellt wird (z.B. Chubb & Moe, 1990, zitiert nach Harker & Tymms, 2004). Letztendlich kann davon ausgegangen werden, dass nicht die Ressourcenerhöhung allein entscheidend ist, sondern vielmehr, wie die erhaltenen Ressourcen von den Schulen eingesetzt werden. In Deutschland wird gegenwärtig eine Vergabe von Sondermitteln nach einem Sozialindex von Schulen etwa durch den Sachverständigenrat deutscher Stiftungen der Integration und Migration (2015) empfohlen, was bereits in einigen Ländern wie Nordrhein-Westfalen zur Praxis gehört (Bonsen et al., 2010; Sachverständigenrat deutscher Stiftungen der Integration und Migration, 2012). Die in

den letzten Jahren zunehmende Schulautonomie in Deutschland ermöglicht zudem eine bewusste Profilbildung an Schulen, die Schulen in Einzugsgebieten eines hohen Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund und bildungsferneren Elternhäusern dazu nutzen können, sich verstärkt für bildungsnähere Eltern interessant zu machen und so einer segregierten Zusammensetzung entgegenzuwirken (Sachverständigenrat deutscher Stiftungen der Integration und Migration, 2012; vgl. *Magnet Schools* in den USA, z.B. Rossell, 1990).

(d) Wilkinson et al. (2000) kommen in ihrem Review des Forschungsstandes zu Effekten der Zusammensetzung zu dem Schluss, dass vor allem Maßnahmen der Veränderung des Unterrichts und des Lehrerverhaltens notwendig seien, um die Effekte abzuschwächen: *„The attention needs to be directed at more careful curriculum specification, higher quality teaching, and higher expectations that students can meet appropriate challenges – and these occur once the classroom door is closed and not by reorganizing which students are behind those doors.“* (S. 83). Dies steht in Einklang mit Hinweisen aus der ersten Teilstudie der vorliegenden Arbeit. Für die pädagogische Praxis sind mindestens zweierlei Handlungsfelder ableitbar, um Effekte der Zusammensetzung durch Einwirken auf den Unterricht abzuschwächen. Zum einen ist eine Beschreibung hoher Leistungserwartungen an alle Schülerinnen und Schüler und das Etablieren allgemeingültiger anspruchsvoller Curricula notwendig (vgl. Rumberger & Palardy, 2005; Schofield, 2006). Es sollen durch Bildungsgerechtigkeit geforderte gleichmäßige Qualitätsansprüche für alle gesetzt werden. In Deutschland wird dies unter anderem durch die im letzten Jahrzehnt verabschiedeten Bildungsstandards angestrebt (KMK, 2005; vgl. Vereinigung der Bayerischen Wirtschaft e.V., 2007). Zum anderen ist eine stärkere Berücksichtigung der Effekte von Schülerzusammensetzung und Unterrichtsqualität in der Aus- und Weiterbildung von Lehrkräften notwendig. Zusammenhänge zwischen Klassenzusammensetzung und Unterrichtsqualität können sowohl durch systematische Unterschiede in der Lehrqualifikation als auch durch differentielle Lehrererwartungen an bestimmte Schülergruppen entstehen (vgl. Abschnitt 8.2 und Abschnitt 8.4). Demnach wären sowohl Ansätze, die sich mit der Kompetenz von Lehrkräften im qualitativ hochwertigen Unterrichten auch (zuwanderungsbezogen) heterogener Schülergruppen auseinandersetzen (vgl. z.B. García, Arias, Harris Murri & Serna, 2010; KMK, 2013b; Sachverständigenrat deutscher Stiftungen für Integration und Migration, 2015; Vedder et al., 2006), als auch Ansätze, die auf die Veränderung von Lehrererwartungen und –

überzeugungen bezüglich verschiedener Schülergruppen abzielen (vgl. Gay, 2010; Hachfeld et al., 2011; Schofield, 2006), notwendig.

Wenn eine bewusst gesteuerte Zuteilung von Schülerinnen und Schülern zu Klassen möglich wäre, ließen sich aus Befunden der vorliegenden Arbeit und aus vorherigen Studien folgende Empfehlungen für eine hypothetische, ideal leistungsförderliche Klassenzusammensetzung für die einzelne Schülerin bzw. den einzelnen Schüler ableiten: Für die individuelle Schülerin bzw. den individuellen Schüler wäre vor allem eine vorteilhafte Leistungsentwicklung in Schulklassen zu erwarten, die sich durch ein möglichst hohes Vorkenntnisniveau, hohes sozioökonomisches Niveau und niedrigen Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund auszeichnen. Diese hypothetische Zusammensetzung ist in der Realität per se nicht für alle Schülerinnen und Schüler umsetzbar. Es ist jedoch davon auszugehen, dass durch eine Schaffung weniger segregierter Zusammensetzungen bezogen auf Leistung und soziale und zuwanderungsbezogene Herkunft allgemein die Vorteile für die Leistungsentwicklung von individuell benachteiligten Schülerinnen und Schülern (d.h. leistungsschwächere Schülerinnen und Schüler mit niedrigem sozioökonomischen Hintergrund und/oder Zuwanderungshintergrund) gegenüber Nachteilen für leistungstärkere Schülerinnen und Schüler überwiegen (vgl. z.B. Kiss, 2013; Van Ewijk & Slegers, 2010b; Zimmer & Toma, 2000).

Für Merkmale zuwanderungsbezogener Klassenzusammensetzung zeigten sich leichte Vorteile heterogen zusammengesetzter Gruppen im mathematischen Bereich bzw. keine geringere sprachliche Leistung in zuwanderungsbedingt heterogenen Klassen. Verbunden mit einer Verringerung des Klassenanteils von Heranwachsenden mit Zuwanderungshintergrund in Bezug auf den Zuwanderungshintergrund stark segregierter Klassen ist zudem eine möglichen Abnahme der Lernmotivation und des Verbundenheitsgefühls mit den Peers bei Schülerinnen und Schülern mit Zuwanderungshintergrund zu erwarten. Es ist jedoch ebenfalls anzunehmen, dass durch die Aufhebung sehr segregierter Klassenzusammensetzungen positive Konsequenzen für die Leistungsentwicklung größeres Gewicht einnehmen als potentielle negative Konsequenzen im psychosozialen Bereich, denen durch zusätzliche flankierende Intervention begegnet werden könnte.

Da eine zentral gesteuerte Zuteilung von Schülerinnen und Schülern zu Klassen kaum in der Praxis realisierbar ist, lässt sich zusammenfassend festhalten, dass eine

möglichst schwache Gliederung des Bildungssystems und möglichst lange Zeit gemeinsamen Lernens segregierten Lernumfeldern und der Entstehung schwieriger Lernmilieus entgegen wirken könnten. Dies sollte einhergehen mit einer allgemeinen Orientierung an hohen Leistungsstandards für alle Schülerinnen und Schüler, die auch in der Lehrerbildung Berücksichtigung finden. Eine unabhängig von Klassenzusammensetzungen hohe Unterrichtsqualität würde dem Ziel der Bildungsgerechtigkeit dienen, das realisierbar wäre, wenn gleichzeitig differentielle Unterstützungsangebote für lernschwächere Schülerinnen und Schüler oder Personen mit besonderen Bedürfnissen bereit gestellt werden.

8.5.2 Implikationen für zukünftige Forschung

Wie bereits teilweise in den vorherigen Abschnitten angesprochen, lassen sich vielfältige Implikationen für zukünftige Forschung aus der vorliegenden Arbeit ableiten. Da davon auszugehen ist, dass auch die Klassenzusammensetzung zu zuwanderungsbezogenen Disparitäten im deutschen Bildungssystem beiträgt (vgl. Abschnitt 8.3), sollten Merkmale der Zusammensetzung auch in zukünftigen Studien, die den Bildungserfolg von Schülerinnen und Schülern mit Zuwanderungshintergrund untersuchen, Berücksichtigung finden. Dabei sollten mehrere Merkmale der Zusammensetzung und weitere Kontextmerkmale, wie der Bildungsgang, simultan hinzugezogen werden, da von einer Konfundierung leistungsbezogener, sozialer und zuwanderungsbezogener Zusammensetzung auszugehen ist. Bei Analyse von Effekten auf die Leistungsentwicklung individueller Schülerinnen und Schüler scheint der Klassenanteil von Schülerinnen und Schülern mit Zuwanderungshintergrund gemeinsam mit anderen Zusammensetzungsmerkmalen wirksam zu sein und kaum einen spezifischen Effekt aufzuweisen (vgl. Ergebnisse Teilstudie eins und drei).

Für psychosoziale Merkmale von Schülerinnen und Schülern sind jedoch spezifische Effekte des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund zu erwarten. Auch in Studien zur psychosozialen Entwicklung sollte also der Lernkontext inklusive der Klassenzusammensetzung berücksichtigt werden. Hierbei wäre eine vertiefte Auseinandersetzung mit und simultane Einbeziehung von weiteren psychosozialen Merkmalen erforderlich, um Prozesse zwischen Schülerinnen und Schülern besser zu verstehen. So kann die ermittelte positive Assoziation zwischen dem Anteil von Schülerinnen und Schülern mit

Zuwanderungshintergrund und der Lernmotivation im Sprachunterricht von Heranwachsenden mit Zuwanderungshintergrund mit einem Verbundenheitsgefühl zu Klassenkameradinnen und -kameraden sowie auch mit vielfältigen Vergleichsprozessen zwischen Schülerinnen und Schülern in Zusammenhang stehen. Klassen mit einem höheren Anteil von Heranwachsenden mit Zuwanderungshintergrund bieten mehr Vergleichsmöglichkeiten mit anderen Schülerinnen und Schülern mit Zuwanderungshintergrund, die unter anderem ein Vorbild für schulische Anpassung und Bildungsaspirationen sein können (vgl. Aboud, 1976). Des Weiteren wäre zu erwarten, dass der (herkunftsspezifische) Zuwanderungshintergrund relevant für die kulturelle Identität der Schülerinnen und Schüler ist und mit Zusammenhängen zwischen zuwanderungsbezogener Klassenzusammensetzung und individuellen psychosozialen Merkmalen in Verbindung steht, was in zukünftiger Forschung berücksichtigt werden sollte.

Bezüglich der vermittelnden Prozesse von Effekten der Zusammensetzung auf die Leistungsentwicklung von Schülerinnen und Schülern geben Befunde der vorliegenden Arbeit Hinweise darauf, dass die Unterrichtsqualität einen Transmissionsweg unter mehreren gleichzeitig ablaufenden Prozessen im Klassenraum darstellen könnte. Zukünftige Forschung in diesem Bereich sollte eine breitere Erfassung von Unterrichtsqualität, insbesondere von Merkmalen eines kognitiv aktivierenden Sprachunterrichts, anstreben. Die angenommene Verbindung von Zusammensetzung der Schülerschaft und Unterricht bedeutet auch, dass Merkmale der Zusammensetzung in der Schuleffektivitätsforschung Berücksichtigung finden sollten (vgl. Baumert et al., 2006; Dumay & Dupriez, 2007).

Was eine differenzierte Auseinandersetzung mit der Messung verschiedener Charakteristika der zuwanderungsbezogenen Zusammensetzung anbelangt, lässt sich eine größere Bedeutung des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund gegenüber Maßen der zuwanderungsbezogenen Heterogenität ableiten. In zukünftigen Studien, deren Autorinnen und Autoren die zuwanderungsbezogene Zusammensetzung zum Beispiel als Kovariate bei Leistungsvorhersagen einbeziehen möchten, ist eine Verwendung des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund als ausreichend einzustufen. Effekte der zuwanderungsbezogenen Heterogenität auf die untersuchten Merkmale von Bildungserträgen waren gering, und verschiedene Heterogenitätsindizes unterschieden sich nicht stark im Ergebnismuster. Inwiefern die zuwanderungsbezogene Heterogenität

in der Klasse jedoch mit weiteren Merkmalen, wie Offenheit gegenüber anderen Kulturen und interkultureller Kompetenz, in Zusammenhang steht, stellt eine weitere zukünftige Forschungsfrage dar. Verschiedene Heterogenitätsmaße sollten dabei zu vergleichbaren Ergebnissen führen, sodass die Verwendung nur eines Index ausreichend ist.

Letztendlich ist davon auszugehen, dass in Zukunft schülerbezogene Heterogenität im Bildungssystem noch stärker an Bedeutung gewinnen wird. Heterogenität bezieht sich dabei nicht nur auf zuwanderungsbezogene Heterogenität, die unter anderem durch veränderte Geburtenraten anteilig im deutschen Bildungssystem zunehmen wird. Steigende Heterogenität innerhalb von Schulen ist auch bezüglich weiterer Merkmale im Leistungsbereich zu erwarten, unter anderem durch Umstrukturierungen im Bildungssystem wie Abschaffung der Hauptschulen und durch Inklusionsprozesse von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf an allgemeinen Schulen. Die Bildungsforschung sollte sich somit mit der Abbildung von Heterogenität anhand multipler Merkmale auseinandersetzen, um deren Verbindung mit der akademischen Entwicklung von Schülerinnen und Schülern zu ermitteln sowie anschließend Einflussmöglichkeiten ableiten zu können. Vorbild kann hier die Forschung zur Heterogenität in Arbeitsgruppen in der Wirtschaftspsychologie darstellen, in der vergleichbare Fragestellungen vorzufinden sind. Analog zur Forschung zur Klassenzusammensetzung wurden in der Forschung zur Teamzusammensetzung zunächst verschiedene Mediatoren und Moderatoren untersucht und verschiedene Maße zur Abbildung ethnischer Teamzusammensetzung analysiert (vgl. z.B. Meyer, in press). Gegenwärtige Entwicklungen in der wirtschaftspsychologischen Forschung kommen zu dem Schluss, dass die Heterogenität der Teamzusammensetzung anhand der simultanen Einbeziehung multipler Merkmale abgebildet werden sollte. Dazu können Verfahren verwendet werden, in denen hypothetische Trennlinien („*faultlines*“) innerhalb von Gruppen basierend auf verschiedenen Personenmerkmalen bestimmt werden (vgl. Meyer, in press; Thatcher & Patel, 2012). Diese Entwicklungen können auch für den pädagogischen Bereich nutzbar gemacht werden, um mit der Erforschung zunehmender Heterogenität im Bildungssystem angemessen auf zukünftige Herausforderungen der Bildungsforschung zu reagieren.

