

**Die Zusammensetzung der Lerngruppe
und ihre Effekte auf psychosoziale Merkmale und Leistung
von Schülerinnen und Schülern**

Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

am Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

vorgelegt von

Franziska Stäbler (M.A.)

Berlin, 2017

Betreuerin

Dr. Hanna Dumont, Deutsches Institut für Internationale Pädagogische Forschung, Berlin

Erstgutachter

Professor Dr. Rainer Watermann, Freie Universität Berlin, Berlin

Zweitgutachter

Professor Dr. Kai Maaz, Deutsches Institut für Internationale Pädagogische Forschung, Berlin

Tag der Disputation: 23.10.2017

DANKSAGUNG

Die vorliegende Dissertation ist im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am Deutschen Institut für Internationale Pädagogische Forschung (DIPF) in Berlin entstanden. Ich möchte hiermit allen Personen herzlich danken, die auf die eine oder andere Weise diesen Weg mit mir gegangen sind und mich unterstützt haben.

Ich danke an erster Stelle Prof. Dr. Kai Maaz, der mir die Rahmenbedingungen für die Erstellung der Dissertation geboten hat sowie die Chance, Teil eines so besonderen und inspirierenden Forschungsumfeldes zu sein. Mein tiefer Dank gilt zudem meiner Betreuerin Dr. Hanna Dumont. Ich danke ihr für die großartige inhaltliche und methodische Betreuung und für die inhaltlichen Freiheiten, die ich bei der Erstellung dieser Arbeit hatte. Unsere vielen Gespräche und ihr stets offenes Ohr waren mir eine unermessliche fachliche sowie persönliche Unterstützung im Auf und Ab der Promotionszeit. Dr. Michael Becker, Dr. Marko Neumann und Dr. Malte Jansen danke ich für die wertvollen methodischen Hinweise und Anregungen sowie für den fachlichen Austausch, der auch über Dissertationsthemen hinausging.

Ebenso bedanke ich mich bei meinen Kolleginnen und Kollegen am DIPF für die gemeinsame Zeit, die tolle Arbeitsatmosphäre sowie die unterhaltsamen Mittagspausen. Insbesondere bei meinen Bürokolleginnen Dr. Susanne Böse, Dr. Denise Klinge und Josefine Lühe bedanke ich mich für den persönlichen Austausch, der den Arbeitsalltag in einer sehr herzlichen Weise bereichert hat. Den studentischen Hilfskräften Freya Müller und Edgar Gerber danke ich sehr für ihre Unterstützung durch Formatierungsaufgaben. Ebenso danke ich meinen Freunden Stefan und Kristin herzlich für die Korrekturen.

Ein großer Dank geht auch an Prof. Dr. Jacquelyne Eccles und Dr. Anna-Lena Dicke für die großartigen Monate in Irvine. Durch sie erhielt ich die Möglichkeit, einen Live-Einblick in die internationale Forschung zu bekommen und mich mit ihnen und ihrem Team auszutauschen. Es war eine unvergessliche Zeit; nicht nur auf fachlicher Ebene. Ich bedanke mich hiermit auch für die Möglichkeit zur Nutzung der Daten für eine Teilstudie der Dissertation.

Mein Dank gilt natürlich auch Prof. Dr. Jürgen Baumert und Prof. Dr. Eckhard Klieme für die Nutzung der Daten für die Analysen der anderen Teilstudien sowie für ihre hilfreichen Hinweise und Anregungen in ihrer Rolle als Co-Autoren bei der Einreichung zweier Dissertationsbeiträge.

Ich würde nicht an dieser Stelle stehen, wenn ich nicht meine Familie zur Seite gehabt hätte. Daher gilt mein besonderer Dank meinen Eltern für die Unterstützung, Aufmunterung und ihr bedingungsloser Glaube an mich. Auch meinem Bruder danke ich zutiefst für die Stärke, die er mir auf meinen Lebensweg mitgegeben hat. Schließlich möchte ich mich bei all meinen Freunden, vor allem bei Kristin, Sandra, Stefan, Luna, Caro und Vicky für die schönen Stunden der Abwechslung von der Dissertation bedanken. Besonders danke ich Kristian, der mir in der aufreibenden Abschlussphase der Promotion ein angenehmes und entspannendes Drumherum bereitet hat.

INHALT

Zusammenfassung	9
Abstract	11
1 Einleitung und theoretischer Rahmen der Arbeit	14
1.1 <i>Schule als Lern- und Entwicklungsort</i>	17
1.1.1 Die gesellschaftlichen Ziele von Schule	17
1.1.2 Multiple Ziele von Schule für Schülerinnen und Schüler	18
1.2 <i>Das Spannungsfeld Person und Umwelt</i>	20
1.2.1 Das Ökosystem Schule – Die Schülerschaft als Mitgestalter der Lernumwelt	21
1.2.2 Die Zusammensetzung der Schülerschaft als Konsequenz der Schulstruktur und gesellschaftlicher Veränderungen	24
1.2.2.1 Die Rolle der Schulstruktur für die Zusammensetzung der Schülerschaft	24
1.2.2.2 Die Rolle aktueller Umstrukturierungen des Schulsystems und gesellschaftlicher Veränderungen für die Zusammensetzung der Schülerschaft	26
1.2.3 Der Einfluss der Schülerzusammensetzung auf die schulische Entwicklung Einzelner .	27
1.3 <i>Peer-Effekte – Die Schülerschaft als aktive Mitgestalter der Lern- und Entwicklungsumwelt</i>	30
1.3.1 Das sozialpsychologische Konzept der normativen und komparativen Einflüsse in Bezugsgruppen	31
1.3.2 Das entwicklungspsychologische Konzept der Peergruppeneffekte	34
1.3.3 Das Konzept des Schul- und Klassenklimas in der Schul- und Unterrichtsforschung	37
1.4 <i>Empirische Befunde zum Einfluss der Schülerzusammensetzung auf die individuelle schulische Entwicklung</i>	38
1.4.1 Empirische Modellierung von Effekten der Schülerzusammensetzung auf die individuelle schulische Entwicklung	39
1.4.2 Effekte der Schülerzusammensetzung auf die individuelle Leistungsentwicklung	42
1.4.2.1 Effekte durch die Komposition der Leistung	42
1.4.2.2 Effekte durch die Komposition des familiären Hintergrunds	43
1.4.2.3 Effekte durch die Komposition psychosozialer Merkmale	44
1.4.3 Effekte der Schülerzusammensetzung auf die individuelle Entwicklung psychosozialer Merkmale	44
1.4.3.1 Effekte durch die Komposition der Leistung	44

1.4.3.2	Effekte durch die Komposition des familiären Hintergrunds.....	46
1.4.3.3	Effekte durch die Komposition psychosozialer Merkmale.....	47
1.4.4	Mediation von Kompositionseffekten	49
1.5	<i>Ableitung der Fragestellung</i>	51
	Literaturverzeichnis I	58
2	Teilstudie 1: Der Effekt der Schulzusammensetzung auf die Entwicklung der Schulfriedenheit und des normverletzenden Verhaltens von Schülerinnen und Schülern.....	78
3	Teilstudie 2: Are jitters contagious? The effect of class-average test anxiety on individual students' test anxiety	108
4	Teilstudie 3: What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept.....	150
5	Teilstudie 4: Ein Blick in die Klassenzimmer Deutschlands – wie heterogen sind Klassen zusammengesetzt?	192
6	Gesamtdiskussion.....	220
6.1	<i>Zusammenfassung und Einordnung der Befunde.....</i>	220
6.1.1	Zentrale Befunde zur Rolle psychosozialer Schülermerkmale als Kriteriumsvariablen von Kompositionseffekten	220
6.1.2	Zentrale Befunde zur Rolle psychosozialer Schülermerkmale als Kompositionsmerkmale	222
6.1.3	Zentrale Befunde zum Zusammenspiel verschiedener Kompositionseffekte	226
6.2	<i>Methodische Bewertung und Grenzen der Dissertation</i>	227
6.3	<i>Implikationen für die pädagogische Praxis und zukünftige Forschung.....</i>	232
6.3.1	Implikationen für die pädagogische Praxis.....	232
6.3.2	Implikationen für die Forschung zu Kompositionseffekten und Ausblick auf zukünftige Forschungsfragen	234
6.4	<i>Schlussfolgerung</i>	238
	Literaturverzeichnis II.....	240
	Abbildungsverzeichnis.....	247
	Tabellenverzeichnis.....	249

ZUSAMMENFASSUNG

Im Rahmen der Schule vollzieht sich die Entwicklung von Kindern und Jugendlichen in der Gemeinschaft Gleichaltriger in einem gemeinsamen Kontext. Dieser konstituiert sich nicht nur durch bildungspolitisch vorgegebene Strukturen des Schulsystems oder durch die Form des Unterrichts, sondern auch durch das Leistungs- und Fähigkeitsniveau sowie psychosoziale Merkmale der Lernenden, die in der Klasse oder Schule zusammenkommen (Eccles & Roeser, 2011). Gemäß dem *bioökologischen Modell* von Bronfenbrenner und Morris (2006) können Lernende somit als Mitgestalter des Schulkontextes gesehen werden. Studien zu sogenannten *Kompositionseffekten* belegen bereits, dass die Schülerschaft in ihrer mittleren leistungs- und herkunftsbezogenen Zusammensetzung auf die individuelle Leistungsentwicklung wirkt (siehe im Überblick Dumont, Neumann, Maaz & Trautwein, 2013). Um die Rolle der Lernenden als Mitgestalter des Schulkontextes noch differenzierter zu ergründen, wurden (1) psychosoziale Schülermerkmale als Merkmale betrachtet, deren Entwicklung im Schulkontext von der Komposition der Lerngruppe bedingt werden kann, (2) psychosoziale Schülermerkmale in ihrer mittleren Ausprägung und in ihrer Verteilung innerhalb der Lerngruppe als Kompositionsmerkmal betrachtet sowie (3) das Zusammenspiel verschiedener Kompositionseffekte. Dieses Vorhaben wurde in vier Teilstudien bearbeitet, für die Datensätze aus verschiedenen Längsschnittstudien genutzt wurden.

In der *ersten Teilstudie* wurde der Frage nachgegangen, ob die individuelle Entwicklung von Schulzufriedenheit und normverletzendem Verhalten in Zusammenhang mit der Schülerkomposition an der Schule steht. Mehrebenenanalysen an einer Stichprobe der BERLIN-Studie ergaben, dass Lernende der 9. Jahrgangsstufe nach einem Schuljahr zufriedener mit der Schule waren, je leistungsstärker die von ihnen besuchte Schule war und je zufriedener die Schülerschaft im Durchschnitt mit der Schule war. Es zeigten sich zudem Hinweise für weitere Effekte auf die Schulzufriedenheit und das normverletzende Verhalten der Lernenden durch die leistungs- und herkunftsbedingte Komposition sowie durch die Komposition der Schülerschaft hinsichtlich der Schulzufriedenheit und des normverletzenden Verhaltens selbst.

In der *zweiten Teilstudie* richtete sich der Fokus auf die psychosoziale Beeinflussung zwischen Jugendlichen im schulischen Kontext. Dazu wurde der Effekt der mittleren Testangst der Klasse auf die Entwicklung der individuellen Testangst anhand einer Stichprobe der MSALT-Studie untersucht. Mehrebenenanalysen zeigten, dass die mittlere Testangst neu formierter

Klassen zu Beginn der 7. Jahrgangsstufe positiv mit der Entwicklung der Testangst der Lernenden vom Ende der sechsten bis zum Beginn der 7. Jahrgangsstufe assoziiert war. Interaktionsanalysen ergaben zudem differentielle Unterschiede zwischen Mädchen und Jungen sowie zwischen Lernenden mit einer geringen und einer hohen vorherigen Testangst.

Im Zentrum der *dritten Teilstudie* stand das Zusammenspiel von Kompositionseffekten auf die Entwicklung der Leistung und auf psychosoziale Merkmale. Dafür wurden anhand der BIJU-Studie die zwei vielfach replizierten Kompositionseffekte auf die individuelle Leistung und auf das individuelle akademische Selbstkonzept, die beide von der mittleren Leistung einer Lerngruppe ausgehen, simultan modelliert. Mehrebenenanalysen inklusive *Cross-Lagged-Pfaden* zwischen der individuellen Leistung und dem akademischen Selbstkonzept ergaben, dass die mittlere Klassenleistung zu Beginn der 7. Jahrgangsstufe noch am Ende des Schuljahrs einen positiven Effekt auf die Individualleistung hatte, negative Effekte auf das akademische Selbstkonzept jedoch nur zu Schuljahresbeginn bestanden. Mediationsanalysen zeigten zudem, dass der leistungsbezogene Kompositionseffekt über die Individualleistung in der Mitte des Schuljahrs mediiert wird, jedoch kein Mediationseffekt vom akademischen Selbstkonzept ausging.

Im Zentrum der *vierten Teilstudie* stand die Komposition der Schülerschaft hinsichtlich der Verteilung verschiedener Schülermerkmale. Dies erfolgte rein deskriptiv an der Stichprobe der DESI-Studie. Die klassenspezifischen Standardabweichungen in den leistungs- und herkunftsbezogenen sowie psychosozialen Schülermerkmalen verweisen darauf, dass die Schülermerkmale in der 9. Jahrgangsstufe an allen Schulformen streuten, sodass trotz der leistungsbezogenen Schulformgliederung heterogene Klassen vorlagen. Zudem ergaben sich aus der Kombination der klassenspezifischen Standardabweichungen und der mittleren Ausprägung eines Merkmals in der Klasse unterschiedliche Profile von heterogenen und homogenen Klassen an den Schulformen.

Die Befunde der Dissertation verweisen auf die Bedeutung der Schülerschaft als aktive Mitgestalter der schulischen Lern- und Entwicklungsumwelt, die für die Entwicklung der Leistung und psychosozialen Merkmale prägend ist. Die auf einer breiteren Konzeption von Kompositionseffekten basierten Befunde eröffnen wichtige Perspektiven für die administrative Ebene des Schulsystems und die pädagogische Praxis in ihrem gemeinsamen Wirken sowie für die zukünftige Forschung zu Effekten der Schülerzusammensetzung auf die individuelle schulische Entwicklung.

ABSTRACT

Within schools, children and adolescents' development occurs in a common context, in a community of same-aged peers. This context is not only constituted by the structure of the school system, as determined by educational policy, or by teaching, but also by the achievement levels of the learners who come together in class or school and by their psychosocial characteristics (Eccles & Roeser, 2011). In line with the *bioecological model* of Bronfenbrenner and Morris (2006), learners can thus be seen as cocreators of the school context. Studies on so-called *composition effects* have already shown that the student body's composition regarding achievement and their family background influences individual achievement development (for an overview, see Dumont, Neumann, Maaz, & Trautwein, 2013). In order to gain a more nuanced understanding of the role of the learner as cocreator of the school context, I considered (1) psychosocial student characteristics as characteristics whose development in the school context may be caused by the composition of the learning group, (2) psychosocial student characteristics at average levels and their distribution within the learning group as a composition characteristic and (3) the interplay of different composition effects. These questions were addressed within four substudies that used data sets from various longitudinal studies.

The *first substudy* addressed the question of whether the individual development of school satisfaction and norm-violating behavior was linked to the student composition of the school. Multilevel analyses of a sample of the BERLIN study showed for 9th Grade students that the greater the school-average achievement and the greater the average satisfaction level of the student body with the school, the greater the individual satisfaction with the school. In addition, there were indications pointing to further effects on learners' school satisfaction and norm-violating behavior by the composition regarding achievement and family background and via the composition of the student body regarding school satisfaction and norm-violating behavior.

In the *second substudy*, the focus was on how psychosocial influence operates between young people in the school context. For this purpose, I studied the effect of the class mean level of test anxiety on the change of individual test anxiety using a sample of the MSALT study. Multilevel analyses showed that the average level of test anxiety of newly formed classes at the beginning of the 7th Grade was positively associated with the change in learners' test anxiety from the end of the 6th Grade to the beginning of the 7th Grade. Interaction analyses also

revealed differences between girls and boys as well as between learners with a low and high prior test anxiety.

At the center of the *third substudy* was the interplay of composition effects on the development of achievement and on psychosocial characteristics. For this purpose, drawing on the BIJU study, I simultaneously modeled two composition effects that have been replicated on multiple occasions—on individual achievement and on individual academic self-concept—both of which proceed from the average achievement of a learning group. Multilevel analyses including *cross-lagged paths* between individual achievement and academic self-concept showed that the class-average achievement at the beginning of the 7th Grade still had a positive effect on individual achievement at the end of the school year, but that the negative effect on academic self-concept was only evident at the beginning of the school year. Mediation analyses additionally showed that the composition effect on individual achievement is mediated by individual achievement in the middle of the school year; yet there was no mediation effect via individual academic self-concept.

At the center of the *fourth substudy* was the composition of the student body with regard to the distribution of various student characteristics. This study was purely descriptive and used a sample from the DESI study. The class-specific standard deviations in achievement, in the social background, and psychosocial student characteristics were distributed in all school tracks in the 9th Grade, meaning classes were heterogeneous despite allocation to school tracks based on achievement. In addition, by combining the class-specific standard deviations and the average levels of a characteristic in the class, different profiles of heterogeneous and homogeneous classes in the different school tracks were found.

The results of the dissertation point to the importance of the student body as a cocreator of the school's learning and development environment, which is influential for the development of achievement and psychosocial characteristics. The findings, which are based on a broader conception of composition effects, open up important perspectives for the joint work of school-system administrators and pedagogic practitioners as well as for future research on effects of the student body composition on the individual school development.

1

Einleitung und theoretischer Rahmen der Arbeit

1 Einleitung und theoretischer Rahmen der Arbeit

“As whole persons encompassing social, emotional, moral, behavioral and cognitive intellectual dimensions of being, humans both learn and are affected by acts of teaching and learning across multiple dimensions – not just the cognitive/intellectual one.”

Eccles & Roeser, 2011, S. 575

Jugendliche verbringen einen Großteil ihres Tages in der Schule. Schule ist so nicht nur ein Ort des Lernens, sondern auch ein Lebensort, an dem Jugendliche sich in ihrer Ganzheit entwickeln. So bedingt der Schulkontext neben außerschulischen Einflüssen die Entwicklung der Jugendlichen hinsichtlich ihrer psychosozialen Merkmale entscheidend mit, was jedoch erst wenig erforscht ist (Eccles & Roeser, 2011; Pekrun & Helmke, 1991). Der jeweilige Schulkontext, in dem sich Lernende befinden, kann dabei ganz unterschiedlich aussehen. So individuell wie jede Schülerin und jeder Schüler, so individuell ist auch jede Klasse und Schule. Die verschiedenen Eigenschaften, die die Lernenden in eine Klasse oder Schule mitbringen, führen zusammengenommen zu einer ganz bestimmten Zusammensetzung der jeweiligen Lerngruppe¹. Im deutschen Bildungssystem zeigt sich dies besonders dahingehend, dass die auf Schülerleistungen basierende Allokation der Schülerinnen und Schüler an verschiedene Schulformen zu einer schulformspezifischen Zusammensetzung der Schülerschaft und somit zu *differenziellen Lern- und Entwicklungsmilieus* geführt hat (Baumert, Stanat & Watermann, 2006). Diese kennzeichnen sich entsprechend der Schulformgliederung durch ihr durchschnittliches Leistungsniveau, aber auch durch die spezifische sozioökonomische und ethnische Zusammensetzung der Schülerschaft. Wie Studien aus der Forschung zu sogenannten Kompositionseffekten einschlägig zeigen konnten, geht die jeweilige Zusammensetzung mit einer bestimmten Leistungsentwicklung der Schülerinnen und Schüler einher (siehe im Überblick Dumont, Neumann, Maaz & Trautwein, 2013; siehe Metaanalyse von van Ewijk & Slegers, 2010a, 2010b; für empirische Studien siehe Burns & Mason, 2002; Danhier & Martin, 2014; Hanushek, Kain, Markman & Rivkin, 2003; Hornstra, van der Veen, Peetsma & Volman, 2015; Justice, Logan, Lin & Kaderavek, 2014; Marks, 2010; Opdenakker, van Damme, de Fraine, van

¹ Der Begriff der Lerngruppe wird in der vorliegenden Dissertation als Oberbegriff für Klassen (auch Kursklassen) und Schulen genutzt.

Landeghem & Onghena, 2002). So ist die individuelle Leistungsentwicklung der Lernenden an Schulen mit einer leistungsschwachen oder sozial benachteiligten Schülerschaft geringer als an Schulen mit einer leistungsstarken und sozial begünstigten Schülerschaft. Zur Frage, inwieweit die Entwicklung psychosozialer Merkmale durch die Komposition der Schülerschaft beeinflusst wird, ist bisher durch den sogenannten *Big-Fish-Little-Pond-Effekt* (BFLPE; Marsh, 1987; siehe im Überblick Marsh, Seaton et al., 2008) bekannt, dass die Einschätzung der eigenen Fachkompetenzen in leistungsstarken Lerngruppen geringer ausfällt als in leistungsschwachen Lerngruppen. Darüber hinaus ist bisher weitgehend ungeklärt, inwiefern sich Schülerinnen und Schüler in Abhängigkeit von der Lerngruppe, in die sie eingebunden sind, in ihren psychosozialen Merkmalen entwickeln. Auch wie zu bewerten ist, dass sowohl erwünschte als auch unerwünschte Effekte von einem Merkmal der Schülerzusammensetzung ausgehen, ist nicht geklärt. Diese Frage ist jedoch von Relevanz, da schulstrukturbezogene Maßnahmen der Bildungspolitik an der Schülerzusammensetzung ansetzen können, um Schule effektiver zu gestalten. In vielen deutschen Bundesländern wurde beispielsweise die Mehrgliedrigkeit der Schulformen hin zu einer Zweigliedrigkeit abgebaut, um die starke Ausdifferenzierung unterschiedlicher Lern- und Entwicklungsmilieus zu hemmen und die Leistungsentwicklung auch an leistungsschwachen Schulen zu verbessern (Neumann, Maaz & Becker, 2013). Dabei wurde ein möglicher negativer Effekt der mittleren Leistung der Schülerschaft auf die individuelle Einschätzung der Fachkompetenzen im Sinne des BFLPE vermutlich nicht in Betracht gezogen. Um abwägen zu können, welchem Effekt die größere Beachtung bei bildungspolitischen Entscheidungen zukommen sollte, sind Informationen über das Zusammenspiel verschiedener Kompositionseffekte nötig.

Mit Steuerungsmaßnahmen wie der Einführung der Zweigliedrigkeit des Schulsystems, um der Kumulation ungünstiger Kompositionsmerkmale an bestimmten Schulformen entgegenzuwirken, geht weiterhin einher, dass sich die Heterogenität innerhalb von Schulen und Klassen erhöht. In Deutschland ist das Thema Heterogenität seit einigen Jahren in aller Munde. In der Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (KMK, 2015) wurde der „Umgang mit Heterogenität“ zu einer der „Schlüsselfragen der Schul- und Unterrichtsentwicklung“ erklärt (KMK, 2015, S. 15). In aktuellen bildungspolitischen und -wissenschaftlichen Diskursen wird unter dem Begriff der Heterogenität diskutiert, ob mit der Heterogenität von Lerngruppen förderliche oder unerwünschte Einflüsse auf die individuelle schulische Entwicklung einhergehen (siehe im Überblick Budde, 2012; sowie Walgenbach,

2014). Dazu wären zunächst Informationen relevant, von welcher Heterogenität innerhalb von Lerngruppen eigentlich ausgegangen werden muss. Trotz des breiten Diskurses über Heterogenität ist wenig darüber bekannt, wie heterogen die Schülerschaft innerhalb deutscher Klassenzimmer tatsächlich ist. Diesbezügliche Studien kamen bisher zu uneinheitlichen Befunden.

Während bildungspolitische Maßnahmen die Schülerzusammensetzung bis zu einem gewissen Grad steuern, können sie am konkreten täglichen Miteinander in der Schule kaum greifen. Nicht zuletzt, weil Schülerinnen und Schüler auch mit ihren psychosozialen Merkmalen zu einer bestimmten Zusammensetzung führen, die vom Schulsystem schwerlich gesteuert werden kann. Hier sind pädagogische Maßnahmen gefragt, die im Schulalltag auf psychosoziale Merkmale der Lernenden einwirken können, sodass eine günstigere Schülerzusammensetzung hinsichtlich ihrer psychosozialen Merkmale entsteht, die wiederum die individuelle schulische Entwicklung begünstigt. Inwiefern psychosoziale Merkmale in ihrer Komposition auf die individuelle schulische Entwicklung Einfluss nehmen, stellt bislang jedoch ein Forschungsdesiderat dar.

Durch eine breitere Auffassung von Kompositionseffekten soll den benannten offenen Fragen zum Einfluss der Schülerzusammensetzung auf die individuelle schulische Entwicklung näher gekommen werden. Die Ziele der vorliegenden Dissertation bestehen damit in der Ausweitung der bislang schmalen empirischen Evidenz (1) zu Effekten der Lerngruppenzusammensetzung auf psychosoziale Merkmale, (2) zu Effekten, die von der psychosozialen Komposition der Lerngruppe ausgehen sowie (3) zum Zusammenspiel verschiedener Kompositionseffekte. Dabei wird die Schülerkomposition sowohl hinsichtlich ihrer mittleren Zusammensetzung als auch hinsichtlich der Verteilung von Schülermerkmalen berücksichtigt. Diesen Zielen wird sich in vier Teilstudien mit je verschiedenem Fokus genähert.

Bevor die benannten Teilstudien vorgestellt werden, werden theoretische Annahmen sowie der Forschungsstand zum Einfluss der Schülerzusammensetzung auf die individuelle Entwicklung der Leistung und psychosozialer Merkmale erörtert. Die theoretischen Annahmen werden zunächst damit eingeleitet, dass die Rolle psychosozialer Merkmale als weitere Zieldimension von Schule neben der Leistungsentwicklung verdeutlicht wird (Kapitel 1.1). Hierauf folgt die Darstellung von Annahmen, inwiefern die Schülerschaft den Lern- und Entwicklungsort Schule mitkreiert und so auf die individuelle schulische Entwicklung der Einzelnen Einfluss nehmen kann (Kapitel 1.2). Anschließend werden Theorien aus der

Perspektive der Sozialpsychologie, der Entwicklungspsychologie und der Schul- und Unterrichtsforschung hinzugezogen, um dem in der Forschung zu Kompositionseffekten angenommenen Phänomen der sozialen Beeinflussung zwischen Lernenden näherzukommen (Kapitel 1.3). Danach wird der aktuelle Forschungsstand zu Kompositionseffekten dargestellt (Kapitel 1.4), der die üblicherweise in der Schuleffektivitätsforschung betrachteten Effekte der leistungs- und herkunftsbezogenen Komposition auf die Leistungsentwicklung um psychosoziale Schülermerkmale als Kriteriums- und Kompositionsmerkmale ergänzt. Ausgehend von den theoretischen Erörterungen und empirischen Befunden, die resümiert werden, erfolgt schließlich die Ableitung der Fragestellungen der vier Teilstudien (Kapitel 1.5).

1.1 Schule als Lern- und Entwicklungsort

Im folgenden Kapitel wird zunächst anhand der gesellschaftlichen Funktionen von Schule beschrieben, inwieweit Schule als Lern- und Entwicklungsort angelegt ist bzw. inwieweit Schule aufgrund gesellschaftlicher Ziele als Lern- und Entwicklungsort bildungspolitisch konstituiert wird. Im zweiten Schritt wird herausgestellt, welche multiplen individuellen Entwicklungsziele durch die verschiedenen gesellschaftlichen Funktionen für Schülerinnen und Schüler hervorgehen.

1.1.1 Die gesellschaftlichen Ziele von Schule

Schule ist nicht nur eine Einrichtung des Lernens zum Zweck des Lernens an sich. Als öffentliche Institution bedient sie den gesellschaftlichen Auftrag, „die Entwicklung der heranwachsenden Generation zu beeinflussen, um den Fortbestand von Gesellschaft und kultureller Menschheitsevolution zu sichern“ (Pekrun & Helmke, 1991, S. 33). Nach Fend's *Neuer Theorie der Schule* (2009) können vier gesellschaftliche Funktionen von Schule unterschieden werden, die gemeinsam zu diesem übergeordneten Ziel führen sollen: kulturelle Reproduktion, Qualifikation, Allokation und Integration. Die Funktion der *kulturellen Reproduktion* beschreibt Fend damit, dass Gesellschaften ihre kulturellen Weltanschauungen, Werteparadigmen und geteilten Symbolsysteme wie Sprach-, Schrift- und Zahlensysteme durch Beschulung an die nächste Generation weitergeben. *Qualifikation* wird als das Ausstatten der Lernenden mit Wissen und Fertigkeiten verstanden, um sie für das spätere Berufsleben vorzubereiten. Moderne Gesellschaften sichern damit ihre Wettbewerbsfähigkeit auf dem internationalen Wirtschaftsmarkt. Mit der *Allokationsfunktion* bezieht sich Fend auf die Zuweisung der Lernenden auf verschiedene Schulformen anhand ihrer schulischen Leistungen.

Damit werde erzielt, dass die Lernenden für ihre spätere berufliche Position vorsortiert werden, um so das Sozialsystem der Gesellschaft zu sichern. Die Verteilung der Lernenden auf verschiedene Bildungsgänge zur Erlangung bestimmter Qualifikationen solle dabei die Grundlage für den späteren Berufsweg bilden. Fend stellt zudem die *Integrations- bzw. Legitimationsfunktion* der Schule fest, über die indirekt das gesellschaftliche Ziel des Zusammenhalts der Gesellschaft sowie der Loyalität gegenüber dem bestehenden politischen Ordnungssystem gesichert werden solle. Die Lernenden werden in die bestehende gesellschaftliche Ordnung integriert, indem Einstellungen, Ansichten und Verhaltensweisen, die in der jeweiligen Gesellschaft als wünschenswert angesehen sind, in der Schule vermittelt und eingeübt werden.

1.1.2 Multiple Ziele von Schule für Schülerinnen und Schüler

Um dem Anspruch der vier in Kapitel 1.1.1 genannten gesellschaftlichen Funktionen von Schule gerecht zu werden, zielt Beschulung laut Fend (2009) auf die langfristig angelegte Entwicklung von Kindern und Jugendlichen in verschiedenen Dimensionen. Aus den *gesellschaftlichen* Funktionen ergeben sich demnach *individuelle* Funktionen von Schule für die Schülerinnen und Schüler. Fend geht davon aus, dass die Vermittlung der grundlegenden kulturellen Sinnsysteme einer Gesellschaft und grundlegender kultureller Fertigkeiten Lernende zur *kulturellen Teilhabe* und zur *Entwicklung einer kulturellen Identität* befähigen sollen. Durch die Qualifikationsfunktion der Schule werde den Lernenden Fähigkeiten und Fertigkeiten vermittelt, die grundlegend für die spätere *Berufsfähigkeit* sind. Für die jeweiligen beruflichen Wege seien die durch die Allokationsfunktion des Bildungssystems definierten Zertifikate relevant, sodass Lernende in meritokratischen Gesellschaften wie in Deutschland ihre berufliche *Lebensplanung* bis zu einem gewissen Grad durch Leistungsbereitschaft strukturieren können. Je nach Internalisierung der in der Schule vermittelten gesellschaftlich erwünschten Einstellungen, Ansichten und Verhaltensweisen werden zudem eine *soziale Identität* erworben sowie die Fähigkeit zur *politischen Teilhabe*.

Um die Ziele des Bildungswesens umzusetzen, müssen die gesellschaftlichen und individuellen Funktionen ineinandergreifen (Fend, 2009). In seiner Neuen Theorie von Schule kommt Fend zu dem Schluss, dass die entscheidende Verbindung zwischen gesellschaftlichen und individuellen Funktionen in der pädagogischen Arbeit liegt, die sich nicht nur auf die Vermittlung von Wissen und dem Aufbau kognitiver Kompetenzen konzentriert. So schreibt er:

„Die Organisation ‚Bildungswesen‘ erfüllt ihre Aufgaben durch die Arbeit an der ‚Seele‘ des Menschen. Sie bearbeitet das Können und die Haltungen lernfähiger junger Menschen. [...] Ihr Arbeitsfeld ist die psychische Verfassung, sind das Können, das Wissen wie auch die seelischen Einstellungen von Kindern und Jugendlichen.“ (Fend, 2009, S. 174). Lernende werden demgemäß in ihrer Ganzheit betrachtet, sodass die Ziele von Beschulung neben der Förderung des Leistungs- und Fähigkeitsniveaus auch in der Entwicklung der Lernenden in psychosozialen Merkmalen wie Einstellungen liegen. Er beschreibt das Ziel von Beschulung konkret als „‚Menschenbildung‘ im Sinne der Förderung von Wissen, Kompetenzen, psychischen Ressourcen und Werten“ (Fend, 2009, S. 169).

Ähnlich wie Fend verschiedene Aspekte der Entwicklung als multiple Ziele der Beschulung anspricht, nennen andere Bildungswissenschaftlerinnen und Bildungswissenschaftler neben dem Ziel der Leistungs- und Fähigkeitsentwicklung der Lernenden weitere Ziele, die die Entwicklung psychosozialer Merkmale betreffen. Im *Kompetenzmodell des deutschen Schulwesens* werden die multiplen Ziele in die vier Kompetenzformen Fach-, Methoden-, Sozial- und Selbstkompetenz nach Lehmann und Nieke (2000) unterschieden. Kunter (2005) gibt als multiple Ziele von Schule Kenntnisse und Fertigkeiten, Persönlichkeitsbildung und fachübergreifende Fähigkeiten an. Eccles und Roeser (2011) unterscheiden die kognitive Entwicklung, sozio-emotionale Entwicklung und Verhaltensentwicklung, wobei sie an anderer Stelle auch von intellektuellen, sozialen, moralischen und emotionalen Entwicklungsbereichen sprechen. Angesichts der Vielzahl an Begrifflichkeiten für Kategorien von Schülermerkmalen, die sich innerhalb von Schule entwickeln, soll die schulische Entwicklung in der vorliegenden Dissertation dahingehend unterteilt werden, ob es sich um die Leistungsentwicklung oder um die Entwicklung psychosozialer Merkmale handelt. Dabei werden fachspezifische Leistungs- und Fähigkeitsniveaus der Lernenden unter dem Begriff der Leistungsentwicklung gebündelt. Unter dem Begriff der psychosozialen Merkmale, der in der Literatur häufig synonym mit den Termini der sozial-emotionalen, nicht-kognitiven, emotional-motivationalen oder psychologisch-pädagogischen Merkmale verwendet wird, sollen motivationale und emotionale Merkmale, Selbstwahrnehmungen und Aspekte der Selbststeuerung, soziale Kompetenzen sowie Verhaltensweisen zusammengefasst werden.