9. Literaturverzeichnis

- About, F. E. (1976). Self-evaluation: Information seeking strategies for interethnic social comparisons. *Journal of Cross-Cultural Psychology*, 7, 289–300.
- Adams, R. J., Wu, M. L., & Carstensen, C. H. (2007). Application of multivariate Rasch models in international large-scale educational assessments. In M. v. Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models* (pp. 271-280). New York: Springer.
- Agirdag, O., Van Houtte, M., & Van Avermaet, P. (2012). Why does the ethnic and socio-economic composition of schools influence math achievement? The role of sense of futility and futility culture. *European Sociological Review*, 28, 366–378. doi: 10.1093/esr/jcq070
- Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, 39, 1490–1528. doi: 10.1177/0149206313478188
- Alba, R. D., Handl, J. & Müller, W. (1994). Ethnische Ungleichheiten im Deutschen Bildungssystem. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 46, 209-237.
- Alexander, K., & Schofield, J. W. (2006). Erwartungseffekte: Wie Lehrerverhalten schulische Leistungen beeinflusst. In J. W. Schofield (Hrsg.), *Migrationshintergrund, Minderheitenzugehörigkeit und Bildungserfolg. Forschungsergebnisse der pädagogischen, Entwicklungs- und Sozialpsychologie. AKI-Forschungsbilanz 5* (S. 47-69). Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).
- Allemann-Ghionda, C., Auernheimer, G., Grabbe, H. & Krämer, A. (2006). Beurteilung von schulischen Leistungen in mehrsprachigen Klassen: Die Kompetenz der Lehrpersonen. *Zeitschrift für Pädagogik*, 51. Beiheft, 250-266.
- Allemann-Ghionda, C. & Pfeiffer, S. (2008). Einleitung. In C. Allemann-Ghionda & S. Pfeiffer (Hrsg.), *Bildungserfolg, Migration und Zweisprachigkeit* (S. 7-11). Berlin: Frank & Timme GmbH.
- Ammermueller, A., & Pischke, J.-S. (2009). Peer effects in European primary schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27, 315-348. doi: 0734-306X/2009/2703-0005\$10.00
- Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, 30, 98–108. doi: 10.1016/j.labeco.2014.05.008
- Artelt, C., Baumert, J., Julius-McElvany, N. & Peschar, J. (2004). *Das Lernen lernen. Voraussetzungen für lebensbegleitendes Lernen. Ergebnisse von PISA 2000*. Paris: OECD.
- Autorengruppe Bildungsberichterstattung (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: W. Bertelsmann Verlag GmbH & Co. KG.

- Babad, E. (1993). Pygmalion - 25 years after interpersonal expectations in the classroom. In P. D. Blanck (Ed.), *Interpersonal expectations: Theory, research, and applications* (pp. 125-153). New York: Cambridge University Press.
- Barker, R. G. (1968). *Ecological Psychology: Concepts and methods for studying the environment of human behaviour*. Stanford: Stanford University Press.
- Barnett, M. D., Sonnert, G., & Sadler, P. M. (2012). More like us: The effect of immigrant generation on college success in mathematics. *International Migration Review*, 46, 891–918. doi: 10.1111/imre.12004
- Barr, R., & Dreeben, R. (1983). *How schools work*. Chicago: University of Chicago Press.
- Barth, J. M., Dunlap, S. T., Dane, H., Lochman, J. E., & Wells, K. C. (2004). Classroom environment influences on aggression, peer relations, and academic focus. *Journal of School Psychology*, 42, 115–133. doi: 10.1016/j.jsp.2003.11.004
- Battistich, V. (2010). School contexts that promote students' positive development. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 111-127). New York, NY: Routledge.
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117, 497–529. doi: 10.1037/0033-2909.117.3.497
- Baumert, J., Blum, W., & Neubrand, M. (2004). Drawing the lessons from PISA 2000 - Long-term research implications: Gaining a better understanding of the relationship between system system inputs and learning outcomes by assessing instructional and learning processes as mediating factors. *Zeitschrift für Erziehungswissenschaft*, 7, 147-157.
- Baumert, J., Carstensen, C. H. & Siegle, T. (2005). Wirtschaftliche, soziale und kulturelle Lebensverhältnisse und regionale Disparitäten des Kompetenzerwerbs. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003, Der zweite Vergleich der Länder in Deutschland - Was wissen und können Jugendliche?* (S. 323–365). Münster: Waxmann
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., . . . Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133-180. doi: 10.3102/0002831209345157
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und Schüler und ihrer Lehrkräfte. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland. Ergebnisse des zweiten internationalen Vergleichs* (S. 314-354). Münster: Waxmann.
- Baumert, J. & Maaz, K. (2012). Migration und Bildung in Deutschland. *Die Deutsche Schule*, 104, 279-302.
- Baumert, J., Maaz, K., Stanat, P. & Watermann, R. (2009). Schulkomposition oder Institution - was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die Deutsche Schule*, 101, 33-46.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000*.

- Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323-410). Opladen: Leske+Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Watermann, R. & Schümer, G. (2003). Disparitäten der Bildungsbeteiligung und des Kompetenzerwerbs. Ein institutionelles und individuelles Mediationsmodell. *Zeitschrift für Erziehungswissenschaft*, 6, 46-72.
- Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration (2010). 8. *Bericht der Beauftragten der Bundesregierung für Migration, Flüchtlinge und Integration über die Lage der Ausländerinnen und Ausländer in Deutschland*. Zugriff am 30.08.2011 unter http://www.bundesregierung.de/Content/DE/___Anlagen/2010/2010-07-07-langfassung-lagebericht-ib,property=publicationFile.pdf
- Becker, B. & Reimer, D. (2010). *Vom Kindergarten bis zur Hochschule. Die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Becker, G. S. (1993). *Human capital. A theoretical and empirical analysis with special reference to education*. Chicago: University of Chicago Press.
- Becker, M. (2009). *Kognitive Leistungsentwicklung in Differenziellen Lernumwelten: Effekte des gegliederten Sekundarschulsystems in Deutschland*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, 104, 682–699. doi: 10.1037/a0027608
- Bellin, N. (2009). *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Benner, A. D., & Crosnoe, R. (2011). The racial/ethnic composition of elementary schools and young children's academic and socioemotional functioning. *American Educational Research Journal*, 48, 621–646. doi: 10.3102/0002831210384838
- Benner, A. D., Graham, S., & Mistry, R. S. (2008). Discerning direct and mediated effects of ecological structures and processes on adolescents' educational outcomes. *Developmental Psychology*, 44, 840-854. doi: 10.1037/0012-1649.44.3.840
- Berkemeyer, N., Bos, W., Manitius, V., Hermstein, B., Bonitz, M. & Semper, I. (2014). *Chancenspiegel 2014. Regionale Disparitäten in der Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme*. Gütersloh: Bertelsmann Stiftung.
- Berkemeyer, N., Bos, W., Manitius, V., Hermstein, B. & Khalatbari, J. (2013). *Chancenspiegel 2013. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme mit einer Vertiefung zum schulischen Ganztag*. Gütersloh: Bertelsmann Stiftung.
- Berry, J. W. (1997). Immigration, acculturation, and adaptation: Lead Article. *Applied Psychology: An International Review*, 46, 5-68.

- Betts, J. R., Rueben, K. S., & Danenberg, A. (2000). *Equal resources, equal outcomes? The distribution of school resources and student achievement in California*. Public Policy Institute of California. Retrieved from http://www.ppic.org/content/pubs/report/r_200jbr.pdf
- Biemann, T., & Kearney, E. (2010). Size does matter: How varying group sizes in a sample affect the most common measures of group diversity. *Organizational Research Methods, 13*, 582-599. doi: 10.1177/1094428109338875
- Blau, P. M. (1960). Structural effects. *American Sociological Review, 25*, 178-193.
- Blau, P. M. (1974). Parameters of social structure. *American Sociological Review, 39*, 615-635.
- Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*. New York, NY: Free Press.
- Bonsen, M., Bos, W., Gröhlich, C., Harney, B. & Imhäuser, K. (2010). *Zur Konstruktion von Sozialindizes. Ein Beitrag zur Analyse sozialräumlicher Benachteiligung von Schulen als Voraussetzung für qualitative Schulentwicklung*. Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Boozer, M. A., & Cacciola, S. E. (2001). *Inside the „Black Box“ of Project STAR: Estimation of peer effects using experimental data*. New Haven: Yale University.
- Borman, K. M., Eitle, T. M., Michael, D., Eitle, D. J., Lee, R., Johnson, L., . . . Shircliffe, B. (2004). Accountability in a postdesegregation era: The continuing significance of racial segregation in Florida's schools. *American Educational Research Journal, 41*, 605-631.
- Bos, W., Lankes, E.-M., Plaßmeier, N. & Schwippert, K. (2004). *Heterogenität. Eine Herausforderung an die empirische Bildungsforschung*. Münster: Waxmann.
- Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.
- Bourdieu, P. (1982). *Die feinen Unterschiede: Kritik der gesellschaftlichen Urteilskraft*. Frankfurt am Main: Suhrkamp.
- Bourdieu, P. & Passeron, J.-C. (1964). *Die Illusion der Chancengleichheit. Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Stuttgart: Klett.
- Bovaird, J. A. (2007). Multilevel structural equation models for contextual factors. In T. D. Little, J. A. Bovaird & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 149-182). Mahwah, NJ: Erlbaum.
- Braun, H. I., Wang, A., Jenkins, F., & Weinbaum, E. (2006). The Black-White achievement gap: Do state policies matter? *Education Policy Analysis Archives, 14*, 1-110.
- Bronfenbrenner, U. (1976). *Ökologische Sozialisationsforschung*. Stuttgart: Klett-Cotta.
- Bronfenbrenner, U. (1996). Ökologische Sozialisationsforschung. In L. Kruse, C.-F. Graumann & E.-D. Lantermann (Hrsg.), *Ökologische Psychologie. Ein Handbuch in Schlüsselbegriffe* (S. 76-79). Weinheim: Beltz PVU.
- Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist, 41*, 1069-1077.

- Brophy, J. (2000). *Teaching*. Brussels, Belgium: International Academy of Education (IAE).
- Büchel, F. & Wagner, G. (1996). Soziale Differenzen der Bildungschancen in Westdeutschland – Unter besonderer Berücksichtigung von Zuwandererkindern. In W. Zapf, J. Schupp & R. Habich (Hrsg.), *Lebenslagen im Wandel. Sozialberichterstattung im Längsschnitt*. Frankfurt am Main: Campus Verlag.
- Budescu, D. V., & Budescu, M. (2012). How to measure diversity when you must. *Psychological Methods, 17*, 215–227. doi: 10.1037/a0027129
- Burns, R. B., & Mason, D. A. (2002). Class composition and student achievement in elementary schools. *American Educational Research Journal, 39*, 207-233. doi: 10.3102/00028312039001207
- Burris, C. C., Heubert, J. P., & Levin, H. M. (2006). Accelerating mathematics achievement using heterogeneous grouping. *American Educational Research Journal, 43*, 105-136. doi: 10.3102/00028312043001105
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema, 20*, 872– 882.
- Byrne, D. (1971). *The attraction paradigm*. New York, NY: Academic Press.
- Caldas, S. J., & Bankston III, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *The Journal of Educational Research, 90*, 269-277.
- Caldas, S. J., & Bankston III, C. (1998). The inequality of separation: Racial composition of schools and academic achievement. *Educational Administration Quarterly, 34*, 533-557.
- Carbonaro, W. J., & Gamoran, A. (2002). The production of achievement inequality in high school english. *American Educational Research Journal, 39*, 801-827. doi: 10.3102/00028312039004801
- Casteel, C. A. (1998). Teacher-student interactions and race in integrated classrooms. *The Journal of Educational Research, 92*, 115-120. doi: 10.1080/00220679809597583
- Chen, C., & Stevenson, H. W. (1995). Motivation and mathematics achievement: A comparative study of Asian-American, Caucasian-American, and East Asian high school students. *Child Development, 66*, 1215–1234. doi: 10.2307/1131808
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.
- Chmielewski, A. K., Dumont, H., & Trautwein, U. (2013). Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept. *American Educational Research Journal, 50*, 925–957. doi: 10.3102/0002831213489843
- Chubb, J., & Moe, T. (1990). *Politics, markets, and America's schools*. Washington, DC: The Brookings Institute.
- Chudgar, A., & Luschei, T. F. (2009). National income, income inequality, and the importance of schools: A hierarchical cross-national comparison. *American Educational Research Journal, 46*, 626-658. doi: 10.3102/0002831209340043
- Cialdini, R. B., Borden, R. J., Thorne, A., Walker, M. R., Freeman, S., & Sloan, L. R. (1976). Basking in reflected glory: Three (football) field studies. *Journal of*

- Personality and Social Psychology*, 34, 366–375. doi: 10.1037/0022-3514.34.3.366
- Coleman, J., Hoffer, T., & Kilgore, S. (1982). *Public, catholic and private schools compared*. New York: Basic Books.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, D. C.: Government Printing Office.
- Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, 119, 51-69. doi: 0033-2909/96/S3.00
- Conchas, G. Q. (2001). Structuring failure and success: Understanding the variability in Latino school engagement. *Harvard Educational Review*, 71, 475–504.
- Crosnoe, R., Johnson, M. K., & Elder, G. H. (2004). Intergenerational bonding in school: The behavioral and contextual correlates of student-teacher relationships. *Sociology of Education*, 77, 60-81.
- Danielsen, A. G., Wiium, N., Wilhelmsen, B. U., & Wold, B. (2010). Perceived support provided by teachers and classmates and students' self-reported academic initiative. *Journal of School Psychology*, 48, 247–267. doi: 10.1016/j.jsp.2010.02.002
- De Fraine, B., Van Damme, J., Van Landeghem, G., Opdenakker, M.-C., & Onghena, P. (2003). The effect of schools and classes on language achievement. *British Educational Research Journal*, 29, 841-859. doi: 10.1080/0141192032000137330
- De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research*, 4, 51-85. doi: 10.1023/A:1011402608575
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory*, 11, 227-268. doi: 10.1207/S15327965PLI1104_01
- Den Brok, P., & Levy, J. (2005). Teacher-student relationships in multicultural classes: Reviewing the past, preparing the future. *International Journal of Educational Research*, 43, 72-88.
- Denson, N., & Chang, M. J. (2009). Racial diversity matters: The impact of diversity-related student engagement and institutional context. *American Educational Research Journal*, 46, 322-353. doi: 10.3102/0002831208323278
- DESI-Konsortium. (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz.
- 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.

- Diefenbach, H. (2010). *Kinder und Jugendliche aus Migrantenfamilien im deutschen Bildungssystem. Erklärungen und empirische Befunde*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dijkstra, P., Kuyper, H., Van der Werf, G., Buunk, A. P., & Van der Zee, Y. G. (2008). Social comparison in the classroom: A review. *Review of Educational Research*, 78, 828–879. doi: 10.3102/0034654308321210
- Dimmick, J. R., & McDonald, D. G. (2001). Network radio as oligopoly, 1926-1956: Rivalrous imitation and program diversity. *Journal of Media Economics*, 14, 197-212. doi: 10.1207/S15327736ME1404_1
- Ditton, H. (1998). *Mehrebenenanalyse: Grundlagen und Anwendungen des Hierarchisch Linearen Modells*. Weinheim: Juventa.
- Ditton, H. & Krüsken, J. (2006). Sozialer Kontext und schulische Leistungen. Zur Bildungsrelevanz segregierter Armut. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, 26, 135-157.
- Ditton, H. & Müller, A. (2011). Schulqualität. In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung. Gegenstandsbereiche* (S. 99-111). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dollmann, J. (2010). *Türkischstämmige Kinder am ersten Bildungsübergang. Primäre und sekundäre Herkunftseffekte*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dollmann, J. (im Druck). Der Übergang von der Primar- in die Sekundarstufe. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 517-542). Wiesbaden: Springer.
- Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, 31, 117–135. doi: 10.1017/S026144480001315X
- Dougherty, K. D., & Huyser, K. R. (2008). Racially diverse congregations: Organizational identity and the accommodations of differences. *Journal for the Scientific Study of Religion*, 47, 23–43. doi: 10.1111/j.1468-5906.2008.00390.x
- Dreeben, R., & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, 61, 129-142.
- Driessen, G. (2002). School composition and achievement in primary education: A large-scale multilevel approach. *Studies in Educational Evaluation*, 28, 347-368. doi: SOISI-491X(02)00043-3
- Dronkers, J., & Levels, M. (2007). Do school segregation and school resources explain region-of-origin differences in the mathematics achievement of immigrant students? *Educational Research and Evaluation: An International Journal on Theory and Practice*, 13, 435 – 462.
- Dumay, X., & Dupriez, V. (2007). Accounting for class effect using the TIMSS 2003 eighth-grade database: Net effect of group composition, net effect of class process, and joint effect. *School Effectiveness and School Improvement*, 18, 383-408. doi: 10.1080/09243450601146371
- Dumont, H., Neumann, M., Becker, M., Maaz, K., & Baumert, J. (2013). Der Übergangsprozess von der Grundschule in die Sekundarstufe I vor und nach der Schulstrukturreform in Berlin: Die Rolle primärer und sekundärer Herkunftseffekte. In K. Maaz, J. Baumert, M. Neumann, M. Becker & H. Dumont

- (Hrsg.), *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen* (S. 133-208). Münster: Waxmann.
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. *Psychologie in Erziehung und Unterricht*, 60, 163-183. doi: 10.2378/peu2013.art14d
- Dumont, H., Neumann, M., Nagy, G., Becker, M., Rose, N. & Trautwein, U. (2013). Einfluss der Klassenkomposition auf die Leistungsentwicklung in Haupt- und Realschulen in Baden-Württemberg. *Psychologie in Erziehung und Unterricht*, 60, 198-213. doi: 10.2378/peu2013.art16d
- Duncan, G. J., Magnuson, K. A., & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1, 59–80. doi: 10.1080/15427609.2004.9683330
- Duncan, G. J., & Raudenbush, S. W. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist*, 34, 29-41.
- Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, 44, 78-89. doi: 10.1080/00461520902832368
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and in families. *American Psychologist*, 48, 90-101. doi: 10.1037/0003-066X.48.2.90
- Eccles, J. S., & Roeser, R. W. (2010). An ecological view of schools and development. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 6-21). New York, NY: Routledge.
- Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, 21, 225-241. doi: 10.1111/j.1532-7795.2010.00725.x
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132. doi: 10.1146/annurev.psych.53.100901.135153
- Edele, A., Stanat, P., Radmann, S. & Segeritz, M. (2013). Kulturelle Identität und Lesekompetenz von Jugendlichen aus zugewanderten Familien. *Zeitschrift für Pädagogik*, 59 .Beiheft, 84–110.
- Edelmann, D. (2008). Lehrer/-innenbildung im Kontext migrationsbedingter Heterogenität. Welche Kompetenzen brauchen Lehrpersonen, damit sie in mehrsprachigen Klassen effektiv unterrichten können? In C. Allemann-Ghionda & S. Pfeiffer (Hrsg.), *Bildungserfolg, Migration und Zweisprachigkeit* (S. 129-139). Berlin: Frank & Timme.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2010). *Statistik und Forschungsmethoden*. Weinheim: Beltz.
- Eitle, T. M., & Eitle, D. J. (2004). Inequality, segregation, and the overrepresentation of African Americans in school suspensions. *Sociological Perspectives*, 47, 269-287.
- Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*, 121–138. doi: 10.1037/1082-989X.12.2.121
- Entwisle, D. R., & Alexander, K. L. (1994). Winter setback: The racial composition of schools and learning to read. *American Sociological Review, 59*, 446-460.
- Esser, H. (2006). *Sprache und Integration. Die sozialen Bedingungen und Folgen des Spracherwerbs von Migranten*. Frankfurt am Main: Campus Verlag.
- Evans, G. W., Yoo, M. J., & Sipple, J. (2010). The ecological context of student achievement: School building quality effects are exacerbated by high levels of student mobility. *Journal of Environmental Psychology, 30*, 239-244. doi: 10.1016/j.jenvp.2010.01.001
- Fager, E. W. (1972). Diversity: A sampling study. *The American Naturalist, 106*, 293-310.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2013). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9. doi: 10.1016/j.learninstruc.2013.07.001
- Federal Statistical Office Germany (Statistisches Bundesamt) (2012). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund - Ergebnisse des Mikrozensus 2011*. Wiesbaden: Statistisches Bundesamt. Heruntergeladen von Zugriff am 01.04.2013 unter <https://www.destatis.de>.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.
- Flam, H. (2007). Sieben Phasen der Einwanderung in die BRD. In H. Flam (Hrsg.), *Migranten in Deutschland. Statistiken – Fakten – Diskurse* (S. 293-297). Konstanz: UVK Verlagsgesellschaft mbH.
- Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction, 17*, 478-493. doi: 10.1016/j.learninstruc.2007.09.001
- Friedrichs, J. (2008). Ethnische Segregation. In F. Kalter (Hrsg.), *Migration und Integration* (S. 380–411). Wiesbaden: VS Verl. für Sozialwissenschaften.
- Fuligni, A. J. (2001). Family obligation and the academic motivation of adolescents from Asian, Latin American, and European backgrounds. *New Directions for Child and Adolescent Development, 94*, 61–75. doi: 10.1002/cd.31
- Fuligni, A. J., Witkow, M., & Garcia, C. (2005). Ethnic identity and the academic adjustment of adolescents from Mexican, Chinese, and European backgrounds. *Developmental Psychology, 41*, 799–811. doi: 10.1037/0012-1649.41.5.799
- Gamoran, A. (1992). The variable effects of high school tracking. *American Sociological Review, 57*, 812-828.
- Ganzeboom, H. B. G. (2010). *A new international socio-economic index [ISEI] of occupational status for the International Standard Classification of Occupation 2008 [ISCO-08] constructed with data from the ISSP 2002-2007; with an analysis of quality of educational measurement in ISSP*. Paper presented at the Annual Conference of International Social Survey Programme, Lisbon. Zugriff am

12.12.2014 unter <http://www.harryganzeboom.nl/Pdf/2010-Ganzeboom-ISEI08-ISSP-Lisbon-%28paper%29.pdf>

- Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D. J., & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*, 1-56. doi: 10.1016/0049-089X(92)90017-B
- Garces, L. M., & Jayakumar, U. M. (2014). Dynamic diversity: Toward a contextual understanding of critical mass. *Educational Researcher, 43*, 115–124. doi: 10.3102/0013189X14529814
- Garcia, E., Arias, M. B., Harris Murri, N. J., & Serna, C. (2010). Developing responsive teachers: A challenge for a demographic reality. *Journal of Teacher Education, 61*, 132-142.
- Gay, G. (2000). *Culturally responsive teaching: Theory, research, and practice*. New York: Teachers College Press.
- Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education, 53*, 106-116.
- Gay, G. (2010). Acting on beliefs in teacher education for cultural diversity. *Journal of Teacher Education, 61*, 143-152.
- Giesinger, J. (2007). Was heißt Bildungsgerechtigkeit? *Zeitschrift für Pädagogik, 53*, 362-381.
- Goldsmith, P. A. (2004). Schools' racial mix, students' optimism, and the Black-White and Latino-White achievement gaps. *Sociology of Education, 77*, 121-147. doi: 10.1177/003804070407700202
- Goldsmith, P. R. (2011). Coleman revisited: School segregation, peers, and frog ponds. *American Educational Research Journal, 48*, 508-535. doi: DOI: 10.3102/0002831210392019
- Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *Journal of Early Adolescence, 13*, 21-43. doi: 10.1177/0272431693013001002
- Gottfredson, N. C., Panter, A. T., Daye, C. E., Allen, W. A., & Wightman, L. F. (2009). The effects of educational diversity in a national sample of law students: Fitting multilevel latent variable models in data with categorical indicators. *Multivariate Behavioral Research, 44*, 305-331. doi: 10.1080/00273170902949719
- Gottfried, M. A. (2014). Peer effects in urban schools: Assessing the impact of classroom composition on student achievement. *Educational Policy, 28*, 607–647. doi: 10.1177/0895904812467082
- Gould, E. D., Lavy, V., & Paserman, M. D. (2009). Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. *The Economic Journal, 119*, 1243–1269. doi: 10.1111/j.1468-0297.2009.02271.x
- Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods, 11*, 323-343. doi: 10.1037/1082-989X.11.4.323
- Gresch, C. (2012). *Der Übergang in die Sekundarstufe I. Leistungsbeurteilung, Bildungsaspiration und rechtlicher Kontext bei Kindern mit Migrationshintergrund*. Wiesbaden: Springer.

- Gresch, C. & Becker, M. (2010). Sozial- und leistungsbedingte Disparitäten im Übergangsverhalten bei türkischstämmigen Kindern und Kindern aus (Spät-)Aussiedlerfamilien. In K. Maaz, J. Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule - Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 181-200). Bonn: BMBF.
- Grundmann, M. & Kunze, I. (2008). Systematische Sozialraumforschung: Urie Bronfenbrenners Ökologie der menschlichen Entwicklung und die Modellierung mikrosozialer Raumgestaltung. In F. Kessl & C. Reutlinger (Hrsg.), *Schlüsselwerke der Sozialraumforschung. Traditionslinien in Text und Kontexten* (S. 172-188). Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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, 124–136. doi: 10.1037/0022-0663.95.1.124
- Guldemon, H., & Bosker, R. J. (2009). School effects on students' progress: A dynamic perspective. *School Effectiveness and School Improvement*, 20, 255-268. doi: 10.1080/09243450902883938
- Gurin, P., Nagda, B. A., & Lopez, G. E. (2004). The benefits of diversity in education for democratic citizenship. *Journal of Social Issues*, 60, 17–34. doi: 10.1111/j.0022-4537.2004.00097.x
- Gurin, P. Y., Dey, E. L., Gurin, G., & Hurtado, S. (2003). How does racial/ethnic diversity promote education? *Western Journal of Black Studies*, 27, 20–29.
- Haag, N., Böhme, K. & Stanat, P. (2012). Zuwanderungsbezogene Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 209-236). Münster: Waxmann.
- Hachfeld, A., Hahn, A., Schroeder, S., Anders, Y., Stanat, P., & Kunter, M. (2011). Assessing teachers' multicultural and egalitarian beliefs: The Teacher Cultural Beliefs Scale. *Teaching and Teacher Education*, 27, 986-996. doi: 10.1016/j.tate.2011.04.006
- Hall, M., & Tideman, N. (1967). Measures of concentration. *Journal of the American Statistical Association*, 62, 162–168.
- Hamre, B. K., & Pianta, R. C. (2010). Classroom environments and developmental processes: Conceptualization and measurement. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 25-41). New York, NY: Routledge.
- Hanushek, E., Kain, J. F., & Rivkin, S. G. (2009). New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement. *Journal of Labor Economics*, 27, 349-383. doi: 0734-306X/2009/2703-0002\$10.00
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19, 141-164.

- Harker, R., & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15, 177-199. doi: 10.1076/sesi.15.2.177.30432
- Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *The Academy of Management Review*, 32, 1199-1228.
- Hartig, J., Jude, N. & Wagner, W. (2008). Methodische Grundlagen zur Messung und Erklärung sprachlicher Kompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 34-54). Weinheim: Beltz Pädagogik.
- Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37, 449-481. doi: 10.1016/S0883-0355(03)00015-6
- Hauser-Cram, P., Sirin, S. R., & Stipek, D. (2003). When teachers' and parents' values differ: Teachers' ratings of academic competence in children from low-income families. *Journal of Educational Psychology*, 95, 813-820.
- Hauser, R. M. (1970). Context and consex: A cautionary tale. *American Journal of Sociology*, 75, 645-664.
- Heckhausen, J. & Heckhausen, H. (2006). Motivation und Handeln: Einführung und Überblick. In J. Heckhausen & H. Heckhausen (Hrsg.), *Motivation und Handeln* (S. 1-9). Heidelberg: Springer.
- Helbig, M. (2010). Neighborhood does matter! Soziostrukturelle Nachbarschaftscharakteristika und Bildungserfolg. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 655–679. doi: 10.1007/s11577-010-0117-y
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4.-12. Klassen, Revision (KFT 4-12+ R)*. Göttingen: Hogrefe.
- Helmke, A. (2010). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Seelze: Klett-Kallmeyer.
- Helmke, A. & Klieme, E. (2008). Unterricht und Lehrerkompetenzen. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 301-312). Weinheim: Beltz.
- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich D Praxisgebiete, Serie I Pädagogische Psychologie, Band 3* (S. 71–176). Göttingen: Hogrefe.
- Helmke, T., Helmke, A., Schrader, F.-W., Wagner, W., Nold, G., & Schröder, K. (2008). Die Videostudie des Englischunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI.-Studie* (S. 345-363). Weinheim: Beltz.
- Heppt, B., Haag, N., Böhme, K., & Stanat, P. (2014). The role of academic-language features for reading comprehension of language-minority students and students from low- SES families. *Reading Research Quarterly*, 50, 61-82. doi: 10.1002/rrq.83
- Herwartz-Emden, L. (2003). Einwandererkinder im deutschen Bildungswesen. In K. S. Cortina, J. Baumert, A. Leschinsky & K.-U. Mayer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 661-709). Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.

- Holzberger, D., Praetorius, A.-K., & Kunter, M. (2015, März). *Unterrichtsqualität: Profile von Lehrkräften und deren Zusammenhang mit der Kompetenzentwicklung von Schülerinnen und Schülern*. Vortrag auf der 3. Jahrestagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Bochum.
- Honneth, A. (2011). *Das Recht der Freiheit. Grundriß einer demokratischen Sittlichkeit*. Berlin: Suhrkamp.
- Hopf, W. (2011). Bildung, chancengleiche Konkurrenz und gleiche gesellschaftliche Teilhabe. *WSI Mitteilungen*, 4/2011, 195-201.
- Hoppe, A., Fujishiro, K., & Heaney, C. (2014). Workplace racial/ethnic similarity, job satisfaction, and lumbar back health among warehouse workers: Asymmetric reactions across racial/ethnic groups. *Journal of Organizational Behavior*, 35, 172–193. doi: 10.1002/job.1860
- Hox, J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.
- Hoxby, C. M. (2000). *Peer effects in the classroom learning from gender and race variation (NBER Working Papers No. 7867)*. Cambridge, MA: National Bureau of Economic Research.
- Hunt, D. E. (1975). Person-environment interaction: A challenge found wanting before it was tried. *Review of Educational Research*, 45, 209-230.
- Hustinx, P. W. J. (2002). School careers of pupils of ethnic minority background after the transition to secondary education: Is the ethnic factor always negative? *Educational Research and Evaluation: An International Journal on Theory and Practice*, 8, 169-195.
- Jencks, C., Smith, M., Acland, H., J., B. M., Cohen, D., H., G., . . . Michelson, S. (1972). *Inequality*. New York: Basic Books.
- Jensen, R., & Würtz Rasmussen, A. (2011). The effect of immigrant concentration in schools on native and immigrant children's reading and math skills. *Economics of Education Review*, 30, 1503– 1515. doi: 10.1016/j.econedurev.2011.08.002
- Johnson, M. K., Crosnoe, R., & Elder, G. H. (2001). Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education*, 74, 318-340.
- Jungbluth, P. (1999). Lehrererwartungen und Ethnizität. Innerschulische Chancendeterminanten bei Migrantenkindern in den Niederlanden. *Zeitschrift für Pädagogik*, 40, 113–125.
- Junge, K. (1994). Diversity of ideas about diversity measurement. *Scandinavian Journal of Psychology*, 36, 16–26.
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131-155. doi: 10.1207/s15327957pspr0902_3
- Juvonen, J., Nishina, A., & Graham, S. (2006). Ethnic diversity and perceptions of safety in urban middle schools. *Psychological Science*, 17, 393-400.
- Kahlenberg, R. D. (2001). *All together now: Creating middle-class schools through public school choice*. Washington, DC: Brookings Institution.