Schließlich ist hervorzuheben, dass die multiplen Ziele von Schule nicht immer unabhängig voneinander zu erzielen sind. Aus pädagogisch-psychologischer Perspektive wird angenommen, dass die Leistungsentwicklung und die Entwicklung psychosozialer Merkmale in wechselseitiger

Beziehung stehen. So werden beispielsweise emotionale Schülermerkmale wie Testangst oder Lernfreude als Reaktionen auf Lern- und Leistungssituationen angesehen, die wiederum nachfolgendes Lernverhalten und Leistungsergebnisse bedingen (Hascher, 2004; Pekrun, 2000; Walberg, 1990). Zudem wird angenommen, dass hohe schulische Leistungen mit einer höheren Motivation einhergehen, wodurch wiederum das Lernverhalten geprägt wird, z. B. indem Motivation mehr Anstrengungsbereitschaft oder Durchhaltevermögen auslöst (Gottfried, 1990; Skaalvik & Valås, 1999; Wigfield, Eccles, Schiefele, Roeser & Davis-Kean, 2006). Für die Selbstwirksamkeit, intrinsische Motivation, Selbstkontrolle und metakognitive Strategien konnten bereits positive Effekte auf die Leistungsentwicklung nachgewiesen werden. Ein prominentes Beispiel für die wechselseitige Beziehung von Leistungsmerkmalen und psychosozialen Merkmalen ist das sogenannte *Reciprocal Effects Model* (siehe im Überblick Marsh & Martin, 2011). Demgemäß besteht eine reziproke Beziehung zwischen Leistungskompetenzen und dem *akademischen Selbstkonzept*, d.h. der Einschätzung der eigenen Leistungen in einer bestimmten Domäne. Bei hohem Leistungsniveau fällt die Einschätzung der eigenen Leistungen demnach entsprechend positiver aus als bei geringem Leistungsniveau (*self enhancement*; Calsyn & Kenny, 1977). In entgegengesetzter Richtung wirkt ein positives akademisches Selbstkonzept bestärkend auf das spätere Leistungsniveau (*skill development*; Calsyn & Kenny, 1977).

1.2 Das Spannungsfeld Person und Umwelt

Da zur Umsetzung der multiplen Ziele von Schule die Lernangebote aktiv von den Lernenden genutzt werden müssen (Fend, 2009; Helmke, 2015) und Lernen im Schulsetting in einem sozialen Rahmen stattfindet, soll im Folgenden die Rolle der Lernenden bei der Konstitution der Lern- und Entwicklungsumwelt Schule beleuchtet werden. Daher soll zunächst die Position der Schülerschaft in der Schulumwelt anhand des bioökologischen Modells der menschlichen Entwicklung (Bronfenbrenner & Ceci, 1994; Bronfenbrenner & Morris, 2006) bzw. anhand seiner Anwendung auf den Schulkontext (Eccles & Roeser, 2011) verortet werden. Anschließend wird eruiert, inwieweit die Rolle der Schülerschaft durch strukturelle Maßnahmen des jeweiligen Schulsystems in vorstrukturierten Bahnen verläuft. Schließlich wird aus der Perspektive der Forschung zu Schuleffektivität erschlossen, inwiefern der Schülerschaft durch ihre spezifische Zusammensetzung auch Einflüsse auf die schulische Entwicklung der einzelnen Lernenden zugesprochen werden können.

1.2.1 Das Ökosystem Schule – Die Schülerschaft als Mitgestalter der Lernumwelt

Menschliche Entwicklung wird im Sinne von Lewin (1952) als eine Funktion der Personenmerkmale und der wahrgenommenen Umwelt gesehen. Das *Ökologische Modell der menschlichen Entwicklung* nach Bronfenbrenner (1979) bzw. das darauf aufbauende, aktuell von ihm postulierte bioökologische Modell der menschlichen Entwicklung nach Bronfenbrenner und Morris (2006) erörtert, dass Menschen auch auf ihre Umwelt wirken. Durch interaktive Prozesse mit den sie umgebenden Umwelten, d.h. der verschiedenen strukturellen und sozialen Kontexte, werden Menschen in ihrer Entwicklung von der Umwelt geprägt. Gleichzeitig prägen sie dabei auch wechselwirkend die Umwelt. Menschen können damit als aktive Gestalter ihrer unmittelbaren Umwelten gesehen werden. Unmittelbare Umwelten sind bei Bronfenbrenner (1979) und Bronfenbrenner und Morris (2006) alle sogenannten Mikrosysteme, in denen Menschen in Beziehung mit anderen Menschen bzw. Menschengruppen stehen, wie z. B. die Schulklasse als ein Mikrosystem oder der Freundeskreis als ein anderes Mikrosystem. Die verschiedenen Mikrosysteme und ihr Zusammenspiel wie z. B. gleichzeitige Einflüsse aus dem Schulkontext und aus dem Freundeskreis, bilden das sogenannte Mesosystem. Auf hierarchischen Ebenen angeordnet, differenzieren Bronfenbrenner und Morris weiterhin das Exosystem und das Makrosystem als Teil der gesamten Umwelt eines Menschen. Auf diese haben der einzelne Mensch keinen direkten Einfluss. Dem Exo- und Makrosystem werden jedoch Einflüsse auf die jeweils untergeordneten Systeme zugeschrieben, wodurch sie indirekt die Entwicklung des Einzelnen beeinflussen. Mit dem Exosystem sind Umwelten gemeint, welche die direkten Umwelten von Menschen formen, wie z. B. das Bildungssystem. Das Makrosystem als oberste Ebene des gesamten *Ökosystems* der Umwelt eines Menschen umfasst die grundlegenden Wertesysteme und Weltanschauungen der Gesellschaft, in der sich der Mensch befindet.

Wird das bioökologische Modell der menschlichen Entwicklung auf den Schulkontext bezogen, so kann der Schulkontext als mehrschichtiges soziales System konzeptualisiert werden (Eccles & Roeser, 2011). Grundlegend geformt vom Makrosystem der Gesellschaft unterteilt sich der Schulkontext gemäß dem *Modell des Schulkontextes* nach Eccles und Roeser (2011) noch einmal in verschiedene hierarchische Ebenen, die dem Mikrosystem der einzelnen Lernenden seine Gestalt geben (siehe Abb. 1.1). Der Schulkontext lässt sich dabei in sieben Ebenen unterteilen, die in interdependenter Beziehung stehen: die unmittelbare Ebene der schulischen Betätigung des Lernenden, die Ebene der Klassen, Kurse und Lerngruppen, die Ebene der

Lehrerschaft, die Ebene des Allokationssystems², die Ebene der Einzelschule als Organisation, die Ebene der Verbindung zwischen außerschulischer Lebenswelt und Schule sowie die Ebene des politischen Systems, in der sich die Schule befindet. Der Schulkontext breitet sich somit über distale Ebenen wie das jeweilige Bildungssystem an sich, quasi-proximale Ebenen wie die Schule als Organisation bis zur proximalen Ebene des unmittelbaren Unterrichtsgeschehens aus, in dem sich eine Schülerin oder ein Schüler befindet. Eigenschaften der distalen Ebenen stellen dabei Einflussfaktoren dar, die nicht unmittelbar auf die schulische Betätigung der Lernenden einwirken. Beispielweise gibt die Art der Allokation des jeweiligen Schulsystems strukturelle Bedingungen für die einzelnen Schulen vor und prägt somit schulform- und schulspezifische Charakteristiken mit. Schulformspezifische Lehrpläne oder Charakteristiken der Einzelschule wie ein spezifisches Schulprogramm können wiederum als quasi-proximale Einflussfaktoren auf die Entwicklung der Lernenden gesehen werden, die Vorgaben für das Unterrichtssetting setzen. Im Unterrichtssetting wird von wechselwirkenden proximalen Beziehungen zwischen Lernenden untereinander und zu Lehrerinnen und Lehrern ausgegangen. Eccles und Roeser (2011) nehmen an, dass Lernende durch das Zusammenspiel der verschiedenen Ebenen nicht nur in ihrer Leistungsentwicklung beeinflusst werden, sondern in ihrer ganzen Person, was auch die soziale und emotionale Entwicklung sowie die Verhaltensentwicklung einschließt. Gleichzeitig werden Lernende als Teil des Systems Schule gesehen, sodass sie den Schulkontext, in dem sie sich befinden, mitgestalten.

² Aus der Perspektive des deutschen Schulsystems mag die Anordnung des Allokationssystems als hierarchische Ebene unterhalb der Schulebene etwas ungewöhnlich sein. Zu beachten ist, dass Eccles und Roeser (2011) dieses Modell vorwiegend aus der Perspektive des amerikanischen Schulsystems entworfen haben. Anders als in Deutschland geschieht die Aufteilung der Lernenden in leistungsdifferente Lerngruppen hier innerhalb der Schulen.

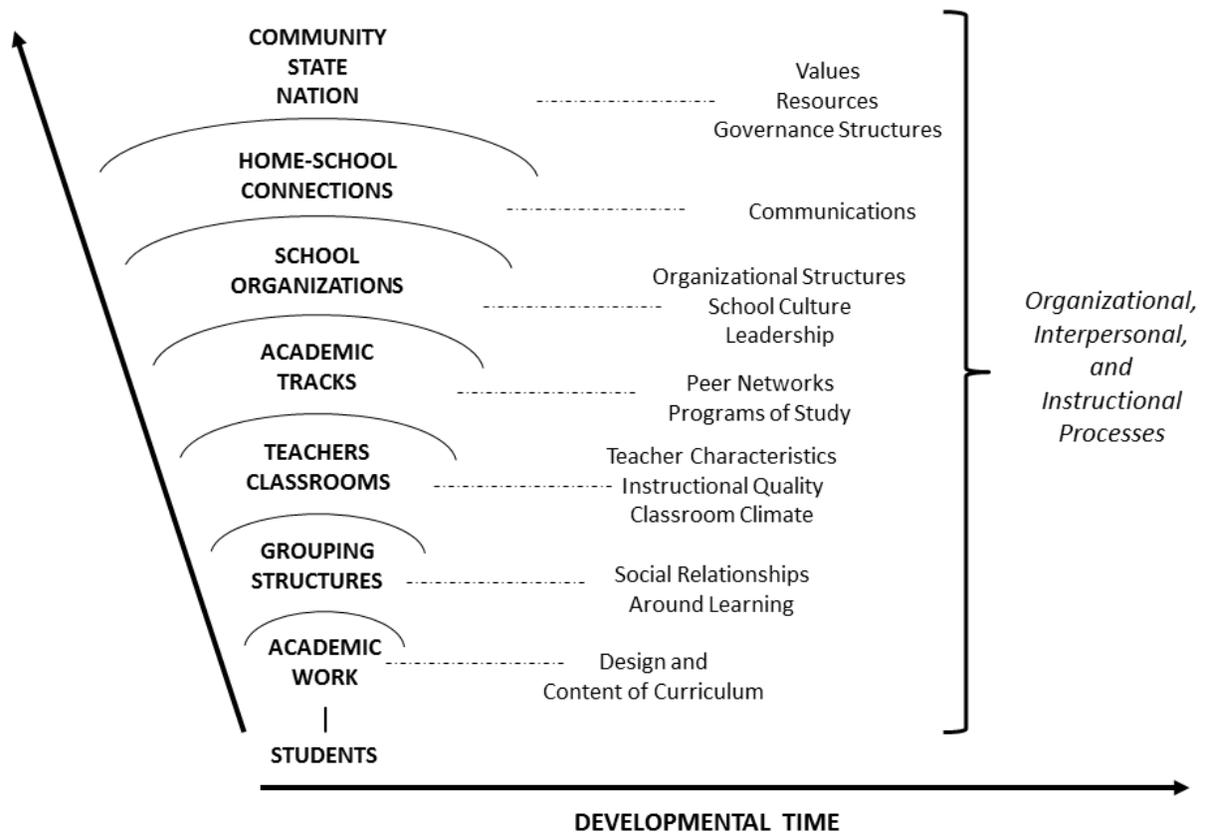


Abbildung 1.1. Modell des Schulkontextes nach Eccles und Roeser (2011, S. 574)

Das Zusammenspiel der einzelnen Ebenen des Schulkontextes erfolgt laut Eccles und Roeser (2011) über *Instruktionsprozesse* sowie *organisatorische* und *soziale Prozesse*, die auf und zwischen den verschiedenen Ebenen des Schulkontextes stattfinden. Als typisches Beispiel für einen *Instruktionsprozess* wird das Erteilen von Unterrichtsinstruktionen durch Lehrkräfte genannt. Die Qualität der eingesetzten Instruktionen sei dabei abhängig von Eigenschaften der Lehrerinnen und Lehrer wie deren professionelle Kompetenzen, aber auch davon, welche Schülerinnen und Schüler jeweils unterrichtet werden. Als Beispiel für *organisatorische Prozesse* kann die Organisation eines Programms zur Verbesserung des Schulklimas durch die Schulleitung genannt werden. Je nachdem, welche Merkmale die Schülerschaft aufweist, trägt sie gemäß dem Modell des Schulkontextes zu einem bestimmten Schulklima bei und bietet Anknüpfungspunkte für organisatorische Maßnahmen zur Verbesserung des Schulklimas. Zudem wird angenommen, dass *soziale Prozesse*, die zwischen Lernenden untereinander und zwischen

Lernenden und Lehrkräften stattfinden, von Merkmalen der Beteiligten geprägt sind, diese aber nachfolgend wiederum von sozialen Prozessen beeinflusst werden.

Angesichts der multiplen Ebenen von Schule und deren Zusammenspiel kann davon ausgegangen werden, dass jede Schule eine spezifische Umwelt zur Entwicklung der Leistung und psychosozialer Merkmale der Lernenden darstellt. Hervorzuheben ist, dass die Leistung der Lernenden und psychosoziale Schülermerkmale dabei nicht nur als Zielfaktoren dargestellt werden, sondern in ihrem kumulierten Auftreten in Lerngruppen auch als Bedingungsfaktoren der Lernentwicklung. Der Lern- und Entwicklungsort Schule wird demnach auch von den Merkmalen der Schülerschaft mitkonstituiert. Diese Gleichzeitigkeit von Schülermerkmalen als Bedingungs- und Zielfaktoren findet sich ebenfalls im *Angebots-Nutzungs-Modell* von Helmke (2006) und in seiner Erweiterung durch Lipowsky (2006). Helmke beschreibt, dass es auf die individuelle Leistung und die individuelle Ausprägung psychosozialer Merkmale ankommt, inwieweit eine Schülerin oder ein Schüler die Lernangebote für sich nutzen kann. In der Erweiterung des Angebots-Nutzungs-Modells stellt Lipowsky heraus, dass Leistung und psychosoziale Schülermerkmale nicht nur auf individueller Ebene den Schulerfolg beeinflussen, sondern in ihrer Komposition innerhalb einer Lerngruppe zusammengenommen das Lehrer- und Schülerverhalten beeinflussen und somit indirekt die schulische Entwicklung der einzelnen Lernenden bedingen.

1.2.2 Die Zusammensetzung der Schülerschaft als Konsequenz der Schulstruktur und gesellschaftlicher Veränderungen

1.2.2.1 Die Rolle der Schulstruktur für die Zusammensetzung der Schülerschaft

Bei der Betrachtung der Schülerschaft als Mitgestalter des Schulkontextes und damit als Bedingungsfaktor für die schulische Entwicklung der einzelnen Lernenden, muss das jeweils bestehende Bildungssystem berücksichtigt werden. Die Zusammensetzung der Schülerschaft an den Schulen geschieht nicht zufällig. Sie kann als Konsequenz von *tracking*, d.h. der jeweiligen Form der Allokation von Schülerinnen und Schülern auf Schulformen, Schulen und/oder Kurse gesehen werden (Dar & Resh, 1986). Daher können vermeintliche Einflüsse der Schülerschaft auf die individuelle schulische Entwicklung der Lernenden mit Einflüssen verwechselt werden, die allein durch die Tendenz entstehen, dass sich Schülerinnen und Schüler einer Lerngruppe aufgrund der gleichen, sie umgebenden Umweltmerkmale ähnlich entwickeln (Manski 1993). Zum einen kommt eine bestimmte Zusammensetzung hinsichtlich der familiären

Hintergrundmerkmale durch das Einzugsgebiet von Schulen zustande – das sogenannte *implicit tracking* (Hallinan, 1994; Trautwein, Lüdtke, Marsh, Köller & Baumert, 2006). Bestimmte Regionen oder Stadtteile sind durch eine spezifische Population gekennzeichnet, sodass auch die ortsansässigen Schulen eine entsprechende Zusammensetzung der Schülerschaft hinsichtlich sozialer und ethnischer Hintergrundmerkmale aufweisen. Schulen mit besonderem Profil oder besonders präferierte Schulen können ebenfalls eine spezifische Zusammensetzung aufweisen, wenn der Schulbesuch hohe finanzielle Ressourcen erfordert oder wenn sie besonders von Eltern mit bestimmten Interessen oder hohen Bildungsaspirationen für ihr Kind nachgefragt werden. Haben die Eltern Möglichkeiten der Mitentscheidung wird diese Form des trackings als *opt-in-tracking* bezeichnet (Maaz, Trautwein, Lüdtke & Baumert, 2008; Trautwein et al., 2006).

Zum anderen entsteht eine bestimmte Zusammensetzung durch das jeweils vorherrschende System der Allokation von Schülerinnen und Schülern in leistungsdifferente Lerngruppen, was unter dem Begriff des *ability groupings* bekannt ist (Ireson & Hallam, 2001; Kulik & Kulik, 1982; Oakes, 1985; Trautwein, Köller, Lüdtke & Baumert, 2005). Das Bilden leistungsdifferenter Lerngruppen kann auf verschiedenen, institutionellen Ebenen geschehen (Chmielewski, Dumont & Trautwein, 2013; Trautwein et al., 2006): auf unterschiedliche Schulformen, die auf ein bestimmtes Leistungsniveau abgestimmt sind (*between-school tracking*), innerhalb von Schulen auf verschiedene Leistungsprofile (*within-school tracking*), innerhalb von Klassen in temporäre Kleingruppen (*within-class tracking*) oder kursweise innerhalb von Schulen für bestimmte Fächer (*course-by-course tracking*). Im Fall des *ability groupings* in leistungsdifferente Schulformen, wird auch von *explicit tracking* gesprochen, da die Zugehörigkeit zu einer leistungsschwachen oder -starken Gruppen hier besonders auffallend ist (Trautwein et al., 2006). Die leistungsdifferenten Schulformen, die durch *ability grouping* entstehen, charakterisieren sich neben dem entsprechend resultierenden mittleren Leistungsniveau der Schülerschaft weiterhin durch schulformspezifische Curricula mit unterschiedlich anspruchsvollen Leistungszielen sowie durch die damit einhergehende Eröffnung bzw. Verwehrung unterschiedlicher Karrierewege (Maaz et al., 2008; Trautwein et al., 2006).

Hervorzuheben ist, dass mit der leistungsbezogenen Schulformgliederung nicht nur die Leistungszusammensetzung an den verschiedenen Schulformen gesteuert wird, sondern aufgrund sogenannter sekundärer Herkunftseffekte auch die sozioökonomische und ethnische Komposition (Maaz et al., 2008). Sekundäre Herkunftseffekte bestehen darin, dass durch die Variation der Übergangsentscheidungen der Eltern, der Schulformempfehlungen durch Lehrerinnen und Lehrer

und der Leistungsbeurteilungen nach sozioökonomischem Hintergrund der Lernenden wesentlich mehr Lernende aus sozial begünstigten deutschstämmigen Familien ans Gymnasium übergehen als Lernende aus sozial benachteiligten Familien mit Zuwanderungshintergrund (Maaz & Nagy, 2009). In Deutschland findet sich daher an Gymnasien, also jenen Schulen mit einer leistungsstarken Schülerschaft, zumeist auch eine sozial begünstigte Schülerschaft und eine Komposition aus wenigen Schülerinnen und Schülern mit Zuwanderungshintergrund vor, während dies an Hauptschulen, als Schulen mit der leistungsschwächsten Schülerschaft, der umgekehrte Fall ist (Baumert et al., 2006). Wird hier betrachtet, inwieweit die individuelle schulische Entwicklung von der Zusammensetzung der Schülerschaft geprägt wird, ist zu beachten, dass Einflüsse durch die Schülerzusammensetzung und durch Merkmale der Schulform miteinander verwoben sind (Baumert et al., 2006). Förderliche Einflüsse auf die individuelle Leistungsentwicklung an Gymnasien sind so beispielsweise durch einen anspruchsvolleren Lehrplan als an den weniger akademisch orientierten Schulformen denkbar, andererseits aber auch durch die leistungsstarke und sozial begünstigte Schülerschaft.

1.2.2.2 Die Rolle aktueller Umstrukturierungen des Schulsystems und gesellschaftlicher Veränderungen für die Zusammensetzung der Schülerschaft

Bei der Betrachtung der Mitgestaltung des Schulkontextes durch die Schülerschaft muss beachtet werden, dass strukturelle Vorgaben des bestehenden Schulsystems sowie aktuelle politische, wirtschaftliche oder gesellschaftliche Veränderungen die Heterogenität in der Schülerschaft vorstrukturieren. Lange ging man davon aus, dass die Zusammenführung ähnlicher Schülerinnen und Schüler die beste Möglichkeit sei, die Lernenden optimal in ihrer Entwicklung zu fördern (LeTendre, Hofer & Shimizu, 2003). Dies zeigt sich historisch nicht nur in der Bildung leistungshomogener Gruppen durch ability grouping, sondern auch in der Bildung altersgleicher Gruppen durch Jahrgangsstufen oder der Separierung von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in spezielle Sonderschuleinrichtungen (Tillmann, 2008). Aktuelle Trends verlaufen in entgegengesetzter Richtung. Im deutschen Schulsystem restrukturierten die meisten Bundesländer das mehrgliedrige Sekundarschulsystem in ein zweigliedriges System (Neumann et al., 2013), was eine verstärkte Heterogenität in den Schulen und Klassen impliziert. Zudem finden sich in vielen Bundesländern Ansätze zur Aufweichung der Jahrgangsstufen im Elementarschulsystem (Carle & Metzen, 2014). Eine weitere Umstrukturierung der Schulsysteme zu mehr Heterogenität löste das Übereinkommen

über die Rechte von Menschen mit Behinderung aus, das 2009 von Deutschland ratifiziert wurde (Bundeskabinett, 2011). Das Übereinkommen beinhaltet unter anderem die Forderung zur Schaffung eines inklusiven Schulsystems, in dem Schülerinnen und Schüler ohne und mit Behinderung in allgemeinbildenden Schulen gemeinsam lernen. Aktuell werden Schulen zudem bezüglich der ethnischen Hintergründe von Lernenden heterogener, da kriegsbedingte Fluchtbewegungen die rasche Integration geflüchteter Kinder und Jugendlicher ins Schulsystem erfordern (KMK, 2016). Von welchem Ausmaß der Heterogenität hinsichtlich der verschiedenen Schülermerkmale innerhalb von Klassen und Schulen in Deutschland dabei eigentlich ausgegangen werden muss, ist noch nicht belegt. Die bisherige Vorstellung der Bildung homogener Lerngruppen durch leistungsdifferente Schulformen bezeichnet Tillmann (2007) zumindest als „Fiktion“ (S. 8). Diese Einschätzung zieht er aus dem Befund der ersten PISA-Studie, dass Leistungsüberschneidungen zwischen Lernenden verschiedener Schulformen bestehen (PISA-Konsortium, 2001). Für die Heterogenität von *Schulen* zeigte sich tatsächlich, dass die Leistungen an deutschen Sekundarschulen im Vergleich zu anderen OECD-Staaten, die an PISA teilnahmen, überdurchschnittlich streuten (Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2016). Wie leistungsheterogen aber *Klassen* sind, in denen soziale Prozesse im Unterricht konkret stattfinden, ist nicht bekannt. Auch bezüglich der Verteilung psychosozialer Merkmale innerhalb von Schulen und Klassen besteht kaum Empirie. Es zeigten sich lediglich Hinweise, dass die Schülerschaft in Gymnasialklassen bezüglich verschiedener Kompetenzbereiche (Prenzel, Sälzer, Klieme & Köller, 2013) sowie im akademischen Selbstkonzept und in der Lernmotivation (Watermann, Klingebiel & Kurtz, 2010) homogener als an anderen Schulformen ist.

1.2.3 Der Einfluss der Schülerzusammensetzung auf die schulische Entwicklung Einzelner

Basierend auf der Durkheim'schen Auffassung, dass soziale Strukturen das menschliche Verhalten leiten (Durkheim, 1894), werden in der Schuleffektivitätsforschung seit Langem auch die Merkmale der Schülerschaft als Bedingungsfaktoren für schulische Entwicklung angesehen (Baumert et al., 2006; Coleman et al., 1966; Dar & Resh, 1986; Dreeben & Barr, 1988; Dumont et al., 2013; Harker & Tymms, 2004; Thrupp, 1999; Wilkinson et al., 2000). Demnach basiert schulische Entwicklung auf individuellen Dispositionen sowie Charakteristiken des Schulkontextes, aber eben auch auf den Merkmalen der Schülerschaft. In der Forschungstradition zur Schuleffektivität werden Merkmale der Schülerschaft unter dem Begriff der *Komposition*

zusammengefasst (siehe im Überblick Harker & Tymms, 2004; sowie Dumont et al., 2013). Um den spezifischen Einfluss der Schülerkomposition von Einflüssen anderer Charakteristiken des Schulkontextes wie den in Kapitel 1.2.2.1 genannten Einflüssen durch die Schulform abzugrenzen, wird der Begriff der *Kompositionseffekte* genutzt (Harker & Tymms, 2004; Wilkinson et al., 2000). Dieser wird zudem als Unterkategorie verschiedener Kontexteffekte verstanden, zu denen auch die als institutionelle Effekte bezeichneten Einflüsse der Schul- oder Schulformcharakteristiken zählen (Dumont et al., 2013). Die Komposition der Schülerschaft kann sich durch die *mittlere Zusammensetzung* der Schülerschaft hinsichtlich verschiedener Schülermerkmale sowie durch die *Verteilung* der verschiedenen Schülermerkmale innerhalb der Lerngruppe kennzeichnen (Dumont et al., 2013). Schülermerkmale, die in ihrer Komposition als relevant behandelt werden sind das Leistungs- und Fähigkeitsniveau, der sozioökonomische Hintergrund, der Bildungshintergrund, die Zugehörigkeit zu einer ethnischen Minderheit sowie das Vorliegen eines Zuwanderungshintergrunds (siehe im Überblick Harker & Tymms, 2004; sowie Baumert et al., 2006; Dumont et al., 2013). Baumert et al. (2006) führen darüber hinaus die Kompositionsmerkmale *belastende Familienverhältnisse* wie instabile Familienstrukturen und die Erwerbslosigkeit des hauptverdienenden Elternteils sowie *lernbiographische Belastungsfaktoren* wie Klassenwiederholungen und der Wechsel auf eine weniger akademisch orientierte Schulform an.

Es wird angenommen, dass diese verschiedenen Kompositionsmerkmale auf die individuelle Leistung sowie auf Schülermerkmale wie Bildungsaspirationen, Leistungsmotivation und Selbsteinschätzungen wirken, die besonders mit der Leistungsentwicklung verbunden sind (Baumert et al., 2006; Wilkinson et al., 2000). Dies wird vermutet, da auch die Schülerschaft mit ihren je spezifischen Merkmalen den Lern- und Entwicklungsort Schule formen dürfte.

Die verschiedenen Wege, über die die Schülerschaft die Schulumwelt mitzugestalten vermag, können grob in drei Kategorien zusammengefasst werden (siehe im Überblick Harker & Tymms, 2004; sowie Dumont et al., 2013). Eine erste Möglichkeit, durch die die Schülerschaft die Schulumwelt mitgestaltet, besteht darin, dass die jeweilige Zusammensetzung der Lernenden die finanziellen und sozialen Ressourcen der Schule bedingt und so die Schulqualität beeinflussen könnte. Besonders deutlich wird dies im amerikanischen Schulsystem, wo Schulen häufig aus Gemeindesteuern finanziert werden und die materielle Ausstattung abhängig von den finanziellen Ressourcen der Gemeinden ist, aus dem die Lernenden kommen (Harker & Tymms, 2004; Harris, 2010). Aber auch in Schulsystemen, in denen dies nicht der Fall ist, könnte durch

die Schulstruktur, das Einzugsgebiet oder an Schulen mit besonderem Profil eine bestimmte Zusammensetzung der Schülerschaft entstehen, die der Schule mehr oder weniger soziale Ressourcen wie elterliches Engagement oder elterliche Leistungs- und Verhaltenserwartungen zur Verfügung stellt (Baumert et al., 2006). Zudem könnten Schulen mit ungünstiger Schülerkomposition und den damit einhergehenden geringen finanziellen, materiellen und/oder sozialen Ressourcen eher weniger hochqualifiziertes Personal anziehen (Harker & Tymms, 2004).

Eine zweite Annahme, inwiefern die Zusammensetzung die Lernumwelt entwicklungsrelevant prägen kann, besteht darin, dass Lehrerinnen und Lehrer sowie die Schule als Organisation auf die Zusammensetzung der Lernenden reagieren, indem sie Maßnahmen an die Merkmale der Lernenden adaptiert (Dumont et al., 2013; Harker & Tymms, 2004). So könnten Lehrkräfte beispielsweise Tempo und Niveau ihres Unterrichts an die Leistungsstärke der Lerngruppe anpassen (Dumont et al., 2013; Hornstra et al., 2015). Auch die Erwartungen der Lehrkräfte an die Leistungen der Schülerinnen und Schüler könnten von der Leistung der Lerngruppe abhängen, die dann wiederum wie selbsterfüllende Prophezeiungen auf die schulische Entwicklung der Lernenden wirken (Dumont et al., 2013; Hornstra et al., 2015). Weiterhin besteht die Möglichkeit, dass die Schulorganisation auf die Zusammensetzung der Lernenden reagiert, indem das Curriculum und die Zuordnung der Schülerinnen und Schüler zu den Klassen auf die Schulklientel abgestimmt werden oder Schulen Maßnahmen für besondere Problemlagen der jeweiligen Schülerklientel entwickeln (Harker & Tymms, 2004).

Drittens wird angenommen, dass sich Schülerinnen und Schüler gegenseitig über gewisse Gruppendynamiken (Dar & Resh, 1986; Dreeben & Barr, 1988; Erbring & Young, 1979; Wilkinson et al., 2000) und duale Schüler-Schüler-Interaktionen (Erbring & Young, 1979; Webb, 1991; Wilkinson et al., 2000) beeinflussen, die zusammengefasst als *Peer-Effekte* bezeichnet werden (Dumont et al., 2013; Harker & Tymms, 2004; Wilkinson et al., 2000). Gruppendynamiken können sich in mehreren Formen äußern. Zum einen, indem Lernende die umgebende Lerngruppe als Referenzgruppe nutzen, mit der sie sich vergleichen und so zu Einschätzungen zu ihren eigenen Kompetenzen kommen. Zum anderen wird angenommen, dass Lernende ihr Lernverhalten und andere Verhaltensformen an die Normen der sie umgebenden Lerngruppe anpassen. Daraus könnten sich bestimmte Muster der Lehrer-Schüler-Interaktionen und Schüler-Schüler-Interaktionen in Kleingruppen, Klassen und Schulen ergeben (Erbring & Young, 1979; Wilkinson et al., 2000). Beispiele für duale Schüler-Schüler-Interaktionen, die auf

die schulische Entwicklung wirken könnten, sind das Geben von Feedback, aufgabenbezogene Unterstützung oder Ermuntern (Webb, 1991). Gruppendynamiken und duale Schüler-Schüler-Interaktionen prägen so die jeweils einzigartige Lern- und Entwicklungsumwelt, in der Schülerinnen und Schüler sozialisiert werden. Der Begriff der Peer-Effekte, der die gegenseitige Beeinflussung der Schülerinnen und Schüler betont, wird daher folgendermaßen von Kompositionseffekten im Allgemeinen abgegrenzt: “Where peers are intimately involved in this socialization process, we construe these dynamics as a 'peer effect'.” (Wilkinson et al., 2000). Wilkinson et al. (2000) siedeln Peer-Effekte dabei auf einer direkteren Wirkungsebene als Kompositionseffekte im Allgemeinen an.

Bisher gelten die Annahmen zu besagten Wirkmechanismen in der Forschung zu Kompositionseffekten noch weitgehend als *Black Box* (Dumont et al., 2013; Jencks & Mayer, 1990; van Ewijk & Sleegers, 2010a). Es wird zudem betont, dass die verschiedenen Mechanismen vermutlich nicht isoliert voneinander verlaufen, sondern gleichzeitig und wechselwirkend (Hallinan, 1988; Nash, 2003).

1.3 Peer-Effekte – Die Schülerschaft als *aktive* Mitgestalter der Lern- und Entwicklungsumwelt

Für pädagogische Maßnahmen, die laut Fend (2009) die gesellschaftlichen Ziele von Beschulung an die Schülerinnen und Schüler vermitteln, dürften Sozialisationsprozesse innerhalb der Schülerschaft einen besonders attraktiven Anknüpfungspunkt der Schülerzusammensetzung darstellen, um sowohl die Entwicklung der Leistung als auch psychosozialer Merkmale zu fördern. Die in der Schuleffektivitätsforschung als Peer-Effekte bezeichnete gegenseitige Beeinflussung zwischen Schülerinnen und Schülern verweist auf die Rolle der Gleichaltrigen in der Schule als Sozialisatoren. In der Sozialisationsforschung wird die Schulumwelt als sekundäre Sozialisationsinstanz konzeptualisiert, die substantiell zur Entwicklung des Menschen beiträgt (Berger & Luckmann, 1969). Menschen wachsen demnach in die Gesellschaft hinein, indem sie sich in den jeweils aktuell zentralen Beziehungsgefügen aktiv mit ihrer Umwelt auseinandersetzen. Im Schulalter stellt die Lerngruppe der eigenen Schule, Klasse oder der Kurse ein zentrales Beziehungsgefüge dar, mit denen Schülerinnen und Schüler in täglicher und überdauernder Interaktion stehen (Burlingame, 1970; Steinberg & Silverberg, 1986). Soziale Wertvorstellungen, Normen und Rollen werden im wechselseitigen Austausch erlernt, sodass das Individuum in seinen Eigenschaften „geformt“ wird. Während solche Sozialisationsprozesse in

der Forschung zu Schuleffektivität hinsichtlich *leistungsbezogener* Werte, Normen und Rollen angenommen werden (Wilkinson et al., 2000), beschäftigen sich andere Disziplinen im übergeordneten Sinn mit Peer-Effekten bzw. gegenseitigen Einflüssen in Bezugsgruppen. Im folgenden Kapitel werden daher Theorien aus der Sozialpsychologie, Entwicklungspsychologie sowie der Schul- und Unterrichtsforschung zusammengetragen, um die in der Forschung zu Kompositionseffekten getätigten Annahmen zur aktiven Rolle der Schülerschaft in der Gestaltung des Schulkontextes zu ergänzen.

1.3.1 Das sozialpsychologische Konzept der normativen und komparativen Einflüsse in Bezugsgruppen

Aus der Perspektive der sozialpsychologischen *Theorie der Bezugs-* oder auch *Referenzgruppe* kann die Schülerschaft einer Schule oder Klasse als Bezugsgruppe definiert werden. Bezugsgruppen werden hierbei als Gruppen verstanden, zu denen Individuen in Beziehung stehen (Kelley, 1968; Sherif, 1953). Nach Kelley (1968) haben Bezugsgruppen zwei Funktionen: eine *normative* und eine *komparative*. Die normative Funktion besteht darin, dass eine Bezugsgruppe Normen für die Mitglieder setzt und diese über Belohnungs- und Sanktionssysteme bei Einhaltung bzw. Nichteinhaltung aufrechterhält. Dies diene dem Individuum als Richtschnur für eigenes Verhalten und eigene Einstellungen. Eine komparative Funktion üben Bezugsgruppen laut Kelley daher aus, da sie einen Maßstab darstellen, an dem Individuen sich selbst und andere vergleichen. Anhand der Merkmale, die Gruppenmitglieder aufweisen und/oder welches Verhalten und welche Einstellungen die Norm sind, werden individuelle Merkmale, Einstellungen und eigenes Verhalten beurteilt. Auch andere Personen werden nach diesen Maßstäben beurteilt. Kelley betont, dass die normative und komparative Funktion als eng miteinander verwoben gesehen werden müssen, da das Anpassen an die Normen nicht ohne eine vorherige Einschätzung der eigenen Merkmale in Relation zur Gruppennorm möglich sei und auch Vergleiche nicht ohne die Wahrnehmung der Norm getätigt werden können. Da dennoch spezifische Phänomene für die Erklärung der jeweiligen Funktionen angenommen werden, sollen sie im Folgenden getrennt voneinander beleuchtet werden.

Die normative Funktion von Bezugsgruppen basiert sozialpsychologisch gesehen auf der kognitiven Repräsentation gruppenspezifischer Richtlinien, die angeben, welche Verhaltensweisen innerhalb einer Gruppe angemessen sind (Schwartz, 1977; Wiswede, 2004). Solche Richtlinien bzw. sozialen Normen bilden sich durch soziales Handeln der Mitglieder

innerhalb einer Bezugsgruppe aus und bedingen wiederum nachfolgendes soziales Handeln (Coleman, 1990).³ Fasst man die verschiedenen Theorien der Sozialpsychologie zusammen, so geschieht das Einhalten bzw. die Aufrechterhaltung der Normen aufgrund des menschlichen Bedürfnisses nach Zugehörigkeit sowie nach Zustimmung und Akzeptanz durch Andere (Cialdini & Trost, 1998; McClelland, 1961). Aus diesen Bedürfnissen ergibt sich nach Ajzen (2005) die Motivation zur Konformität, die leitend für geplantes Handeln ist, um innerhalb des vorgegebenen Rahmens der Gruppennormen zu agieren. Je nach Ausprägung des Normbewusstseins nehmen Menschen die impliziten Normen, die innerhalb einer bestimmten Gruppe gelten, mehr oder weniger wahr (Ajzen, 2005). Welches Verhalten der Norm entspricht, wird durch sogenanntes *soziales Lernen* angeeignet (Bandura, 1979, 1986). Demnach reagiert die Bezugsgruppe auf gezeigtes Verhalten mit sozialen Konsequenzen wie Anerkennung, Zuwendung oder auch Ablehnung und andere Arten von Sanktionen, die als *soziale Verstärker* bezeichnet werden. Diese können entweder selbst erfahren werden (*direkte Verstärkung*) oder als Reaktion der Gruppe auf das Verhalten anderer beobachtet werden (*stellvertretende Verstärkung, Lernen am Model*; Bandura, 1979, 1986). Durch die sozialen Verstärker ergibt sich ein spezifisches Belohnungs- und Sanktionssystem von Gruppen, das die Ausübung normkonformen Verhaltens über Konditionierung fördert (Bandura, 1979, 1986). Dies führt bezüglich einiger Gruppennormen dazu, dass sie internalisiert werden und unbewusst handlungsleitend werden, ohne dass die Konformitätserwartung sowie Sanktionen bzw. Belohnungen explizit ausgesprochen werden oder implizite Erwartungen bewusst wahrgenommen werden (Scott, 1971). Vielmehr wird die Kontrollinstanz dabei ebenfalls verinnerlicht, sodass die Person diese Rolle selbst übernimmt. In anderen Fällen wiederum werden Normen und Erwartungen explizit an die Mitglieder herangetragen, indem die Gruppe aktiv einen gewissen Druck zur Konformität durch Androhung von Sanktionen oder Versprechen von Belohnung ausübt (Forgas & Williams, 2001).

Die komparative Funktion von Bezugsgruppen besteht darin, dass Vergleiche der eigenen Merkmale anhand des Maßstabs von Gruppennormen der Validierung der eigenen Merkmale

³ Die Bezugsgruppentheorie nach Kelley bezieht sich sowohl auf Gruppen, zu denen eine Person zugehörig gilt, als auch auf solche, denen eine Person gern angehören möchte (Kelley, 1968). Zudem kann zwischen *In- und Outgroups* unterschieden werden (Tajfel, 1982), wobei die Gruppe, zu der eine Person gehört oder angehören möchte (*Ingroup*) von der Gruppe unterschieden wird, zu der eine Person sich nicht zugehörig fühlt oder nicht gehören möchte (*Outgroup*). Auch mit Personen von *Outgroups* finden Vergleiche statt, wobei vor allem Differenzen im Vordergrund stehen (Mussweiler & Bodenhausen, 2002). Da es in der vorliegenden Arbeit um den Einfluss der Mitschülerinnen und Mitschüler geht, werden nur Intragruppenprozesse betrachtet. Funktionen von *Outgroups* sollen hier nicht behandelt werden.

dienen und so auch Aufschlüsse über die eigene soziale Position in der Gruppe oder über die eigene Situation geben. Unabhängig von der Theorie zu Funktionen von Bezugsgruppen, entwickelte sich nahezu zeitgleich die *Theorie sozialer Vergleichsprozesse* (Festinger, 1954), die sich inhaltlich stark mit dem Aspekt der komparativen Funktion von Gruppen überschneidet. Obgleich sie bisher nicht integriert wurden (Fischer & Wiswede, 2009), liefert die Theorie sozialer Vergleichsprozesse weitergehende Annahmen über die komparative Funktion von Gruppen, die in dieser Tiefe nicht in der Bezugsgruppentheorie behandelt wurden. Der Grundgedanke der Theorie sozialer Vergleichsprozesse liegt darin, dass Menschen eine Basismotivation nach sozialen Vergleichen haben um ihre eigenen Meinungen und Fähigkeiten zu validieren (*Validierungsbedürfnis*; Festinger, 1954). Frey, Dauenheimer, Parge und Haisch (1993) erweiterten diesen Grundgedanken dahingehend, dass soziale Vergleiche für jegliche Zustände und Eigenschaften von Menschen relevant sein können. Hierarchisch dem Validierungsbedürfnis nachgeordnet, nimmt Festinger (1954) drei weitere Bedürfnisse an, die durch soziale Vergleiche ihre Befriedigung erhalten sollen. So dienen Vergleichsprozesse dem *Bedürfnis zur Unsicherheitsreduktion* in der Beurteilung von Meinungen, Fähigkeiten bzw. generell Merkmalen anderer Personen und der eigenen Person. Weiterhin können Vergleiche dem *Bedürfnis nach Selbstverbesserung* dienlich sein, wenn durch sogenannte Aufwärtsvergleiche mit Personen, die in einem bestimmten Merkmal besser eingeschätzt werden, die Motivation erhöht wird, die eigene Ausprägung in dem jeweiligen Merkmal zu steigern. Ein drittes Bedürfnis, welches über soziale Vergleiche bedient wird, bezeichnet Festinger als *Bedürfnis nach Demonstration*. Durch Abwärtsvergleiche oder die Absicherung, dass die eigene Merkmalsausprägung der Gruppennorm entspricht, wird das eigene Selbstbild erhöht bzw. gesichert. Vergleiche finden dabei vornehmlich mit Personen statt, die ähnliche relevante Attribute aufweisen (Goethals & Darley, 1977), d.h. die in bestimmten Merkmalen, die für das entscheidende Personenmerkmal für den Vergleichenden von kausaler Relevanz sind, ähnliche Ausprägungen aufweisen.

In Studien zu Effekten der Komposition der Schülerschaft auf die individuelle schulische Entwicklung wurde die normative und komparative Funktion von Bezugsgruppen bisher vor allem hinsichtlich leistungsrelevanter Schülermerkmale thematisiert (Baumert et al., 2006; Marsh, 2005; Zurbriggen, 2016). Es wird angenommen, dass je nach Komposition der Lerngruppe eine bestimmte normative Kultur hinsichtlich leistungsrelevanter Einstellungen und Verhaltensweisen besteht, die die individuelle Leistungsentwicklung bedingt (Baumert et al.,

2006; Zurbriggen, 2016). Als Beispiel für einen komparativen Einfluss durch die Lerngruppe ist das *referenzgruppentheoretische Modell* von Marsh (1987) zu nennen. Demnach gelangen Lernende zu Einschätzungen ihrer akademischen Kompetenzen, indem sie diese mit den jeweiligen Kompetenzen der anderen Schülerinnen und Schüler in der Lerngruppe vergleichen. Durch die Tendenz zu Aufwärtsvergleichen bei sozialen Vergleichen der Leistung (Festinger, 1954), fällt die Einschätzung der eigenen Leistungskompetenzen bzw. das Niveau der akademischen Selbstkonzepte in leistungsstarken Lerngruppen ungünstiger aus als in leistungsschwachen Lerngruppen.

1.3.2 Das entwicklungspsychologische Konzept der Peergruppeneffekte

Werden Einflüsse der Schülerschaft aus entwicklungspsychologischer Perspektive betrachtet, so können Theorien zu *Peergruppeneffekten* hinzugezogen werden (Berndt & Ladd, 1989; Breidenstein, 2008). Vornehmlich werden Freundeskreise von Gleichaltrigen mit dem Begriff der *Peergruppe* bzw. *Peergroup* verbunden. Im engeren Sinn geht es jedoch um Personen, die ein alltägliches Umfeld teilen und so eine gewisse Bezugsgruppe bilden, an der sich Personen orientieren und sich in Bezug setzen (Berndt & Ladd, 1989; Breidenstein, 2008). Dabei können *informelle*, grundsätzlich freiwillig gewählte und gepflegte Peergruppen von *formalen* bzw. unfreiwilligen Peergruppen wie die Lerngruppe der Klasse oder Schule differenziert werden. Formale Peerbeziehungen zeichnen sich dadurch aus, dass sie nicht von den Jugendlichen selbst gesteuert werden (Brown, 2013). So werden Schülerschaften einer Schule oder Klasse als formale Peergruppen bezeichnet, da sie durch die übergeordneten Allokationsprozesse zu einer Bezugsgruppe zusammengeschlossen wurden, deren Alltag von Lehrerinnen und Lehrern begleitet und vom Schulstundenplan gesteuert wird (Brown, 2013; Juvonen & Galván, 2008; Ulich, 2001).

Aus einer entwicklungspsychologischen Perspektive sind gegenseitige Einflüsse Gleichaltriger besonders im frühen Jugendalter und somit auch für den Schulkontext zentral. Im Laufe der Kindheit steigt die Beeinflussbarkeit durch Peergruppen bis sie in der Pubertät – etwa im Alter zwischen 11 und 16 Jahren – einen Höhepunkt erreicht (siehe im Überblick Müller & Minger 2013). Während die Eltern in der Kindheit meist die wichtigste Bezugsquelle darstellen, wird im Jugendalter nach größerer Unabhängigkeit von den Eltern bei gleichzeitig verstärkter Orientierung an Gleichaltrigen gestrebt (Burlingame, 1970; Steinberg & Silverberg, 1986). Diese Umorientierung entspricht der *Life Course Theory* (Giele & Elder 1998; Kellam & van Horn,

1997), die hervorhebt, dass in jeder Altersphase bestimmte soziale Anforderungen im Vordergrund stehen, die zu bewältigen sind. Durch strukturelle Veränderungen in den Lebensphasen wie z. B. der Eintritt in die Sekundarstufe, wechselt auch die soziale Umwelt, die in den jeweiligen Phasen im Fokus steht und die Einübung neuer Rollen fordert. Eine der zentralen Entwicklungsaufgaben der Pubertät kann in der Identitätsbildung gesehen werden (Erikson, 1968). Dabei wird angenommen, dass die eigene Identität geformt wird, in dem sich Jugendliche an Verhaltensweisen, sozialen Kompetenzen sowie Einstellungen orientieren, die als geteilte Normen in der Peergruppe aufgefasst werden oder die von allgemein beliebten Peers gezeigt werden (Brown, Mory & Kinney, 1994; Miller & Prentice, 1994). Verhaltensweisen, soziale Kompetenzen und Einstellungen, die unbeliebte Peers aufweisen, werden dagegen gemieden, um die eigene Identität von diesen zu distanzieren (Prinstein & Dodge, 2008). In Rückgriff auf die sozialpsychologische Annahme, dass Menschen zumeist eine positiv angesehene Identität anstreben (Tajfel, 1984), kann das Angleichen an die wahrgenommenen sozialen Normen durch Jugendliche begründet werden (Stone & Brown, 1999).

Brown, Bakken, Ameringer und Mahon (2008) vermuten fünf Formen von Peereinflüssen im Jugendalter: *behavioral display*, *behavioral reinforcement*, *peer pressure*, *antagonistic behaviors* und *structuring opportunities*. Bei den ersten drei Formen von Peereinflüssen werden die sozialpsychologischen Theorien zur sozialen Verstärkung, zum Lernen am Modell sowie zum Gruppendruck (zur Erläuterung siehe Kapitel 1.3.1) auf Peergruppen von Jugendlichen übertragen. *Behavioral display* beschreibt, dass Jugendliche die Verhaltensweisen von Peers sowie die darauffolgenden Konsequenzen beobachten und diesen Zusammenhang kognitiv verarbeiten. Ergeben sich positive Folgen aus einem Verhalten, z. B. in dem die beobachtete Person dadurch Anerkennung durch die Peergruppe erfährt, so wird versucht, das beobachtete Verhalten selbst zu reproduzieren, um die erwünschten Folgen herbeizuführen. Der Begriff des *behavioral reinforcements* basiert auf der Theorie der sozialen Verstärkung durch die Peergruppe (Bandura 1979, 1986). Wird gezeigtes Verhalten durch positive verbale Reaktionen wie Zuspruch oder nonverbale Reaktionen wie Lächeln durch Peers belohnt, regt dies wiederum zum erneuten Zeigen des Verhaltens an. Ist ein Verhalten konstituierend für eine bestimmte Peergruppe, so ergeben sich Gruppennormen, die sich diesen anzupassen um dazuzugehören (*peer pressure*). Bei gruppennormwidrigen Verhaltensweisen kann es zu den sogenannten *antagonistic behaviors* durch die Peers kommen, die von Sticheleien bis zu aggressiven Formen wie *Bullying* reichen können. Der Mechanismus *structuring situations* meint das Herbeiführen von Situationen, in

denen bestimmte Verhaltensweisen, die durch Erwachsene verboten sind, erst ermöglicht werden. Wird ein Rahmen ohne Aufsicht durch Erwachsene geschaffen, kann auch ohne direkte Aufforderung der Peers das Ausprobieren verbotener Verhaltensweisen erleichtert werden.

Wie stark Jugendliche von ihren Peers beeinflusst werden, hängt entwicklungspsychologisch gesehen von multiplen Faktoren wie Merkmalen der Beeinflussten selbst, Merkmalen der Peergruppe, der Qualität der Beziehungen zwischen Peers sowie der Salienz des Merkmals ab (Dishion & Tipsord, 2011; Prinstein, 2007). Es wird angenommen, dass besonders Jugendliche mit geringem Selbstbewusstsein, hoher sozialer Angst und wenig stabilen Freundschaften anfällig für Peereinflüsse sind (Adams, Bukowski & Bagwell 2005; Bukowski, Velasquez & Brendgen, 2010). Lernen am Modell dürfte sich stärker vollziehen, wenn die Person aufmerksam beobachtet, beim Nachahmen genügend Erinnerungsvermögen aufweist und die Kompetenz besitzt, das Gesehene auch nachahmen zu können (Brown, 1989). Auf der Seite der Peergruppe, die die Norm in gewisser Weise vorgibt, spielen vor allem Eigenschaften von besonders prominenten Peers eine Rolle. Diese umfassen einen hohen sozialen Status in der Gruppe, Beliebtheit, Intelligenz, hohe soziale Kompetenz und Attraktivität, aber auch besonders dominantes Auftreten oder antisoziales Verhalten (Allan & Antonishak, 2008; Brown, 1989; Cillessen & Rose, 2005; LaFontana & Cillessen, 2002). Die Qualität der Peerbeziehung spielt vor allem hinsichtlich der Kohäsion zwischen den Peers eine Rolle (Minsel & Roth, 1978). Eine stärkere Identifikation mit der Peergruppe oder mit einem bestimmten Peer, z. B. aufgrund des gleichen Geschlechts, der gleichen Herkunft oder einfach durch eine vertrautere Beziehung im Lauf der Zeit, führt dazu, diese eher als normativ richtungsweisend anzusehen (Brown, 1989; Bukowski et al., 2010; Prinstein, 2007). Saliente Merkmale, also solche, die stärker wahrnehmbar sind, werden dabei eher als normativ leitend angesehen (Noack, 2011).

Aufgrund der verschiedenen Einflüsse zwischen jugendlichen Peers wird eine mehr oder minder unbewusst stattfindende Übertragung („Ansteckung“) von Verhalten und Einstellungen zwischen den Jugendlichen angenommen, die als *Peer-Contagion-Effekt* bezeichnet wird (Dishion & Tipsord, 2011). Peer-Contagion-Effekte wurden bisher fast ausschließlich für informelle Peergruppen wie dyadische Freundschaften und Freundeskreise thematisiert (siehe im Überblick Dishion & Tipsord, 2011), nur selten für den schulischen Kontext (siehe auch Müller, Hofmann, Fleischli & Studer, 2015). Dabei sind diese gerade für die Sekundarstufe denkbar, wenn in der Peergruppe der Mitschülerinnen und Mitschüler Lernende zusammentreffen, die sich im Stadium der Pubertät befinden und somit geistige Reifeprozesse durchlaufen, die mit der

Ausbildung der Identität einhergehen (Brown, 1990). Mit dem Beginn der Pubertät, der in vielen Ländern mit dem Übergang von der Elementarstufe in die Sekundarstufe zusammenfällt, sind viele Änderungen in den Sozialstrukturen verbunden (Brown, 2013; Wigfield & Eccles, 1990). So verlassen Lernende beim Übergang von der Elementarschule in die Sekundarstufe auch ihr bekanntes Klassengefüge, in dem sie über Jahre gemeinsam gelernt haben, sowie Freunde in der Elementarschulklasse, die nicht an die gleiche Sekundarschule überwechseln. Weiterhin wird das feste Klassengefüge durch Wahlkurse oder within-school-tracking von wechselnden Klassengefügen abgelöst. Dies alles führt zu neuen Herausforderungen für die Lernenden, zu deren Bewältigung die Orientierung an Gleichaltrigen dienen kann (Brown, Mory & Kinney, 1994; Miller & Prentice, 1994).

1.3.3 Das Konzept des Schul- und Klassenklimas in der Schul- und Unterrichtsforschung

Weitere Hinweise für den Einfluss der Schülerschaft auf die Lernumwelt Schule können aus dem Konzept des *Schul- und Klassenklimas* gezogen werden, das in der Schul- und Unterrichtsforschung verankert ist. Das Schul- bzw. Klassenklima wird als die von den „Betroffenen wahrgenommene Konfiguration bedeutsamer Merkmale innerhalb der jeweiligen schulischen Umwelt“ (Eder, 1998, S. 424) definiert. Es wird angenommen, dass dabei nicht nur Merkmale der Schulorganisation und der Unterrichtspraktiken relevant sind, sondern auch Merkmale, die die Gemeinschaft der Lernenden charakterisiert (Libbey 2004). Dies betrifft geteilte Normen, Überzeugungen und Einstellungen sowie typische Merkmale der Beziehungen und Interaktionen zwischen Lernenden (siehe im Überblick Cohen, McCabe, Michelli & Pickeral, 2009; sowie Dreesmann et al., 1992; Eder, 2002; Ellis, 1988; Pekrun, 1985). Geteilte Normen, Überzeugungen und Einstellungen innerhalb der Schülerschaft können sowohl explizit zwischen den Lernenden vereinbart werden als auch in Form von „ungeschriebenen Gesetzen“ bestehen (Stewart, 2003). Als Beziehungs- und Interaktionsmerkmale zwischen Lernenden, die das Schulklima konstituieren, werden unter anderem die Qualität des sozialen Zusammenhalts und der sozialen Unterstützung, Spannungen zwischen den Lernenden, typische Kommunikationsmuster, Konkurrenz zwischen den Lernenden sowie das Vorhandensein von Gewalt gesehen (Batanova & Loukas, 2016; Berkowitz et al., 2015; Cohen et al., 2009; Eder, 2002; Götz, Frenzel & Pekrun, 2008; Stewart, 2003; Thapa, Cohen, Guffey & Higgins-D’Alessandro, 2013).

Einige Forschende spezifizieren den Begriff des Schulklimas entlang der Ebenen von Schule weiter aus. Götz et al. (2008) sowie Eder (2002) unterscheiden dabei das Schulklima vom Klassen- und Unterrichtsklima. Auf den verschiedenen Ebenen kann das Zusammenspiel der Merkmale der Schülerschaft mit den Merkmalen der Schulorganisation und der Unterrichtspraktiken zu einem je spezifischen Klima führen. Es wird dabei von Wechselwirkungen zwischen Schul-, Klassen- und Unterrichtsklima ausgegangen. Andere Forschende heben zudem das soziale Miteinander hervor und nutzen den Begriff des Sozialklimas (Fend, 1977; Götz et al., 2008; Maschke & Stecher, 2010; von Salderm & Littig 1987). Nach Götz et al. (2008) formiert sich dieses aus interpersonalen Aspekten wie typische Umgangsformen der Höflichkeit oder der Streitkultur in der Klasse oder Schule, aus Verhaltensaspekten wie das Vorkommen von aggressivem Handeln, aus motivationalen Aspekten wie Lerninteresse, aus kognitiven Aspekten wie fachspezifische Wertüberzeugungen sowie aus emotionalen Aspekten wie dem Testangstniveau der Schülerschaft.

Die Relevanz des Schulklimas wird in der Schul- und Unterrichtsforschung damit begründet, dass dem Schulklima ein Einfluss auf die Entwicklung der Leistung und psychosozialer Merkmale der Lernenden zukommt. So wird angenommen, dass ein positives Schulklima förderlich auf psychosoziale Merkmale wie Selbstwert, allgemeines Wohlbefinden, Lernmotivation, Lernzufriedenheit und Lernengagement wirkt und Aggression, Gewaltverhalten und Schulabsentismus mildert (siehe im Überblick Cohen et al., 2009; sowie Eder, 2002; Götz et al., 2008; Thapa et al., 2013). Auch die Leistungen der Schülerinnen und Schüler könnten durch ein positives Klassen- oder Schulklima gesteigert werden (siehe im Überblick Cohen et al., 2009; Eder, 2002; Götz et al., 2008; Thapa et al., 2013). Aufgrund der Vielfalt an Konzeptualisierungen von Schulklima, d.h. welche Merkmale konstituierend für das Schulklima angesehen werden und aufgrund der dadurch entstehenden Vielzahl an eingesetzten Instrumenten, liegt bisher keine einheitliche empirische Befundlage vor für den spezifischen Einfluss der Schülerschaft, die das Schulklima mitkonstituiert (siehe auch Götz et al., 2008).

1.4 Empirische Befunde zum Einfluss der Schülerzusammensetzung auf die individuelle schulische Entwicklung

Im vorherigen Kapitel wurden theoretische Annahmen zusammengetragen, inwiefern Merkmale der Schülerschaft die individuelle schulische Entwicklung mitbedingen könnten. Im folgenden Kapitel soll nun die empirische Befundlage zu Einflüssen der

Schülerzusammensetzung auf die individuelle Entwicklung der Leistung und psychosozialer Merkmale aufgezeigt werden. Dazu werden einleitend methodische Aspekte der Modellierung von solchen Kompositionseffekten angeführt.

1.4.1 Empirische Modellierung von Effekten der Schülerzusammensetzung auf die individuelle schulische Entwicklung

Um Effekte der Schülerzusammensetzung auf die individuelle schulische Entwicklung empirisch zu untersuchen, ist es üblich, den Effekt eines auf Schul- oder Klassenebene aggregierten Schülermerkmals unter Kontrolle desselbigen Merkmals auf der Schülerebene auf ein weiteres oder dasselbe Schülermerkmal zum späteren Zeitpunkt zu untersuchen (Harker & Tymms, 2004; siehe Abb. 1.2). Meist wird der Aggregatwert dabei als *mittlere Zusammensetzung* operationalisiert, indem der arithmetische Mittelwert oder der prozentuale Anteil von Lernenden mit einer bestimmten Merkmalsausprägung generiert wird. Der Aggregatwert wird teilweise aber auch als Standardabweichung oder Varianz eines Schülermerkmals innerhalb einer Lerngruppe gebildet um die Komposition einer Schülerschaft unter dem Aspekt der *Verteilung* von Schülermerkmalen innerhalb einer Lerngruppe zu betrachten (siehe Dumont et al., 2013).

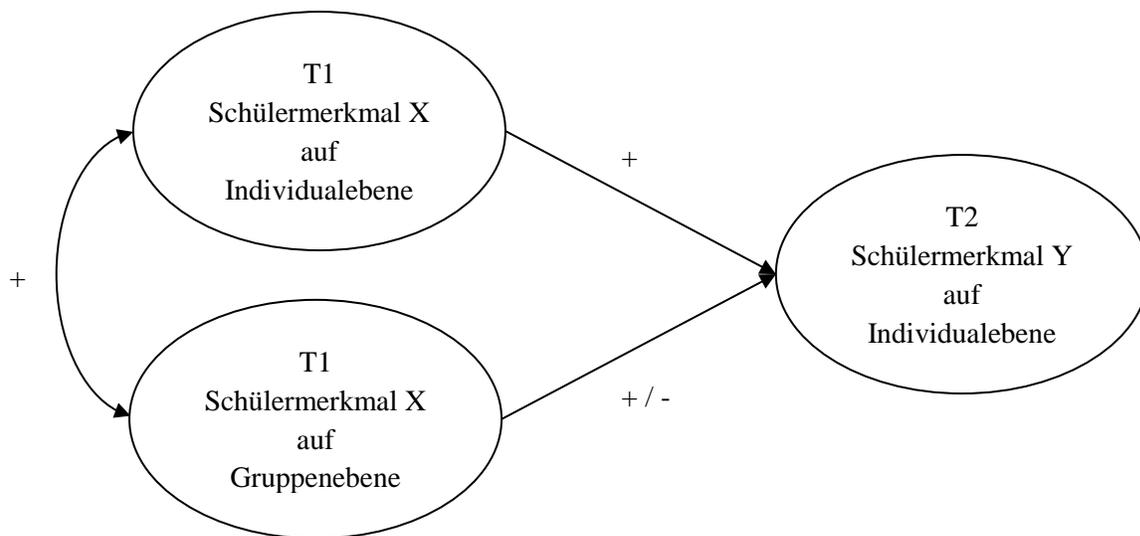


Abbildung 1.2. Klassische pfadanalytische Darstellung des leistungsbezogenen Kompositionseffekts.

Zur Prüfung von Kompositionseffekten muss mindestens ein Zwei-Ebenen-Modell spezifiziert werden, dessen Ebene 1-Gleichung für die vorherzusagende individuelle Ausprägung eines Schülermerkmals y_{ij} wie folgt lautet:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij} \quad (1)$$

Hierbei repräsentiert X_{ij} die Ausprägung eines Merkmals der Schülerin oder des Schülers i in Gruppe j (Klasse oder Schule j) und r_{ij} das individuelle Residuum. Der Achsenabschnittsparameter β_{0j} stellt den aggregierten Wert für Gruppe j dar, korrigiert um den X -Wert auf Individualebene. β_{1j} ist der Regressionskoeffizienten für Gruppe j und stellt den Effekt von X auf Y dar.

Da β_{0j} und β_{1j} zwischen den Gruppen variieren können, werden sie auf Gruppenebene modelliert, sodass sich folgende Gleichungen für diese Parameter auf Gruppenebene ergeben:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\bar{X}_{1j} + u_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (3)$$

In diesem Gleichungsbeispiel wird die mittlere Zusammensetzung der Schülerschaft hinsichtlich eines Merkmals zur Veranschaulichung gewählt, sodass γ_{00} den Gesamtmittelwert des Schülermerkmals Y über alle Gruppen hinweg darstellt, \bar{X}_{1j} den Mittelwert von X in Klasse j und u_{0j} das gruppenspezifische Residuum. Bei Zentrierung von X am Mittelwert der Gesamtstichprobe (*Grandmean-Zentrierung*; Enders & Tofighi, 2007) repräsentiert γ_{01} den Effekt des Gruppenmittelwerts \bar{X} auf Y und stellt somit den Schätzer des Kompositionseffekts dar. Der Regressionskoeffizient β_{1j} stellt den Effekt innerhalb der Gruppen dar, wobei dieser zwischen den Klassen mit einem Fehlerterm u_{1j} variieren kann. Durch Einsetzen der Gleichungen (2) und (3) in Gleichung (1) ergibt sich für die Schätzung der individuellen Ausprägung eines Schülermerkmals Y_{ij} folgende Gleichung:

$$Y_{ij} = \gamma_{00} + \gamma_{01}\bar{X}_{1j} + \gamma_{10}X_{ij} + u_{1j}X_{ij} + r_{ij} + u_{0j} \quad (4)$$

Zur Schätzung solcher Kompositionseffekte werden Mehrebenenanalysen im Längsschnitt vorausgesetzt (Baumert et al., 2006; Dumont et al., 2013; Thrupp, Lauder & Robinson, 2002).

Die Bedingung der Mehrebenenmodellierung wird immanent durch die Konzeption von Kompositionseffekten als Einfluss eines Schülermerkmals in einer Gruppe über das individuelle Merkmal hinaus vorgegeben, sodass die Ebene des Individuums und die Ebene der Klasse oder Schule differenziert werden muss. Als geeignete Aggregatebene für die Bestimmung der Komposition von Lerngruppen wird die Klasse empfohlen, in der die unmittelbaren Interaktionen mit den Mitschülerinnen und Mitschülern und mit Lehrkräften stattfinden (van Ewijk & Sleegers, 2010a). Hinweise dazu ergaben sich in einer Metaanalyse von van Ewijk & Sleegers (2010a), da Kompositionseffekte auf Klassenebene größer ausfielen als jene auf Schulebene. Zu beachten ist, dass hierbei die mittlere sozioökonomische Komposition betrachtet wurde, sodass dies nicht auf alle Kompositionsmerkmale generalisiert werden kann. In einer Studie zur Komposition hinsichtlich des mittleren aggressiven Verhaltens zeigte sich beispielsweise, dass das mittlere Aggressionsniveau der Schule einen verstärkenden Effekt für das individuelle Aggressionsniveau hatte, vom Aggressionsniveau der Klasse ging dagegen ein geringerer Effekt aus (Gommans, Stevens, ter Bogt & Cillessen, 2016).

Die Analyse von Kompositionseffekten im Längsschnitt ist notwendig, um für die Kriteriumsvariable zum vorherigen Zeitpunkt kontrollieren zu können (Baumert et al., 2006; Dumont et al., 2013; Thrupp et al., 2002). Da das individuelle vorherige Niveau des entwicklungsrelevanten Merkmals den stärksten Prädiktor darstellt, kann es zur Überschätzung der Kompositionseffekte kommen, wenn es nicht kontrolliert wird. Wird beispielsweise bei der Analyse des Effekts der mittleren sozioökonomischen Komposition auf die individuelle Leistung die vorherige individuelle Leistung nicht berücksichtigt, wie es bei früheren Querschnittsanalysen zu Kompositionseffekten häufig der Fall war, wird vom *Endogenitätsproblem* oder *Selection Bias* gesprochen (Baumert et al., 2006; sowie Duncan, Magnuson & Ludwig, 2004; Heckman, 1979). Die individuelle Leistung als nicht spezifizierter Störfaktor ist durch die nicht zufällige Allokation von Lernenden auf die verschiedenen Schulen und Schulformen sowohl mit der erklärenden Variablen des sozioökonomischen Hintergrunds korreliert als auch prädiktiv für die abhängige Variable Leistung. Zu beachten ist allerdings, dass die Spezifizierung zu vieler Kontrollvariablen, die mit der abhängigen Variable kovariieren, wiederum zur Überkontrolle und damit zum sogenannten *mediation bias* führen kann (Baumert et al., 2006).