- Kalorides, D., & Loeb, S. (2013). Different teachers, different peers: The magnitude of student sorting within schools. *Educational Researcher*, 42, 304-316. doi: 10.3102/0013189X13495087
- Kao, G., & Tienda, M. (1995). Optimism and achievement: The educational performance of immigrant youth. *Social Science Quarterly*, 76, 1-19.
- Karsten, S., Felix, C., Ledoux, G., Meijnen, W., Roeleveld, J., & Van Schooten, E. (2006). Choosing segregation or integration? The extent and effects of ethnic segregation in Dutch cities. *Education and Urban Society*, 38, 228-247. doi: 10.1177/0013124505282606
- Kiss, D. (2013). The impact of peer achievement and peer heterogeneity on own achievement growth: Evidence from school transitions. *Economics of Education Review*, 37, 58–65. doi: 10.1016/j.econedurev.2013.08.002
- Klieme, E., Jude, N., Rauch, D., Ehlers, H., Helmke, A., Eichler, W., . . . Willenberg, H. (2008). Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts. In DESI-Konsortium (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 319-344). Weinheim und Basel: Beltz.
- Klieme, E., Lipowsky, F., Rakoczy, K. & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. M. Prenzel & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms*. (S. 128–146). Münster: Waxmann.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54, 222-237.
- Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: Aufgabenkultur und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (Hrsg.), *TIMSS - Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43-57). München: Medienhaus Biering.
- Klieme, E., Steinert, B. & Hochweber, J. (2010). Zur Bedeutung der Schulqualität für Unterricht und Lernergebnisse. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S. 231-255). Münster: Waxmann.
- KMK (2005). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung*. München: Luchterhand.
- KMK (2012). *Basic structure of the education system in the Federal Republic of Germany*. Zugriff am 11.12.2012 unter http://www.kmk.org/fileadmin/doc/Dokumentation/Bildungswesen_en_pdfs/en-2012.pdf
- KMK (2013a). *The education system in the Federal Republic of Germany 2011/2012. A description of the responsibilities, structures and developments in education policy for the exchange of information in Europe*. Zugriff am 24.07.2013 unter http://www.kmk.org/fileadmin/doc/Dokumentation/Bildungswesen_en_pdfs/dossier_en_ebook.pdf
- KMK. (2013b). *Interkulturelle Bildung und Erziehung in der Schule (Beschluss der Kultusministerkonferenz vom 25.10.1996 i. d. F. vom 05.12.2013)*. Zugriff am

- 26.06.2015 unter <http://www.kmk.org/bildung-schule/allgemeine-bildung/migration-integration.html>
- Knigge, M. (2009). *Hauptschüler als Bildungsverlierer? Eine Studie zu Stigma und selbstbezogenem Wissen bei einer gesellschaftlichen Problemgruppe*. Münster: Waxmann.
- Köller, O. (2004). *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- Köller, O. & Baumert, J. (2002). Entwicklung schulischer Leistungen. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 756–786). Weinheim: Beltz.
- Köller, O., Schütte, K., Zimmermann, F., Retelsdorf, J. & Leucht, M. (2013). Starke Klasse, hohe Leistungen? Die Rolle der Leistungsstärke der Klasse für die individuellen Mathematik- und Leseleistungen in der Sekundarstufe I. *Psychologie in Erziehung und Unterricht*, 60, 184-197. doi: 10.2378/peu2013.art15d
- Konan, P. N. D., Chatard, A., Selimbegović, L., & Mugny, G. (2010). Cultural diversity in the classroom and its effects on academic performance: A cross-national perspective. *Social Psychology*, 41, 230-237. doi: 10.1027/1864-9335/a000031
- Korte, E. (1990). Die Rückkehrorientierung im Eingliederungsprozess der Migrantenfamilien. In H. Esser & J. Friedrich (Hrsg.), *Generation und Identität. Theoretische und empirische Beiträge zur Migrationssoziologie* (S. 207–259). Opladen: Westdeutscher Verlag.
- Kozaki, Y., & Ross, S. J. (2011). Contextual dynamics in foreign language learning motivation. *Language Learning*, 61, 1328-1354. doi: 10.1111/j.1467-9922.2011.00638.x
- Kramer, R.-T. & Helsper, W. (2010). Kulturelle Passung und Bildungsungleichheit – Potenziale einer an Bourdieu orientierten Analyse der Bildungsungleichheit. In H.-H. Krüger, U. Rabe-Kleberg, R.-T. Kramer & J. Budde (Hrsg.), *Bildungsungleichheit revisited. Bildung und soziale Ungleichheit vom Kindergarten bis zur Hochschule* (S. 103-125). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krapp, A. (1993). Die Psychologie der Lernmotivation – Perspektiven der Forschung und Probleme ihrer pädagogischen Rezeption. *Zeitschrift für Pädagogik*, 39, 187–206.
- Kristen, C. (2008). Primary school choice and ethnic school segregation in German elementary schools. *European Sociological Review*, 24, 495–510.
- Kristen, C. & Dollmann, J. (2009). Sekundäre Effekte der ethnischen Herkunft? Kinder aus türkischen Familien am ersten Bildungsübergang. *Zeitschrift für Erziehungswissenschaft, Sonderheft 12*, 205-229.
- Kristen, C., Edele, A., Kalter, F., Kogan, I., Schulz, B., Stanat, P., & Will, G. (2011). The education of migrants and their children across the life course. *Zeitschrift für Erziehungswissenschaft*, 14, 121–137. doi: 10.1007/s11618-011-0194-3
- Kuhl, P., Siegle, T. & Lenski, A. (2013). Soziale Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 275-296). Münster: Waxmann.
- Kumar, R. (2006). Students experiences of home–school dissonance: The role of school academic culture and perceptions of classroom goal structures. *Contemporary Educational Psychology*, 31, 253–279. doi: 10.1016/j.cedpsych.2005.08.002

- Kumar, R., & Maehr, M. L. (2010). Schooling, cultural diversity, and student motivation. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 308-324). New York, NY: Routledge.
- Kunter, M., & Baumert, J. (2006a). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, *9*, 231–251. doi: 10.1007/s10984-006-9015-7
- Kunter, M., & Baumert, J. (2006b). Linking TIMSS to research on learning and instruction: A reanalysis of the German TIMSS and TIMSS video data. In S. J. Howie, & T. Plomp (Eds.), *Learning mathematics and science: Lessons learned from TIMSS* (pp. 335–351). London: Routledge.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster: Waxmann.
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction*, *17*, 494-509. doi: 10.1016/j.learninstruc.2007.09.002
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*, 805–820. doi: 10.1037/a0032583
- Kvålseth, T. O. (1991). Note on biological diversity, evenness, and homogeneity measures. *Oikos*, *62*, 123–127.
- Lanahan, L., McGrath, D. J., McLaughlin, M., Burian-Fitzgerald, M., & Salganik, L. (2005). *Fundamental problems in the measurement of instructional processes: estimating reasonable effect sizes and conceptualizing what is important to measure*. Washington: American Institutes for Research.
- Lauen, D. L., & Gaddis, S. M. (2013). Exposure to classroom poverty and test score achievement: Contextual effects or selection? *American Journal of Sociology*, *118*, 943-979. doi: 0002-9602/2013/11804-0003\$10.00
- Lazear, E. P. (1999). Culture and language. *Journal of Political Economy*, *107*, 95-126.
- Lee, H. (2007). The effects of school racial and ethnic composition on academic achievement during adolescence. *Journal of Negro Education*, *76*, 154-172.
- Lee, V. E., Smerdon, B. A., Alfeld-Liro, C., & Brown, S. L. (2000). Inside large and small high schools: Curriculum and social relations. *Educational Evaluation and Policy Analysis*, *22*, 147–171. doi: 10.3102/01623737022002147
- Lee, V. E., & Smith, J. B. (1999). Social support and achievement for young adolescents in Chicago: The role of school academic press. *American Educational Research Journal*, *36*, 907-945.
- Leenen, W. R., Grosch, H. & Kreidt, U. (1990). Bildungsverständnis, Platzierungsverhalten und Generationenkonflikt in türkischen Migrantenfamilien. *Zeitschrift für Pädagogik*, *36*, 753–771.
- Legewie, J. (2012). Die Schätzung von kausalen Effekten: Überlegungen zu Methoden der Kausalanalyse anhand von Kontexteffekten in der Schule. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *64*, 123–153. doi: 10.1007/s11577-012-0158-5

- Lehmann, R. (2008). *Erhebung zum Lese- und Mathematikverständnis: Entwicklungen in den Jahrgangsstufen 4 bis 6 in Berlin*. Version: 1. IQB – Institut zur Qualitätsentwicklung im Bildungswesen. Datensatz. doi: 10.5159/IQB_ELEMENT_v1
- Les, M., & Maher, C. (1998). Measuring diversity: Choice in local housing markets. *Geographical Analysis*, 30, 172–190. doi: 10.1111/j.1538-4632.1998.tb00395.x
- Levy, J., Wubbels, T., Brekelmans, M., & Morganfield, B. (1997). Language and cultural factors in students' perception of teacher communication style. *International Journal of Intercultural Relations*, 21, 29-56. doi: S0147-1767(96)00005-3
- Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- Lewin, K. (1963). *Feldtheorien in den Sozialwissenschaften*. Bern: Huber.
- Lieberson, S. (1969). Measuring population diversity. *American Sociological Review*, 34, 850–862. doi: 10.2307/2095977
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19, 527-537. doi: 10.1016/j.learninstruc.2008.11.001
- Lorenz, G., Gentrup, S., Kristen, C., Stanat, P., & Kogan, I. (eingereicht). Stereotype bei Lehrkräften? Eine Untersuchung systematisch verzerrter Lehrererwartungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*.
- Lubienski, S. T., & Lubienski, C. (2006). School sector and academic achievement: A multilevel analysis of NAEP mathematics data. *American Educational Research Journal*, 43, 651-698. doi: 10.3102/00028312043004651
- Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2 x 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, 16, 444-467. doi: 10.1037/a0024376
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-Level effects in contextual studies. *Psychological Methods*, 13, 203-229.
- Lüdtke, O., Robitzsch, A. & Köller, O. (2002). Statistische Artefakte bei Kontexteffekten in der pädagogisch-psychologischen Forschung. *Zeitschrift für Pädagogische Psychologie*, 16, 217-231. doi: 10.1024//1010-0652.16.34.217
- Lüdtke, O., Robitzsch, A., Trautwein, U., & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology*, 34, 120-131. doi: 10.1016/j.cedpsych.2008.12.001
- Maaz, K., Baumert, J., Gresch, C. & McElvany, N. (2010). *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten*. Berlin: BMBF.
- Maaz, K., Baumert, J., Neumann, M., Becker, M. & Dumont, H. (2013). *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen*. Münster: Waxmann.

- Maaz, K., Baumert, J. & Trautwein, U. (2010). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? In K. Maaz, J. Baumert, C. Gresch & N. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule – Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 27-63). Berlin: BMBF.
- Maaz, K., Hausen, C., McElvany, N. & Baumert, J. (2006). Stichwort: Übergänge im Bildungssystem. *Zeitschrift für Erziehungswissenschaft*, 9, 299-327.
- Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, 2, 99–106.
- MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews*, 40, 510–533. doi: 10.1111/j.1469-185X.1965.tb00815.x
- Maestri, V. (2011). Can ethnic diversity have a positive effect on school achievement? Retrieved from <http://mpira.ub.uni-muenchen.de/33455/>
- Maggi, S., Hertzman, C., Kohen, D., & D'Angiulli, A. (2004). Effects of neighborhood socioeconomic characteristics and class composition on highly competent children. *Journal of Educational Research*, 98, 109-114.
- Makles, A. & Schneider, K. (2012). Freie Wahl der Grundschule: Wie entscheiden sich Eltern und welche Konsequenzen hat die Schulwahl für die Segregation? *Die Deutsche Schule*, 104, 332-346.
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23, 129–149.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic selfconcept. *Journal of Educational Psychology*, 79, 280–295. doi: 10.1037/0022-0663.79.3.280
- Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Longitudinal multilevel modeling of the big fish little pond effect on academic self-concept: Counterbalancing social comparison and reflected glory effects in Hong Kong high schools. *Journal Personality and Social Psychology*, 78, 337–349.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J. S., Abduljabbar, A. S., & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, 47, 106–124. doi: 10.1080/00461520.2012.670488
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research*, 44, 764–802. doi: 10.1080/00273170903333665
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x
- McCann, K. (2000). The diversity–stability debate. *Nature*, 405, 228–233. doi: 10.1038/35012234

- McDonald, D. G., & Dimmick, J. (2003). The conceptualization and measurement of diversity. *Communication Research*, 30, 60-79. doi: 10.1177/0093650202239026
- Meece, J. L., & Schaefer, V. A. (2010). Schools as context of human development. In J. L. Meece & J. S. Eccles (Eds.), *Handbook of research on schools, schooling, and human development* (pp. 3-5). New York, NY: Routledge.
- Meisel, J. C., & Blumberg, C. J. (1990). The social comparison choices of elementary and secondary school students: The influence of gender, race, and friendship. *Contemporary Educational Psychology*, 15, 170–182.
- Merkens, H. (2005). *Schulkarrieren von Kindern mit Migrationshintergrund in den ersten drei Jahren der Grundschule. Ergebnisse aus dem Projekt BeLesen: Berliner Längsschnittstudie zur Lesekompetenzentwicklung von Grundschulkindern. (Bericht aus der Arbeit des Arbeitsbereichs Empirische Erziehungswissenschaft Nr. 43)*. Berlin: Freie Universität Berlin.
- Meyer, B. (in press). Team diversity: A review of the literature. In R. Rico (Ed.), *The Wiley Blackwell handbook of the psychology of teamwork and collaborative processes*. Chichester, UK: Wiley-Blackwell.
- Mickelson, R. A., Bottia, M. C., & Lambert, R. (2013). Effects of school racial composition on K-12 mathematics outcomes : A metaregression analysis. *Review of Educational Research*, 83, 121-158. doi: 10.3102/0034654312475322
- Muijs, D., Campbell, J., Kyriakides, L., & Robinson, W. (2005). Making the case for differentiated teacher effectiveness: An overview of research in four key areas. *School Effectiveness and School Improvement*, 16, 51-70. doi: 10.1080/09243450500113985
- Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology*, 25, 3-53. doi: 10.1006/ceps.1999.1019
- Muthén, L. K., & Muthén, B. O. (1998-2010). Mplus (Version 6.1). Los Angeles, CA: Muthén & Muthén.
- Nagengast, B., & Marsh, H. W. (2012, April 9). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, Advance online publication. doi: 10.1037/a0027697
- Neal, L. I., McCray, A. D., Webb-Johnson, G., & Bridgest, S. T. (2003). The effects of African American movement styles on teachers' perceptions and reactions. *The Journal of Special Education*, 37, 49-57.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom : Applying self-determination theory to educational practice. *Theory and Research in Education*, 7, 133-143. doi: 10.1177/1477878509104318
- Nikolova, R. (2011). *Grundschulen als differenzielle Entwicklungsmilieus. Objektive und subjektive Kontextmerkmale der Schülerzusammensetzung und deren Auswirkung auf die Mathematik- und Leseleistungen*. Münster: Waxmann.
- O'Connor, C. (1999). Race, class and gender in America: Narrative opportunity among low-income African American Youth. *Sociology of Education*, 72, 137-157.
- OECD (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: OECD.