In Schulsystemen mit expliziter Schulformgliederung wie in Deutschland ist zur Vermeidung des Endogenitätsproblems die zusätzliche Kontrolle der Schulform erforderlich, sodass Kompositionseffekte und institutionelle Effekte weitestgehend differenziert werden können

(Baumert et al., 2006). Einschränkend muss gesagt werden, dass eine komplette Differenzierung von Kompositionseffekten und institutionellen Effekten jedoch nicht realisierbar ist. So zeigten Baumert et al. (2006), dass für die Lesekompetenz jeweils spezifische Varianzanteile von der mittleren Leistungskomposition der Schule und von der Schulform aufgeklärt werden konnten, der größte Anteil der Varianz jedoch auf eine konfundierte Komponente aus Schulform, mittlerer Leistungsfähigkeit der Schule sowie dem Anteil von Lernenden mit Eltern ohne Berufsausbildung zurückging. Aus diesem Ergebnis geht auch die weitere methodische Empfehlung hervor, zur Prävention von Überschätzungen verschiedene Kompositionsmerkmale simultan zu schätzen. Wie sich auch in dem Beispiel zur Lesekompetenz bei Baumert et al. (2006) zeigte, sind Leistungskomposition, soziale und ethnische Komposition meist stark miteinander konfundiert (Baumert et al., 2006; Dumont et al., 2013; Thrupp et al., 2002).

1.4.2 Effekte der Schülerzusammensetzung auf die individuelle Leistungsentwicklung

1.4.2.1 Effekte durch die Komposition der Leistung

Untersuchungen zum Einfluss der leistungsbezogenen Komposition einer Klasse oder Schule kommen weitgehend zu dem Ergebnis, dass sich die Zugehörigkeit zu einer leistungsstarken Lerngruppe förderlich auf die individuelle Leistungsentwicklung auswirkt, während der Besuch einer Lerngruppe mit niedriger mittlerer Leistung ungünstig für die eigene Leistungsentwicklung ist (siehe im Überblick Dumont et al., 2013; für empirische Studien siehe Danhier & Martin, 2014; Hanushek et al., 2003; Justice et al., 2014; Marks, 2010; Opdenakker et al., 2002). Pionierarbeiten für Kompositionseffekte im deutschen Schulsystem wurden von der Arbeitsgruppe um Jürgen Baumert geleistet (Baumert & Schümer, 2001; Baumert et al. 2006; Baumert, Trautwein & Artelt, 2003). In ihren Untersuchungen anhand von PISA-Daten stellte sich die mittlere Leistung der Schule als das bedeutendste Kompositionsmerkmal für die individuelle Leistungsentwicklung heraus. Diese im Querschnitt gefundenen Effekte zeigten sich in darauffolgenden Studien auch im Längsschnitt (Bellin, 2009; Köller, Schütte, Zimmermann, Retelsdorf & Leucht, 2013; Nikolova, 2011; Gröhlich, Guill, Scharenberg & Bos, 2010; Stanat, Schwippert & Gröhlich, 2010). Obgleich die mittlere Leistung der Schule oder Klasse in Deutschland durch die auf Schülerleistungen basierende Schulformgliederung stark mit der Schulform konfundiert, zeigten sich diese Effekte trotz Kontrolle der Schulform. Die Studien umfassen Kompositionseffekte auf Klassen- und Schulebene sowie den Elementar- und Sekundarschulbereich.

Studien, die die Leistungszusammensetzung der Klasse oder Schule über die Varianz oder Standardabweichung eines Merkmals untersuchen, kommen dagegen bisher zu uneinheitlichen Befunden. Teilweise fanden sich positive Effekte im Sinne einer höheren individuellen Leistungsentwicklung in leistungsheterogenen *Klassen*, z. B. für das Leseverständnis sowie für mathematische Kompetenzen in Elementarschulen (Lehmann, 2006; Rjosk, Richter, Lüdtke & Eccles, 2017 für mathematische Kompetenzen; Scharenberg, 2012). In anderen Untersuchungen zeigten sich wiederum negative (für die Elementarstufe: Künsting, Post, Greb, Faust & Lipowsky, 2010 für die 2. Jahrgangsstufe; für die Sekundarstufe: Ding & Lehrer, 2007) oder nicht signifikante Effekte (für die Elementarstufe: Bos & Scharenberg, 2010; Künsting et al., 2010 für die 1. Jahrgangsstufe; Leiter, 1983; Rjosk et al., 2017 für die Lesekompetenz; für die Sekundarstufe: Kiss, 2013). Bei Hanushek et al. (2003), die die Varianz der Leistung auf *Schulebene* untersucht haben, zeigte sich kein signifikanter Effekt auf die Leistungsentwicklung der Elementarschülerinnen und -schüler.

1.4.2.2 Effekte durch die Komposition des familiären Hintergrunds

Als Hintergrundmerkmale der Schülerschaft spielen in der Forschung zu Kompositionseffekten vor allem der Sozial- und Bildungsstatus der Eltern als auch der Zuwanderungshintergrund und die Zugehörigkeit zu ethnischen Minderheiten eine Rolle (Dumont et al., 2013). Bezüglich der Komposition des mittleren sozioökonomischen Status einer Klasse oder Schule der Elementar- und Sekundarstufe kommen eine Vielzahl internationaler Studien zum Ergebnis eines positiven Effekts auf die individuelle Leistungsentwicklung (siehe Metaanalyse von van Ewijk & Slegers, 2010a; siehe im Überblick Dumont et al., 2013; für empirische Studien siehe Benson & Borman, 2010; Hornstra et al., 2015; Lauder, Kounali, Robinson & Goldstein, 2010; Peetsma, van der Veen, Koopman & van Schooten, 2006; Rangel & Lleras, 2010). Demnach entwickelt sich die Leistung von Schülerinnen und Schülern in Klassen und Schulen mit einer sozial begünstigten Schülerschaft besser als in jenen mit einer sozial benachteiligten Schülerschaft. Bei Kontrolle des individuellen Leistungs- und Fähigkeitsniveaus und dem mittleren Leistungs- und Fähigkeitsniveau der Lerngruppe fallen diese Kompositionseffekte zumeist geringer aus (siehe im Überblick Dumont et al., 2013).

Für die Komposition der Schülerschaft in Klassen und Schulen der Elementar- und Sekundarstufe hinsichtlich ihres Anteils an Lernenden mit Zuwanderungshintergrund oder aus einer ethnischen Minderheit wurden in internationalen Studien kleine negative Effekte auf die

individuelle Leistungsentwicklung gefunden (siehe Metaanalyse von van Ewijk & Slegers, 2010b; für empirische Studien siehe Benson & Borman, 2010; Goldsmith, 2011; Hornstra et al., 2015; Mickelson, Bottia & Lambert, 2013). Demnach ist es ungünstig für die individuelle Lernentwicklung, in einer Klasse unterrichtet zu werden, die einen erhöhten Anteil von Schülerinnen und Schülern aufweist, die selbst oder deren Eltern zugewandert sind oder die einer ethnischen Minderheit angehören. Bei Kontrolle der sozialen Zusammensetzung fand sich dies in einigen Studien nicht mehr (Rumberger & Palardy, 2005; van der Slik, F, Driessen & Bot, 2006). Insbesondere in deutschen Studien fand sich über die soziale und leistungsbezogene Zusammensetzung hinaus häufig kein eigenständiger Effekt für den ethnischen Hintergrund (Bellin, 2009; Ditton & Krüsken, 2006; Gröhlich et al., 2010; Stanat, 2006; Stanat et al., 2010; Walter, 2008). In einer Studie an Elementarschulen, in der die *Verteilung* der ethnischen Herkunft innerhalb der Schule über den Simpson's *index of diversity* (1949) operationalisiert wurde, zeigte sich dagegen ein positiver Effekt der Heterogenität in der Herkunft auf Mathematik- und Lesekompetenzen (Benner & Crosnoe, 2011).

1.4.2.3 Effekte durch die Komposition psychosozialer Merkmale

Zum Einfluss der mittleren Zusammensetzung der Schülerschaft hinsichtlich bestimmter psychosozialer Merkmale auf die individuelle Leistungsentwicklung liegen bisher kaum Studien vor. Eine der wenigen Befunde kommt aus einer Sekundarschul-Studie, die positive Einflüsse der mittleren Leistungsmotivation der Schule auf die Sprachkompetenzen in der Landessprache fanden, nicht aber für Mathematikkompetenzen (van Landeghem, Pustjens, van Damme & Munter, 2007). In einer Studie an Elementarschulen wurde die durchschnittliche Unpünktlichkeit als psychosoziales Merkmal untersucht. Die mittlere Rate der Schülerinnen und Schüler in der Klasse an verspätetem Erscheinen zum Unterricht wirkte sich negativ auf die individuellen Lese- und Mathematikkompetenzen aus (Gottfried, 2014).

1.4.3 Effekte der Schülerzusammensetzung auf die individuelle Entwicklung psychosozialer Merkmale

1.4.3.1 Effekte durch die Komposition der Leistung

Oft nicht als Kompositionseffekt bezeichnet, aber methodisch und inhaltlich als Kompositionseffekt einzuordnen, ist der Big-Fish-Little-Pond-Effekt als prominentester Effekt der Komposition auf ein psychosoziales Merkmal zu nennen (Marsh, 1987; siehe im Überblick Marsh, Seaton et al., 2008; für empirische Studien siehe Marsh et al., 2014; Marsh, Kong & Hau,

2000; Marsh & O'Mara, 2010; Marsh, Trautwein, Lüdtke, Baumert & Köller, 2007; Nagengast & Marsh, 2011; Seaton, Marsh & Craven, 2010). Wie häufig repliziert wurde wirkt sich das mittlere Leistungsniveau einer Klasse oder Schule negativ auf die Einschätzung des individuellen akademischen Selbstkonzepts aus. Von zwei Lernenden mit gleichen Leistungen weist dementsprechend derjenige, der eine leistungsschwache Klasse besucht, ein höheres akademisches Selbstkonzept als derjenige in einer leistungsstarken Klasse auf. Dies wurde über die verschiedenen Schulstufen hinweg gefunden (siehe im Überblick Marsh, 2005). Der BFLPE wurde zumeist in Bezug zum akademischen Selbstkonzept betrachtet. Darüber hinaus wurden in einzelnen Studien negative Effekte der mittleren Schulleistung auf das Selbstvertrauen und die Kontrollerwartung von Sekundarschülerinnen und -schülern (Marsh, Trautwein, Lüdtke & Köller, 2008) sowie auf das fachspezifische Interesse von Schülerinnen und Schülern der Sekundarstufe II gefunden (Köller, Baumert & Schnabel, 2000). Für die mittlere Klassenleistung ergaben sich zudem negative Effekte auf die Übergangsempfehlung von Elementarschülerinnen und -schülern (Gröhlich & Guill, 2009; Tiedemann & Billmann-Mahecha, 2007; Trautwein & Baeriswyl, 2007), dem tatsächlichen erfolgten Übergang (Trautwein & Baeriswyl, 2007) sowie auf Berufsaspirationen von Sekundarschülerinnen und -schülern (Zurbriggen, 2016). Gottfried (2015) operationalisierte die Leistungskomposition über den prozentualen Anteil von Lernenden unter dem Klassenniveau der Mathematikkompetenz. Er fand einen positiven Effekt auf externalisierendes und internalisierendes Verhalten sowie negative Effekte auf Selbstkontrolle, lerntheoretische Fähigkeiten (*approaches to learning*) und interpersonale Skills. Ein hohes mittleres Leistungsniveau der Klasse kann allerdings auch positiv auf die individuelle psychosoziale Entwicklung wirken. So fanden van Landeghem, van Damme, Opdenakker, de Fraine und Onghena (2002), dass die mittlere Leistung der Klasse einen positiven Effekt bei Sekundarschülerinnen und -schülern auf das Gefühl ausübte, sich in der Schule zu Hause zu fühlen und das Beste für die Schule zu geben.

Für das Ausmaß der Leistungsverteilung innerhalb von Sekundarschulklassen fanden Evertson, Sanford und Emmer (1981), dass sich eine hohe Standardabweichung in den Eingangsleistungen der Schülerinnen und Schüler ungünstig für die individuelle Aufgabenorientierung und Kooperation auswirkte. Hier spielte jedoch auch die Qualität des Unterrichts eine Rolle, indem diese negativen Effekte bei Lehrerinnen und Lehrern mit höheren Klassenmanagement-Fähigkeiten geringer ausfielen.

1.4.3.2 Effekte durch die Komposition des familiären Hintergrunds

Bisherige Studien zur herkunftsbedingten Zusammensetzung der Schülerschaft auf psychosoziale Merkmale ergaben sowohl positive, nicht signifikante als auch negative Effekte. Scharenberg (2016) fand an einer Stichprobe von Sekundarschülerinnen und -schülern, dass diese zufriedener mit der Schule waren, wenn die Schülerschaft der Klasse vorwiegend aus sozial begünstigten Elternhäusern stammte. Peetsma et al. (2006) fanden wiederum keine Effekte der mittleren Zusammensetzung von Klassen hinsichtlich der sozialen und ethnischen Herkunft auf das Wohlbefinden. Sie betrachteten den Anteil von Lernenden mit Zuwanderungshintergrund und den Anteil von Lernenden aus sozioökonomisch benachteiligten Familien in Klassen an einer Stichprobe von Elementarschülerinnen und -schülern.

Eine Reihe anderer Studien zeigten dagegen, dass Lerngruppen mit einer sozial benachteiligten Schülerschaft oder mit einem hohen Anteil von Lernenden mit Zuwanderungshintergrund günstige Effekte auf die psychosoziale Entwicklung haben. Van Landeghem et al. (2002) fanden für Sekundarschülerinnen und -schüler, dass Lernende in Klassen mit sozial benachteiligter Schülerschaft ein höheres Level angaben, das Beste für die Schule zu geben als in Klassen mit sozial begünstigter Schülerschaft. Für den mittleren sozioökonomischen Status der Schule zeigte sich zudem ein tendenzieller negativer Kompositionseffekt auf das Gefühl, sich in der Schule zu Hause zu fühlen. Kiwanuka et al. (2016) fanden für Sekundarschülerinnen und -schüler einen negativen Effekt des mittleren sozioökonomischen Status der Klasse auf die Wahrnehmung der Nützlichkeit von Mathematik. Hornstra et al. (2015) zeigten für Elementarschülerinnen und -schüler, dass die Aufgabenorientierung und die Motivation in Klassen mit einem höheren Anteil von Lernenden mit sozial benachteiligtem Hintergrund stärker stiegen als in Klassen mit einem geringeren Anteil. Für den Anteil von ethnischen Minderheiten fanden sie einen positiven Effekt auf die Selbstwirksamkeit und die Motivation. Zu letzterem Befund kamen auch Rjosk, Richter, Hochweber, Lüdtke und Stanat (2015) für Sekundarschülerinnen und -schüler, als sie die ethnische Komposition der Klassen über die Erstsprache betrachteten. Wies ein höherer Anteil der Lernenden in der Klasse eine andere Erstsprache als Deutsch auf oder hatte zusätzlich zu Deutsch eine weitere Erstsprache erlernt, stieg die intrinsische Motivation der Schülerinnen und Schüler im Fach Deutsch innerhalb eines Schuljahrs stärker. Scharenberg (2016) fand zudem einen positiven Effekt der zuwanderungsbezogenen Komposition der Klasse auf die Schulzufriedenheit von Sekundarschülerinnen und -schülern.

Der Befund, dass mit der ethnischen Komposition positive Einflüsse auf psychosoziale Merkmale einhergehen, wird versucht mit der *specialization hypothesis* zu erklären (Hornstra et al., 2015). Es wird demzufolge angenommen, dass Lehrerinnen und Lehrer in Klassen mit einem höheren Anteil von Lernenden mit Zuwanderungshintergrund, aus ethnischen Minderheiten oder mit einer anderen Erstsprache als die Mehrheitssprache eines Landes, möglicherweise stärker gefordert sind und dadurch auch kompetenter darin werden, ihren Unterricht an die individuellen Bedürfnisse aller Lernenden anzupassen.

In der bereits oben erwähnten Elementarschulstudie von Benner und Croesnoe (2011), in der die Schülerkomposition hinsichtlich der Verteilung der ethnischen Herkunft über Simpson's index of diversity (1949) operationalisiert wurde, wurde auch der Einfluss der Heterogenität auf psychosoziale Schülermerkmale untersucht. Es zeigte sich kein statistisch signifikanter Effekt der Heterogenität auf externalisierendes Problemverhalten und interpersonale Skills.

1.4.3.3 Effekte durch die Komposition psychosozialer Merkmale

Untersuchungen zum Einfluss psychosozialer Merkmale in ihrer mittleren Ausprägung in der Lerngruppe auf die individuelle Entwicklung psychosozialer Merkmale kamen erst in den letzten Jahren auf. Die meisten davon entspringen der Forschungsdisziplin zu Peer-Contagion-Effekten oder sind an diese angelehnt. Peer-Contagion-Effekte sind bereits vielfach für *außerschulische* jugendliche Peergruppen gezeigt wurden (siehe im Überblick Dishion & Tipsord, 2011), jedoch nur selten für Peergruppen *innerhalb des Schulkontextes*. Die wenigen Studien zu Peergruppen im Schulkontext gehen bei der Modellierung des Peer-Contagion-Effekts gemäß der Modellierung von Kompositionseffekten in der Tradition der Schuleffektivitätsforschung vor. Sie untersuchen den Effekt der mittleren Zusammensetzung hinsichtlich eines psychosozialen Merkmals unter Kontrolle desselbigen Merkmals auf der Schülerebene auf die individuelle Entwicklung dieses psychosozialen Merkmals. Peer-Contagion-Effekte solcher Art können somit ebenfalls als Kompositionseffekte bezeichnet werden. Die mittlere Zusammensetzung hinsichtlich der psychosozialen Merkmale in der Klasse wurde dabei unterschiedlich operationalisiert. In einigen Studien wurden die Lernenden gebeten, jeden ihrer Mitschülerinnen und Mitschüler auf einer Skala bezüglich der spezifischen Merkmalsausprägung einzuschätzen (Kellam, Ling, Merisca, Brown & Ialongo, 2000; Thomas & Bierman, 2006; Thomas, Bierman & Powers, 2011; Warren, Schoppelrey, Moberg & McDonald, 2005). In anderen Studien wurden die Lernenden gebeten, sich selbst einzuschätzen (Müller et al., 2015). Der Durchschnitt dieser

Ratings ergab dann den Klassenmittelwert. Andere Studien generierten den Mittelwert dieser Ratings unter Ausschluss der Ratings für das jeweilige Individuum, sodass für jeden Lernenden ein individualisierter Klassenmittelwert gebildet wurde (Thomas & Bierman, 2006).

Studien zu Peer-Contagion-Effekten im schulischen Bereich folgen der Tradition der Forschung zu Peer-Contagion-Effekten im außerschulischen Bereich indem sie vor allem unerwünschte Verhaltensweisen und Einstellungen wie Aggressivität und Akzeptanz von Aggressivität fokussieren. Sie zeigen, dass ein hoher Anteil von Lernenden mit aggressivem Verhalten in Klassen steigernd auf das individuelle aggressive Verhalten von Elementarschülerinnen und -schülern wirkt (für die Elementarstufe: Kellam et al., 1998; Thomas & Bierman 2006; Thomas et al., 2011; Warren et al. 2005). Müller et al. (2015) fanden diesen positiven Effekt speziell für störendes Verhalten im Unterricht in Sekundarschulklassen. Weiterhin weisen Studien auf eine „Ansteckung“ hinsichtlich normativer Überzeugungen zwischen Mitschülerinnen und Mitschülern hin. So kamen Henry et al. (2000) für Elementarschulen sowie Hofmann und Müller (2016) für Sekundarschulen zum Ergebnis, dass die auf Klassenebene gemittelten normativen Überzeugungen zur Angemessenheit oder *Coolness* von physisch oder verbal aggressivem Verhalten die entsprechenden individuellen Überzeugungen positiv beeinflussten. Busching und Krahe (2015) fanden diesen Effekt ausschließlich für die Klassennormen gleichgeschlechtlicher Mitschüler bzw. Mitschülerinnen in der Sekundarstufe. So wurden Mädchen in ihren individuellen Normen gegenüber Aggressivität von der Klassennorm ihrer Mitschülerinnen positiv beeinflusst und Jungen entsprechend von der Klassennorm der Mitschüler. Solch einen Peer-Contagion-Effekt durch Normen gleichgeschlechtlicher Mitschülerinnen bzw. Mitschüler fanden Stearns, Dodge und Nicholson (2008) ebenfalls für normverletzendes Verhalten von Elementarschülerinnen und -schülern. Über Untersuchungen zu unerwünschten Verhaltensweisen und Einstellungen hinaus, betrachten erst wenige Studien „ansteckende“ Effekte von *wünschenswerten* psychosozialen Merkmalen. An einer Stichprobe an Sekundarschulklassen ergab sich ein Peer-Contagion-Effekt bezüglich des wahrgenommenen Nutzens von Mathematik und der Freude an Mathematik (Kiwanuka et al., 2016).

Neben Befunden zum Effekt der mittleren Zusammensetzung der Lerngruppe hinsichtlich eines psychosozialen Merkmals auf das *gleiche* psychosoziale Merkmal auf der Individualebene, bestehen auch Studien, die den Einfluss eines psychosozialen Merkmals in seiner mittleren Ausprägung auf ein *anderes* psychosoziales Merkmal untersuchten. So können die mittleren

normativen Einstellungen der Klasse zu aggressivem Verhalten nicht nur die entsprechenden individuellen Einstellungen positiv beeinflussen, sondern auch die praktische Ausübung von aggressivem Verhalten (Busching & Krahe, 2015; Henry et al., 2000). Yudron und Jones (2016) ermittelten zudem, dass sich eine hohe Ausprägung an Fehlverhalten innerhalb der Lerngruppe auch positiv auswirken kann; in dem Fall auf die Entwicklung der Sozialkompetenz von Vorschulkindern. Abgesehen von nicht wünschenswertem Verhalten, betrachteten van Landeghem et al. (2002) wünschenswerte psychosoziale Kompositionsmerkmale auf Klassen- und Schulebene: Auf Klassenebene führte eine höhere mittlere Leistungsmotivation zu einer höheren sozialen Integration der Sekundarschülerinnen und -schüler. Eine höhere mittlere Immunität gegenüber Stress innerhalb der Klasse wirkte dagegen ungünstig auf das allgemeine Selbstkonzept der Schülerinnen und Schüler. Die mittlere Immunität gegenüber Stress innerhalb der Schule zeigte tendenziell einen günstigen Effekt auf das individuelle Gefühl, sich in der Schule zu Hause zu fühlen. Für die mittlere Leistungsmotivation an der Schule gab es Hinweise, dass es die Lernenden positiv beeinflusst, das Beste für die Schule zu geben.

1.4.4 Mediation von Kompositionseffekten

Bisher liegen kaum Studien vor, die die Wirkmechanismen von Kompositionseffekten untersuchen. Eine der wenigen Ausnahmen stellt die Studie von Zurbriggen (2016) an Elementarschulklassen dar. Es konnte gezeigt werden, dass der Effekt der mittleren Leistung auf die individuelle Leistung positiv über die mittlere emotionale Beteiligung der Klasse, das mittlere Ausmaß der Annäherung-Zielorientierung der Klasse (inwieweit es den Lernenden wichtig ist, ihre Leistungen zu demonstrieren) sowie der Bezugsnormorientierung der Lehrerinnen und Lehrer mediiert wurde. Die Effekte waren jedoch gering und nur signifikant, wenn alle drei Mediatoren in einem Modell spezifiziert wurden. Baumert et al. (2006) konnten dagegen keine Hinweise dafür finden, dass der Effekt der mittleren Leistung auf die individuelle Leistung über die normative Lernkultur vermittelt wird. Sie untersuchten dies jedoch für die eher distale Ebene der Schule. Es zeigte sich kein mediiender Effekt über den Anteil gewaltbereiter Schülerinnen und Schüler, die mittlere Schulzufriedenheit oder über die mittlere Akzeptanz von Leistungsnormen in der Schule. Als Nebenbefund ergab sich das interessante Ergebnis, dass der Anteil gewaltbereiter Lernender an der Schule weder von der leistungs- noch von der herkunftsbedingten Komposition abhängig war. Nikolova (2011) fand an einer Stichprobe an Elementarschulen Hinweise für einen sogenannten *Pygmalioneffekt*, d.h. dass Lehrkräfte von

leistungsstarken Gruppen mehr erwarten und sich diese Erwartungen dann an die Lernenden übertragen, wodurch indirekt die Leistungsentwicklung gefördert wird. Multiple Regressionsanalysen ergaben, dass Lernende in Klassen, bei denen der Anteil von leistungsstarken Lernenden von Lehrerinnen und Lehrern als hoch eingeschätzt wurde, bei gleicher mittlerer Leistung ein Jahr später eine höhere individuelle Leistung aufwiesen als Klassen, bei denen der Anteil von leistungsstarken Lernenden von Lehrerinnen und Lehrern als durchschnittlich eingeschätzt wurde. Allerdings konnte dies nur für das Fach Deutsch gezeigt werden, nicht für das Fach Mathematik. Boonen et al. (2014) stellten für die Erklärung der Effekte der leistungsbedingten Komposition und der sozioökonomischen Komposition auf die individuelle Leistungsentwicklung heraus, dass dies über den Kontakt zwischen Schule und Eltern positiv mediiert wurde. So waren Elementarschulen mit einer leistungsstarken oder sozial begünstigten Schülerschaft stärker in Kontakt mit Eltern, was begünstigend auf die individuelle Leistungsentwicklung wirkte.

Für den Effekt der Komposition des familiären Hintergrunds in der Schülerschaft auf die individuelle Leistungsentwicklung fanden sich bei Liu, van Damme, Gielen und van den Noortgate (2015) positive Mediationseffekte über das Niveau an Disziplinproblemen, der Schülermoral sowie das Schülerverhalten. So gaben die Lernenden an Sekundarschulen mit einer sozial begünstigten Schülerschaft weniger Disziplinprobleme im Unterricht an und die Schulleiterinnen und Schulleiter berichteten von einer positiver ausgeprägten Schülermoral sowie von positiverem Verhalten der Lernenden als an Schulen mit sozial benachteiligten Schülerschaften, was sich positiv auf die Leistungsentwicklung auswirkte. Anders sahen die Befunde von Berkowitz et al. (2015) aus, die ebenfalls die Mediation des Kompositionseffekts des mittleren sozioökonomischen Status auf die individuelle Leistungsentwicklung über Schulklimamerkmale untersuchten. Sie fanden heraus, dass das Schulklima, welches sie aus den Einschätzungen der Lernenden zur Qualität der Schüler-Lehrer-Beziehungen, zu riskantem Verhaltensnormen in der Schule sowie zu Gewalt an der Schule konstituierten, einen moderierenden, aber keinen mediiierenden Effekt hatte. Goldsmith (2011), der den Kompositionseffekt des Anteils an Lernenden aus einer ethnischen Minderheit an der Schule in der 8. Jahrgangsstufe auf deren spätere Bildungsabschlüsse bis zur Erreichung des 26. Lebensjahres untersuchte, identifizierte einen positiven mediiierenden Effekt über das Niveau an Schulproblemen und die Anzahl an Freunden mit Schulproblemen und/oder geringen Bildungsaspirationen. Rjosk et al. (2014) fanden für den Effekt der sozioökonomischen

Zusammensetzung auf die individuelle Deutschleistung Hinweise einer Mediation über das Fokussieren der Lehrkraft auf die Sprache während des Unterrichts, jedoch keine Mediationseffekte durch ein schülerorientiertes Klima oder durch ein strukturiertes Klassenmanagement.

Bezüglich der Mediation von *Peer-Contagion-Effekten* lässt sich eine Studie von Müller et al. (2015) für den Sekundarschulbereich nennen. Sie stellten für den Effekt des mittleren Niveaus an unterrichtsstörendem Verhalten auf das individuelle unterrichtsstörende Verhalten heraus, dass dieser durch die individuelle Wahrnehmung des Klassenniveaus an unterrichtsstörendem Verhalten positiv mediiert wurde. In Klassen mit höherem mittlerem Niveau an unterrichtsstörendem Verhalten, nahmen die Lernenden dies entsprechend stärker wahr, wodurch sie selbst mehr unterrichtsstörendes Verhalten zeigten.

1.5 Ableitung der Fragestellung

In Anbetracht multipler Ziele von Beschulung hat Schule gleichsam den Auftrag, die Lernenden in ihrer Entwicklung der Leistung und psychosozialer Merkmale zu unterstützen (Fend, 2009). Der Einfluss von Schule auf die psychosoziale Entwicklung der Lernenden ist jedoch erst wenig erforscht (Eccles & Roeser, 2011; Pekrun & Helmke, 1991). Obgleich in der Schule vorrangig Leistungsziele verfolgt werden, ist das Ausmaß der Unterrepräsentanz von Studien zum Einfluss der Schule auf psychosoziale Merkmale erstaunlich. Die psychosoziale Entwicklung der Lernenden wird nicht nur für sich gesehen als eines der multiplen Ziele von Schule als relevant erachtet (Eccles & Roeser, 2011; Fend, 2009; Kunter, 2005). Zudem kann als nachgewiesen gelten, dass sich eine positive Entwicklung in zahlreichen psychosozialen Merkmalen wie intrinsische Motivation oder Selbstkontrolle günstig auf die Leistungsentwicklung auswirkt (Gutman & Schoon, 2013).

Auch bei der Untersuchung des Einflusses der Schülerschaft als Teil der Schulumwelt, die die individuelle schulische Entwicklung mitbedingt, steht die psychosoziale Entwicklung bisher im Schatten der Leistungsentwicklung. Studien zum Einfluss der Schülerschaft in ihrer jeweils spezifischen leistungs- und herkunftsbezogenen Komposition ergeben eine breite Evidenz zum Effekt der leistungs- und herkunftsbezogenen Zusammensetzung auf die individuelle *Leistung*; (siehe im Überblick Dumont et al., 2013; sowie van Ewijk & Sleegers, 2010a, 2010b). Zu Effekten der leistungs- und herkunftsbezogenen Komposition auf *psychosoziale Merkmale* ist

jedoch abgesehen vom vielfach rezipierten BFLPE (Marsh, 1987; Marsh, Seaton et al., 2008) erst wenig bekannt. An diesem Desiderat soll die vorliegende Dissertation ansetzen.

Damit geht ein weiterer Punkt einher, der ebenfalls im Rahmen dieser Dissertation untersucht werden soll: Anknüpfend an Theorien zur Funktion von Bezugsgruppen, zu Peereinflüssen und zum Schulklima (siehe Kapitel 1.3) könnten sich Erkenntnisse zum Einfluss der Schülerkomposition auf die psychosoziale Entwicklung ergeben, wenn neben den leistungs- und herkunftsbezogenen Kompositionsmerkmalen auch die Komposition der Lernenden hinsichtlich psychosozialer Merkmale berücksichtigt werden würde. Die Schule oder Klasse, zu der man gehört, kann im sozialpsychologischen Sinn als Bezugsgruppe verstanden werden, die dem Einzelnen als Orientierungspunkt dient (Kelley, 1968). Demnach spielt es für die individuelle Entwicklung der Einzelnen eine Rolle, welche psychosozialen Merkmale die Mitschülerinnen und Mitschüler aufweisen, von denen sie täglich umgeben sind und mit denen sie in täglichem Kontakt stehen. Als Bezugsgruppe hat die Schülerschaft der Klasse oder Schule zwei Funktionen inne (Kelley, 1968): Lernende passen sich an die Normen der Gruppe an (normative Funktion) und sie vergleichen sich mit den anderen Lernenden in ihrer Gruppe (komparative Funktion). Auch aus der Perspektive der Forschung zu Peereinflüssen in der Jugend und zum Schulklima lässt sich ableiten, dass die psychosoziale Zusammensetzung der Lerngruppe konstituierend ist für Normen in der Lerngruppe, an denen sich Lernende orientieren und an die sie sich angleichen. Bedenkt man den „für die Sozialklimaforschung zentrale[n] Gedanke[n], dass individuelles Verhalten eine Funktion von Person- und Umweltmerkmalen ist“ (Götz et al., 2008, S. 503), dürften die Muster des Miteinanders eng mit der psychosozialen Zusammensetzung der Schülerschaft verknüpft sein. Je nachdem, wie die Schülerschaft der Klasse oder Schule zusammengesetzt ist, die die Peergruppe im Schulkontext bildet, können Peereinflussprozesse wie Lernen am Modell oder soziale Verstärkung stärker oder schwächer in bestimmte Richtungen auftreten. Befindet sich beispielsweise niemand in der Klasse mit besonders guten Argumentationskompetenzen ist kein Modell vorhanden, dessen Argumentationskompetenzen übernommen werden können. Oder gehen die in der Klasse oder Schule geteilten Peernormen eher in Richtung einer starken Leistungsorientierung, so würde die Angleichung der Peers gemäß Brown et al. (2008) auch in diese Richtung gehen. Befunde aus der Forschung zu Peer-Contagion-Effekten verweisen auf die Möglichkeit, dass von der psychosozialen Zusammensetzung möglicherweise auch ein eigenständiger Einfluss auf die schulische Entwicklung der Schülerinnen und Schüler ausgeht. Hier wurde gezeigt, dass sich

Jugendliche in Peergruppen bezüglich verschiedener psychosozialer Merkmale angleichen (siehe im Überblick Dishion & Tipsord, 2011). Erste Studien zeigen dies auch für das Setting der Schulklasse (z. B. Müller et al., 2015; Thomas et al., 2011; Warren et al., 2005). Psychosoziale Merkmale sind damit nicht nur als ein Aspekt der multiplen Ziele von Beschulung neben der Leistungsentwicklung zu betrachten, sie spielen zudem in ihrer Komposition in der Schülerschaft eine Rolle für die psychosoziale Entwicklung der Schülerinnen und Schüler.

Ein weiterer Faktor, der hinsichtlich der Komposition der Lerngruppe eine Rolle für die psychosoziale Entwicklung spielen könnte, liegt in der *Verteilung* von Schülermerkmalen innerhalb einer Lerngruppe. Gemäß der theoretischen Konzeption von Kompositionseffekten kann die Komposition der Schülerschaft nicht nur hinsichtlich der *mittleren Zusammensetzung* der Schülerschaft relevant für die schulische Entwicklung sein, sondern auch hinsichtlich der Verteilung von Schülermerkmalen innerhalb einer Lerngruppe (siehe im Überblick Dumont et al., 2013). Lange Zeit wurde versucht, die Verteilung von Schülermerkmalen aufgrund der Annahme einer leichteren Beschulung in homogenen Lerngruppen gering zu halten (LeTendre et al., 2003). So sind durch die in vielen Ländern bestehende leistungsbezogene Schulformgliederung relativ leistungshomogene Lerngruppen entstanden. Fraglich ist aber, ob dadurch auch Homogenität hinsichtlich psychosozialer Schülermerkmale erzielt wurde. Lediglich für die Wirkung der leistungshomogenen bzw. -heterogenen Komposition auf die Leistungsentwicklung liegen Befunde vor; diese sind jedoch inkonsistent (Ding & Lehrer, 2007; Hanushek et al., 2003; Kiss, 2013; Künsting et al., 2010; Lehmann, 2006; Leiter, 1983; Scharenberg, 2012). Zu Einflüssen der Verteilung psychosozialer Schülermerkmale ist die Befundlage noch weitaus geringer. Um hier ansetzen zu können wäre es zunächst notwendig zu ermitteln, wie heterogen die Schülerschaft innerhalb deutscher Klassenzimmer tatsächlich ist, auch hinsichtlich psychosozialer Merkmale.

In der Theorie zu Kompositionseffekten kommen psychosoziale Merkmale der Lerngruppe bisher nur indirekt als Wirkmechanismen zur Sprache (siehe im Überblick Dumont et al., 2013; sowie Harker & Tymms, 2004). Zur Erklärung leistungs- und herkunftsbedingter Kompositionseffekte auf die Leistungsentwicklung der Schülerinnen und Schüler wird die normative Lernkultur in der Lerngruppe herangezogen (Baumert et al., 2006; Harker & Tymms, 2004; Harris, 2010; Zurbriggen, 2016). So könnte in einer Schülerschaft aus leistungsstarken Lernenden mit hohen Bildungszielen eine anregende Lernkultur entstehen, die sich auf die einzelnen Schülerinnen und Schüler überträgt (Harker & Tymms, 2004; Harris, 2010). Gleichzeitig wird aufgrund der komparativen Funktion der Lerngruppe angenommen, dass

Lernende sich mit leistungsstärkeren Lernenden vergleichen, was zu einer geringeren Einschätzung der eigenen Fachkompetenzen führt. Damit wird der negative BFLPE der Leistungskomposition auf das akademische Selbstkonzept erklärt. Normative und komparative Phänomene können somit gegenläufige Effekte desselben Kompositionsmerkmals auf die schulische Entwicklung der Lernenden bewirken: in dem Fall einen günstigen Effekt auf die Leistung und einen ungünstigen auf das akademische Selbstkonzept. Angesichts der Wechselwirkungen zwischen individueller Leistung und akademischem Selbstkonzept bleibt die Frage, wie normative und komparative Funktion im Wechselspiel stehen, noch unbeantwortet.

In Anbetracht der normativen und komparativen Prozesse, die für die Erklärung von Kompositionseffekten angenommen werden und die sich durch das Zusammentreffen von Lernenden mit je bestimmten psychosozialen Merkmalen ergeben, liegt der übergeordnete Gegenstand der vorliegenden Dissertation in der Erweiterung der Forschung zu Kompositionseffekten hinsichtlich der Rolle psychosozialer Schülermerkmale. Es ergeben sich dazu folgende drei Ansätze für die vorliegende Dissertation: (1) Psychosoziale Schülermerkmale werden als Merkmale betrachtet, deren Entwicklung im Schulkontext von der Komposition der Lerngruppe bedingt werden kann. (2) Psychosoziale Schülermerkmale werden in ihrer mittleren Ausprägung und in ihrer Verteilung innerhalb der Lerngruppe als Kompositionsmerkmal betrachtet. (3) Es wird das Zusammenspiel verschiedener Kompositionseffekte betrachtet. Diesem breiten Zielvorhaben nähert sich die vorliegende Dissertation in vier Teilstudien, in denen spezifische psychosoziale Merkmale in jeweils unterschiedlichen Rollen untersucht werden. Grundlage der Analysen bilden repräsentative Stichproben verschiedener Längsschnittstudien. Deren Datensätze umfassen verschiedene Altersgruppen der Sekundarschule I: die 6., 7., 9. und 10. Jahrgangsstufe. Die Analysen zu Kompositionseffekten wurden jeweils entsprechend der Modellierung von Kompositionseffekten in der Tradition der Schuleffektivitätsforschung (Harker & Tymms, 2004) umgesetzt. Die vier Teilstudien werden im Folgenden kurz skizziert.

In der ersten Teilstudie (*Der Effekt der Schulzusammensetzung auf die Entwicklung der Schulfriedenheit und des normverletzenden Verhaltens von Schülerinnen und Schülern*) wird exemplarisch an den zwei psychosozialen Merkmalen Schulfriedenheit und normverletzendes Verhalten untersucht, inwiefern die psychosoziale Entwicklung durch Merkmale der Schülerschaft beeinflusst werden kann. Die beiden Merkmale werden dabei sowohl als Kriteriumsvariablen betrachtet als auch selbst als Kompositionsmerkmale. Angesichts der

bekanntem Kompositionseffekte ist denkbar, dass die leistungsbedingte und sozioökonomische Zusammensetzung der Schülerschaft einer Schule auch mit der individuellen Entwicklung von Schulzufriedenheit und normverletzendem Verhalten in Zusammenhang steht. Da diese beiden psychosozialen Merkmale Schülermerkmale darstellen, die stark zum Schulklima einer Schule beitragen (siehe im Überblick Cohen et al., 2009), ist ein normativer Einfluss zwischen den Peers der Mitschülerinnen und Mitschüler denkbar, bei dem sich Lernende im Sinne eines Peer-Contagion-Effekts einander angleichen. Da vermutet wird, dass mit einem positiven Schulklima auch die Leistungsentwicklung begünstigt werden kann (siehe im Überblick Thapa et al., 2013), ist zudem ein positiver Einfluss der mittleren Schulzufriedenheit und ein negativer Effekt des mittleren normverletzenden Verhaltens auf die Lernentwicklung möglich. Als Datengrundlage dient eine Stichprobe von 2 193 Schülerinnen und Schülern aus 116 Schulen der Stadt Berlin. Es handelt sich um eine längsschnittlich angelegte Studie mit einer Befragung und Leistungstestung in der 9. Jahrgangsstufe und einer erneuten Befragung ein Jahr später in der 10. Jahrgangsstufe. Zur Analyse der Fragestellungen wird das Verfahren der *hierarchisch linearen Modellierung* angewandt (Raudenbush & Bryk, 2002).

Die zweite Teilstudie (*Are jitters contagious? The influence of class-average test anxiety on individual students' test anxiety*) untersucht den Einfluss der psychosozialen Komposition auf die psychosoziale Entwicklung am Beispiel der Testangst. Obgleich die Testangst eine der am intensivsten erforschten Leistungsemotionen darstellt (Zeidner, 2014), ist wenig darüber bekannt, welche Rolle die soziale Umwelt für Testangst spielt (siehe im Überblick Zeidner, 2014). Angesichts der Forschungsergebnisse zu Peer-Contagion-Effekten (siehe im Überblick Dishion & Tipsord, 2011), ist denkbar, dass sich Testangst gerade im Jugendalter in der Schule zwischen Lernenden überträgt. Auf Basis der längsschnittlichen Studie MSALT (Eccles et al., 1989) wird daher untersucht, ob auch die mittlere Testangst der Klasse in positivem Zusammenhang mit der individuellen Entwicklung der Testangst steht und Testangst somit zwischen den Lernenden einer Peergruppe übertragen wird. Da für solche Peereinflüsse angenommen wird, dass individuelle Faktoren wie das Ausgangsniveau des betreffenden Merkmals die Anfälligkeit für Peereinflüsse bedingen (Dishion & Tipsord, 2011; Prinstein, 2007), werden differenzielle Effekte für die vorherige Testangst sowie für das Geschlecht untersucht. Dazu wurde in 95 Schulklassen die mittlere Testangst der Lernenden zu Beginn der 7. Jahrgangsstufe direkt nach dem Übergang auf die Sekundarschule ermittelt. Von 1 811 Lernenden lagen Angaben vor, wie testängstlich sie am Ende der 6. Jahrgangsstufe sowie zu Beginn und am Ende der 7. Jahrgangsstufe waren. Die

statistische Auswertung erfolgt mittels hierarchisch linearer Modellierung (Raudenbush & Bryk, 2002).

In der dritten Teilstudie (*What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept*) wird das Zusammenspiel von Kompositionseffekten auf die leistungsbezogene und psychosoziale Entwicklung untersucht. Dies geschieht am Beispiel der simultanen Betrachtung des positiven Effekts der mittleren Klassenleistung auf die individuelle Leistung und des negativen Effekts der mittleren Klassenleistung auf das individuelle akademische Selbstkonzept. Bisherige Studien betrachteten die Wechselwirkungen zwischen akademischem Selbstkonzept und Leistung entweder ausschließlich auf der Individualebene (z. B. Pinxten, de Fraine, van Damme & D'Haenens, 2013; Seaton, Parker, Marsh, Craven & Yeung, 2014) oder untersuchten den BFLPE und Kompositionseffekte im Längsschnitt getrennt voneinander (z. B. Köller & Baumert, 2001; Köller, Trautwein, Lüdtke & Baumert, 2006). Daher werden in dieser Teilstudie die zwei gegenläufigen Effekte der mittleren Leistungskomposition auf die Entwicklung der Leistung und psychosozialer Merkmale simultan über mehrere Zeitpunkte hinweg betrachtet. Als Datenbasis dient die BIJU-Studie (Bildungsverläufe und psychosoziale Entwicklung im Jugendalter). Sie ist besonders für diese Frage geeignet, da sowohl das akademische Selbstkonzept als auch Leistungsmaße im entsprechenden Fach über drei Messzeitpunkte hinweg erhoben wurden. So wurde eine Stichprobe von 6 207 Schülerinnen und Schülern aus 271 Klassen zu Beginn, in der Mitte sowie am Ende der 7. Jahrgangsstufe hinsichtlich ihres mathematischen Selbstkonzepts befragt und in ihren Fähigkeiten in Mathematik getestet. Die statistische Auswertung erfolgt mittels hierarchisch linearer Modellierung mit *Cross-Lagged-Pfaden* (Helmke & van Aken, 1995; Skaalvik & Valås, 1999) auf Schülerebene, wobei die unabhängigen Variablen auf der Individualebene am Gesamtmittelwert zentriert werden.

Mit der vierten Teilstudie (*Ein Blick in die Klassenzimmer Deutschlands – wie heterogen sind Klassen zusammengesetzt?*) wird der Fokus auf die *Verteilung* von Schülermerkmalen anstatt auf die *mittlere Zusammensetzung* der Schülerschaft verschoben. Im Gegensatz zu Kompositionseffekten durch die mittlere Zusammensetzung und trotz des breiten Diskurses über Heterogenität ist wenig darüber bekannt, wie Schülermerkmale in ihrer Verteilung innerhalb der Lerngruppe auf die schulische Entwicklung wirken. Um der Klärung des Einflusses der Heterogenität von Lerngruppen näherzukommen, soll erst einmal deskriptiv betrachtet werden, von welcher Heterogenität der Schülermerkmale ausgegangen werden kann. Dazu werden anhand

der DESI-Studie (Klieme, 2008; Klieme & Beck, 2007) verschiedene Schülermerkmale in ihrer Verteilung betrachtet, um einen Einblick zu gestatten, wie heterogen die Schülerschaft innerhalb von Klassenzimmern in Deutschland tatsächlich ist und ob dies nach Schulform variiert. Aufgrund der theoretisch postulierten Rolle psychosozialer Merkmale der Lerngruppe für die individuelle Entwicklung werden nicht nur leistungs- und herkunftsbezogene Schülermerkmale in ihrer Verteilung innerhalb der Klasse eruiert, sondern auch die Verteilung psychosozialer Merkmale. Die DESI-Studie stellt dazu eine geeignete Datenbasis dar, da komplette Schulklassen befragt wie getestet wurden und die Stichprobe repräsentativ für Deutschland ist. Mit den Angaben von 10 062 Neuntklässlern in 427 Klassen werden umfassende deskriptive Betrachtungen vorgenommen, um die Heterogenität der Schülerschaft in ihren verschiedenen Facetten zu illustrieren. Zudem wird nach Schulform differenziert.

Literaturverzeichnis I

- Adams, R. E., Bukowski, W. M. & Bagwell, C. (2005). Stability of aggression during early adolescence as moderated by reciprocated friendship status and friend's aggression. *International Journal of Behavioral Development*, 29(2), 139–145.
- Ajzen, I. (2005). *Attitudes, personality, and behavior* (2nd ed.). Berkshire: Open University Press.
- Allan, J.P. & Antonishak, J. (2008). Understanding adolescent peer influences: Beyond the dark side. In M.J. Prinstein & K.S. Dodge (Eds.), *Peer Influence Among Youth* (pp. 141–160). New York, NY: Guilford Press.
- Bandura, A. (1979). The social learning perspective: Mechanisms of aggression. In H. Toch (Ed.), *Psychology of crime and criminal justice* (pp. 193–236). Prospect Heights, IL: Waveland Press.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive perspective*. Englewood Cliffs, NJ: Prentice-Hall.
- Batanova, M. & Loukas, A. (2016). Empathy and effortful control effects on early adolescents' aggression: When do students' perceptions of their school climate matter? *Applied Developmental Science*, 20(2), 79–93.
- Baumert, J. & Schümer, G. (2001). Schulformen als selektionsbedingte Lernmilieus. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, et al. (Hrsg.), *PISA 2000* (S. 454–467). Opladen: Leske und Budrich.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.) (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., Trautwein, U. & Artelt, C. (2003). Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, et al. (Hrsg.), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 261–331). Opladen: Leske und Budrich.
- Bellin, N. (Hrsg.). (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(3), 621–646.

- Benson, J. & Borman, G. D. (2010). Family, neighborhood, and school settings across seasons. When do socioeconomic context and racial composition matter for the reading achievement growth of young children? *Teachers College Record*, 112 (5), 1338–1390.
- Berger, P. L. & Luckmann, T. (1969). Die soziale Konstruktion der Wirklichkeit. *Eine Theorie der Wissensgesellschaft*. Frankfurt a. M.: Fischer.
- Berkowitz, R., Glickman, H., Benbenishty, R., Ben-Artzi, E., Raz, T., Lipshtat, N. et al. (2015). Compensating, mediating, and moderating effects of school climate on academic achievement gaps in Israel. *Teachers College Record*, 117(7), 1–34.
- Berndt, T. J. & Ladd, G. W. (Eds.) (1989). *Peer relationships in child development*. New York, NY: Wiley.
- Boonen, T., Speybroeck, S., Bilde, J. de, Lamote, C., van Damme, J. & Onghena, P. (2014). Does it matter who your schoolmates are? An investigation of the association between school composition, school processes and mathematics achievement in the early years of primary education. *British Educational Research Journal*, 40(3), 441–466.
- Bos, W. & Scharenberg, K. (2010). Lernentwicklung in leistungshomogenen und -heterogenen Schulklassen. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (S. 173–194). Münster: Waxmann.
- Breidenstein, G. (2008). Peer-Interaktion und Peer-Kultur. In W. Helsper & J. Böhme (Hrsg.), *Handbuch der Schulforschung* (S. 945–964). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bronfenbrenner, U. (1979). *The ecology of human development. Experiments in nature and design*. Cambridge, MA: Harvard University Press.
- Bronfenbrenner, U. & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, 101(4), 568–586.
- Bronfenbrenner, U. & Morris, P. A. (2006). The bioecological model of human development. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology* (Theoretical models of human development, Vol. 1, 6th ed., pp. 793–828). New York, NY: John Wiley.
- Brown, B. B. (1989). The role of peer groups in adolescents' adjustment to secondary school. In T. J. Berndt & G. W. Ladd (Eds.), *Peer relationships in child development* (pp. 188–215). New York, NY: Wiley.

- Brown, B. B. (1990). Peer Groups and Peer Cultures. In S. S. Feldman & G. R. Elliot (Ed.), *At the threshold. The developing adolescent* (pp. 171–547). Cambridge, MA: Harvard University Press.
- Brown, B. B. (2013). Adolescents' relationships with Peers. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology: Contextual influences on adolescent development* (Vol. 2, pp. 363–394). Hoboken, NJ: Wiley & Sons.
- Brown, B. B., Bakken, J. P., Ameringer, S. W. & Mahon, S. D. (2008). A comprehensive conceptualization of the peer influence process in adolescence. In M. J. Prinstein & K. A. Dodge (Eds.), *Understanding Peer Influence in Children and Adolescents* (pp. 17–44). New York, NY: The Guilford Press.
- Brown, B. B., Mory, M. S. & Kinney, D. (1994). Casting adolescent crowds in a relational perspective: Caricature, channel, and context. In R. Montemayor, G. R. Adams & T. P. Gulotta (Eds.), *Personal relationships during adolescence* (Advances in adolescent development, Vol. 6, pp. 123–167). Newbury Park, CA: Sage Publications.
- Budde, J. (2012). *Problematisierende Perspektiven auf Heterogenität als ambivalentes Thema der Schul- und Unterrichtsforschung*. Weinheim: Beltz Juventa.
- Bukowski, W. M., Velasquez, A. M. & Brendgen, M. (2010). Variation in patterns of peer Influence. considerations of self and other. In M. J. Prinstein & K. A. Dodge (Eds.), *Understanding peer influence in children and adolescents* (pp. 125–140). New York, NY: The Guilford Press.
- Bundeskabinett. (2011). Übereinkommen der Vereinten Nationen über Rechte von Menschen mit Behinderungen. UN-Behindertenrechtskonvention.
- Burlingame, W. V. (1970). The youth culture. In E.D. Evans (Ed.), *Adolescents: Readings in behavior and development* (pp. 131–149). Hinsdale, IL: Dryden.
- Burns, R. & Mason, D. A. (2002). Class composition and student achievement in elementary schools. *American Educational Research Journal*, 39(1), 207–233.
- Busching, R. & Krahe, B. (2015). The girls set the tone. *Personality and Social Psychology Bulletin*, 41(5), 659–676.
- Calsyn, R. J. & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology*, 69(2), 136.
- Carle, U. & Metzen, H. (2014). Wie wirkt Jahrgangsübergreifendes Lernen. *Internationale Literaturübersicht zum Stand der Forschung, der praktischen Expertise und der*

pädagogischen Theorie. Eine wissenschaftliche Expertise des Grundschulverbandes.

Frankfurt a. M.: Grundschulverband.

- 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(5), 925–957.
- Cialdini, R. B. & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology* (Vol. 2, pp. 151–192). New York, NY: McGraw-Hill.
- Cillessen, A. H. & Rose, A. J. (2005). Understanding popularity in the peer system. *Current Directions in Psychological Science*, 14(2), 102–105.
- Cohen, J., McCabe, E. M., Michelli, N. M. & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record*, 111(1), 180–213.
- Coleman, J. (1990). *Foundations of social theory*. Cambridge, MA: Belknap.
- 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, DC: US Government Printing Office.
- Danhier, J. & Martin, É. (2014). Comparing compositional effects in two education systems: The case of the Belgian communities. *British Journal of Educational Studies*, 62(2), 171–189.
- Dar, Y. & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 23(3), 357–374.
- Ding, W. & Lehrer, S. (2007). Do peers affect student achievement in China's secondary schools? *Review of Economics and Statistics*, 89(2), 300–312.
- Dishion, T. J. & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, 62, 189–214.
- Ditton, H. & Krüsken, J. (2006). Der Übergang von der Grundschule in die Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 9(3), 348–372.
- Dreeben, R. & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education*, 61(3), 129.
- Dreesmann, H., Eder, F., Fend, H., Pekrun, R., von Saldern, M. & Wolf, B. (1992). Schulklima. In K. Ingenkamp, R. S. Jäger, H. Petillon & B. Wolf (Hrsg.), *Empirische Pädagogik 1970–1990. Eine Bestandsaufnahme der Forschung in der Bundesrepublik, Band 2* (S. 655–682). Weinheim: Deutscher Studien Verlag.

- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60(3), 163–183.
- Duncan, G. J., Magnuson, K. A. & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1(1/2), 59–80.
- Durkheim, E. (1894). *Les Règles de la méthode sociologique*. Paris: Presses Universitaires de France.
- Eccles, J. S. & Roeser, R. W. (2011). School and Community Influences on Human Development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental Science. An Advanced Textbook* (6th ed., pp. 571–643). New York, NY: Psychology Press.
- Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D. A. & Yee, D. (1989). Self-concepts, domain values, and self-esteem. relations and changes at early adolescence. *Journal of Personality*, 57(2), 283–310.
- Eder, F. (1998). Schul- und Klassenklima. In Rost, D. H. (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 424–430). Weinheim: Psychologie Verlags Union.
- Eder, F. (2002). Unterrichtsklima und Unterrichtsqualität. *Unterrichtswissenschaft Zeitschrift für Lernforschung*, 30(3), 213–229.
- Ellis, T. I. (1988). School Climate. *Research Roundup*, 4(2), 1–6.
- Enders, C. K. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121–138.
- Erbring, L. & Young, A. A. (1979). Individuals and social structure: Contextual effects as endogenous feedback. *Sociological Methods & Research*, 7(4), 396–430.
- Erikson, E. (1968). *Identity, youth, and crisis*. New York, NY: Morton Inc.
- Evertson, C. M., Sanford, J. P. & Emmer, E. T. (1981). Effects of class heterogeneity in Junior-High School. *American Educational Research Journal*, 18(2), 219–232.
- Fend, H. (1977). *Schulklima. Soziale Einflussprozesse in der Schule*. Weinheim: Beltz.
- Fend, H. (2009). *Neue Theorie der Schule: Einführung in das Verstehen von Bildungssystemen* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.
- Fischer, L. & Wiswede, G. (2009). *Grundlagen der Sozialpsychologie* (3. Aufl.). München: Oldenbourg Wissenschaftsverlag.

- Forgas, J. P. & Williams, K. D. (2001). *Social influence: Direct and indirect processes* (3rd ed.). Philadelphia, PA: Psychology Press.
- Frey, D., Dauenheimer, D., Parge, O. & Haisch, J. (1993). Die Theorie sozialer Vergleichsprozesse. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie* (S. 81–121). Bern: Huber.
- Giele, J. Z. & Elder, G. H. (1998). *Methods of life course research: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage Publications.
- Goethals, G. R. & Darley, J. M. (1977). Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives* (pp. 259–278). Washington, D. C.: Hemisphere.
- Goldsmith, P. R. (2011). Coleman revisited: School segregation, peers, and frog ponds. *American Educational Research Journal*, 48(3), 508–535.
- Gommans, R., Stevens, G. W. J. M., ter Bogt, T. F. M. & Cillessen, A. H. N. (2016). Adolescent substance use, aggressive behaviors, and peer context behavioral norms. *Journal of Cognitive Education and Psychology*, 15(2), 185–205.
- Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology*, 82(3), 525–538.
- Gottfried, M. A. (2014). The achievement effects of tardy classmates: Evidence in urban elementary schools. *School Effectiveness and School Improvement*, 25(1), 3–28.
- Gottfried, M. A. (2015). Does classmate ability influence students' social skills? *School Effectiveness and School Improvement*, 26(4), 554–585.
- Götz, T., Frenzel, A. & Pekrun, R. (2008). Sozialklima in der Schule. In W. Schneider & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie* (Bd. 10, S. 503–514). Göttingen: Hogrefe.
- Gröhlich, C. & Guill, K. (2009). Wie stabil sind Bezugsgruppeneffekte der Grundschulempfehlungen für die Schulformzugehörigkeit in der Sekundarstufe. *Journal for Educational Research Online*, 1(1), 154–171.
- Gröhlich, C., Guill, K., Scharenberg, K. & Bos, W. (2010). Differenzielle Lern- und Entwicklungsmilieus beim Erwerb der Lesekompetenz in den Jahrgangsstufen 7 und 8. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (S. 100–106). Münster: Waxmann.

- Gutman, L. M. & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people. Literature review*. London: Education Empowerment Foundations.
- Hallinan, M. (1988). School composition and learning. A critique of Dreeben-Barr model. *Sociology of Education*, 61(3), 143–146.
- Hallinan, M. (1994). Tracking. From theory to practice. *Sociology of Education*, 67(2), 79–84, 89–91.
- Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527–544.
- Harker, R. & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15(2), 177–199.
- Harris, D. N. (2010). How do school peers influence student educational outcomes? Theory and evidence from economics and other social sciences. *Teachers College Record*, 112(4), 1163–1197.
- Hascher, T. (Hrsg.) (2004). *Schule positiv erleben. Erkenntnisse und Ergebnisse zum Wohlbefinden von Schülerinnen und Schülern*. Bern: Haupt AG.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.
- Helmke, A. (2006). *Unterrichtsqualität: Erfassen, Bewerten, Verbessern* (4. Aufl.). Seelze: Kallmeyersche Vertragsbuchhandlung.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (6. Aufl.). Seelze: Klett-Kallmeyer.
- Helmke, A. & van Aken, M. A. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, 87(4), 624–637.
- Henry, D., Guerra, N., Huesmann, R., Tolan, P., VanAcker, R. & Eron, L. (2000). Normative influences on aggression in urban elementary school classrooms. *American Journal of Community Psychology*, 28(1), 59–81.
- Hofmann, V. & Müller, C. M. (2016). Attitude transference among adolescents: The relationship between injunctive classroom norms and individual antisocial attitudes. *Journal of Cognitive Education and Psychology*, 15(2), 206–224.
- Hornstra, L., van der Veen, I., Peetsma, T. & Volman, M. (2015). Does classroom composition make a difference: effects on developments in motivation, sense of classroom belonging, and

achievement in upper primary school. *School Effectiveness and School Improvement*, 26 (2), 1–28.

Ireson, J. & Hallam, S. (2001). *Ability grouping in education*. London: Sage Publications.

Jencks, C. & Mayer, S. E. (1990). The social consequences of growing up in a poor neighborhood. In L. E. Lynn & M. F. H. McGuey (Eds.), *Inner-city poverty in the United States* (pp. 111-186). Washington, DC: National Academy Press.

Justice, L. M., Logan, J. A. R., Lin, T.-J. & Kaderavek, J. N. (2014). Peer effects in early childhood education: Testing the assumptions of special-education inclusion. *Psychological Science*, 25(9), 1722–1729.

Juvonen, J. & Galván, A. (2008). Peer influence in involuntary social groups. Lessons from research on bullying. In M. J. Prinstein & K. A. Dodge (Eds.), *Understanding peer influence in children and adolescents* (pp. 225–244). New York, NY: Guilford Press.

Kellam, S. G., Ling, X., Merisca, R., Brown, C. H. & Jalongo, N. (2000). Erratum: The effect of the level of aggression in the first grade classroom on the course and malleability of aggressive behavior into middle school. *Development and Psychopathology*, 12(1), 107.

Kellam, S. G. & van Horn, Y. V. (1997). Life course development, community epidemiology, and preventive trials: A scientific structure for prevention research. *American Journal of Community Psychology*, 25(2), 177–187.

Kelley, H. H. (1968). Two functions of reference groups. In H.H. Hyman & E. Singer (Eds.), *Readings in reference group theory and research* (pp. 77–83). New York, NY: The Free Press.

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.

Kiwanuka, H. N., van Damme, J., van Den Noortgate, Wim, Anumendem, D. N., Vanlaar, G., Reynolds, C. et al. (2016). How do student and classroom characteristics affect attitude toward mathematics? A multivariate multilevel analysis. *School Effectiveness and School Improvement*, 28(1), 1–21.

Klieme, E. (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz Pädagogik.

Klieme, E. & Beck, B. (2007). *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band 1*. Weinheim: Beltz Pädagogik.

- Köller, O. & Baumert, J. (2001). Leistungsgruppierungen in der Sekundarstufe I. Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für pädagogische Psychologie*, 15(2), 99–110.
- Köller, O., Baumert, J. & Schnabel, K. (2000). Zum Zusammenspiel von schulischem Interesse und Lernen im Fach Mathematik: Längsschnittanalysen in den Sekundarstufen I und II. In U. Schiefele & K. Wild (Hrsg.), *Interesse und Lernmotivation - Untersuchungen zu Entwicklung, Förderung und Wirkung* (S. 163-182). Münster: Waxmann.
- 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(3), 184–197.
- Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für pädagogische Psychologie*, 20(1-2), 27–39.
- Kulik, C. L. C. & Kulik, J. A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal*, 19(3), 415–428.
- Künsting, J., Post, S., Greb, K., Faust, G. & Lipowsky, F. (2010). Leistungsheterogenität im mathematischen Anfangsunterricht. Ein Risiko für die Leistungsentwicklung? *Zeitschrift für Grundschulforschung*, 3(1), 46–64.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht*. Münster: Waxmann.
- LaFontana, K. M. & Cillessen, A. H. (2002). Children's perceptions of popular and unpopular peers: a multimethod assessment. *Developmental Psychology*, 38(5), 635.
- Lehmann, G. & Nieke, W. (2000). *Zum Kompetenz-Modell*. Zugriff am 22.05.2017, von <http://www.bildungsserver-mv.de/download/material/text-lehmann-nieke.pdf>.
- Lehmann, R. H. (2006). Zur Bedeutung der kognitiven Heterogenität von Schulklassen für den Lernstand am Ende der Klassenstufe 4. In A. Schröder-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung: Migration, Leistungsangst und Schulübergang* (S. 109–121). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Leiter, J. (1983). Classroom composition and achievement gains. *Sociology of Education*, 56(3), 126–132.

- LeTendre, G. K., Hofer, B. K. & Shimizu, H. (2003). What is tracking? Cultural expectations in the United States, Germany, and Japan. *American Educational Research Journal*, 40(1), 43–89.
- Lewin, K. (1952). Constructs in field theory [1944]. In D. Cartwright (Ed.), *Field theory in social science: Selected theoretical papers by Kurt Lewin* (pp. 30–42). London: Tavistock.
- Libbey, H. P. (2004). Measuring student relationships to school: Attachment, bonding, connectedness, and engagement. *Journal of School Health*, 74(7), 274–283
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern* (S. 47-70). Weinheim: Beltz.
- Liu, H., van Damme, J., Gielen, S. & van Den Noortgate, Wim. (2015). School processes mediate school compositional effects: Model specification and estimation. *British Educational Research Journal*, 41(3), 423–447.
- Lauder, H., Kounali, D., Robinson, T., & Goldstein, H. (2010). Pupil composition and accountability: An analysis in English primary schools. *International Journal of Educational Research*, 49(2 – 3), 49–68.
- Maaz, K. & Nagy, G. (2009). Der Übergang von der Grundschule in die weiterführenden Schulen des Sekundarschulsystems: Definition, Spezifikation und Quantifizierung primärer und sekundärer Herkunftseffekte. In J. Baumert, K. Maaz & U. Trautwein (Hrsg.), *Bildungsentscheidungen* (Zeitschrift für Erziehungswissenschaft, Sonderheft 12, S. 153-182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- 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(2), 99–106.
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3), 531–542.
- Marks, G. N. (2010). What aspects of schooling are important? School effects on tertiary entrance performance. *School Effectiveness and School Improvement*, 21(3), 267–287.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79(3), 280–295.

- Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift für Pädagogische Psychologie*, 19(3), 119–127.
- Marsh, H. W., Abduljabbar, A. S., Parker, P. D., Morin, A. J. S., Abdelfattah, F. & Nagengast, B. (2014). The big-fish-little-pond effect in mathematics: A cross-cultural comparison of U.S. and Saudi Arabian TIMSS responses. *Journal of Cross-Cultural Psychology*, 4(5), 777–804.
- Marsh, H. W., Kong, C. K. & Hau, K. T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept. Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology*, 78(2), 337–349.
- Marsh, H. W. & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81(1), 59–77.
- Marsh, H. W. & O'Mara, A. J. (2010). Long-term total negative effects of school-average ability on diverse educational outcomes. *Zeitschrift für pädagogische Psychologie*, 24(1), 51–72.
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J. et al. (2008). The big-fish-little-pond effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20(3), 319–350.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J. & Köller, O. (2007). The big-fish-little-pond effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, 44(3), 631–669.
- Marsh, H. W., Trautwein, U., Lüdtke, O. & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs. Role of generalized and specific others. *Journal of Educational Psychology*, 100(3), 510–524.
- Maschke, S. & Stecher, L. (2010). *In der Schule. Vom Leben, Leiden und Lernen in der Schule*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: Van Nostrand Co.
- 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(1), 121–158.
- Minsel, B. & Roth, W. K. (1978). *Soziale Interaktion in der Schule*. München: Urban & Schwarzenberg.
- Miller, D. T. & Prentice, D. A. (1994). Collective errors and errors about the collective. *Personality and Social Psychology Bulletin*, 20(5), 541–550.