- OECD. (2006). *Where immigrant students succeed - A comparative review of performance and engagement in PISA 2003*. Paris: OECD.
- OECD (2007). *Jobs for Immigrants (Vol. 1): Labour Market Integration in Australia, Denmark, Germany and Sweden*. Paris: OECD.
- OECD (2010). *PISA 2009 Results: Overcoming Social Background: Equity in learning opportunities and outcomes (volume II)*.
<http://dx.doi.org/10.1787/9789264091504-en>
- Ogbu, J. U. (1987). Variability in minority school performance: A problem in search of an explanation. *Anthropology and Education Quarterly*, 18, 312-334. doi: 10.1525/aeq.1987.18.4.04x0022v
- Ogbu, J. U. (2004). Collective identity and the burden of “acting White” in Black history, community, and education. *The Urban Review*, 36, 1–35. doi: 10.1023/B:URRE.0000042734.83194.f6
- Ogbu, J. U., & Simons, H. D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology and Education Quarterly*, 29, 155-188.
- Okagaki, L. (2001). Triarchic model of minority children’s school achievement. *Educational Psychologist*, 36, 9–20.
- Okullo, P., & Moe, S. R. (2012). Large herbivores maintain termite-caused differences in herbaceous species diversity patterns. *Ecology*, 93, 2095–2103. doi: 10.1890/11-2011.1
- Opendakker, M.-C., 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, 95–119. doi: 10.1080/09243453.2011.619198
- Opendakker, M.-C., & Minnaert, A. (2011). Relationship between learning environment characteristics and academic engagement. *Psychological Reports*, 109, 259-284. doi: 10.2466/09.10.11.PR0.109.4.259-284
- Opendakker, M.-C., & Van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal*, 27, 407-432. doi: 10.1080/0141192012007143 4
- Opendakker, M.-C., & Van Damme, J. (2007). Do school context, student composition and school leadership affect school practice and outcomes in secondary education? *British Educational Research Journal*, 33, 179-206. doi: 10.1080/01411920701208233
- Opendakker, M.-C., Van Damme, J., De Fraine, B., Van Landeghem, G., & Onghena, P. (2002). The effect of schools and classes on mathematics achievement. *School Effectiveness and School Improvement*, 13, 399-427. doi: 10.1076/sesi.13.4.399.10283
- Orfield, G. (2014). Tenth annual Brown Lecture in education research. A new civil rights agenda for American education. *Educational Researcher*, 43, 273–292. doi: 10.3102/0013189X14547874
- Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: A multiple group, multilevel latent growth curve analysis.

- School Effectiveness and School Improvement*, 19, 21-49. doi: 10.1080/09243450801936845
- Palardy, G. J. (2013). High school socioeconomic segregation and student attainment. *American Educational Research Journal*, 50, 714-754. doi: 10.3102/0002831213481240
- Palardy, G. J. (2015). Classroom-based inequalities and achievement gaps in first grade: The role of classroom context and access to qualified and effective teachers. *Teachers College Record*, 117, 1-48.
- Park, R. E., & Burgess, E. W. (1928). *Introduction to the science of sociology*. Chicago: University Press.
- Peetsma, T., van der Veen, I., Koopman, P., & van Schooten, E. (2006). Class composition influences on pupils' cognitive development. *School Effectiveness and School Improvement*, 17, 275-302. doi: 10.1080/13803610500480114
- Pekrun, R. & Zirngibl, A. (2004). Schülermerkmale im Fach Mathematik. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (S. 191-210). Münster: Waxmann.
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Oxford, England: Viking.
- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109-119. doi: 10.3102/0013189X09332374
- PISA-Konsortium Deutschland (2006). *PISA 2003: Dokumentation der Erhebungsinstrumente* Münster: Waxmann.
- Pöhlmann, C., Haag, N. & Stanat, P. (2013). Zuwanderungsbezogene Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (pS. 297-329). Münster: Waxmann.
- Postmes, T., & Branscombe, N. R. (2002). Influence of long-term racial environmental composition on subjective well-being in African Americans. *Journal of Personality and Social Psychology*, 83, 735-751. doi: 10.1037//0022-3514.83.3.735
- Praetorius, A.-K., Lenske, G., & Helmke, A. (2012). Observer ratings of instructional quality: Do they fulfill what they promise? *Learning and Instruction*, 22, 387- 400. doi: 10.1016/j.learninstruc.2012.03.002
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Newbury Park , CA: Sage.
- Raudenbush, S. W., Fotiu, R. P., & Cheong, Y. F. (1998). Inequality of access to educational resources: A national report card for eighth-grade math. *Educational Evaluation and Policy Analysis*, 20, 253-267. doi: 10.3102/01623737020004253
- Raudenbush, S. W., Rowan, B., & Cheong, Y. F. (1993). Higher order instructional goals in secondary schools: Class, teacher, and school influences. *American Educational Research Journal*, 30, 523-553. doi: 10.3102/00028312030003523

- Raykov, T., & Marcoulides, G. A. (2004). Using the delta method for approximate interval estimation of parameter functions in SEM. *Structural Equation Modeling, 11*, 621-637. doi: 10.1207/s15328007sem1104_7
- Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal, 48*, 335-360. doi: 10.3102/0002831210374874
- Retelsdorf, J., Becker, M., Köller, O., & Möller, J. (2012). Reading development in a tracked school system: A longitudinal study over 3 years using propensity score matching. *British Journal of Educational Psychology, 82*, 647-671. doi: 10.1111/j.2044-8279.2011.02051.x
- Rheinberg, F. (2006). *Motivation*. Stuttgart: Kohlhammer.
- Richter, D., Engelbert, M., Böhme, K., Haag, N., Hannighofer, J., Reimers, H., . . . Stanat, P. (2012). Anlage und Durchführung des Ländervergleichs. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs 2011* (S. 85-102). Münster: Waxmann.
- Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 191-207). Münster: Waxmann.
- Rindermann, H. (2007). Die Bedeutung der mittleren Klassenfähigkeit für das Unterrichtsgeschehen und die Entwicklung individueller Fähigkeiten. *Unterrichtswissenschaft, 35*, 68-89.
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E., & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32*, 63-72. doi: 10.1016/j.learninstruc.2014.01.007
- Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., & Stanat, P. (2015, May 4). Classroom composition and language minority students' motivation in language lessons. *Journal of Educational Psychology*. (Advance online publication)
- Roebbers, C. M., Mecheril, A. & Schneider, W. (1998). Migrantenkinder in deutschen Schulen. Eine Studie zur Persönlichkeitsentwicklung. *Zeitschrift für Pädagogik, 5*, 723-736.
- Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (2000). School as a context of early adolescents' academic and social-emotional development: A summary of research findings. *The Elementary School Journal, 100*, 443-471. doi: 10.1086/499650
- Roick, T., Meier, C., Henschel, S. & Frederking, V. (2014, März). *Entwicklung der literarischen Textverstehenskompetenz zum Ende der Sekundarstufe I – Erste Befunde und Schlussfolgerungen*. Vortrag auf der 2. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), Frankfurt am Main.
- Roscigno, V. J. (2000). Family/school inequality and African-American/Hispanic achievement. *Social Problems, 47*, 266-290. doi: 10.1525/sp.2000.47.2.03x0291e
- Rosenbaum, J. (1991). Black pioneers—do their moves to the suburbs increase economic opportunity for mothers and children? *Housing Policy Debate, 2*, 1179-1213.

- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York, NY: Holt, Rinehart & Winston.
- Rossell, C. (1990). *The carrot or the stick for school desegregation policy: Magnet schools or forced busing*. Philadelphia, PA: Temple University Press.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.
- Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record, 107*, 1999-2045. doi: 10.1111/j.1467-9620.2005.00583.x
- Ryan, A. M. (2000). Peer groups as a context for the socialization of adolescents' motivation, engagement, and achievement in school. *Educational Psychologist, 35*, 101-111.
- Sachverständigenrat deutscher Stiftungen der Integration und Migration (2015). *Unter Einwanderungsländern: Deutschland im internationalen Vergleich. Jahresgutachten 2015*. Berlin: SVR GmbH.
- Sachverständigenrat deutscher Stiftungen der Integration und Migration (2012). *Integration im föderalen System: Bund, Länder und die Rolle der Kommunen. Jahresgutachten 2012 mit Integrationsbarometer*. Berlin: SVR GmbH.
- Salikutluk, Z. (2013). *Immigrants' aspiration paradox. Theoretical explanations and determinants of the aspiration gap between native and immigrant students*. Mannheim: Mannheimer Zentrum für Europäische Sozialforschung.
- Sanders, H. L. (1968). Marine benthic diversity: A comparative study. *The American Naturalist, 102*, 243-282.
- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie*. Münster: Waxmann.
- Schiefele, U. & Köller, O. (2006). Intrinsische und extrinsische Motivation. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 303-310). Weinheim: Beltz PVU.
- Schiefele, U. & Streblow, L. (2005). Intrinsische Motivation – Theorien und Befunde. In R. Vollmeyer & J. Brunstein (Hrsg.), *Motivationspsychologie und ihre Anwendung* (S. 39-58). Stuttgart: Kohlhammer.
- Schofield, J. W. (1995). Improving intergroup relations among students. In J. A. Banks & C. A. M. Banks (Eds.), *Handbook of research on multicultural education* (pp. 635-646). New York, NY: Simon & Schuster Macmillan.
- Schofield, J. W. (2006). *Migrationshintergrund, Minderheitenzugehörigkeit und Bildungserfolg. Forschungsergebnisse der pädagogischen, Entwicklungs- und Sozialpsychologie, AKI-Forschungsbilanz 5*. Berlin: Wissenschaftszentrum Berlin.
- Schümer, G. (2004). Zur doppelten Benachteiligung von Schülern aus unterprivilegierten Gesellschaftsschichten im deutschen Schulwesen. In G. Schümer, K.-J. Tillmann & M. Weiß (Eds.), *Die Institution Schule und die Lebenswelt der Schüler. Vertiefende Analysen der PISA-2000-Daten zum Kontext von Schülerleistungen* (S. 73-114). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Seaton, E. K., & Yip, T. (2009). School and neighborhood contexts, perceptions of racial discrimination, and psychological well-being among African American

- adolescents. *Journal of Youth and Adolescence*, 38, 153–163. doi: 10.1007/s10964-008-9356-x
- Segeritz, M., & Pant, H. A. (2013). Do they feel the same way about math? Testing measurement invariance of the PISA “Students’ Approaches to Learning” instrument across immigrant groups within Germany. *Educational and Psychological Measurement*, 73, 601–630. doi: 10.1177/0013164413481802
- Segeritz, M., Walter, O. & Stanat, P. (2010). Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation? *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 62, 113–138. doi: 10.1007/s11577-010-0094-1
- Seibert, W. (2012). *Geschichte der Zuwanderung nach Deutschland nach 1950*. Bundeszentrale für Politische Bildung. Zugriff am 10.05.2015 unter <http://www.bpb.de/politik/grundfragen/deutsche-verhaeltnisse-eine-sozialkunde/138012/geschichte-der-zuwanderung-nach-deutschland-nach-1950?p=all>
- 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, 454–499. doi: 10.3102/0034654307310317
- Shajek, A., Lüdtke, O. & Stanat, P. (2006). Akademische Selbstkonzepte bei Jugendlichen mit Migrationshintergrund. *Unterrichtswissenschaft*, 36, 125–145.
- Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.
- Shemla, M., Meyer, B., Greer, L., & Jehn, K. A. (2014). A review of perceived diversity in teams: Does how members perceive their team’s composition affect team processes and outcomes? *Journal of Organizational Behavior*, online publication 7 October 2014. doi: 10.1002/job.1957
- Siahaan, F., Lee, D. Y., & Kalist, D. E. (2014). Educational attainment of children of immigrants: Evidence from the National Longitudinal Survey of Youth. *Economics of Education Review*, 38, 1–8. doi: 10.1016/j.econedurev.2013.10.001
- Simpson, E. H. (1949). Measurement of diversity. *Nature*, 163, 688.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. London: Sage
- Stanat, P. (2006a). Disparitäten im schulischen Erfolg: Analysen zur Rolle des Migrationshintergrunds. *Unterrichtswissenschaft*, 43, 98 - 124.
- Stanat, P. (2006b). Schulleistungen von Jugendlichen mit Migrationshintergrund: Die Rolle der Zusammensetzung der Schülerschaft. In J. Baumert, P. Stanat & R. Watermann (Eds.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (pp. 189–219). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Stanat, P., & Christensen, G. (2006). Students’ interest and motivation in mathematics. In OECD (Ed.), *Where immigrant students succeed—A comparative review of performance and engagement in PISA 2003* (pp. 88–91). Paris: OECD Publishing.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.

- Stanat, P., Schwippert, K. & Gröhlich, C. (2010). Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb: Längsschnittliche Überprüfung eines umstrittenen Effekts. *Zeitschrift für Pädagogik*, 56 (55. Beiheft), 147-164.
- Stanat, P., Segeritz, M. & Christensen, G. (2010). Schulbezogene Motivation und Aspiration von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung* (S 31–57). Münster: Waxmann.
- Stanat, P., Weirich, S. & Radmann, S. (2012). Sprach- und Leseförderung. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik: Ergebnisse des IQB-Ländervergleichs* (S. 251–276). Münster: Waxmann.
- Steele, C. M. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69, 797-811.
- Stipek, D. (2004). Teaching practices in kindergarten and first grade: Different strokes for different folks. *Early Childhood Research Quarterly*, 19, 548–568. doi: 10.1016/j.ecresq.2004.10.010
- Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society Interface*, 4, 707-719. doi: 10.1098/rsif.2007.0213
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago, IL: Nelson-Hall.
- Tam, M. Y. S., & Bassett, G. W. (2004). Does diversity matter? Measuring the impact of high school diversity on freshman GPA. *The Policy Studies Journal*, 32, 129-143.
- Tanner, D. (1973). Inequality misconstrued? *Educational Leadership*, 30, 703-705.
- Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal*, 104, 3-28. doi: 0013-5984/2004/10401-0001\$05.00
- Teachman, J. D. (1980). Analysis of population diversity: Measures of qualitative variation. *Sociological Methods & Research*, 8, 341–362.
- Tenenbaum, H. R., & Ruck, M. D. (2007). Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99, 253-273. doi: 10.1037/0022-0663.99.2.253
- Thatcher, S. M. B., & Patel, P. C. (2012). Group faultlines: A review, integration, and guide to future research. *Journal of Management*, 38, 969–1009. doi: 10.1177/0149206311426187
- Thrupp, M. (1995). The school mix effect: The history of an enduring problem in educational research, policy and practice. *British Journal of Sociology of Education*, 16, 183-203. doi: 10.1080/0142569950160204
- Thrupp, M. (1999). *Schools making a difference: Let's be realistic! School mix, school effectiveness and the social limits of reform*. Buckingham: Open University Press.