- Müller, C. M., Hofmann, V., Fleischli, J. & Studer, F. (2015). „Sag‘ mir, was deine Klassenkameraden tun und ich sage dir, was du tun wirst“? Zum Einfluss der Klassenzusammensetzung auf die Entwicklung von schulischem Problemverhalten. *Zeitschrift für Erziehungswissenschaft*, 18(3), 569–589.
- Müller, C. M. & Minger, M. (2013). Welche Kinder und Jugendliche werden am stärksten durch die Peers beeinflusst? Eine systematische Übersicht für den Bereich dissozialen Verhaltens. *Empirische Sonderpädagogik*, 5(2), 107–129
- Mussweiler, T. & Bodenhausen, G. V. (2002). I know you are, but what am I? Self-evaluative consequences of judging in-group and out-group members. *Journal of Personality and Social Psychology*, 82(1), 19–32.
- Nagengast, B. & Marsh, H. W. (2011). The negative effect of school-average ability on science self-concept in the UK, the UK countries and the world: The big-fish-little-pond effect for PISA 2006. *Educational Psychology*, 31(5), 629–656.
- Nash, R. (2003). Inequality/difference in education: Is a real explanation of primary and secondary effects possible? *The British Journal of Sociology*, 54(4), 433–451.
- Neumann, M., Maaz, K. & Becker, M. (2013). Die Abkehr von der traditionellen Dreigliedrigkeit im Sekundarschulsystem: Auf unterschiedlichen Wegen zum gleichen Ziel? *Recht der Jugend und des Bildungswesens*, 61(3), 274–292.
- Nikolova, R. (2011). *Grundschulen als differenzielle Entwicklungsmilieus. Objektive Kontextmerkmale der Schülerzusammensetzung und deren Auswirkungen auf die Mathematik- und Leseleistungen*. Münster: Waxmann.
- Noack, P. (2011). Socialization. In B.B. Brown & M. Prinstein (Eds.), *Encyclopedia of adolescence* (Vol 2, pp. 352–359). London: Elsevier Inc.
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. Binghamton, NY: Vail-Ballou Press.
- Opdenakker, 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(4), 399–427.
- 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(3), 275–302.

- Pekrun, R. (1985). Classroom climate and test anxiety: Developmental validity of expectancy-value theory of anxiety. In R. Schwarzer, C. D. Spielberger & H. M. van der Ploeg (Eds.), *Advances in test anxiety research* (Vol. 4, pp. 147-158). Lisse: Swets & Zeitlinger.
- Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development: Developing motivation and motivating development* (pp. 143–163). Amsterdam: Elsevier.
- Pekrun, R. & Helmke, A. (1991). Schule und Persönlichkeitsentwicklung: Theoretische Perspektiven und Forschungsstand. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung: Ein Resümee der Längsschnittforschung (Der Mensch als soziales und personales Wesen* (Bd. 11, S. 33–56). Stuttgart: Enke.
- Pinxten, M., de Fraine, B., van Damme, J. & D’Haenens, E. (2013). Student achievement and academic self-concept among secondary students in Flanders: Gender and changes over time. *Irish Educational Studies*, 32(2), 157–178.
- PISA-Konsortium, D. (Hrsg.) (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske und Budrich.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Prinstein, M. J. (2007). Moderators of peer contagion: A longitudinal examination of depression socialization between adolescents and their best friends. *Journal of Clinical Child & Adolescent Psychology*, 36(2), 159–170.
- Prinstein, M. J. & Dodge, K. A. (Hrsg.) (2008). *Understanding peer influence in children and adolescents*. New York; NY: The Guilford Press.
- Rangel, C., & Lleras, C. (2010). Educational inequality in Colombia: Family background, school quality and student achievement in Cartagena. *International Studies in Sociology of Education*, 20(4), 291–317.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models*. Thousand Oaks, CA: Sage Publications.
- Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E. & Köller, O. (Hrsg.) (2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation*. Münster: Waxmann.
- 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(0), 63–72.

- 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*, 4(107), 1171–1185.
- Rjosk, C., Richter, D., Lüdtke, O. & Eccles, J. S. (2017). Ethnic composition and heterogeneity in the classroom: Their measurement and relationship with student outcomes. *Journal of Educational Psychology*. Advance online publication.
- Rumberger, R. W. & Palardy, G. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, 107(9), 1999–2045.
- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie*. Münster: Waxmann.
- Scharenberg, K. (2016). The interplay of social and ethnic classroom composition, tracking, and gender on students' school satisfaction. *Journal of Cognitive Education and Psychology*, 15(2), 320–346.
- Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221–279), New York, NY: Academic Press.
- Scott, J. F. (1971). *Internalization of norms*. Englewood Cliffs, N.J.: Prentice-Hall.
- Seaton, M., Marsh, H. W. & Craven, R. G. (2010). Big-fish-little-pond effect. Generalizability and moderation – Two sides of the same coin. *American Educational Research Journal*, 47(2), 390–433.
- Seaton, M., Parker, P., Marsh, H. W., Craven, R. G. & Yeung, A. S. (2014). The reciprocal relations between self-concept, motivation and achievement. Juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*, 34(1), 49–72.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring* (Beschluss der 350. Kultusministerkonferenz vom 11.06.2015).
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2016). *Bericht der Kultusministerkonferenz zur Integration von jungen Geflüchteten durch Bildung* (Beschluss der Kultusministerkonferenz vom 06.10.2016).

- Sherif, M. (1953). The concept of reference groups in human relations. In M. Sherif & M. V. Wilson (Eds.), *Group relations at the crossroads* (pp. 203–231). New York, NY: Harper and Row.
- Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*, 688.
- Skaalvik, E. M. & Valås, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *The Journal of Experimental Education*, *67*(2), 135–149.
- Stanat, P. (2006). Disparitäten im schulischen Erfolg. Forschungsstand zur Rolle des Migrationshintergrunds. *Unterrichtswissenschaft*, *36*(2), 98–124.
- Stanat, P., Schwippert, K. & Gröhlich, C. (2010). Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb. Längsschnittliche Überprüfung eines umstrittenen Effekts. In C. P. Allemann-Ghionda, P. Stanat, K. Göbel & C. Röhner (Hrsg.), *Migration, Sprache, Identität* (Zeitschrift für Pädagogik, 55. Beiheft, S. 147–164), Weinheim: Beltz.
- Stearns, E., Dodge, K. A. & Nicholson, M. (2008). Peer contextual influences on the growth of authority-acceptance problems in early elementary school. *Merrill-Palmer Quarterly*, *54*(2), 208–231.
- Steinberg, L. & Silverberg, S. B. (1986). The vicissitudes of autonomy in early adolescence. *Child Development*, *57*(4), 841–851.
- Stewart, E. A. (2003). School social bonds, school climate, and school misbehavior: A multilevel analysis. *Justice Quarterly*, *20*(3), 575–604.
- Stone, M. R. & Brown, B. B. (1999). Identity claims and projections: Descriptions of self and crowds in secondary school. *New Directions for Child and Adolescent Development*, *84*, 7–20.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*(1), 1–39.
- Tajfel, H. (1984) Intergroup relations, social myths and social justice. In H. Tajfel, C. Fraser & J. Jaspars (Eds.), *The social dimension: European developments in social psychology* (Vol. 2, pp. 695–715). Cambridge: Cambridge University Press.
- Thapa, A., Cohen, J., Guffey, S. & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, *83*(3), 357–385.

- Thomas, D. E. & Bierman, K. L. (2006). The impact of classroom aggression on the development of aggressive behavior problems in children. *Development and Psychopathology*, 18(2), 471–487.
- Thomas, D. E., Bierman, K. L. & Powers, C. J. (2011). The influence of classroom aggression and classroom climate on aggressive-disruptive behavior. *Child Development*, 82(3), 751–757.
- Thrupp, M. (1999). *Schools making a difference: School mix, school effectiveness, and the social limits of reform*. Buckingham, UK: Open University Press.
- Thrupp, M., Lauder, H. & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, 37(5), 483–504.
- Tiedemann, J. & Billmann-Mahecha, E. (2007). Zum Einfluss von Migration und Schulklassenzugehörigkeit auf die Übergangsempfehlung für die Sekundarstufe I. *Zeitschrift für Erziehungswissenschaft*, 10(1), 108–120.
- Tillmann, K.-J. (2007). Heterogenität – Homogenität. Was wir darüber wissen. *Schul-Management*, 38(5), 8–10.
- Tillmann, K.-J. (2008). Viel Selektion – wenig Leistung: Erfolg und Scheitern in deutschen Schulen. In R. Lehberger & U. Sandfuchs (Hrsg.), *Schüler fallen auf. Heterogene Lerngruppen in Schule und Unterricht* (S. 62–78). Bad Heilbrunn: Klinkhardt.
- Trautwein, U. & Baeriswyl, F. (2007). Wenn leistungsstarke Klassenkameraden ein Nachteil sind. Referenzgruppeneffekte bei Übertrittsentscheidungen. *Zeitschrift für pädagogische Psychologie*, 21(2), 119–133.
- Trautwein, U., Köller, O., Lüdtke, O. & Baumert, J. (2005). Student tracking and the powerful effects of opt-in courses on self-concept. Reflected-glory effects do exist after all. In H. W. Marsh, R. G. Craven & D. M. McInerney (Eds.), *New frontiers for self-research*, (Vol.2, pp. 307–327). Greenwich, CT: Information Age.
- 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. *The Journal of Educational Psychology*, 98(4), 788–806.
- Ulich, K. (2001). *Einführung in die Sozialpsychologie der Schule*. Weinheim: Beltz.

- van der Slik, F., Driessen, G. & Bot, K. de. (2006). Ethnic and socioeconomic class composition and language proficiency. A longitudinal multilevel examination in Dutch elementary schools. *European Sociological Review*, 22(3), 293–308.
- van Ewijk, R. & Sleegers, P. (2010a). The effect of peer socioeconomic status on student achievement. A meta-analysis. *Educational Research Review*, 5(2), 134–150.
- van Ewijk, R. & Sleegers, P. (2010b). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, 21(3), 237–265.
- van Landeghem, G., Pustjens, H., van Damme, J. & Munter, A. de. (2007). Impact of students' and their schoolmates' achievement motivation on the status and growth in math and language achievement of boys and girls across grades 7 through 8. *Psychologica Belgica*, 47(1), 5–29.
- van Landeghem, G., van Damme, J., Opdenakker, M.-C., de Fraine, B. & Onghena, P. (2002). The effect of schools and classes on noncognitive outcomes. *School Effectiveness and School Improvement*, 13(4), 429–451.
- von Saldern, M. & Littig, K.E. (1986). *Landauer Skalen zum Sozialklima*. Weinheim: Beltz.
- Walberg, H. (1990). A theory of educational productivity: Fundamental substance and method. In P. Vedder (Ed.), *Fundamental studies in educational research* (pp. 214–229). Lisse: Swets & Zeitlinger.
- Walgenbach, K. (2014). *Heterogenität – Intersektionalität – Diversity in der Erziehungswissenschaft*. Stuttgart: UTB/ Budrich.
- Walter, O. (2008). Ethno-linguale Kompositionseffekte in neunten Klassen. Befunde aus der Klassenstichprobe von PISA 2006. In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (Zeitschrift für Erziehungswissenschaft, Sonderheft 10, S. 169–184). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Warren, K., Schoppelrey, S., Moberg, D. P. & McDonald, M. (2005). A model of contagion through competition in the aggressive behaviors of elementary school students. *Journal of Abnormal Child Psychology*, 33(3), 283–292.
- Watermann, R., Klingebiel, F. & Kurtz, T. (2010). Die motivationale Bewältigung des Grundschulübergangs aus Schüler- und Elternsicht. In K. Maaz, J. Baumert, C. Gresch & M. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule – Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 355–384). Bonn: BMBF.

- Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, 22, 366–389.
- Wigfield, A. & Eccles, J. S. (1990). Test anxiety in the school setting. In M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 237–250). Boston, MA: Springer US.
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. & Davis-Kean, P. (2006). Development of achievement motivation. In W. Damon, R. M. Lerner & N. Eisenberg (Eds.), *Handbook of child psychology, social, emotional, and personality development* (Vol. 6, 3rd ed., pp. 933–1002). Hoboken, NJ: Wiley.
- Wilkinson, I. A., Hattie, J., Parr, J. M., Townsend, Michael A. R., Fung, I. Y., Ussher, C. et al. (2000). *Influence of peer effects on learning outcomes. A review of the literature*. Auckland: Auckland UniServices Limited.
- Wiswede, G. (2004). *Sozialpsychologie-Lexikon*. Berlin: Walter de Gruyter GmbH & Co KG.
- Yudron, M. & Jones, S. M. (2016). Developmental trajectories of children's social competence in early childhood: The role of the externalizing behaviors of their preschool peers. *Journal of Cognitive Education and Psychology*, 15(2), 268–292.
- Zeidner, M. (2014). Anxiety in education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International Handbook of Emotions in Education* (pp. 265–288). New York, NY: Routledge.
- Zurbriggen, C. (2016). *Schulklasseneffekte: Schülerinnen und Schüler zwischen komparativen und normativen Einflüssen*. Wiesbaden: Springer VS.

2

Teilstudie 1:

Der Effekt der Schulzusammensetzung auf die Entwicklung der Schulzufriedenheit und des normverletzenden Verhaltens von Schülerinnen und Schülern

Hinweis: Dies ist die erste Autorenversion des Manuskripts, welche zur Publikation im *Journal for Educational Research Online/Journal für Bildungsforschung Online* eingereicht wurde.

Stäbler, F. & Dumont, H. (under review). Der Effekt der Schulzusammensetzung auf die Entwicklung der Schulzufriedenheit und des normverletzenden Verhaltens von Schülerinnen und Schülern.

2 Der Effekt der Schulzusammensetzung auf die Entwicklung der Schulzufriedenheit und des normverletzenden Verhaltens von Schülerinnen und Schülern

Zusammenfassung

In der ersten Teilstudie wurde an einer Stichprobe von 2 193 Schülerinnen und Schülern untersucht, ob die individuelle Entwicklung von Schulzufriedenheit und normverletzendem Verhalten mit der Zusammensetzung der Schülerschaft einer Schule in Zusammenhang steht. Es wurde sowohl die leistungsbezogene und soziale Zusammensetzung betrachtet, als auch die mittlere Schulzufriedenheit sowie das mittlere Ausmaß an normverletzendem Verhalten an einer Schule. Die Ergebnisse der Mehrebenenanalysen weisen auf verschiedene Kompositionseffekte hin. Beispielsweise waren Lernende nach einem Schuljahr zufriedener mit der Schule, je leistungsstärker die von ihnen besuchte Schule war. Für die soziale Komposition der Schule ließ sich tendenziell beobachten, dass sich Schülerinnen und Schüler weniger normverletzend verhielten, je höher der mittlere HISEI war. Wurden die beiden psychosozialen Merkmale selbst als Kompositionsvariablen konzeptualisiert, ergab sich, dass die Schulzufriedenheit der Lernenden von der 9. zur 10. Jahrgangsstufe stieg, je zufriedener die Mitschülerinnen und Mitschüler im Durchschnitt mit der Schule waren. Ein hohes Maß an normverletzendem Verhalten an der Schule erhöhte tendenziell das normverletzende Verhalten der einzelnen Schülerinnen und Schüler im Laufe eines Schuljahres.

Schlagworte: Kompositionseffekte, Normverletzendes Verhalten, Peer contagion, Schulzufriedenheit

Einleitung

Forschungsanlass und Ziele

Ähnlich wie sich Eltern Gedanken darüber machen, mit wem ihre Kinder ihre Freizeit verbringen, beschäftigt es Eltern auch, von welchen Mitschülerinnen und Mitschülern ihre Kinder in der Schule umgeben sind. Wenn es darum geht, eine Schule auszusuchen, werden in der Regel Schulen bevorzugt, die von leistungsstarken Schülerinnen und Schülern besucht werden und in denen möglichst wenige Kinder mit Zuwanderungshintergrund oder aus sozial schwachen Elternhäusern lernen (Jurczok & Lauterbach, 2014; Klinge, 2016). Auch wenn diese Einstellung einer gesellschaftlich wünschenswerten sozialen Durchmischung entgegensteht, lassen sich für die Annahmen vieler Eltern, dass sich ihre Kinder an Schulen mit einer leistungsstarken und sozial begünstigten Schülerschaft besser entwickeln, tatsächlich empirische Befunde finden. So wurde im Rahmen der Forschung zu sogenannten *Kompositionseffekten* vielfach empirisch belegt, dass Schülerinnen und Schüler, die von einer leistungsstarken und sozial privilegierten Schülerschaft umgeben sind, bessere Leistungen zeigen als Lernende, deren Mitschülerinnen und Mitschüler aus sozial benachteiligten Familien stammen und eher schwache Leistungen aufweisen (siehe im Überblick Dumont, Neumann, Maaz & Trautwein, 2013; siehe Metaanalyse von van Ewijk & Slegers, 2010b; für empirische Studien siehe Burns & Mason, 2002; Danhier & Martin, 2014; Hanushek, Kain, Markman & Rivkin, 2003; Hornstra, van der Veen, Peetsma & Volman, 2014; Justice, Logan, Lin & Kaderavek, 2014; Marks, 2010; Opdenakker, van Damme, de Fraine, van Landeghem & Onghena, 2002). Im gegliederten deutschen Bildungssystem spielen solche Kompositionseffekte eine besondere Rolle, da die Verteilung von Schülerinnen und Schülern auf verschiedene Schulformen auf Basis ihrer Leistungen auch im Hinblick auf die soziale Herkunft zu großen Unterschieden in der Schülerschaft führt. Während es zu den Auswirkungen solcher Kompositionsunterschiede auf die Leistungsentwicklung schon eine Reihe von empirischen Befunden gibt, existiert bislang vergleichsweise wenig empirisches Wissen über die Auswirkungen einer (un-) günstigen Komposition einer Schule auf die *psychosoziale Entwicklung* von Schülerinnen und Schülern, wie bereits an anderer Stelle angemerkt wurde (Belfi, Goos, de Fraine & van Damme, 2012; Hornstra et al., 2014). Allerdings ist davon auszugehen, dass insbesondere die psychosoziale Entwicklung im Schulkontext stark von sozialen Interaktionsprozessen und damit von den Mitschülerinnen und Mitschülern geprägt wird (siehe Müller & Zurbriggen, 2016). Vor diesem Hintergrund ist weiterhin denkbar, dass nicht nur die leistungs- und herkunftsbezogene Komposition, sondern

auch die Zusammensetzung der Schülerschaft hinsichtlich psychosozialer Charakteristika auf die psychosoziale Entwicklung wirkt.

In der vorliegenden Studie wird daher die Analyse von Kompositionseffekten durch eine psychosoziale Perspektive erweitert. Dazu werden zwei für den Schulkontext zentrale psychosoziale Variablen betrachtet: Schulzufriedenheit und normverletzendes Verhalten. Betrachtet werden diese Variablen zum einen als individuelle Merkmale von Schülerinnen und Schülern, die von der Zusammensetzung der Schule beeinflusst werden können. Zum anderen werden sie in aggregierter Form als Kompositionsmerkmale konzeptualisiert, um zu überprüfen, inwiefern sie in kumulierter Weise einen Effekt auf die individuelle Entwicklung ausüben. Darüber hinaus werden auch die „klassischen“ Kompositionsmerkmale, d.h. die Zusammensetzung in Hinblick auf Leistung und familiärer Herkunft, als auch die Lernentwicklung der Kinder in Form von Noten in den Blick genommen. Die Analyse der Kompositionseffekte erfolgt jeweils unter besonderer Berücksichtigung der unterschiedlichen Schulformen.

Effekte der Klassen- und Schulkomposition auf die individuelle Leistungsentwicklung

So individuell wie jedes Kind ist, so individuell sind auch Klassen und Schulen. Sie bilden, je nachdem, welche Eigenschaften die Lernenden mitbringen, ganz bestimmte Charakteristiken aus, die sie von anderen Klassen und Schulen unterscheiden. Weist ein Individualmerkmal über seinen intraindividuellen Effekt hinaus, eine Wirkung durch ein kumulatives Auftreten innerhalb einer Klasse oder Schule auf, wird dies als Kompositionseffekt bezeichnet (Harker & Tymms, 2004).

Als relevante Kompositionseffekte haben sich der Effekt der leistungsbezogenen Komposition sowie der sozioökonomischen und zuwanderungsbezogenen Zusammensetzung einer Klasse oder Schule auf die individuelle Leistung herausgestellt. So kommen diesbezügliche Untersuchungen weitgehend übereinstimmend zu dem Ergebnis, dass sich die Zugehörigkeit zu einer leistungsstarken Klasse positiv auf die individuelle Leistungsentwicklung auswirkt, während das Umfeld einer Klasse mit niedriger mittlerer Leistung mit negativen Effekten auf die eigene Leistungsentwicklung einhergeht (Danhier & Martin, 2014; Hanushek et al., 2003; Marks, 2010). Im Falle der sozioökonomischen und zuwanderungsbezogenen Zusammensetzung zeigen sich negative Effekte auf die Leistungsentwicklung von Schülerinnen und Schülern, wenn sie in einer Klasse oder Schule lernen, in denen die Schülerschaft überwiegend aus sozial benachteiligten Familien stammt (van Ewijk & Slegers, 2010b). Für die zuwanderungsbezogene Zusammensetzung der Schülerschaft ist die Befundlage weniger eindeutig; u. a. deswegen, weil sie häufig mit der

sozialen Zusammensetzung konfundiert ist. Hier gibt es sowohl empirische Belege für Kompositionseffekte (van Ewijk & Sleegers, 2010a), als auch Studien, die keinen Einfluss der ethnischen und zugewanderungsbezogenen Komposition finden (Rumberger & Palardy, 2005; van der Slik, Driessen & Bot, 2006).

Für das deutsche Bildungssystem sind Kompositionseffekte aufgrund der Mehrgliedrigkeit des Sekundarschulsystems von besonderer Bedeutung. Die Aufteilung der Schülerinnen und Schüler auf verschiedene Schulformen auf Basis ihrer Leistungsfähigkeit hat aufgrund des empirisch hohen Zusammenhangs zwischen Leistung und familiärer Herkunft zur Folge, dass sich die Schülerschaften von Schulen unterschiedlicher Schulformen nicht nur hinsichtlich ihrer Leistungsstärke, sondern auch hinsichtlich ihres sozioökonomischen Hintergrundes und ihres Migrationshintergrundes stark unterscheiden. Neben Kompositionseffekten können Schulformen jedoch auch durch sogenannte *Institutionseffekte* einen Einfluss auf die Lern- und Leistungsentwicklung von Lernenden haben. So unterscheiden sich Schulformen nicht nur hinsichtlich ihrer Schülerschaft, sondern auch hinsichtlich ihrer Stundentafeln, Lehrpläne oder durch die unterschiedliche Ausbildung der Lehrerinnen und Lehrer (Baumert, Maaz, Stanat & Watermann, 2009). Diese institutionellen Unterschiede sind häufig konfundiert mit der Zusammensetzung der Schülerschaft, sodass Kompositionseffekte nur schwer von institutionellen Effekten zu trennen sind. Entsprechend sollte bei der Analyse von Kompositionseffekten im deutschen Schulsystem immer die Schulform mitberücksichtigt werden (Dumont et al., 2013).

Effekte der Klassen- und Schulkomposition auf die individuelle psychosoziale Entwicklung

Wie die berichteten Befunde über relevante Kompositionsmerkmale zeigen, wurde in der Kompositionseffektforschung bisher hauptsächlich fokussiert, wie die individuelle *Leistung* durch die Komposition beeinflusst wird. Nur wenige Studien betrachteten Kompositionseffekte auf die *psychosoziale* Entwicklung der Schülerinnen und Schüler in der Schule (siehe z. B. Hornstra et al., 2014; Rjosk, Richter, Hochweber, Lüdtke & Stanat, 2015; Scharenberg, 2016). Da das Ziel der Beschulung nicht nur in der Förderung der Leistung besteht, sondern multiple Ziele verfolgt werden sollen, ist es wichtig, neben den Effekten der Komposition auf die Leistungsentwicklung solche Effekte auch auf andere Merkmale wie die psychosoziale Entwicklung zu untersuchen (siehe auch Belfi et al., 2012; Hornstra et al., 2014).

Einige Studien haben sich bereits dieser Fragestellung gewidmet. Sie geben Hinweise darauf, dass eine Schülerschaft mit hohen Leistungen und begünstigtem sozioökonomischen

Hintergrund zwar die Leistungsentwicklung der einzelnen Lernenden fördern, die psychosoziale Entwicklung jedoch beeinträchtigen kann. Ein bekanntes Phänomen in dieser Hinsicht ist der sogenannte *Big-Fish-Little-Pond-Effekt* (BFLPE; Marsh, 1987; siehe im Überblick Marsh et al. 2008), der methodisch gesehen auch als Kompositionseffekt beschrieben werden kann, da er den Effekt der *mittleren Leistung der Schülerschaft* auf das individuelle akademische Selbstkonzept bei Kontrolle der individuellen Leistung darstellt. Er besagt, dass Lernende, die in leistungsstarken Klassen unterrichtet werden, aufgrund von sozialen Vergleichsprozessen ihre eigenen Fachleistungen geringer einschätzen als Lernende des gleichen Leistungsniveaus in leistungsschwächeren Klassen. Der BFLPE wird vor allem in Bezug auf das akademische Selbstkonzept betrachtet. Es gibt jedoch auch empirische Hinweise darauf, dass dieser Effekt auf andere psychosoziale Variablen übertragbar ist, wie z. B. auf das fachspezifische Interesse oder die beruflichen Aspirationen von Schülerinnen und Schülern (Köller, Schnabel & Baumert, 2000; Marsh, 1991).

Für die Wirkung der *sozioökonomischen Komposition* gibt es sowohl empirische Befunde für positive Effekte als auch Befunde für negative Effekte auf psychosoziale Schülermerkmale. Scharenberg (2016) ermittelte, dass Lernende zufriedener mit der Schule waren, wenn die Schülerschaft vorwiegend aus sozial privilegierten Elternhäusern stammte. In einer Studie von Hornstra et al. (2014) konnte dagegen gezeigt werden, dass Lernende in Bezug auf ihre Aufgabenorientierung profitierten, wenn mehr Mitschülerinnen und Mitschüler aus sozial benachteiligten Schichten kamen. Das Interesse am eigentlichen Wissenszuwachs beim Lösen von Aufgaben (statt die Aufgabe nur besser als andere lösen zu wollen) war stärker in Klassen, in denen ein großer Teil der Schülerschaft aus sozial benachteiligten Familienhäusern stammte als in Klassen mit einem hohen Anteil von Lernenden aus sozial privilegierten Elternhäusern.

Für den Effekt der *zuwanderungsbezogenen Komposition* zeigten sich in bisherigen Studien positive Effekte auf psychosoziale Schülermerkmale. Beispielsweise waren der bereits genannten Studie von Scharenberg (2016) zufolge, Schülerinnen und Schüler im Verlauf der 6. zur 8. Jahrgangsstufe zufriedener in bzw. mit der Schule, wenn sie eine Klasse mit einem höheren Anteil von Lernenden mit Zuwanderungshintergrund besuchten. In einer Studie von Rjosk et al. (2015) entwickelte sich die intrinsische Motivation der Schülerinnen und Schüler vom Beginn bis zum Ende der 9. Jahrgangsstufe stärker, wenn ein höherer Anteil der Lernenden in der Klasse eine andere Erstsprache als Deutsch aufwies oder zusätzlich zu Deutsch eine weitere Erstsprache erlernt hatte. Diese positiven Effekte der Komposition hinsichtlich der Erstsprache – über die eigene Erstsprache hinaus – auf psychosoziale

Merkmale, wurden damit erklärt, dass Lehrpersonen in Klassen mit einer solchen Komposition möglicherweise stärker gefordert sind und dadurch auch kompetenter darin werden, ihren Unterricht an die individuellen Bedürfnisse aller Lernenden anzupassen (als „specialization hypothesis“ bekannt; Hornstra et al., 2014).

Effekte der psychosozialen Komposition der Klasse oder Schule

Ähnlich wie die individuelle Entwicklung von psychosozialen Merkmalen im Kontext von Kompositionseffekten bislang kaum betrachtet wurde, ist auch die Untersuchung von psychosozialen Merkmalen als Kompositionsmerkmale selten. In einer der wenigen Studien mit diesem Fokus betrachteten van Landeghem, van Damme, Opdenakker, de Fraine und Onghena (2002) die mittlere Leistungsmotivation der Klasse und der Schule als Beispiel für psychosoziale Kompositionsmerkmale: Auf Klassenebene führte eine höhere mittlere Leistungsmotivation zu einer höheren sozialen Integration der Schülerinnen und Schüler; für die mittlere Leistungsmotivation der Schule gab es Hinweise, dass es die Lernenden positiv beeinflusst, das Beste für die Schule zu geben. Müller, Hofmann, Fleischli und Studer (2015) fanden des Weiteren heraus, dass Lernende vom Beginn bis zum Ende der 7. Jahrgangsstufe mehr schulisches Fehlverhalten zeigten, wenn sie sich in einer Klasse befanden, in der das Ausmaß solchen Verhaltens insgesamt hoch war. Eine andere Studie liefert Hinweise dafür, dass dieser negative Effekt auf das individuelle schulische Fehlverhalten stärker vom Ausmaß des schulischen Fehlverhaltens auf *Schulebene* als vom Ausmaß auf *Klassenebene* ausgeht (Gommans, Stevens, ter Bogt & Cillessen, 2016). Yudron und Jones (2016) ermittelten für die Vorschule, dass sich eine hohe Ausprägung an Fehlverhalten innerhalb der Lerngruppe (hier die Klasse) auch positiv auswirken kann; in dem Fall auf die Entwicklung der Sozialkompetenz der Schülerinnen und Schüler.

Bei diesen Befunden handelt es sich um erste Ergebnisse, die Anlass zur Annahme geben, dass verschiedenste psychosoziale Merkmale der Schülerschaft einen Einfluss auf die schulische Entwicklung der Lernenden nehmen können. Dieser Einfluss könnte z. B. dadurch entstehen, dass das Lehrerverhalten durch die psychosoziale Konstitution der Gruppe verändert wird und dieses veränderte Lehrerverhalten wiederum auf die individuelle leistungsbezogene und psychosoziale Entwicklung wirkt (Müller & Zurbriggen, 2016).

Des Weiteren sind sogenannte *Peer-Effekte* durch die psychosoziale Komposition als Erklärungsmechanismen denkbar. Schule ist nicht nur ein Ort des Lernens, sondern auch ein Ort der Gemeinschaft, an dem Kinder und Jugendliche einen Großteil ihres Tages verbringen und miteinander interagieren. Durch verschiedene soziale Interaktionsmechanismen können

sich Schülerinnen und Schüler gegenseitig in Verhaltensweisen und normativen Einstellungen gegenüber der Schule beeinflussen (Bandura, 1986). Ein zentraler Mechanismus ist der sogenannte *Peer-Contagion-Effekt*, der als mehr oder weniger unbewusst stattfindende Übertragung („Ansteckung“) von Verhalten und Einstellungen zwischen Kindern und Jugendlichen definiert wird (Dishion & Tipsord, 2011). Ist beispielsweise in der Peergruppe ein bestimmtes psychosoziales Merkmal wie aggressives Verhalten stark ausgeprägt, so sind Kinder und Jugendliche eher geneigt, dieses Verhalten zu übernehmen (Müller, Hofmann & Arm, 2016). Peer-Contagion-Effekte wurden bisher vor allem für normbrechendes Verhalten in selbstgewählten Freundeskreisen untersucht (Dishion & Tipsord, 2011). Wenig Evidenz existiert über solche Prozesse innerhalb des Schulsettings, in dem nicht alle Kontakte selbstgewählt sind. Besonders rar sind hier Befunde für das Jugendalter sowie für positive Verhaltensweisen und Einstellungen gegenüber der Schule (siehe auch Gehrman, 2015; Müller et al., 2016).

Schulklima als Ausdruck kollektiver Normen, Einstellungen und gängiger Verhaltensweisen

Weitere Hinweise dafür, dass psychosoziale Merkmale als Kompositionsmerkmale über die Individualebene hinaus Einfluss auf individuelle Schülermerkmale haben, kommen aus der Schulklimaforschung. Unter dem Begriff des *Schulklimas* wird die spezifische Atmosphäre einer Schule bezeichnet, die sich aus den jeweils vorherrschenden kollektiven Normen, Einstellungen gegenüber der Schule und Gebräuchen des Miteinanders ergibt (Berkowitz et al., 2015; de Pedro, Gilreath & Berkowitz, 2016). Es konnte gezeigt werden, dass ein positives Schulklima einen förderlichen Effekt auf Noten (Wang et al., 2014), Leistungen in standardisierten Tests (siehe im Überblick Espelage, Hong, Rao & Low, 2013; siehe auch Thapa, Cohen, Guffey & Higgins-D'Alessandro, 2013) und auf das Wohlbefinden der Lernenden haben kann (Ruus et al., 2007; Shochet, Dadds, Ham & Montague, 2006). Auch aggressives Verhalten kann durch ein positives Schulklima vermindert werden (siehe Metaanalyse von Steffgen, Recchia & Viechtbauer, 2013). Als Erklärung für Effekte eines positiven Schulklimas wird angenommen, dass Lernende über mehr Ressourcen für Lerntätigkeiten und für lernbezogene Interaktionen mit Lehrpersonen und Mitlernenden verfügen, wenn sie die Atmosphäre in der Schule, z. B. hinsichtlich ihrer Sicherheit oder ihrer Beziehungen, positiv wahrnehmen (Benbenishty, Astor, Roziner & Wrabel, 2016).

Operationalisiert wird die Einschätzung des Schulklimas in der Regel über die individuelle Einschätzung des Klimas an der eigenen Schule durch die Schülerinnen und Schüler (siehe im Überblick Cohen, McCabe, Michelli & Pickeral, 2009; siehe im Überblick

auch Eder, 2002). Dabei spielt unter anderem eine große Rolle, wie zufrieden Lernende in der Schule sind oder wie häufig normverletzendes Verhalten in der Schule geschieht (Berkowitz et al., 2015; Fraser 1982; Eder & Mayr, 2000; von Saldern & Littig, 1987).

Die vorliegende Studie

Da Schule nicht nur die Lernentwicklung, sondern auch die psychosoziale Entwicklung von Lernenden unterstützen soll, ist das übergeordnete Ziel der vorliegenden Studie, die Analyse von Kompositionseffekten um psychosoziale Aspekte zu erweitern. Dies soll am Beispiel von *Schulzufriedenheit* und *normverletzendem Verhalten* geschehen, da sich die beiden in ihrer Kumulation in einer Lerngruppe als geteilte Einstellungen gegenüber Schule und als gängige Verhaltensweisen formieren. Somit können sie als zwei Variablen betrachtet werden, für die davon auszugehen ist, dass sie für das Schulklima relevant sind. Dabei wird die Entwicklung von Schulzufriedenheit und normverletzendem Verhalten von der 9. zur 10. Jahrgangsstufe in den Blick genommen.

Im Rahmen der ersten Fragestellung wird untersucht, ob die individuelle Entwicklung der Schulzufriedenheit und des normverletzenden Verhaltens mit der Schülerschaft einer Schule in Zusammenhang steht. Die Schülerschaft wird dabei hinsichtlich der Mathematikleistung, des sozioökonomischen Status der Eltern und des Zuwanderungshintergrundes betrachtet. Im Rahmen der zweiten Fragestellung werden die beiden psychosozialen Merkmale Schulzufriedenheit und normverletzendes Verhalten zur Charakterisierung der Schülerschaft herangezogen. Untersucht wird, ob auch von der mittleren Schulzufriedenheit und dem Ausmaß an normverletzendem Verhalten an einer Schule ein Kompositionseffekt auf die individuelle Entwicklung der Schulzufriedenheit und des normverletzenden Verhaltens ausgeht.

Ergänzend wird für beide Fragestellungen auch der Effekt der Kompositionsmerkmale auf die Entwicklung der Mathematiknote der Schülerinnen und Schüler von der 9. zur 10. Jahrgangsstufe analysiert. Aufgrund der hohen Konfundierung der Komposition der Schülerschaft einer Schule mit der Schulformzugehörigkeit wird für beide Fragestellungen zudem untersucht, ob die Kompositionseffekte bestehen bleiben, wenn für die Schulformzugehörigkeit zusätzlich kontrolliert wird. Da davon auszugehen ist, dass die Entwicklung der Schulzufriedenheit, des normverletzenden Verhaltens und der Schulnoten nicht nur von der Schülerschaft der Schule abhängt, sondern auch von weiteren individuellen Merkmalen, werden die Kompositionseffekte auch unter Kontrolle einer Reihe von Schülermerkmalen untersucht.

Methode

Stichprobe

Die Daten der vorliegenden Stichprobe stammen aus einer Kohorte einer längsschnittlich angelegten Multikohortenstudie in der Stadt Berlin (für weiterführende Informationen zur Studie siehe Maaz, Baumert, Neumann, Becker & Dumont, 2013). Die 2 193 Schülerinnen und Schüler der betreffenden Kohorte wurden in der 9. und 10. Jahrgangsstufe (in den Schuljahren 2010/11 sowie 2011/12) befragt. Sie besuchten 116 Schulen: 29 Hauptschulen, 23 Realschulen, 35 Gesamtschulen und 29 Gymnasien. Die Stichprobenziehung erfolgte nach einem mehrstufig stratifizierten randomisierten Clusterdesign, wobei zunächst Schulen innerhalb der verschiedenen Schulformen und in einem weiteren Schritt Schülerinnen und Schüler der 9. Jahrgangsstufe innerhalb von Schulen gezogen wurden. In der vorliegenden Stichprobe wurden im Durchschnitt 25 Schülerinnen und Schüler pro Schule untersucht.

Instrumente

Kompetenzen im Fach Mathematik. Die schulischen Kompetenzen im Fach Mathematik wurden in der 9. Jahrgangsstufe anhand eines standardisierten Leistungstests erhoben, dessen Aufgaben aus der PISA-Studie von 2006 stammen (Prenzel et al., 2007, 2008). Die 48 Testaufgaben, die in zwei 60-minütigen Testsitzungen zu bearbeiten waren, wurden im Multi-Matrix-Design administriert. Als Leistungstestcores dienten Weighted-Likelihood-Schätzer (WLE; Warm, 1989), die über ein probabilistisches Item-Response-Modell skaliert wurden (Rasch, 1960). Die Leistungstestcores wurden für die vorliegenden Analysen zur besseren Interpretierbarkeit z-standardisiert.

Noten im Fach Mathematik. Sowohl für das Schuljahresende der 9. als auch der 10. Jahrgangsstufe lagen die Mathematiknoten der Lernenden vor. Die Endjahresnoten der 9. Jahrgangsstufe basierten auf Angaben der Schulen ($M_{9. Jgst.} = 3.55$, $SD_{9. Jgst.} = 1.13$), während die Noten der 10. Jahrgangsstufe direkt den Endjahreszeugnissen entnommen wurden ($M_{10. Jgst.} = 3.56$, $SD_{10. Jgst.} = 1.16$). Die Intraklassenkorrelation lag bei $ICC = .17$ (9. Jgst.) und $ICC = .20$ (10. Jgst.). Zur besseren Interpretierbarkeit wurden die Noten in den vorliegenden Analysen z-standardisiert und so transformiert, dass höhere Werte bessere Noten indizieren.

Sozioökonomischer Status der Eltern. Als Indikator für die sozioökonomische Herkunft der Lernenden wurde aus den schülerperzipierten Angaben zur derzeitigen Berufstätigkeit ihrer Eltern der *International Socio-Economic Index* gebildet (ISEI; Ganzeboom, de Graaf & Treiman, 1992), wobei von den Elternteilen der jeweils höhere sozioökonomische Status

(HISEI) herangezogen wurde. Der ISEI bzw. der HISEI hat einen Wertebereich von 16 bis 90, wobei höhere Werte einem höheren sozioökonomischen Status entsprechen. In der untersuchten Stichprobe stammten die Lernenden zu einem überwiegenden Teil aus Familien mit mittlerem Status ($M = 45.63$, $SD = 20.97$). Für die Analysen wurde der HISEI z-standardisiert.

Zuwanderungshintergrund. Aus den Angaben der Lernenden, in welchem Land ihre Eltern geboren waren, wurde eine dichotome Variable zur Erfassung des Zuwanderungshintergrundes der Kinder gebildet. Sie gibt an, ob mindestens ein Elternteil in einem anderen Land als Deutschland geboren wurde; als Referenzkategorie dienten Lernende, deren Eltern in Deutschland geboren wurden. 52.80% der Lernenden in der untersuchten Stichprobe hatten einen Zuwanderungshintergrund.

Schulzufriedenheit. Die Schulzufriedenheit der Lernenden wurde an beiden Messzeitpunkten über jeweils vier Items gemessen (Beispielitem: „In unserer Schule fühle ich mich wohl.“, zu beantworten auf einer vierstufigen Likert-Skala von 1 = *stimmt überhaupt nicht* bis 4 = *stimmt völlig*). Die Items sind an die Operationalisierung von Bos und Pietsch (2006) angelehnt. Die Skala aus den vier Items wies sowohl in der 9. als auch in der 10. Jahrgangsstufe eine zufriedenstellende interne Konsistenz auf (Cronbachs $\alpha_{9. Jgst.} = .83$, Cronbachs $\alpha_{10. Jgst.} = .86$). In der betrachteten Stichprobe entsprach der Mittelwert der Schulzufriedenheit ungefähr dem theoretischen Mittelwert ($M_{9. Jgst.} = 2.57$, $SD_{9. Jgst.} = .74$; $M_{10. Jgst.} = 2.32$, $SD_{10. Jgst.} = .79$). Die Intraklassenkorrelation lag bei ICC = .06 (9. Jgst.) und ICC = .07 (10. Jgst.). Für die Analysen wurde die Schulzufriedenheit z-standardisiert.

Normverletzendes Verhalten. Das normverletzende Verhalten wurde auf Basis von Schülerangaben erhoben. Die Lernenden wurden zu beiden Messzeitpunkten gefragt, ob sie schon einmal Schuleigentum zerstört, eine Remperei begonnen, mit einer Lehrperson gestritten oder wegen einer Störung einen Klassenbucheintrag bekommen haben (zu beantworten auf einer fünfstufigen Likert-Skala von 1 = *noch nie* bis 5 = *ja, häufiger*). Die vier Items, die für die vorliegende Studie entwickelt wurden, wurden an beiden Messzeitpunkten zu einer Skala zusammengefasst, die jeweils eine zufriedenstellende interne Konsistenz aufwies (Cronbachs $\alpha_{9. Jgst.} = .83$, Cronbachs $\alpha_{10. Jgst.} = .83$). Das Niveau des normverletzenden Verhaltens ist in der betrachteten Stichprobe eher als gering einzustufen, da der Mittelwert in der 9. Jahrgangsstufe bei 1.86 ($SD = .85$) und in der 10. Jahrgangsstufe bei 1.79 ($SD = .87$) lag. Die Intraklassenkorrelation lag bei ICC = .10 (9. Jgst.) und ICC = .09 (10. Jgst.). Gemäß der ICCs gingen 10% (9. Jgst.) bzw. 9% (10. Jgst.) der Varianz im

normverletzenden Verhalten auf Unterschiede zwischen Schulen zurück. Auch das normverletzende Verhalten wurde für die bessere Interpretierbarkeit z-standardisiert.

Kompositionsmerkmale. Die Mathematikleistung, der HISEI, der Zuwanderungshintergrund sowie die beiden psychosozialen Merkmale Schulzufriedenheit und normverletzendes Verhalten wurden durch Aggregation auf Schulebene zusätzlich auch als Kompositionsmerkmale betrachtet. Dazu wurde jeweils die mittlere Ausprägung des Merkmals auf Schulebene berechnet. Eine Ausnahme bildet die zuwanderungsbezogene Zusammensetzung der Schule, die in der Weise generiert wurde, dass sie den Anteil der Kinder mit Zuwanderungshintergrund pro Schule darstellt. Zur besseren Interpretation erfolgte die Aggregation auf Basis der z-standardisierten Merkmale, wobei nach der Aggregation erneut auf Schulebene z-standardisiert wurde. Zwischen den berücksichtigten Kompositionsmerkmalen bestanden mittlere bis hohe Korrelationen (siehe Tabelle 2.1).

Tabelle 2.1

Korrelationsmatrix der betrachteten Variablen auf Schulebene

Variablen	1	2	3	4
1 Matheleistung 9. Jgst.	1			
2 HISEI	.90***	1		
3 Zuwanderungshintergrund	-.66***	-.68***	1	
4 Schulzufriedenheit 9. Jgst.	.57***	.55***	-.34***	1
5 Normverletzendes Verhalten 9. Jgst.	-.71***	-.63***	.42***	-.52***

Anmerkungen. * $p < .05$, ** $p < .01$, *** $p < .001$; Jgst. = Jahrgangsstufe.

Schulform. Es wurden drei Dummy-Variablen generiert, die jeweils indizieren, ob die Lernenden eine Realschule, eine Gesamtschule oder ein Gymnasium besuchen. Die Hauptschule dient dabei als Referenzgruppe.

Kontrollvariablen. In den erweiterten Modellen zur Analyse der Kompositionseffekte wurden die oben bereits beschriebenen Variablen Mathematikleistung, HISEI und Zuwanderungshintergrund als Kontrollvariablen auf Schülerebene eingeführt. Zudem wurde für das Geschlecht (46.10% Mädchen) kontrolliert, da Unterschiede zwischen den Geschlechtern hinsichtlich der Schulzufriedenheit und dem normverletzenden Verhalten zu vermuten sind (Hascher & Hagenauer, 2011; Müller et al., 2016).

Analysestrategie

Um die einzelnen Kompositionseffekte zu ermitteln, die jeweils im Fokus der beiden Fragestellungen stehen, wurden Mehrebenenanalysen (Raudenbush & Bryk, 2002) in Mplus 7.11 (Muthén & Muthén, 1998–2010) mit Schülerinnen und Schülern auf Ebene 1 und mit Schulen auf Ebene 2 berechnet. Entsprechend der Definition von Kompositionseffekten nach Harker und Tymms (2004) wurde der Effekt eines auf Schulebene aggregierten Individualmerkmals (unter Kontrolle des gleichen Merkmals auf Individualebene) auf das gleiche Merkmal zu einem späteren Messzeitpunkt bzw. auf ein weiteres Individualmerkmal untersucht. Bei Letzterem wurde immer auch der Ausgangswert der abhängigen Variable kontrolliert. Dieser Logik folgend wurden für die erste Fragestellung pro abhängiger Variable (Schulzufriedenheit, normverletzendes Verhalten und Mathematiknote) jeweils drei Kompositionsmodelle geschätzt: der Kompositionseffekt von der mittleren Mathematikleistung, vom mittleren HISEI und vom Anteil an Lernenden mit Zuwanderungshintergrund an der Schule. Für die zweite Fragestellung wurde jeweils der Kompositionseffekt von der mittleren Schulzufriedenheit und dem Ausmaß an normverletzendem Verhalten an der Schule auf die drei abhängigen Variablen (Schulzufriedenheit, normverletzendes Verhalten, und Mathematiknote) geschätzt. Von einem Modell, in dem alle Kompositionsmerkmale gleichzeitig spezifiziert werden, wurde aufgrund der hohen Interkorrelationen abgesehen (siehe Tabelle 2.1).

Die so berechneten Kompositionseffekte sind in den Modellen mit der Kennzeichnung *a* in den Tabellen 2.2 bis 2.4 zu finden. Da Schulform und Komposition im deutschen Schulsystem hoch konfundiert sind (Dumont et al., 2013), wurde in den Modellen mit der Kennzeichnung *b* in den Tabellen 2.2 bis 2.4 zusätzlich die Schulform mit der Hauptschule als Referenzkategorie auf Ebene zwei aufgenommen. Eine noch strengere Prüfung der Kompositionseffekte findet in den Modellen mit der Kennzeichnung *c* in den Tabellen 2.2 bis 2.4 statt, in denen zusätzlich zur Schulform die Schülermerkmale Mathematikleistung, HISEI, Zuwanderungshintergrund und Geschlecht auf Ebene eins als Kontrollvariable spezifiziert wurden.

Der Anteil fehlender Werte der einbezogenen Variablen belief sich auf eine Spanne von 10.50% in den Mathematiknoten der 9. Jahrgangsstufe und 38.10% im HISEI. Um fehlende Angaben zu ersetzen, wurde der Ansatz der *Multivariate Imputation by Chained Equations* (MICE; van Buuren & Oudshoorn, 1999) angewandt, der im Programm R implementiert ist. Die Schulform wurde dabei als eine der Prädiktorvariablen für die Schätzung fehlender Werte

einbezogen. So entstanden fünf imputierte Datensätze, die gemäß der Formel von Rubin (1987) in Mplus integriert wurden.

Ergebnisse

Der Effekt der leistungs- und zuwanderungsbezogenen sowie sozialen Zusammensetzung der Schule auf die psychosoziale Entwicklung der Lernenden

Im Rahmen der ersten Fragestellung wurde untersucht, ob die leistungs- und zuwanderungsbezogene sowie soziale Zusammensetzung der Schülerschaft einer Schule mit der Entwicklung der Schulzufriedenheit (Tabelle 2.2, Modelle 1, 2 und 3) und dem normverletzenden Verhalten (Tabelle 2.3, Modelle 1, 2 und 3) von Schülerinnen und Schülern zusammenhing. Ergänzend wurde darüber hinaus der Effekt auf die Mathematiknote (Tabelle 2.4, Modelle 1, 2 und 3) der Schülerinnen und Schüler untersucht.

Schulzufriedenheit. Wie aus Modell 1a in Tabelle 2.2 hervorgeht, fand sich ein positiver Kompositionseffekt für die mittlere Mathematikleistung der Schule in der 9. Jahrgangsstufe auf die Schulzufriedenheit der Schülerinnen und Schüler in der 10. Jahrgangsstufe (Modell 1a: $\beta = .09^{**}$, siehe Tabelle 2.2). Das heißt, bei gleicher Mathematikleistung in der 9. Jahrgangsstufe waren Lernende nach einem Schuljahr zufriedener mit der Schule, je leistungsstärker die Schule war, die sie besuchten. Dieser Kompositionseffekt blieb auch bestehen, wenn zusätzlich für die Schulform (Modell 1b) und für weitere Merkmale der Schülerinnen und Schüler (Modell 1c) kontrolliert wurde. Von der Zusammensetzung der Schülerschaft hinsichtlich des mittleren HISEI ging dagegen kein statistisch signifikanter Effekt auf die Schulzufriedenheit der Schülerinnen und Schüler aus (siehe Modell 2a bis 2c in Tabelle 2.2). Ein geringerer Anteil von Lernenden mit Zuwanderungshintergrund begünstigte die Schulzufriedenheit der Lernenden (Modell 3a: $\beta = -.05^*$, siehe Tabelle 2.2), war jedoch unter Kontrolle der Schulform (Modell 3b) und weiterer Schülermerkmale (Modell 3c) nicht länger statistisch signifikant.

Tabelle 2.2

Mehrebenenmodelle verschiedener Kompositionseffekte auf die individuelle Schulzufriedenheit in der 10. Jahrgangsstufe

	a		b		c	
	β	SE	β	SE	β	SE
Modell 1: Mathematikleistung						
Schulzufriedenheit 9. Jgst. (Ebene 1)	.55***	.02	.55***	.02	.53***	.02
Matheleistung 9. Jgst. (Ebene 1)	-.02	.03	-.02	.03	-.01	.03
Matheleistung 9. Jgst. (Ebene 2)	.09**	.03	.12**	.05	.10*	.05
R ² Ebene 1/ Ebene 2	.31/.37		.31/.59		.33/.47	
Modell 2: HISEI						
Schulzufriedenheit 9. Jgst. (Ebene 1)	.54***	.02	.54***	.02	.54***	.02
HISEI 9. Jgst. (Ebene 1)	.06**	.02	.06**	.02	.07**	.02
HISEI 9. Jgst. (Ebene 2)	.05	.03	.04	.04	.05	.04
R ² Ebene 1/ Ebene 2	.32/.15		.32/.36		.33/.37	
Modell 3: Zuwanderungshintergrund						
Schulzufriedenheit 9. Jgst. (Ebene 1)	.55***	.02	.55***	.02	.54***	.02
Zuwanderungsh. 9. Jgst. (Ebene 1)	.04	.04	.04	.04	.07	.04
Zuwanderungsh. 9. Jgst. (Ebene 2)	-.05*	.09	-.04	.02	-.02	.03
R ² Ebene 1/ Ebene 2	.32/.15		.31/.45		.33/.33	
Modell 4: Schulzufriedenheit						
Schulzufriedenheit 9. Jgst. (Ebene 1)	.54***	.02	.54***	.02	.52***	.02
Schulzufriedenheit 9. Jgst. (Ebene 2)	.08***	.02	.07**	.02	.06**	.02
R ² Ebene 1/ Ebene 2	.31/.34		.31/.57		.32/.49	
Modell 5: Normverletzendes Verhalten						
Schulzufriedenheit 9. Jgst. (Ebene 1)	.54***	.02	.54***	.02	.54***	.02
Normverletzendes Verh. 9. Jgst. (Ebene 1)	-.04	.03	-.04	.03	-.04	.03
Normverletzendes Verh. 9. Jgst. (Ebene 2)	-.05	.03	-.05	.04	-.05	.03
R ² Ebene 1/ Ebene 2	.32/.12		.32/.42		.32/.35	

Anmerkungen. a = ohne Kontrollvariablen, b = unter Kontrolle der Schulform (Referenzgruppe: Hauptschule), c = unter Kontrolle der Schulform und Schülermerkmale (Geschlecht, HISEI, Zuwanderungshintergrund, Mathematikleistung); Jgst. = Jahrgangsstufe.

Normverletzendes Verhalten. Die mittlere Mathematikleistung der Schule zeigte auf normverletzendes Verhalten der Schülerinnen und Schüler einen negativen Kompositionseffekt ($\beta = -.07^{**}$, siehe Tabelle 2.3, Modell 1a): Bei gleicher Mathematikleistung in der 9. Jahrgangsstufe verhielten sich Lernende an Schulen mit hoher mittlerer Mathematikleistung weniger normverletzend in der 10. Jahrgangsstufe als an Schulen mit niedrigerer mittlerer Mathematikleistung. Dieser Effekt blieb jedoch statistisch nicht signifikant, wenn zusätzlich die Schulform und weitere Schülermerkmale als Kontrollvariablen aufgenommen wurden (siehe Modell 1b und 1c). Vom mittleren Schul-HISEI ging ebenfalls ein negativer Kompositionseffekt auf das normverletzende Verhalten aus ($\beta = -.11^{***}$, siehe Tabelle 2.3, Modell 2a): Je höher der durchschnittliche HISEI war, desto geringer fiel das normverletzende Verhalten der Schülerinnen und Schüler ein Jahr später aus, unabhängig davon, wie hoch der HISEI in der eigenen Familie war. Unter Kontrolle der Schulform zeigte sich dieser Kompositionseffekt weiterhin statistisch signifikant (Modell 2b), nicht jedoch unter Kontrolle weiterer Schülermerkmale (Modell 2c). Der Anteil von Lernenden mit Zuwanderungshintergrund an der Schule hatte keinen Effekt auf das normverletzende Verhalten der Schülerinnen und Schüler (siehe Modell 3a bis 3c in Tabelle 2.3).

Tabelle 2.3

Mehrebenenmodelle verschiedener Kompositionseffekte auf das individuelle normverletzende Verhalten in der 10. Jahrgangsstufe

	a		b		c	
	β	SE	β	SE	β	SE
Modell 1: Mathematikleistung						
Normverletzendes Verh. 9. Jgst. (Ebene 1)	.60***	.02	.60***	.03	.56***	.03
Matheleistung 9. Jgst. (Ebene 1)	-.05	.03	-.05	.03	-.07*	.03
Matheleistung 9. Jgst. (Ebene 2)	-.07**	.02	-.06	.04	-.05	.05
R ² Ebene 1/ Ebene 2	.39/.91		.39/.94		.41/.90	
Modell 2: HISEI						
Normverletzendes Verh. 9. Jgst. (Ebene 1)	.61***	.02	.60***	.03	.56***	.03
HISEI 9. Jgst. (Ebene 1)	.04	.02	.04	.02	.05*	.02
HISEI 9. Jgst. (Ebene 2)	-.11***	.02	-.08*	.03	-.05	.03
R ² Ebene 1/ Ebene 2	.38/.89		.37/.97		.41/.91	
Modell 3: Zuwanderungshintergrund						
Normverletzendes Verh. 9. Jgst. (Ebene 1)	.61***	.02	.60***	.03	.56***	.03
Zuwanderungsh. 9. Jgst. (Ebene 1)	.07	.05	.08	.04	.06	.04
Zuwanderungsh. 9. Jgst. (Ebene 2)	.04	.02	.01	.02	.00	.02
R ² Ebene 1/ Ebene 2	.39/.35		.38/.93		.41/.79	
Modell 4: Schulzufriedenheit						
Normverletzendes Verh. 9. Jgst. (Ebene 1)	.60***	.03	.59***	.03	.55***	.03
Schulzufriedenheit 9. Jgst. (Ebene 1)	-.05*	.03	-.06*	.03	-.05*	.02
Schulzufriedenheit 9. Jgst. (Ebene 2)	-.03*	.02	.01	.02	.02	.02
R ² Ebene 1/ Ebene 2	.39/.26		.38/.89		.41/.76	
Modell 5: Normverletzendes Verhalten						
Normverletzendes Verh. 9. Jgst. (Ebene 1)	.60***	.03	.60***	.03	.56***	.03
Normverletzendes Verh. 9. Jgst. (Ebene 2)	.08***	.02	.03	.03	.03	.03
R ² Ebene 1/ Ebene 2	.37/.81		.38/.92		.41/.84	

Anmerkungen. a = ohne Kontrollvariablen, b = unter Kontrolle der Schulform (Referenzgruppe: Hauptschule), c = unter Kontrolle der Schulform und Schülermerkmale (Geschlecht, HISEI, Zuwanderungshintergrund, Mathematikleistung); Jgst. = Jahrgangsstufe.

Mathematiknoten. Hinsichtlich der mittleren Mathematikleistung der Schule und der individuellen Mathematikleistung der Schülerinnen und Schüler ließ sich ein positiver Kompositionseffekt feststellen. Lernende mit gleicher Mathematikleistung in der 9. Jahrgangsstufe erhielten ein Jahr später bessere Noten, wenn die Schule, die sie besuchten, leistungsstärker war ($\beta = .06^*$, siehe Tabelle 2.4, Modell 1a). Dieser Effekt blieb auch unter Kontrolle der Schulform signifikant (Modell 1b), nicht aber, wenn zusätzlich die Kontrollvariablen auf Schülerebene aufgenommen wurden (Modell 1c). Auch ein hoher mittlerer HISEI an der Schule begünstigte die Mathematiknote der Lernenden ($\beta = .12^{***}$, siehe Tabelle 2.4, Modell 2a), auch unabhängig von der Schulform und weiteren Kontrollvariablen auf Schülerebene (siehe Modell 2b und 2c). Von der Komposition der Schule hinsichtlich des Zuwanderungshintergrundes auf die Mathematiknote der Schülerinnen und Schüler ergab sich ein negativer Effekt ($\beta = -.10^{***}$, siehe Tabelle 2.4, Modell 3a): Die Mathematiknoten der Schülerinnen und Schüler in der 10. Jahrgangsstufe waren desto besser, je mehr Lernende mit Zuwanderungshintergrund an einer Schule waren. Dieser Effekt fand sich auch unter Kontrolle der Schulform (siehe Modell 3b) und weiterer Schülermerkmale (siehe Modell 3c).

Tabelle 2.4

Mehrebenenmodelle verschiedener Kompositionseffekte auf die individuelle Mathematiknote in der 10. Jahrgangsstufe

	a		b		c	
	β	SE	β	SE	β	SE
Modell 1: Mathematikleistung						
Mathenote 9. Jgst. (Ebene 1)	.64***	.02	.64***	.02	.64***	.02
Matheleistung 9. Jgst. (Ebene 1)	.15***	.03	.15***	.03	.14***	.03
Matheleistung 9. Jgst. (Ebene 2)	.06*	.03	.12**	.04	.07	.04
R ² Ebene 1/ Ebene 2	.56/.39		.56/.48		.57/.33	
Modell 2: HISEI						
Mathenote 9. Jgst. (Ebene 1)	.68***	.02	.68***	.02	.63***	.02
HISEI 9. Jgst. (Ebene 1)	.05*	.02	.05*	.02	.03	.02
HISEI 9. Jgst. (Ebene 2)	.12***	.02	.12***	.02	.06*	.03
R ² Ebene 1/ Ebene 2	.53/.68		.53/.73		.57/.36	
Modell 3: Zuwanderungshintergrund						
Mathenote 9. Jgst. (Ebene 1)	.69***	.02	.68***	.02	.64***	.02
Zuwanderungsh. 9. Jgst. (Ebene 1)	-.12**	.04	-.12**	.04	-.07*	.03
Zuwanderungsh. 9. Jgst. (Ebene 2)	-.10***	.02	-.07**	.02	-.04*	.02
R ² Ebene 1/ Ebene 2	.54/.47		.53/.72		.57/.32	
Modell 4: Schulzufriedenheit						
Mathenote 9. Jgst. (Ebene 1)	.69***	.02	.69***	.02	.63***	.02
Schulzufriedenheit 9. Jgst. (Ebene 1)	.02	.02	.02	.02	.02	.02
Schulzufriedenheit 9. Jgst. (Ebene 2)	.07***	.02	.02	.13	.00	.02
R ² Ebene 1/ Ebene 2	.53/.21		.53/.49		.58/.10	
Modell 5: Normverletzendes Verhalten						
Mathenote 9. Jgst. (Ebene 1)	.69***	.02	.69***	.02	.63***	.02
Normverletzendes Verh. 9. Jgst. (Ebene 1)	-.02	.02	-.03	.02	-.02	.02
Normverletzendes Verh. 9. Jgst. (Ebene 2)	-.10***	.03	-.06*	.03	-.03	.03
R ² Ebene 1/ Ebene 2	.53/.36		.53/.52		.57/.14	

Anmerkungen. a = ohne Kontrollvariablen, b = unter Kontrolle der Schulform (Referenzgruppe: Hauptschule), c = unter Kontrolle der Schulform und Schülermerkmale (Geschlecht, HISEI, Zuwanderungshintergrund, Mathematikleistung); Jgst. = Jahrgangsstufe.

Der Effekt psychosozialer Kompositionsmerkmale auf psychosoziale Individualmerkmale und Mathematiknoten der Lernenden

Im Rahmen der zweiten Fragestellung wurde untersucht, ob die psychosoziale Komposition der Lernenden einen eigenständigen Effekt auf die Entwicklung der Lernenden hatte, die über die Ausprägung des psychosozialen Merkmals auf Individualebene hinausging. Dazu wurden in den Modellen 4 und 5 der Tabellen 2.2 bis 2.4 Schulzufriedenheit und normverletzendes Verhalten als Kompositionsmerkmale auf die gleichen abhängigen Variablen wie in der ersten Fragestellung betrachtet.

Schulzufriedenheit. Es zeigte sich, dass die durchschnittliche Schulzufriedenheit der Mitschülerinnen und Mitschüler in der 9. Jahrgangsstufe positiv auf die eigene Schulzufriedenheit in der 10. Jahrgangsstufe wirkte, unabhängig davon, wie hoch die vorherige eigene Schulzufriedenheit war ($\beta = .08^{***}$, siehe Tabelle 2.2, Modell 4a). Dieser Effekt blieb auch unter Kontrolle der Schulform (Modell 4b) sowie der Schulform und weiterer Merkmale der Schülerinnen und Schüler (Modell 4c) statistisch signifikant. Das Ausmaß des normverletzenden Verhaltens in der Schule hatte über das eigene Niveau an normverletzenden Verhalten hinaus keinen Effekt auf die Schulzufriedenheit der Schülerinnen und Schüler (siehe Modell 5a bis 5c in Tabelle 2.2).

Normverletzendes Verhalten. Bei gleicher Schulzufriedenheit in der 9. Jahrgangsstufe verhielten sich Schülerinnen und Schüler in der 10. Jahrgangsstufe weniger normverletzend, wenn sich die Schülerschaft insgesamt zufriedener an der Schule fühlte ($\beta = -.03^*$, siehe Tabelle 2.3, Modell 4a). Auch das mittlere Ausmaß des normverletzenden Verhaltens an der Schule hatte einen Effekt auf das normverletzende Verhalten der Schülerinnen und Schüler ($\beta = .08^{***}$, siehe Tabelle 2.3, Modell 5a): Wer eine Schule besuchte, in der sich die Schülerschaft in der 9. Jahrgangsstufe insgesamt weniger normverletzend verhielt, zeigte ein Jahr später selbst seltener normverletzendes Verhalten als vergleichbare Schülerinnen und Schüler an Schulen, an denen sich die Schülerschaft im Durchschnitt stärker normverletzend verhielt. Beide Kompositionseffekte waren jedoch nicht länger statistisch signifikant, wenn für die Schulform (Modell 4b und 5b) oder zusätzlich für weitere Merkmale auf Schülerebene kontrolliert wurde (Modell 4c und 5c).

Mathematiknoten. Bezüglich der psychosozialen Kompositionsmerkmale und der Mathematiknote (Modelle 4 und 5 in Tabelle 2.4) ließ sich ein positiver Effekt der mittleren Schulzufriedenheit auf die Mathematiknote feststellen. Je zufriedener Mitschülerinnen und Mitschüler mit der Schule waren, desto bessere Mathematiknoten erhielten die Lernenden in der 10. Jahrgangsstufe, unabhängig davon, wie zufrieden die Lernenden selbst mit der Schule

in der 9. Jahrgangsstufe waren ($\beta = .07^{***}$, siehe Tabelle 2.4, Modell 4a). Dieser Effekt blieb statistisch signifikant, wenn zusätzlich die Schulform ins Modell aufgenommen wurde (Modell 4b), nicht aber, wenn weitere Schülermerkmale kontrolliert wurden (Modell 4c). Auch das Ausmaß des normverletzenden Verhaltens an Schulen hatte einen Effekt auf die Mathematiknote in der 10. Jahrgangsstufe ($\beta = -.10^{***}$, siehe Tabelle 2.4, Modell 5a). Je weniger normverletzendes Verhalten in der Schülerschaft verbreitet war, desto besser waren die Mathematiknoten der Schülerinnen und Schüler in der 10. Jahrgangsstufe, unabhängig vom eigenen Niveau an normverletzendem Verhalten in der 9. Jahrgangsstufe. Unter Kontrolle der Schulform und weiterer Merkmale auf Schülerebene war dieser Effekt jedoch nicht statistisch signifikant (Modell 5b und 5c).

Diskussion

Ziel des vorliegenden Beitrages war die Erweiterung der empirischen Erkenntnisse über Kompositionseffekte um psychosoziale Aspekte am Beispiel von Schulzufriedenheit und normverletzendem Verhalten. Bezüglich der ersten Fragestellung, ob die individuelle Entwicklung der Schulzufriedenheit und des normverletzenden Verhaltens mit der Schülerschaft einer Schule in Zusammenhang steht, stach besonders ein Befund hervor: Lernende waren bei gleicher Mathematikleistung in der 9. Jahrgangsstufe ein Schuljahr später, also in der 10. Jahrgangsstufe, desto zufriedener mit der besuchten Schule, je leistungsstärker deren Schülerschaft war. Darüber hinaus ergaben sich Hinweise darauf, dass die mittlere Leistung der Schule auch normverletzendes Verhalten verringerte. Für die soziale Komposition der Schule ließ sich tendenziell beobachten, dass Schülerinnen und Schüler zufriedener waren, je weniger Lernende mit Zuwanderungshintergrund die Schule besuchten. Schülerinnen und Schüler verhielten sich weniger normverletzend, je höher der mittlere HISEI war. Im Rahmen der zweiten Forschungsfrage konnte gezeigt werden, dass auch von der mittleren Schulzufriedenheit und dem Ausmaß an normverletzendem Verhalten an einer Schule ein Kompositionseffekt ausging. So waren die Schülerinnen und Schüler nach einem Schuljahr umso zufriedener, je zufriedener ihre Mitschülerinnen und Mitschüler im Durchschnitt mit der Schule waren. Darüber hinaus ergaben sich Hinweise, dass eine hohe durchschnittliche Schulzufriedenheit an der Schule die Schulnoten begünstigte und das normverletzende Verhalten der einzelnen Schülerinnen und Schüler milderte. Ein hohes Ausmaß an normverletzendem Verhalten an der Schule beeinträchtigte dagegen tendenziell die Schulnoten und verstärkte das normverletzende Verhalten der einzelnen Schülerinnen und Schüler.