- Thrupp, M., Lauder, H., & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, 37, 483-504. doi: 10.1016/S0883-0355(03)00016-8
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology*, 98, 788–806. doi: 10.1037/0022-0663.98.4.788
- Turner, J. C., & Meyer, D. K. (2000). Studying and understanding the instructional contexts of classrooms: Using our past to forge our future. *Educational Psychologist*, 35, 69-85. doi: 10.1207/S15326985EP3502_2
- Urdu, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology*, 44, 331-349. doi: 10.1016/j.jsp.2006.04.003
- Van Ackeren, I. (2006). Freie Wahl der Grundschule? Zur Aufhebung fester Schulbezirke und deren Folgen. *Die Deutsche Schule*, 98, 301-310.
- Van der Veen, I. (2002, May). Stemmen scholen en leerkrachten hun aanbod af op de samenstelling van de leerlingpopulatie? [Do schools and teachers adapt teaching activities to the composition of the school population?]. Paper presented at the Onderwijs Research Dagen [Days of Educational Research]. *Antwerp, Belgium*.
- Van der Werf, G., Opdenakker, M.-C., & Kuyper, H. (2008). Testing a dynamic model of student and school effectiveness with a multivariate multilevel latent growth curve approach. *School Effectiveness and School Improvement*, 19, 447-462. doi: 10.1080/09243450802535216
- Van Ewijk, R., & Slegers, P. (2010a). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5, 134-150. doi: 10.1016/j.edurev.2010.02.001
- Van Ewijk, R., & Slegers, P. (2010b). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, 21, 237-265. doi: 10.1080/09243451003612671
- Van Houtte, M. (2004). Tracking effects on school achievement: A quantitative explanation in terms of the academic culture of school staff. *American Journal of Education*, 110, 354-388.
- Van Houtte, M., & Stevens, P. A. J. (2009). School ethnic composition and students' integration outside and inside schools in Belgium. *Sociology of Education*, 82, 217–239.
- Vause, A., Dupriez, V., & Dumay, X. (2010). L'efficacité différenciée des pratiques pédagogiques sur les performances en mathématiques des élèves de sixième primaire en Belgique francophone. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 42, 234-246. doi: 10.1037/a0018038
- Vedder, P., Horenczyk, G., Liebkind, K., & Nickmans, G. (2006). Ethno-culturally diverse education settings: Problems, challenges and solutions. *Educational Research Review*, 1, 157-168. doi: 10.1016/j.edurev.2006.08.007
- Vereinigung der Bayerischen Wirtschaft e.V (2007). *Bildungsgerechtigkeit. Jahresgutachten 2007*. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Vervoort, M. H. M., Scholte, R. H. J., & Scheepers, P. L. H. (2011). Ethnic composition of school classes, majority–minority friendships, and adolescents' intergroup attitudes in the Netherlands. *Journal of Adolescence*, *34*, 257-267.
- Wagner, W., Göllner, R., Helmke, A., Trautwein, U., & Lüdtke, O. (2013). Construct validity of student perceptions of instructional quality is high, but not perfect: Dimensionality and generalizability of domain-independent assessments. *Learning and Instruction*, *28*, 1-11.
- Walsh, J. A., & Taylor, R. B. (2007). Predicting decade-long changes in community motor vehicle theft rates: Impact of structure and surround. *Journal of Research in Crime and Delinquency*, *44*, 64–90. doi: 10.1177/0022427806295552
- Walter, O. (2006). Die Entwicklung der mathematischen und der naturwissenschaftlichen Kompetenz von Jugendlichen mit Migrationshintergrund im Verlauf eines Schuljahres. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003 - Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres* (S. 249-275). Münster: Waxmann.
- Walter, O. (2008a). Herkunftsassoziierte Disparitäten im Lesen, der Mathematik und den Naturwissenschaften: Ein Vergleich zwischen PISA 2000, PISA 2003 und PISA 2006 *Zeitschrift für Erziehungswissenschaft, Sonderheft 10*, 149-168.
- Walter, O. (2008b). Ethno-linguale Kompositionseffekte in neunten Klassen: Befunde aus der Klassenstichprobe von PISA 2006. *Zeitschrift für Erziehungswissenschaft, 10. Jahrg., Sonderheft 10*, 169-184.
- Walter, O., & Stanat, P. (2008). Der Zusammenhang des Migrantenanteils in Schulen mit der Lesekompetenz: Differenzierte Analysen der erweiterten Migrantenstichprobe von PISA 2003. *Zeitschrift für Erziehungswissenschaft, 11*, 84-105. doi: 10.1007/s11618-008-0005-7
- Walter, O. & Taskinen, P. (2008). Naturwissenschaftsbezogene Motivationen und Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund in Deutschland: Der Einfluss der Generation, der Herkunft und des Elternhauses *Zeitschrift für Erziehungswissenschaft–Sonderheft 10*, 185–203.
- Wang, M.-T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, *28*, 12-23. doi: 10.1016/j.learninstruc.2013.04.002
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*, 249-294. doi: 10.3102/00346543063003249
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*, 427– 450. doi: 10.1007/ BF02294627
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*, 89–122.
- Weinert, F. E., Schrader, F.-W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, *13*, 895-914. doi: 10.1016/0883-0355(89)90072-4
- Wilkinson, I. A. G., Hattie, J. A., Parr, J. M., & Townsend, M. A. R. (2000). *Influence of peer effects on learning outcomes: A review of the literature*. Wellington, NZ: Ministry of Education.

- Wilkinson, I. A. G., Parr, J. M., Fung, I. Y. Y., Hattie, J. A. C., & Townsend, M. A. R. (2002). Discussion: Modeling and maximizing peer effects in school. *International Journal of Educational Research*, 37, 521–535. doi: 10.1016/S0883-0355(03)00018-1
- Willenberg, H. (2007). Lesen. In B. Beck & E. Klieme (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band 1*. (S. 122-135). Weinheim: Beltz Pädagogik.
- Wilson, A. B. (1959). Residential segregation of social classes and aspirations of high school boys. *American Sociological Review*, 24, 836-845.
- Winkelmann, H., van den Heuvel-Panhuizen, M., & Robitzsch, A. (2008). Gender differences in the mathematics achievements of German primary school students: Results from a German large-scale study. *ZDM Mathematics Education*, 40, 601–616. doi: 10.1007/s11858-008-0124-x
- Wubbels, T., & Brekelmans, M. (2005). Two decades of research on teacher–student relationships in class. *International Journal of Educational Research*, 43, 6–24. doi: 10.1016/j.ijer.2006.03.003
- Zimmer, R. W., & Toma, E. F. (2000). Peer effects in private and public schools across countries. *Journal of Policy Analysis and Management*, 19, 75 – 92.

10. Abbildungsverzeichnis

Abbildung 2.1:	Modell des Zusammenhangs zwischen Struktur- und Prozessmerkmalen der familiären Lebensverhältnisse und Bildungsbeteiligung bzw. Kompetenzerwerb (entnommen aus Baumert et al, 2003, S. 56)	13
Abbildung 2.2:	Rahmenkonzept für die Analyse von Indikatoren des Lehrens und Lernens (aus: Baumert et al., 2004, S. 146)	30
Abbildung 2.3:	Vermittlungsmodell für Kontexteffekte von Baumert et al. (2006, S. 126)	41
Abbildung 2.4:	Conceptual model of peer influences on learning von Wilkinson et al. (2002, S. 524)	42
Abbildung 2.5:	Conceptual framework for mechanisms through which socioeconomic composition impacts educational attainment von Palardy (2013, S. 724)	43
Abbildung 2.6:	Grunddimensionen der Unterrichtsqualität und deren vermutliche Wirkung (entnommen aus Klieme & Rakoczy, 2008, S. 228)	44
Abbildung 2.7:	Schematische Darstellung von Effekten innerhalb und zwischen Klassen (entnommen und leicht angepasst aus Lüdtke, Robitzsch, Trautwein & Kunter, 2009, S. 129 und Raudenbush & Bryk, 2002, S. 140)	77
Abbildung 3.1:	Arbeitsmodell der vorliegenden Arbeit	87
Abbildung 3.2:	Überblick über Teilstudie 1	88
Abbildung 3.3:	Überblick über Teilstudie 2	90
Abbildung 3.4:	Überblick über Teilstudie 3	92
Figure 4.1:	Schematic depiction of the mediation model using the feature focus on language as an example	110
Figure 5.1:	Hypothesized model of classroom composition effects on students' motivation including Research Questions (RQ) 1 to 3	130
Figure 5.2:	Distribution of language minority students (LMiSs) across the 352 sampled classrooms.	134
Figure 5.3:	Differential effect of the proportion of language minority students (LMiSs) on the level of motivation for LMiSs and students with only German native language (Model 3 in Table 5.3).	146
Figure 6.1:	Joint distribution of the proportion of ethnic minority students in a classroom and Simpson's D for N = 903 classrooms. (Students with only German background, i.e. majority students, were counted as one ethnic category when calculating Simpson's D)	173
Abbildung 8.1:	Beispielhafte schematische Darstellung möglicher Zusammenhänge zwischen Klassenzusammensetzung, Unterrichtsqualität und Bildungsergebnissen der Schülerinnen und Schüler	214
Figure A2.1:	Description of additional classroom level covariates by quartiles of proportion of LMiSs in a classroom	256
Abbildung A4.1:	(Teilstudie eins). Vergleich der mittleren Lesekompetenz zum ersten Messzeitpunkt in Klassen, die sich im Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund unterscheiden	269
Abbildung A4.2:	(Teilstudie eins). Vergleich des mittleren sozioökonomischen Status (HISEI) in Klassen, die sich im Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund unterscheiden	270

11. Anhang

11.1 Anhang A1 zu Teilstudie 1

Table A1.1. Results of Multilevel Structural Equation Models with the Reading Achievement at Measurement Point Two as Outcome Variable and Features of Instructional Quality as Mediating Variables (Student Level Results), Continuation of Table 4.2

	Model 1 (without instruction)	Model 2 (focus on language)	Model 3 (student-oriented climate)	Model 4 (structured classroom management)
	$\beta(SE)$	$\beta(SE)$	$\beta(SE)$	$\beta(SE)$
A) Reading achievement (t_2) on student background variables				
GLL	-.17*** (.03)	-.17*** (.03)	-.17*** (.03)	-.17*** (.03)
SES	.02 (.01)	.02* (.01)	.02 (.01)	.02 (.01)
Reading achievement t_1	.76*** (.01)	.76*** (.01)	.76*** (.01)	.76*** (.01)
Female	.12*** (.03)	.11*** (.03)	.11*** (.02)	.11*** (.02)
B) Instruction on student background variables				
GLL	-	-.01 (.07)	-.03 (.04)	.06 (.05)
SES	-	-.05 (.03)	-.04 (.02)	-.04 (.02)
Reading achievement t_1	-	.12** (.04)	.12*** (.03)	.06 (.03)
Female	-	.32*** (.05)	.31*** (.04)	.29*** (.04)
C) Reading achievement (t_2) on instruction (Level 1)				
Instruction	-	.06* (.03)	.04* (.02)	.07** (.03)
Fit statistics				
$\chi^2 (df)$	0.001 (0)	85.720 (12)	452.986 (83)	224.157 (43)
RMSEA	0	.028	.024	.023
CFI	1	.982	.977	.969

SRMR _{within}	0	.013	.015	.016
SRMR _{between}	0	.024	.030	.050

Note. * $p < .05$ ** $p < .01$ *** $p < .001$; Level 1 = student level, GLL = German language learner, SES = socioeconomic status. Regression coefficients were standardized by the total variance of student level outcome variable (see Marsh, et al., 2012). Model 1: The model is saturated so that $df = 0$ and the fit is perfect.

11.2 Anhang A2 zu Teilstudie 2

Table A2.1 Results of Separate Preliminary Multilevel Analyses with Independent and Control Variables Predicting Motivation at T₂ Controlling for T₁ Motivation

	Model A-1	Model A-2	Model A-3	Model A-4	Model A-5
	β (SE)				
Motivation (T ₂) on composition variables (L2)					
LMiSs %	.06** (.02)	-	-	-	-
LMiSs (%)	-.01 (.01)	-	-	-	-
quadratic term					
SES (Mean)	-	.01 (.03)	-	-	-
Female %	-	-	-.04** (.02)	-	-
German	-	-	-	-.15** (.03)	-
achievement T1					
(Mean)					
Academic track	-	-	-	-	.14* (.06)
Vocational track	-	-	-	-	.13* (.06)
Motivation (T ₂) on student background variables (L1)					
LMiS	-.05 (.05)	-	-	-	-
SES	-	.00 (.02)	-	-	-
Female	-	-	.30** (.04)	-	-
German	-	-	-	.14** (.02)	-
achievement T1					
Motivation T1	.70** (.03)	.70** (.03)	.67** (.03)	.68** (.03)	.70** (.03)
Explained variance					
R ² between	.078	.001	.055	.185	.039
R ² within	.535	.537	.526	.534	.537
Fit statistics					
χ^2 (df)	82.764 (18)	70.716 (16)	138.966 (16)	138.568 (16)	54.358 (13)
RMSEA	.022	.021	.032	.032	.020
CFI	.981	.984	.964	.964	.988
SRMR _{between}	.076	.063	.024	.022	.066
SRMR _{within}	.017	.018	.040	.031	.014
N (L2)	352	352	352	352	352
N (L1)	7615	7620	7615	7615	7615

Note. L2 = classroom level, L1 = student level, LMiS = Language Minority Student, SES = socioeconomic status. Regression coefficients were standardized by the total variance (within + between) of the outcome variable. * $p < .05$, ** $p < .01$

Table A2.2 Results of Multilevel Structural Equation Models Predicting Motivation in German Lessons at T₂ (Student Level Results), Continuation of Table 5.3

	Model 1		Model 2		Model 3		Model 4	
	β	(SE)	β	(SE)	β	(SE)	β	(SE)
Motivation (T ₂) on student background variables								
LMiS	-.03	(.05)	-.03	(.04)	-.31**	(.07)	-.23**	(.07)
Female	.28**	(.04)	.19**	(.03)	.27**	(.04)	.19**	(.03)
German achievement	.09**	(.03)	.05*	(.02)	.10**	(.03)	.06**	(.02)
T ₁								
Motivation T ₁	.66**	(.03)	.51**	(.03)	.66**	(.03)	.51**	(.03)
Instructional climate on student background variables								
LMiS	–		-.02	(.04)	–		-.02	(.04)
Female	–		.19**	(.04)	–		.19**	(.04)
German achievement	–		.08**	(.02)	–		.08**	(.02)
T ₁								
Motivation T ₁	–		.34**	(.03)	–		.35**	(.03)
Motivation (T ₂) on instructional climate								
Instructional climate	–		.35**	(.03)	–		.34**	(.03)
Fit statistics								
–2 log likelihood (FIML)	84595.808		168029.982		84581.606		168014.442	
χ^2 (df)	260.348 (36)		1072.787 (148)		– ^a		– ^a	
RMSEA	.029		.029		–		–	
CFI	.942		.950		–		–	
SRMR _{between}	.032		.039		–		–	
SRMR _{within}	.028		.032		–		–	
N (Level 1)	7615		7650		7615		7650	

Note. Level 1 = student level, Level 2 = classroom level; LMiS = Language Minority Student. Regression coefficients were standardized using the total variance (within + between) of the outcome variable. Note that interaction analyses (Model 3 and Model 4) change the meaning of main effects (i.e., LMiS, LMiSs %) representing an estimate of the regression in case that the other variables equal to zero (Hox, 2010). ^a Fit statistics such as χ^2 are not reported in Mplus for random slope models. * $p < .05$, ** $p < .01$.

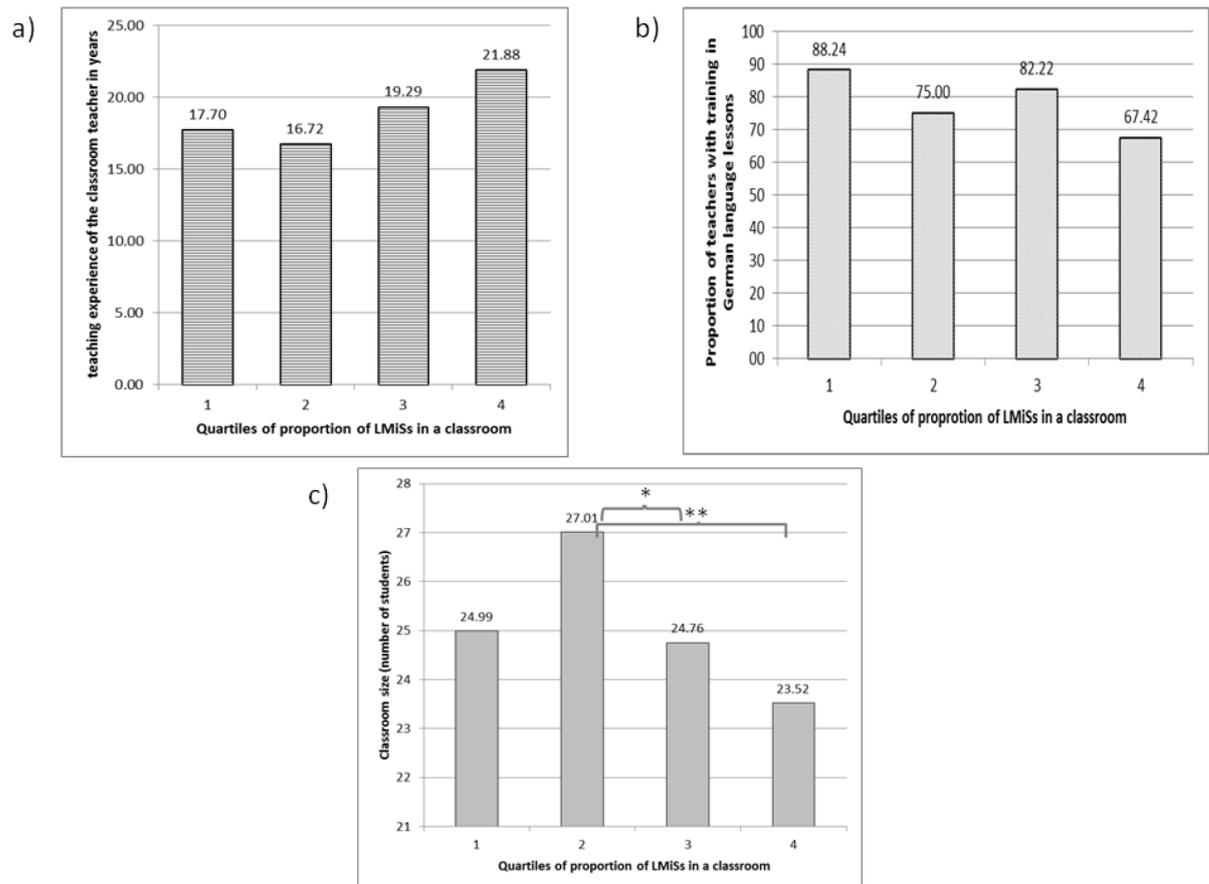


Figure A2.1: Description of additional classroom level covariates by quartiles of proportion of LMISs in a classroom (first quartile: < 9.68%, second quartile: 9.68 – 17.64%, third quartile: 17.65 – 30.77%, fourth quartile: > 30.77%). a) classroom teachers' years of teaching experience: no significant mean difference between the quartiles of proportion of LMISs in a classroom ($F(3, 295) = 2.60, p = .052, \omega^2 = .02$). b) proportion of teachers who completed training at a university in German language lessons: no significant differences between the quartiles of proportion of LMISs in a classroom ($\chi^2(3) = 5.02, p = .17$). c) class size: significant mean difference between second and third quartile and between second and fourth quartile ($F(3, 348) = 7.05, p < .001, \omega^2 = .05$, Bonferroni's test for post hoc comparisons between second and third quartile $p = .021$, between second and fourth quartile $p < .001$).

11.3 Anhang A3 zu Teilstudie 3

Table A3.1. Overview of Measures for the Ethnic Makeup of Classrooms Derived from Different Disciplines

name of measure	formula	characteristics	used for instance in study by
proportion of ethnic minority students	$\text{prop} = p_{\text{minority}}$ or $\text{prop} = (1 - p_{\text{majority}})$	<ul style="list-style-type: none"> – focus on the proportion of students belonging to a particular (majority or minority) group in a given context – most used measure for the ethnic makeup in educational research – also called “simplistic minority-majority approach” by Budescu and Budescu (2012) – insensitivity to the distribution of the ethnic groups that compose the population of ethnic minority students – when there are only two ethnicities ($c=2$) this measure is sufficient to describe diversity (see Budescu & Budescu, 2012) 	Agirdag et al. (2012); Stipek (2004)
variance of a binomial distribution	$\text{var} = p_{\text{majority}} p_{\text{minority}}$	<ul style="list-style-type: none"> – variance implied by the broad distinction between ethnic minority and majority students – maximal diversity is present at $\text{var} = 0.25$ (when $p_{\text{majority}}=0.5$ and $p_{\text{minority}}=0.5$) and minimum diversity at $\text{var} = 0$ – insensitivity to the distribution of the ethnic groups that compose the population of ethnic minority students (see Budescu & Budescu, 2012) 	Walsh & Taylor (2007)
number of categories	$N_{\text{cat}} = c$	<ul style="list-style-type: none"> – Counting of ethnicities present in a classroom (see MacArthur, 1965) – Single concept measure – Addresses variety as a property of diversity (see Stirling, 2007) 	Driessen (2002)

Simpson's D (1949)	$D = 1 - \sum_{i=1}^c p_i^2$	<ul style="list-style-type: none"> – D represents the probability that two students selected at random from a classroom belong to different ethnicities (the greater the value of D the greater the diversity) – probability-based measure – equals the trace of the covariance matrix, i.e. sum of the variances of the c categories (see Budescu & Budescu, 2012) – corresponds for instance to “generalized variance GV” (Budescu & Budescu, 2012), “Blau’s index” (1977), “Herfindahl’s index”(1950), “Herfindahl-Hirschman’s index” (see Biemann & Kearney, 2010) – dual concept measure that addresses the diversity properties variety and disparity (see Stirling, 2007) – its minimum equals $D = 0$ (i.e., all students of a classroom have the same ethnic background), and its maximum equals $D = (c - 1)/c$, that is, it is achieved when the distribution over the c ethnic groups in the classroom is uniform – D is sensitive to the number of categories (c) in a particular setting; in order to compare distributions with different numbers of categories one has to normalize the index (i.e., divide D by $(1 - (1/c))$) (see Budescu & Budescu, 2012) – for higher-order variants of D representing the probability that more than two students pertain to different ethnicities see Biswas and Mandal (2010) – for possibilities to adjust the measure for group size see Biemann and Kearney (2010) 	Benner & Crosnoe (2011); Tam & Bassett (2004)
Les and Maher's Ω (1998)	$\Omega = \sum_{i=1}^c p_i (1 - p_i)$	<ul style="list-style-type: none"> – probability-based measure – dual concept measure – focus primarily on the proportions across categories (see McDonalds & Dimmick, 2003) – identical to the Simpson’s D in value and interpretation and therefore not 	Les & Maher (1998)

included into the present study

Junge's H (1994)	$H = (1 - \sqrt{c})(\sqrt{c-1} - \sqrt{c \sum_{i=1}^c p_i^2 - 1})$	<ul style="list-style-type: none"> - measure of concentration (higher values represent lower diversity) - probability-based measure - dual concept measure - focus primarily on the proportions across categories (see McDonalds & Dimmick, 2003) - particularly sensitive to changes in c (see McDonalds & Dimmick, 2003) - addresses the diversity properties variety, balance, and disparity (see Stirling, 2007) 	Stirling (2007)
Kvålseth's OD (1991)	$OD = \left(1 - \sum_{i=1}^c p_i^2\right)^{-1}$	<ul style="list-style-type: none"> - probability-based measure - dual concept measure - focus primarily on the proportions across categories (see McDonalds & Dimmick, 2003) - close in agreement to Simpson's D (see McDonald & Dimmick, 2003) 	Okullo & Moe (2012)
Fager's S (1972)	$S = \left(\frac{c \sum_{i=1}^c p_i^2 - N^2}{c(c-1)} \right)^{\frac{1}{2}}$	<ul style="list-style-type: none"> - measure of concentration (higher values represent lower diversity) - probability-based measure - dual concept measure - focus primarily on the proportions across categories (see McDonalds & Dimmick, 2003) - close in agreement to Simpson's D (see McDonald & Dimmick, 2003) 	McDonalds & Dimmick (2003)

Shannon's H (Shannon & Weaver, 1949)	$H = -\sum_{i=1}^c p_i \ln(p_i)$	<ul style="list-style-type: none"> – logarithm-based measure; weighted sum of the probabilities where the weights are their logarithms (typically base 2 or natural logarithm) (see Budescu & Budescu, 2012) – corresponds for instance to “entropy” (Budescu & Budescu, 2012), “Teachman’s (1980) index”, and “Shannon Wiener index” (Fager, 1972) – dual concept measure that addresses the diversity properties variety and disparity (see Stirling, 2007) – its minimum equals $H = 0$ (i.e., all students of a classroom have the same ethnic background), and its maximum equals $H = \ln(c)$, that is, it is achieved when the distribution over the c ethnic groups in the classroom is uniform – H is sensitive to the number of categories (c) in a particular setting; in order to compare distributions with different numbers of categories one has to normalize the index (see Budescu & Budescu, 2012) – for possibilities to adjust the measure for group size see Biemann and Kearney (2010) 	Dougherty & Huyser (2008)
Gleason's D (Sanders, 1968)	$D = \frac{(c-1)}{\ln(N)}$	<ul style="list-style-type: none"> – logarithm-based measure – number of categories (i.e., ethnicities) minus 1 divided by the logarithm of the number of elements (i.e., students in a classroom) – dual concept measure 	Fager (1972)

Note. c = number of categories (i.e., ethnicities), p_i = relative frequency of category i , N = number of elements (i.e., number of students in a classroom), \ln = natural logarithm

Table A3.2. Preliminary Analyses: Correlations between Measures of Ethnic Composition Resp. Ethnic Diversity in the Classroom (N = 903 Classrooms)

	1)	2)	3)	4)	5)	6)	7)	8)	9)
1) % minority students	1								
2) N categories	.62**	1							
3) Variance	.34**	.53**	1						
4) Simpson's D	.92**	.76**	.66**	1					
5) Junge's H^a	.94**	.76**	.60**	1.00**	1				
6) Kvålseth's OD	.92**	.70**	.39**	.92**	.95**	1			
7) Fager's S^a	.94**	.75**	.57**	.99**	1.00**	.97**	1		
8) Shannon's H	.89**	.86**	.61**	.98**	.98**	.93**	.98**	1	
9) Gleason's D	.65**	.98**	.52**	.78**	.78**	.73**	.77**	.87**	1

^a Measures of concentration have been multiplied by -1 to represent diversity. German students (i.e. majority students) were counted as one ethnic category when calculating the measures of diversity.

Table A3.3. Correlations between Optimized Measures of Ethnic Composition Resp. Ethnic Diversity among Minority Students in the Classroom (N = 903 Classrooms)

	1)	2)	3)	4)	5)	6)	7)	8)	9)	$r_{\text{mathematics achievement}}$	$r_{\text{reading achievement}}$	$r_{\text{belonging}}$	r_{SES}	$r_{\text{prior achievement}}$
1) % minority students	1									-.41**	-.32**	-.10*	-.04	.05
2) N categories	.63**	1								-.07	-.26**	-.07	-.06	.01
3) Variance	.34**	.52*	1							-.10	-.15*	-.06	.25**	.02
4) Simpson's D	.49**	.85*	.51*	1						-.04	-.22**	-.02	.01	.03
5) Junge's H^a	.49**	.88*	.49*	1.00	1					-.03	-.23**	-.03	-.02	.03
6) Kvålseth's OD	.44**	.88*	.40*	.88*	.93*	1				.01	-.22**	-.04	-.09*	.03
7) Fager's S^a	.48**	.89*	.47*	.98*	1.00	.96*	1			-.02	-.23**	-.03	-.04	.03
8) Shannon's H	.55**	.94*	.52*	.97*	.99*	.94*	.99*	1		-.04	-.24**	-.04	-.03	.03
9) Gleason's D	.22**	.83*	.36*	.86*	.87*	.84*	.88*	.89*	1	.11	-.13	-.05	-.07	-.02

^a Measures of concentration have been multiplied by -1 to represent diversity. German students (i.e. majority students) not counted as a category when calculating the measures of diversity. The measures thus represent diversity within the proportion of ethnic minority students.

Table A3.4 Results of Multilevel Structural Equation Models Predicting Mathematics Achievement (Student Level Results, Continuation of Table 6.3)

	Model M.1	Model M.1b	Model M.2 (Ncat)	Model M.3 (SimD)	Model M.4 (ShaH)	Model M.5 (Ncat)	Model M.6 (SimD)	Model M.7 (ShaH)	Model M.8 (Ncat)	Model M.9 (SimD)	Model M.10 (ShaH)
Turkish background	-.29**	-.29**	-.33**	-.34**	-.33**	-.29**	-.29**	-.29**	-.29**	-.29**	-.29**
Former USSR background	-.08**	-.08**	-.11**	-.11**	-.11**	-.08**	-.08**	-.08**	-.08**	-.08**	-.08**
Polish background	-.06	-.06	-.07	-.08*	-.08	-.06	-.06	-.06	-.07	-.06	-.07
Former Yugoslavia background	-.19**	-.19**	-.21**	-.22**	-.22**	-.19**	-.19**	-.19**	-.19**	-.19**	-.19**
Other background	-.13**	-.13**	-.16**	-.16**	-.16**	-.13**	-.13**	-.13**	-.14**	-.14**	-.14**
unidentifiable	-.32**	-.32**	-.35**	-.35**	-.35**	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**
SES	.18**	.18**	.18**	.18**	.18**	.18**	.18**	.18**	.17**	.17**	.17**
female	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**	-.32**
prior achievement	.48**	.48**	.48**	.48**	.48**	.48**	.48**	.48**	.48**	.48**	.48**

Note. * $p < .05$, ** $p < .01$, Ncat = number of categories (i. e. ethnicities), SimD = Simpson's D , ShaH = Shannon's H , SES = socioeconomic status. Regression coefficients were standardized by the total variance (within + between) of the outcome variable.

Table A3.5. Results of Multilevel Structural Equation Models Predicting German Reading Achievement (Student Level Results, Continuation of Table 6.4)

	Model R.1	Model R.1b	Model R.2 (Ncat)	Model R.3 (SimD)	Model R.4 (ShaH)	Model R.5 (Ncat)	Model R.6 (SimD)	Model R.7 (ShaH)	Model R.8 (Ncat)	Model R.9 (SimD)	Model R.10 (ShaH)
Turkish background	-.34**	-.34**	-.38**	-.39**	-.38**	-.34**	-.34**	-.34**	-.34**	-.34**	-.34**
Former USSR background	-.14**	-.14**	-.17**	-.17**	-.17**	-.14**	-.14**	-.14**	-.14**	-.14**	-.14**
Polish background	-.10*	-.10*	-.11*	-.12*	-.11*	-.10*	-.10*	-.10*	-.10*	-.10*	-.10*
Former Yugoslavia background	-.21**	-.21**	-.24**	-.25**	-.24**	-.21**	-.21**	-.21**	-.22**	-.22**	-.22**
Other background	-.12**	-.12**	-.15**	-.15**	-.15**	-.12**	-.12**	-.12**	-.13**	-.13**	-.13**
unidentifiable	-.35**	-.35**	-.38**	-.38**	-.38**	-.35**	-.35**	-.35**	-.35**	-.35**	-.35**
SES	.21**	.21**	.21**	.21**	.21**	.21**	.21**	.21**	.20**	.20**	.20**
female	.15**	.15**	.15**	.15**	.15**	.15**	.15**	.15**	.15**	.15**	.15**
prior achievement	.40**	.40**	.40**	.40**	.40**	.40**	.40**	.40**	.40**	.40**	.40**

Note. * $p < .05$, ** $p < .01$, Ncat = number of categories (i. e. ethnicities), SimD = Simpson's D , ShaH = Shannon's H , SES = socioeconomic status. Regression coefficients were standardized by the total variance (within + between) of the outcome variable.

Table A3.6. Results of Multilevel Structural Equation Models Predicting Feeling of Belonging with one's Peers (Student Level Results, Continuation of Table 6.5)

	Model B.1	Model B.1b	Model B.2 (Ncat)	Model B.3 (SimD)	Model B.4 (ShaH)	Model B.5 (Ncat)	Model B.6 (SimD)	Model B.7 (ShaH)	Model B.8 (Ncat)	Model B.9 (SimD)	Model B.10 (ShaH)
Turkish background	.01	.01	-.04	-.05	-.04	.01	.01	.01	.01	.01	.01
Former USSR background	-.22**	-.22**	-.25**	-.26**	-.26**	-.22**	-.22**	-.22**	-.22**	-.22**	-.22**
Polish background	-.06	-.07	-.08	-.09	-.09	-.07	-.07	-.07	-.07	-.07	-.07
Former Yugoslavia background	.00	.01	-.02	-.04	-.03	.00	.00	.00	.00	.00	.00
Other background	-.02	-.02	-.05	-.05	-.05	-.02	-.02	-.02	-.03	-.03	-.03
unidentifiable	-.13**	-.13**	-.16**	-.17**	-.17**	-.13**	-.13**	-.13**	-.13**	-.13**	-.13**
SES	.09**	.09**	.09**	.09**	.09**	.09**	.09**	.09**	.09**	.09**	.09**
female	.21**	.21**	.21**	.21**	.21**	.21**	.21**	.21**	.21**	.21**	.21**
prior achievement	.05**	.05**	.06**	.06**	.06**	.05**	.05**	.05**	.04**	.04**	.04**

Note. * $p < .05$, ** $p < .01$, Ncat = number of categories (i. e. ethnicities), SimD = Simpson's D , ShaH = Shannon's H , SES = socioeconomic status. Regression coefficients were standardized by the total variance (within + between) of the outcome variable.

Table A3.7. Results of Multilevel Structural Equation Models Predicting Mathematics Achievement (Classroom Level Results) Using Additional Diversity Measures

	Model M.11 (var)	Model M.12 (JungH) ^a	Model M.13 (KvalOD)	Model M.14 (FagS) ^a	Model M.15 (GleaD)	Model M.17 (JungH) ^a	Model M.18 (KvalOD)	Model M.19 (FagS) ^a	Model M.20 (GleaD)
minority students %	-	-	-	-	-	-.14**	-.13**	-.14**	-.13**
minority students %: quadratic term	-	-	-	-	-	-.00	-.01	-.00	-.01
diversity measure	-.02	-.05**	-.05**	-.05**	-.00	.03*	.02	.02*	.03*
prior achievement (mean)	-	-	-	-	-	.07**	.07**	.07**	.07**
SES (mean)	-	-	-	-	-	.06**	.06**	.06**	.06**
R ² L2	0.01	0.02	0.02	0.02	0.00	0.39	0.39	0.39	0.39
R ² L1	0.35	0.35	0.35	0.35	0.35	0.34	0.34	0.34	0.34

Note. ^aMeasures of concentration have been multiplied by -1 to represent diversity. *p < .05, **p < .01, var = variance, JungH = Junge's *H*, KvalOD = Kvålseth's *OD*, FagS = Fager's *S*, GleaD = Gleason's *D*, SES = socioeconomic status, L1 = student level, L2 = classroom level; line "diversity measure" shows coefficients of the respective measure indicated in the first line. Covariates at the student level: ethnic background, SES, proxy for prior achievement, and gender. Regression coefficients were standardized by the total variance (within + between) of the outcome variable. Model M.16 including the proportion of ethnic minority students and the variance of this binomial distribution simultaneously did not converge, most probably due to multicollinearity, and was excluded from the table.

Table A3.8. Results of Multilevel Structural Equation Models Predicting German Reading Achievement (Classroom Level Results) Using Additional Diversity Measures

	Model R.11 (var)	Model R.12 (JungH) ^a	Model R.13 (KvalOD)	Model R.14 (FagS) ^a	Model R.15 (GleaD)	Model R.17 (JungH) ^a	Model R.18 (KvalOD)	Model R.19 (FagS) ^a	Model R.20 (GleaD)
minority students %	-	-	-	-	-	-.10**	-.10**	-.10**	-.10**
minority students %: quadratic term	-	-	-	-	-	-.01	-.01	-.01	-.01
diversity measure	-.02	-.06**	-.06**	-.06**	-.03*	.00	-.00	-.00	.00
prior achievement (mean)	-	-	-	-	-	.05**	.05**	.05**	.05**
SES (mean)	-	-	-	-	-	.07**	.07**	.07**	.07**
R ² L2	0.00	0.04	0.04	0.04	0.01	0.30	0.30	0.30	0.30
R ² L1	0.26	0.26	0.26	0.26	0.26	0.25	0.25	0.25	0.25

Note. ^aMeasures of concentration have been multiplied by -1 to represent diversity. *p < .05, **p < .01, var = variance, JungH = Junge’s *H*, KvalOD = Kvålseth’s *OD*, FagS = Fager’s *S*, GleaD = Gleason’s *D*, SES = socioeconomic status, L1 = student level, L2 = classroom level; line “diversity measure” shows coefficients of the respective measure indicated in the first line. Covariates at the student level: ethnic background, SES, proxy for prior achievement, and gender. Regression coefficients were standardized by the total variance (within +between) of the outcome variable. Model R.16 including the proportion of ethnic minority students and the variance of this binomial distribution simultaneously did not converge, most probably due to multicollinearity, and was excluded from the table.

Table A3.9. Results of Multilevel Structural Equation Models Predicting Feeling of Belonging with one’s Peers (Classroom Level Results) Using Additional Diversity Measures

	Model B.11 (var)	Model B.12 (JungH) ^a	Model B.13 (KvalOD)	Model B.14 (FagS) ^a	Model B.15 (GleaD)	Model B.17 (JungH) ^a	Model B.18 (KvalOD)	Model B.19 (FagS) ^a	Model B.20 (GleaD)
minority students %	-	-	-	-	-	-.07**	-.07**	-.07**	-.06**
minority students %: quadratic term	-	-	-	-	-	.01	.00	.00	-.00
diversity measure	-.02	-.02	-.02	-.02	-.01	.02	.01	.02	.00
prior achievement (mean)	-	-	-	-	-	.08**	.08**	.08**	.08**
SES (mean)	-	-	-	-	-	-.02	-.02	-.02	-.02
R ² L2	0.01	0.01	0.01	0.01	0.00	0.19	0.19	0.19	0.18
R ² L1	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03

Note. ^aMeasures of concentration have been multiplied by -1 to represent diversity. *p < .05, **p < .01, var = variance, JungH = Junge’s *H*, KvalOD = Kvålseth’s *OD*, FagS = Fager’s *S*, GleaD = Gleason’s *D*, SES = socioeconomic status, L1 = student level, L2 = classroom level; line “diversity measure” shows coefficients of the respective measure indicated in the first line. Covariates at the student level: ethnic background, SES, proxy for prior achievement, and gender. Regression coefficients were standardized by the total variance (within +between) of the outcome variable. Model B.16 including the proportion of ethnic minority students and the variance of this binomial distribution simultaneously did not converge, most probably due to multicollinearity, and was excluded from the table.

11.4 Anhang A4: ergänzender Anhang zur vorliegenden Arbeit

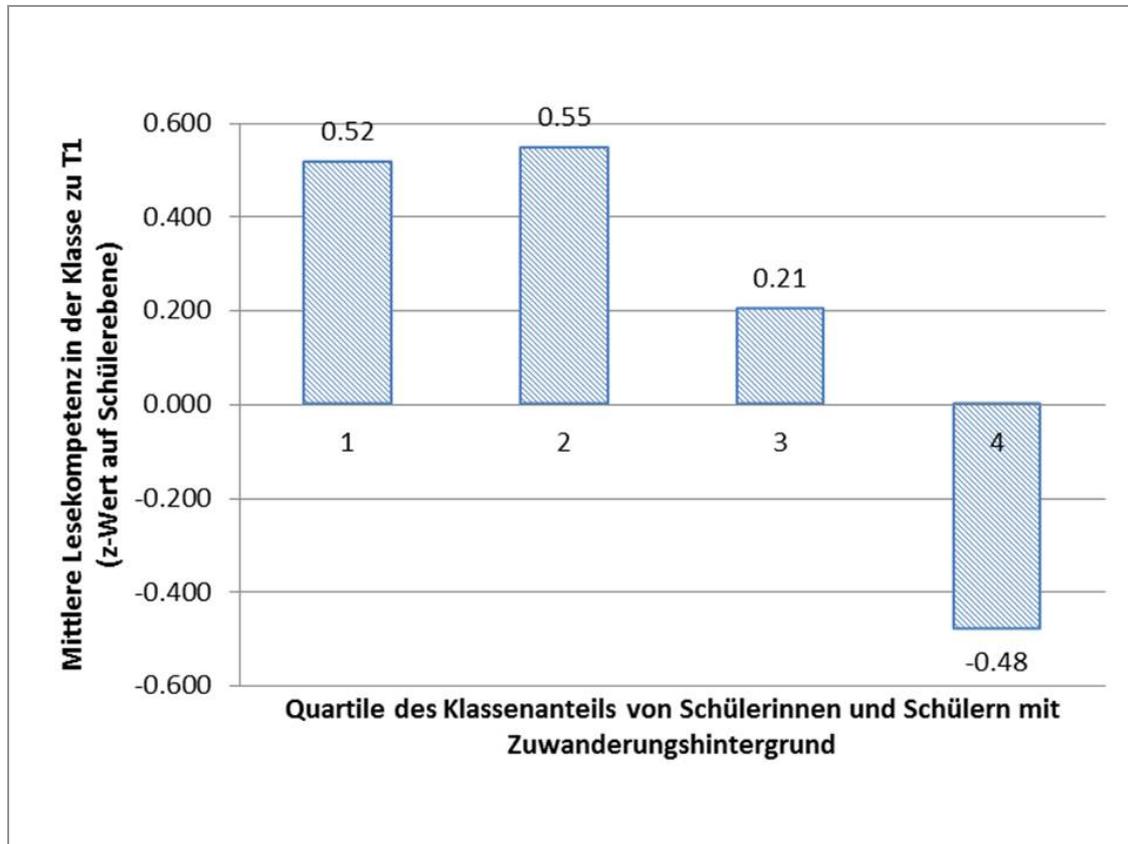


Abbildung A4.1: (Teilstudie eins) Vergleich der mittleren Lesekompetenz zum ersten Messzeitpunkt in Klassen, die sich im Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund unterscheiden (erstes Quartil < 9,68%, zweites Quartil: 9,68 – 17,64%, drittes Quartil: 17,65 – 30,77%, viertes Quartil: > 30,77%). Varianzanalysen mit Post-Hoc-Test nach Bonferroni verweisen auf statistisch bedeutsame Mittelwertsunterschiede zwischen allen Quartilen mit Ausnahme des ersten und zweiten Quartils ($F(3, 348) = 41,37, p < .001$).

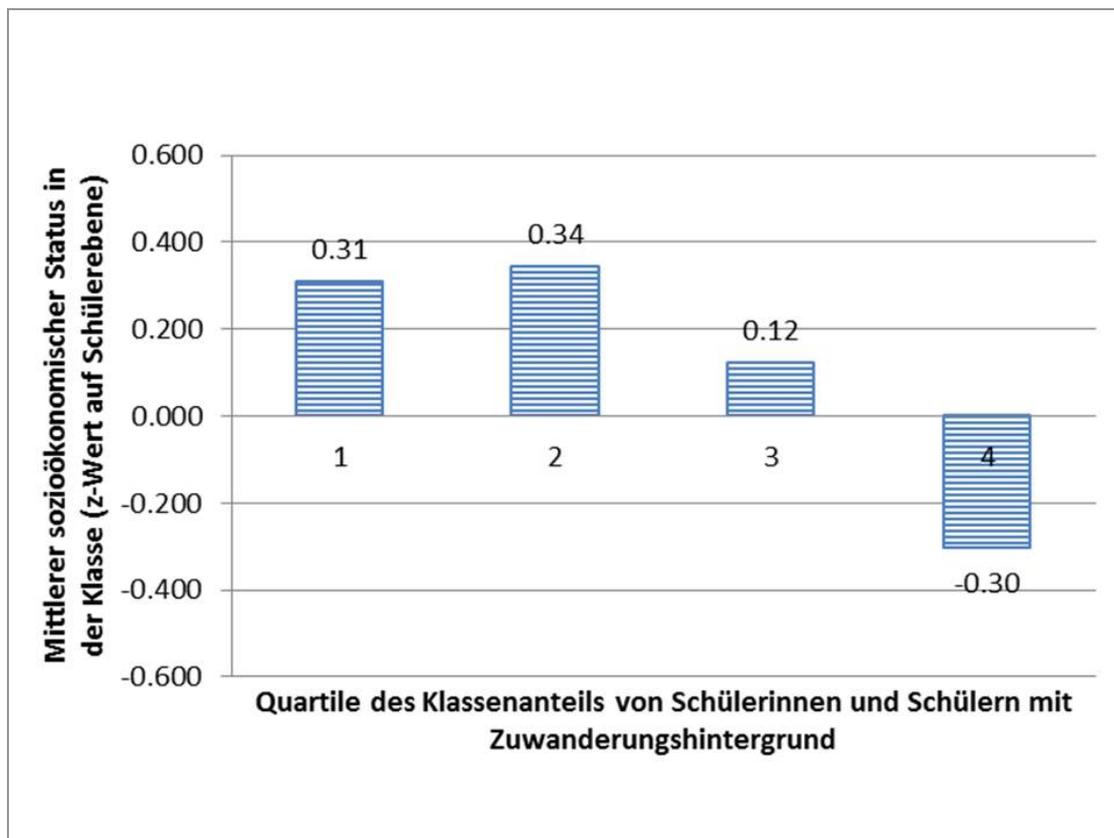


Abbildung A4.2: (Teilstudie eins). Vergleich des mittleren sozioökonomischen Status (HISEI) in Klassen, die sich im Anteil von Schülerinnen und Schülern mit Zuwanderungshintergrund unterscheiden (erstes Quartil < 9,68%, zweites Quartil: 9,68 – 17,64%, drittes Quartil: 17,65 – 30,77%, viertes Quartil: > 30,77%). Varianzanalysen mit Post-Hoc-Test nach Tamhane verweisen auf statistisch bedeutsame Mittelwertsunterschiede zwischen allen Quartilen mit Ausnahme des ersten und zweiten Quartils ($F(3, 348) = 39,11, p < .001$).