Der Befund, dass die Schulzufriedenheit und das normverletzende Verhalten als psychosoziale Merkmale der Lernenden in aggregierter Form auf Schulebene auch selbst als Kompositionsmerkmale wirkten, knüpft an Befunde aus Studien zu Peer-Contagion-Effekten an, die bereits für den außerschulischen Bereichen zeigten, dass sich Jugendliche in Peergroups gegenseitig beeinflussen (Dishion & Tipsord, 2011). Wie auch Müller et al. (2015) für den Effekt des mittleren schulischen Fehlverhaltens zeigen konnten, liefern die vorliegenden Befunde Hinweise dafür, dass sich nicht nur die Mitglieder einer selbstgewählten Peergroup, sondern auch Mitschülerinnen und Mitschüler einer Schule sich gegenseitig hinsichtlich der Schulzufriedenheit oder des normverletzenden Verhaltens „anstecken“. Auch bestätigen die vorliegenden Befunde, dass eine höhere mittlere Schulzufriedenheit sowie ein geringeres Ausmaß an normverletzendem Verhalten an der Schule positiv für die eigene Lernentwicklung sind, Annahmen und Forschungsergebnisse der Schulklimaforschung (siehe im Überblick Espelage et al., 2013). Schließlich liefern die Befunde interessante Hinweise auf den Effekt der Leistungsstärke einer Lerngruppe auf die psychosoziale Entwicklung von Schülerinnen und Schülern. So ergaben sich hier im Gegensatz zu den Befunden zum BFLPE (Marsh, 1987) Hinweise darauf, dass eine leistungsstarke Mitschülerschaft nicht nur Einbußen im akademischen Selbstkonzept nach sich zieht, sondern auch bereichernd für die psychosoziale Entwicklung der Schülerinnen und Schüler sein kann, wenn andere Merkmale, wie hier die Schulzufriedenheit und das normverletzende Verhalten, betrachtet werden.

Limitationen

Für eine adäquate Interpretation der Ergebnisse ist es notwendig, auf Grenzen der Studie hinzuweisen und diese zu diskutieren. Generell muss angemerkt werden, dass die beobachteten Kompositionseffekte relativ gering ausfielen. Ursachen für die geringen Koeffizienten könnten darin liegen, dass die Entwicklung der hier betrachteten Variablen Schulzufriedenheit, normverletzendes Verhalten und Mathematiknote von einer Vielzahl von Faktoren abhängig ist, wovon die Komposition der Mitschülerschaft nur einen darstellt (zur Multideterminiertheit psychologischer Variablen siehe Ahadi & Diener, 1989). Weiterhin sind grundsätzlich geringere Effekte hinsichtlich der psychosozialen Komposition im Vergleich zur leistungsbezogenen und sozialen Komposition zu erwarten, da Unterschiede zwischen Schulen in Bezug auf psychosoziale Variablen geringer ausfallen, als es für die Schulleistung aufgrund der leistungsbezogenen Allokationspraxis in Deutschland der Fall ist. Schließlich könnte auch der Zeitpunkt der Untersuchung eine Erklärung für die geringen Effektgrößen geben. Da sich Lernende in der 9. und 10. Jahrgangsstufe bereits mehrere Jahre auf der Schule

befinden, könnten Effekte der Komposition der Schülerschaft auf Schulzufriedenheit, normverletzendes Verhalten und Schulnoten der Lernenden schon vorher stattgefunden haben, sodass von der 9. zur 10. Jahrgangsstufe nur noch geringe Effekte von der Komposition ausgehen. Würde man hingegen die gesamte Schullaufbahn in der Sekundarstufe betrachten, wäre von größeren Effekten bzw. einer Kumulation kleiner Effekte auszugehen. Weiterhin stellt sich die Frage, wie interpretiert werden sollte, dass viele Kompositionseffekte in den ergänzenden Analysen unter Kontrolle der Schulform und weiterer Variablen auf Schülerebene nicht mehr statistisch signifikant waren. Damit wird eine grundsätzliche Frage berührt: die nach der „richtigen“ Aufnahme von Kontrollvariablen, mit denen man sich der Kausalität nähern möchte. Werden zu wenige Kontrollvariablen berücksichtigt, treten möglicherweise Effekte hervor, die nur aufgrund von Drittvariablen existieren. Auf der anderen Seite besteht jedoch auch die Gefahr einer Überkontrolle (Baumert, Stanat & Watermann, 2006). Das ist insbesondere dann der Fall, wenn Kontrollvariablen mit der interessierenden Prädiktorvariable hoch konfundiert sind, wie dies in der vorliegenden Studie zwischen der Komposition der Schülerschaft und der Schulformzugehörigkeit der Fall ist. Um diese Problematik transparent zu machen, wurden bewusst Modelle mit verschiedenen Kontrollvariablen spezifiziert, sodass die Ergebnisse miteinander verglichen werden können. Gleichzeitig wird jedoch hier deutlich, dass es sich um korrelative Zusammenhänge handelt und keine Aussagen zur Kausalität getroffen werden können. Eine zusätzliche Limitation besteht darin, dass die Mathematikleistung durch standardisierte Tests nur zu einem Zeitpunkt gemessen wurde, sodass für die Leistungsentwicklung nur die Noten der Schülerinnen und Schüler betrachtet werden konnten. Schließlich ist darauf hinzuweisen, dass auf Basis der vorliegenden Befunde keine Aussagen über dahinterliegende Wirkmechanismen wie Peer-Contagion-Effekte getroffen werden können. Diese können nur vermutet werden.

Ausblick

Anknüpfend an die zuvor genannten Limitationen, wäre es wünschenswert, wenn zukünftige Untersuchungen sich verstärkt der Frage nach den zugrundeliegenden Wirkmechanismen und dem Zusammenspiel von Kompositionseffekten widmen würden. Denkbar ist beispielsweise, dass Lernende an leistungsschwachen Schulen ihr Grundbedürfnis nach Kompetenz und sozialer Anerkennung nicht befriedigen können und diese Befriedigung dann in andere Bereiche verlagern, z. B. indem sie ihren eigenen sozialen Status in der Peergruppe durch normverletzendes Verhalten zu stärken versuchen (Eccles & Roeser, 2011).

Dadurch würde ein Kreislauf aus geringer Leistungsentwicklung und lernungünstigem Verhalten entstehen. Im positiven Sinne besteht durch diese enge Verzahnung zwischen Leistungskomposition und psychosozialer Komposition die Möglichkeit, pädagogisch nicht nur auf die Leistung zu reagieren, sondern auch auf das psychosoziale Befinden der Schülerinnen und Schüler. Dies könnte dann langfristig wiederum die Leistungsentwicklung positiv beeinflussen (siehe im Überblick Gutman & Schoon, 2013). An welchen „Stellschrauben“ wie „zu drehen ist“, um dem Einfluss einer ungünstigen Komposition der Schülerschaft einer Schule zu begegnen, sollte dementsprechend in zukünftigen Studien detaillierter nachgegangen werden. Die vorliegende Studie liefert Hinweise darauf, dass es sich lohnt, die Schülerschaft einer Schule nicht nur im Hinblick auf die leistungsbezogene und soziale, sondern auch im Hinblick auf ihre psychosoziale Komposition zu untersuchen.

Literaturverzeichnis

- Ahadi, S. & Diener, E. (1989). Multiple determinants and effect size. *Journal of Personality and Social Psychology*, 56(3), 398–406.
- Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Baumert, J., Maaz, K., Stanat, P. & Watermann, R. (2009). Schulkomposition oder Institution - was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die deutsche Schule*, 101(1), 33–46.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.). (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Belfi, B., Goos, M., de Fraine, B. & van Damme, J. (2012). The effect of class composition by gender and ability on secondary school students' school well-being and academic self-concept. A literature review. *Educational Research Review*, 7(1), 62–74.
- Benbenishty, R., Astor, R. A., Roziner, I. & Wrabel, S. L. (2016). Testing the causal links between school climate, school violence, and school academic performance: A cross-lagged panel autoregressive model. *Educational Researcher*, 45(3), 197–206.
- Berkowitz, R., Glickman, H., Benbenishty, R., Ben-Artzi, E., Raz, T., Lipshtat, N. et al. (2015). Compensating, mediating, and moderating effects of school climate on academic achievement gaps in Israel. *Teachers College Record*, 117(7), 1–34.
- Bos, W. & Pietsch, M. (Hrsg.). (2006). *KESS 4 - Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen*. Münster: Waxmann.
- Burns, R. & Mason, D. A. (2002). Class composition and student achievement in elementary schools. *American Educational Research Journal*, 39(1), 207–233.
- Cohen, J., McCabe, E.M., Michelli, N.M. & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record*, 111(1), 180–213.
- Danhier, J. & Martin, É. (2014). Comparing Compositional Effects in Two Education Systems: The Case of the Belgian Communities. *British Journal of Educational Studies*, 62(2), 171–189.
- de Pedro, K. T., Gilreath, T. & Berkowitz, R. (2016). A latent class analysis of school climate among middle and high school students in California public schools. *Children and Youth Services Review*, 63, 10–15.

- Dishion, T. J. & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual review of psychology*, 62, 189–214.
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60(3), 163–183.
- Eder, F. (2002). Unterrichtsklima und Unterrichtsqualität. *Unterrichtswissenschaft*, 30(3), 213–229.
- Eder, F. & Mayr, J. (2000). *Linzer Fragebogen zum Schul- und Klassenklima für die 4. Bis 8. Klasse (LFSK 4-8)*. Göttingen: Hogrefe.
- Eccles, J.S. & Roeser, R.W. (2015). School and community influences on human development. In L. Mayes & M. Lewis (Eds.) *The Cambridge Handbook of Environment in Human Development (Cambridge Handbooks in Psychology)* (S. 645–728). Cambridge: Cambridge University Press.
- Espelage, D. L., Hong, J. S., Rao, M. A. & Low, S. (2013). Associations between peer victimization and academic performance. *Theory Into Practice*, 52(4), 233–240.
- Fraser, B. J. (1982). Development of short forms of several classroom environment scales. *Journal of Educational Measurements*, 19(3), 221–227.
- Ganzeboom, H. B. G., de Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- Gehrmann, M.-L. (2015). Soziale Ansteckung im Klassenzimmer. Das Risiko kennen, das Potenzial nutzen. *Potsdamer Zentrum für empirische Inklusionsforschung (ZEIF)* (5), 1–8.
- Gommans, R., Stevens, G. W. J. M., ter Bogt, T. F. M. & Cillessen, A. H. N. (2016). Adolescent substance use, aggressive behaviors, and peer context behavioral norms. *Journal of Cognitive Education and Psychology*, 15(2), 185–205.
- Gutman, L.M. & Schoon, I., 2013. *The impact of non-cognitive skills on outcomes for young people. Literature Review*. London: University of London.
- Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5), 527–544.
- Harker, R. & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15(2), 177–199.
- Hascher, T. & Hagenauer, G. (2011). Schulisches Wohlbefinden im Jugendalter - Verläufe und Einflussfaktoren. In A. Ittel, H. Merckens & L. Stecher (Hrsg.), *Jahrbuch Jugendforschung* (S. 15–45). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Hornstra, L., van der Veen, I., Peetsma, T. & Volman, M. (2014). Does classroom composition make a difference: Effects on developments in motivation, sense of classroom belonging, and achievement in upper primary school. *School Effectiveness and School Improvement*, 26(2), 1–28.
- Jurczok, A. & Lauterbach, W. (2014). Schulwahl von Eltern. Zur Geografie von Bildungschancen in benachteiligten städtischen Bildungsräumen. In P. A. Berger, C. Keller, A. Klärner & R. Neef (Hrsg.), *Urbane Ungleichheiten. Neue Entwicklungen zwischen Zentrum und Peripherie* (S. 135–155). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Justice, L. M., Logan, J. A. R., Lin, T.-J. & Kaderavek, J. N. (2014). Peer Effects in Early Childhood Education: Testing the Assumptions of Special-Education Inclusion. *Psychological Science*, 25(9), 1722–1729.
- Klinge, D. (2016). *Die elterliche Übergangentscheidung nach der Grundschule. Werte, Erwartungen und Orientierungen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Köller, O., Schnabel, K. & Baumert, J. (2000). Der Einfluss der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32(2), 70–80.
- Maaz, K., Baumert, J., Neumann, M., Becker, M. & Dumont, H. (Hrsg.) (2013). *Die Berliner Schulstrukturreform. Bewertung durch die beteiligten Akteure und Konsequenzen des neuen Übergangsverfahrens von der Grundschule in die weiterführenden Schulen*. Münster: Waxmann.
- Marks, G. N. (2010). What aspects of schooling are important? School effects on tertiary entrance performance. *School Effectiveness and School Improvement*, 21(3), 267–287.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79(3), 280–295.
- Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal*, 28(2), 445–480.
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J. et al. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20 (3), 319–350.
- Müller, C. M., Hofmann, V. & Arm, S. (2016). Susceptibility to classmates influence on delinquency during early adolescence. *The Journal of Early Adolescence*. Published online first.

- Müller, C. M., Hofmann, V., Fleischli, J. & Studer, F. (2015). „Sag‘ mir, was deine Klassenkameraden tun und ich sage dir, was du tun wirst“? Zum Einfluss der Klassenzusammensetzung auf die Entwicklung von schulischem Problemverhalten. *Zeitschrift für Erziehungswissenschaft*, 18(3), 569–589.
- Müller, C. M. & Zurbriggen, C. L. A. (2016). An overview of classroom composition research on social-emotional outcomes: Introduction to the special issue. *Journal of Cognitive Education and Psychology*, 15(2), 163–184.
- Muthén, B. O. & Muthén, L. K. (1998-2010). *Mplus user's guide*. Los Angeles, CA: Author.
- Opdenakker, 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(4), 399–427.
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. et al. (Hrsg.) (2007). *PISA 2006. Die Ergebnisse der dritten internationalen Vergleichsstudie* (1. Aufl.). Münster: Waxmann.
- Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. et al. (Hrsg.) (2008). *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich*. Münster: Waxmann.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago, IL: University of Chicago Press.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models*. Thousand Oaks, CA: Sage Publications.
- 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.
- Rumberger, R. & Palardy, G. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers College Record*, 107(9), 1999–2045.
- Ruus, V.-R., Veisson, M., Leino, M., Ots, L., Pallas, L., Sarv, E.-S. et al. (2007). Students' well-being, coping, academic success, and school climate. *Social Behavior and Personality*, 35(7), 919–936.
- Scharenberg, K. (2016). The interplay of social and ethnic classroom composition, tracking, and gender on students' school satisfaction. *Journal of Cognitive Education and Psychology*, 15(2), 320–346.

- Shochet, I. M., Dadds, M. R., Ham, D. & Montague, R. (2006). School connectedness is an underemphasized parameter in adolescent mental health: Results of a community prediction study. *Journal of Clinical Child and Adolescent Psychology*, 35(2), 170–179.
- Steffgen, G., Recchia, S. & Viechtbauer, W. (2013). The link between school climate and violence in school: A meta-analytic review. *Aggression and Violent Behavior*, 18(2), 300–309.
- Thapa, A., Cohen, J., Guffey, S. & Higgins-D'Alessandro, A. (2013). A review of school climate research. *Review of Educational Research*, 83(3), 357–385.
- van Buuren, S. & Oudshoorn, K. (1999). *Flexible multivariate Imputation by MICE*. Leiden: TNO Prevention and Health.
- van der Slik, F., Driessen, G. & Bot, K. de. (2006). Ethnic and socioeconomic class composition and language proficiency. A longitudinal multilevel examination in Dutch elementary schools. *European Sociological Review*, 22(3), 293–308.
- van Ewijk, R. & Slegers, P. (2010a). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, 21(3), 237–265.
- van Ewijk, R. & Slegers, P. (2010b). The effect of peer socioeconomic status on student achievement. A meta-analysis. *Educational Research Review*, 5(2), 134–150.
- van Landeghem, G., van Damme, J., Opdenakker, M.-C., de Fraine, B. & Onghena, P. (2002). The effect of schools and classes on noncognitive outcomes. *School Effectiveness and School Improvement*, 13(4), 429–451.
- von Saldern, M. & Littig, K. E. (1987). *Landauer Skalen zum Sozialklima (IASSO)*. Weinheim: Beltz.
- Wang, W., Vaillancourt, T., Brittain, H. L., McDougall, P., Krygsman, A., Smith, D. et al. (2014). School climate, peer victimization, and academic achievement: Results from a multi-informant study. *School Psychology Quarterly*, 29(3), 360–377.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Yudron, M. & Jones, S. M. (2016). Developmental trajectories of children's social competence in early childhood: The role of the externalizing behaviors of their preschool peers. *Journal of Cognitive Education and Psychology*, 15(2), 268–292.

3

Teilstudie 2:

Are jitters contagious? The effect of class-average test anxiety on individual students' test anxiety

3 Are jitters contagious? The effect of class-average test anxiety on individual students' test anxiety

Abstract

In test situations, students are surrounded by their classmates. Thus, the influence of the social environment, i.e. peers, on achievement emotions such as test anxiety has been discussed theoretically (Pekrun, 2000). With regard to general anxiety, it has been shown already that one's own level of anxiety can increase by being in contact with anxious people. Especially in adolescence, these so-called peer contagion effects for several socio-emotional characteristics were found (for an overview, see Dishion & Tipsord, 2011). However, peer contagion effects have not been studied for test anxiety at school. The present study aims to scrutinize if class-average test anxiety predicts students' individual test anxiety and if this varies by students' initial level of test anxiety and gender.

Based on a sample of 1,811 students in 95 classrooms, class-average test anxiety at the beginning of Grade 7 was measured directly after the transition to middle school. Hierarchical multiple regression models showed a positive effect from class-average test anxiety on changes in students' individual test anxiety, especially during the initial formation of a new classroom at the transition to middle school. Furthermore, interaction effect models revealed that students with high test anxiety at the end of Grade 6 and girls were more vulnerable to class-average test anxiety than boys in the change of their test anxiety than students with a lower test anxiety at the end of Grade 6 or boys.

In terms of contagion effects, the results indicate that test anxiety is "contagious". Thus, the study contributes to the little evidence on peer contagion effects in the school setting that exists so far. As socio-emotional student characteristics such as test anxiety can influence subsequent academic achievement (for meta-analysis, see Byron & Khazanchi, 2011) and academic trajectories (Betz & Hackett, 1983; Wirtz et al., 2003), it is important to shed light on disparities in academic development between classes.

Keywords: achievement emotions, compositional effects, peer contagion, test anxiety

Introduction

Research on test anxiety has not only shown that test anxiety debilitates students in applying their potential in test situations (for meta-analysis, see Byron & Khazanchi, 2011; e.g., Cassady, 2004; Chapell et al., 2005; Eysenck, Derakshan, Santos, & Calvo, 2007; Hodapp, 1982; Zeidner, 1998), but also that test anxiety might affect decision-making processes (e.g. Betz & Hackett, 1983; Fennema, 1980; Wirtz, Kruger, Scollon, & Diener, 2003). Thus, students who persistently perceive anxiety in test situations in a specific domain might aspire to lower career goals in the fear-laden domain. Students, however, do not experience achievement situations in isolation, but instead when they are typically surrounded by peers. The influence of the social environment, i.e. peers, on achievement emotions such as test anxiety has been discussed theoretically (Pekrun, 2000), but has not been studied in detail empirically (Zeidner, 2014).

With regard to general anxiety, it has been shown that one's own level of anxiety can be influenced by being in contact with anxious people, that is, an anxious mood can transfer from one person to another (e.g. Behnke, Sawyer, & King, 1994; Bonaccio & Reeve, 2010; Donner & Schonfield, 1975; Gump & Kulik, 1997; Gurtman, Martin, & Hintzman, 1990). Furthermore, there are studies revealing that having a close friend who is persistently anxious raises one's own anxiety level, probably due to mutual rumination about anxiety-inducing circumstances (e.g. Rose, 2002; Schwartz-Mette & Rose, 2012). Especially in adolescence, when peers play such an important role, youth may be able to "infect" each other in several socio-emotional characteristics in out-of-school as well as school settings. These so-called *peer contagion effects* have already been shown for characteristics such as depressive symptoms, violence, and drug use (for an overview, see Dishion & Tipsord, 2011). In daily school life, one salient peer group is one's classmates. Spending a crucial amount of time together, students experience each other in these achievement settings that include the expressions of such achievement emotions as test anxiety. Given possible contagion effects in general anxiety and several other socio-emotional characteristics, it is likely that classmates influence students' individual experiences of test anxiety.

To test this prediction, we investigate the extent to which the class-average level of trait test anxiety predicts individual student's levels of test anxiety using data from a longitudinal study. To gain a better understanding of how contagion unfolds over time, the influence of the class-average level of trait test anxiety on the development of individual student's trait test anxiety were estimated over two time points: (1) a few weeks after the transition from elementary to middle school right after students were assigned to new classrooms and

classmates and (2) at the end of the first year of middle school after students have spent one school year together in the same class environment. Given documented stable individual and gender differences in trait test anxiety (Bonaccio & Reeve, 2010; Goetz, Bieg, Lüdtke, Pekrun, & Hall, 2013; Wong, 2008; Zeidner, 1997), we also investigated if initial levels of test anxiety and gender moderated students' susceptibility to classmates' influence on time 2 trait test anxiety.

Test Anxiety

Test anxiety is classified as an achievement emotion that arises in particular achievement situations (for an overview, see Pekrun, 2006). It is assumed that students feel test anxious when they expect failure in a test that is of personal importance to them and they are not sure whether they are competent enough to pass the test successfully (Pekrun, 2006). The interaction between a low domain specific self-concept and a high utility and intrinsic task value in the respective domain has been shown to lead to a high domain-specific test anxiety (Lauermann, Eccles, & Pekrun, 2017).

According to Pekrun & Goetz (2006), test anxiety is comprised of four components: an affective, a physiological, a cognitive, and a motivational component. Thus, indicators of test anxiety include a) nervous affective arousal, b) somatic reactions such as a heightened heart rate, shaking hands, or sweating, c) cognitive concerns about the test and the consequences of failure, and d) a motivation to escape or to avoid the fear-laden situation. Another popular conceptualization of test anxiety focuses two components: a physiological (i.e., somatic symptoms such as a heightened heart rate) and a cognitive component (i.e., concerned thoughts about the test; Liebert & Morris, 1967). Although these two factors are related, they have been shown to be empirically distinct (for an overview, see Zeidner, 2014).

When anxious reactions occur in a specific test situation such as a math test on day X, they are indicators of *state* test anxiety (Cattell, 1950; Spielberger, 1972). If test anxiety persists beyond single evaluation situations, i.e. with regard to math tests in general, it is called *trait* test anxiety (Cattell, 1950; Spielberger, 1972). Trait test anxiety is more persistent than state test anxiety, but it can change over time and situations (Pekrun, 2006; Spielberger & Vagg, 1995). It is assumed to be related to state test anxiety. That is, the experience of test anxiety in a specific test feeds back onto the general tendency of experiencing test anxiety. Conversely, students who perceive themselves to be generally anxious in testing situations will also be more prone than others to test anxiety in specific test situations (Wine, 1980; Zeidner, 2010, 2014). Further, students high in general anxiety, i.e. with a tendency to react anxious across several occasions (Zeidner, 2014), are more vulnerable to be trait test anxious

as well. General anxiety, which is influenced by a mixture of genetic predispositions, parenting practices, family socialization, and the accumulation of failure experiences (for an overview, see Zeidner, 2014) is conceptualized as more persistent than trait test anxiety.

As early as the beginning of the school career in elementary school, children experience test anxiety (Connor, 2003; for meta-analysis, see Hembree, 1988). The developmental course of test anxiety across school grades remains unclear. Some studies have shown an increase of test anxiety across school years (Hill & Sarason, 1966; King et al., 1989; Ollendick, King, & Frary, 1989), while other studies have found that test anxiety stabilizes in high school (for meta-analysis, see Hembree, 1988; Putwain, 2007; Wigfield & Eccles, 1989). The transition from elementary to secondary school represents an especially critical time point in the school career. Test anxiety appears to increase during that time period, likely due to more rigid grading practices, stricter teaching styles, more frequent tests in secondary school and potential lower standing in the ability class rank within the new class room (Lapointe, Legault, & Batiste, 2005; Wigfield & Eccles, 1990).

Environmental Antecedents of Test Anxiety

The level of trait test anxiety depends on multiple factors. According to the social-cognitive model of the development of achievement emotions, next to individual antecedents, characteristics of the learning environment such as teaching style or learning culture in the classroom are assumed to influence achievement emotions (Goetz, Zirngibl, & Pekrun, 2004). Studies on the influence of teaching style on the development of test anxiety have found that the use of fear appeals (Putwain & Symes, 2011) and a teaching style characterized by high demands, achievement-related pressure, and harshness (Frenzel, Pekrun, & Goetz, 2007; Hembree, 1988) lead to increases in students' test anxiety. In contrast, a feedback style based on an individual reference norm (for an overview, see Rheinberg, 1983) and high student-perceived quality of instruction is associated with decreases in test anxiety (Frenzel et al., 2007). Evidence on the influence of the learning culture in a classroom on test anxiety is less prevalent: Some studies found competition between classmates to be positively related to test anxiety (Frenzel et al., 2007; Lapointe et al., 2005). In contrast, a study by Frenzel et al. (2007) found that individual test anxiety in mathematics was lower in classes with a learning culture characterized by a high intrinsic value for mathematics.

In school settings, classmates' characteristics are a very salient social environmental factor that could influence students' individual trait anxiety in testing situations. Combining the individual characteristics of classmates will create a certain learning culture that could

affect individual students' experience of test anxiety. Pointing to the importance of the classroom composition, Zeidner and Schleyer (1998) found that students attending a special gifted class exhibited a higher individual test anxiety than students taught in a class with mixed ability students. As discussed by Zeidner (2014), as well as by Frenzel et al. (2007), more research is needed on which characteristics of the social environment might influence students' individual trait test anxiety.

Emotional and Peer Contagion Effects

One aspect of a class's learning culture might be the classroom composition of trait test anxiety, which is defined as the average level of trait test anxiety experienced in a classroom. Evidence on so-called contagion effects, i.e. the transfer from characteristics between humans in an immediate social environment (for an overview, see Dishion & Tipsord, 2011; see also, Hatfield, Cacioppo, & Rapson, 1994), points in this direction.

Emotional contagion is defined as "the tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person's and, consequently, to converge emotionally" (Hatfield et al., 1994, pp. 153–154). It has been found to be relevant for general state anxiety. Participants in an experimental study (Gurtman et al., 1990) reported an anxious mood after watching a video of an anxious person. In another study with solely female participants that exposed participants to real persons, participants also reported higher state anxiety when the other person seemed more anxious (Gump & Kulik, 1997). Furthermore, contagion has also been found in real-life social contexts, e.g., from patients' reports to their therapists regarding their anxious mood (Donner & Schonfield, 1975), between undergraduate student roommates concerning their state test anxiety (Bonaccio & Reeve, 2010), and between undergraduate students presenting a speech in terms of their state anxiety (Behnke et al., 1994).

Characteristics of others may be particularly relevant for individuals' emotions during early adolescence (Warr, 2002). A large body of research has shown that belonging to a group or having close friends exhibiting a high level of detrimental behaviors such as aggression, drug use, or depressive symptoms leads to increases in adolescents' individual level in these behaviors (for an overview, see Dishion & Tipsord, 2011; see also Müller, Hofmann, & Arm, 2016; Müller, Hofmann, Fleischli, & Studer, 2015; Snyder, Horsch, & Childs, 1997). This *peer contagion effect*, i.e., the unconscious transmission of characteristics between persons in the same peer group, has also been found with regard to general anxiety: Having a close friend with high general anxiety or social anxiety predicts increases in the individual level of general or social anxiety (Prinstein, 2007; Rose, 2002; Schwartz-Mette & Rose, 2012). It is

assumed that adolescents in close dyadic friendships influence each other due to co-ruminating, i.e., extensive and repetitive talking about their negative emotions and anxiety-inducing problems. Given these findings, it seems likely that classmates' level of test anxiety could influence the level of test anxiety expressed by individual students in the classroom.

The above-mentioned studies on contagion effects in regard to anxiety mostly focus on the cognitive component of test anxiety or use a general scale of test anxiety merging the cognitive and the physiological component of test anxiety. No studies on contagion effects on test anxiety could be found that solely measured the physiological component of anxiety. Furthermore, there is a lack of studies on contagion effects for trait test anxiety in school settings. Studies on emotional contagion effects relate to momentary state anxiety as they describe an automatic reaction. Although the existent studies on peer contagion effects take into account trait anxiety, they have focused on general or social anxiety, but not on trait test anxiety in the school setting.

Furthermore, with respect to contagion effects in the school setting, the prevalence of a contagion effect might differ depending on the specific school setting and age of the students and might be more influential during certain developmental periods than others. However, studies on contagion effects have not explored the role of timing so far.

A new focus on trait test anxiety in school settings requires a measure of trait test anxiety in classrooms. A common operationalization of class composition is the class-average level of the characteristic of interest. In studies on so-called compositional effects, the effect of a class-aggregated characteristic on a student's individual outcome over and above the effect of the respective student characteristic at the individual level is analyzed (Harker & Tymms, 2004). Typically, the class-mean is used to generate the class-aggregate. Thus, one possible operationalization of the class composition of trait test anxiety is the average level of trait test anxiety within the classroom.

Individual Differences in Test Anxiety and Differential Peer Contagion Effects

When studying contagion effects of trait test anxiety, it is likely that peer contagion does not work the same way for all students. According to Prinstein (2007), not only peer group characteristics are relevant for the occurrence of peer contagion. People likely vary in their susceptibility to peer contagion. One characteristic that predicts the proneness to peer contagion effects is the individual's prior level of the respective characteristic. For instance, a highly aggressive peer group is especially detrimental for an adolescent's individual aggressiveness if the adolescent was already exhibiting a high level of aggression (for an

overview, see Dishion & Tipsord, 2011). The same moderation effects might hold true for contagion effects on trait test anxiety. The individual level of trait test anxiety predicts individual differences in the vulnerability to recognize appraisal-relevant cues in one's learning environment and to interpret these cues as frightening (e.g., Boekaerts, 2007; Bonaccio & Reeve, 2010; Wong, 2008). Thus, highly anxious students might also be more sensitive to recognize anxiety symptoms of their classmates and more vulnerable to being “infected” when their classmates are highly anxious.

Another individual predictor of trait test anxiety is gender, particularly for mathematics with females reporting higher levels of math test anxiety than males (Zeidner, 1997). Perhaps girls might also be more prone to experience an increase in their anxiety than boys when they are surrounded by highly anxious classmates.

The Present Study

Drawing on theoretical conceptions that define the learning culture between classmates as an antecedent of achievement emotions (Goetz et al., 2004; Zeidner, 2014) and on previous research on contagion effects, we investigated the influence of class-average trait test anxiety on changes in students' individual trait test anxiety. We studied two indicators of trait test anxiety that might be salient in a classroom, thus creating a certain classroom culture of test anxiety: somatic signs (i.e., physiological reactions) and worry (i.e., cognitive concerns). To investigate how peer contagion effects unfold over time, we examined the influence of a peer contagion effect on the change in students' individual trait test anxiety across the first school year after the transition from elementary school to middle school. This allowed us to separate two effects: the initial influence of peers' attitudes right after the formation of a new classroom composition and its continued influence throughout the first school year. We further focused on the domain of mathematics because this allowed us to put our findings into context with previous findings.

Based on this approach, we addressed the following research questions:

1. Does class-average trait test anxiety (measured as somatic signs and worry) predict students' individual trait test anxiety ...
 - a. ...at the beginning of the first school year after the transition from elementary school to middle school?
 - b. ...throughout the first school year after the transition from elementary school to middle school?

2. Does the extent to which class-average trait test anxiety (measured as somatic signs and worry) predict the development of students' individual trait test anxiety vary...
 - a. ...by individual students' initial level of trait test anxiety?
 - b. ...by individual students' gender?

For the first research question, we hypothesized a peer contagion effect of class-average test anxiety on students' individual test anxiety. As there are no studies on peer contagion effects regarding anxiety that discriminate between somatic signs and worry, it was explored, if students are more contagious with regards to the physiological component of test anxiety or the cognitive component of test anxiety. For the second research question we predict higher contagion effects in test anxiety for students with a higher prior level of test anxiety than for students with a lower prior level of test anxiety, as they might be more prone to be "infected" by classmates' test anxiety. According to findings for gender differences in the level of test anxiety, we further predicted stronger contagion effect of class-average test anxiety for girls than for boys.

Method

Sample

In order to examine our research questions, we draw on data from the longitudinal Michigan Study of Adolescent Life Transitions (MSALT), which investigated the educational and life experiences of a cohort of sixth-grade students in ten predominantly white, middle or working class districts in southeastern Michigan across (see Eccles et al., 1989 for full sample details).

Data used in the current study was collected from participants at the end of Grade 6 (Spring 1984) and after the transition from elementary school to middle school at the beginning and the end of Grade 7 (Fall and Spring of 1984/1985). At all three measurement points, students provided information about their trait test anxiety with regard to both somatic signs and worry. To ensure a reliable estimate of class-average achievement, only classrooms with ten or more participating students were included in the present study. This resulted in a final sample of 1,811 students from 95 classrooms (50.9% female; average age at the beginning of Grade 7: $M = 11.6$ years, $SD = .58$). Due to the characteristics of the sampled districts in Michigan, 92.4% of these students were European-American. In addition, the majority of students came from similar academic backgrounds: 98.4% of participating students had a parent with at least a high school degree and 86.2 % had some type of college degree.

Instruments

Trait test anxiety in mathematics. Two components of trait test anxiety were assessed: somatic signs and worry. *Somatic signs* captured the physiological component of students' trait test anxiety in mathematics and was assessed using three items (see Appendix A). In the questionnaire section focused on students' mathematics lessons, students were asked about somatic signs they experience during test situations in mathematics (e.g., "Does the hand you write with shake when you are taking a test?"). All items required students to use a seven-point Likert scale to indicate their agreement (1 = *not at all* to 7 = *a lot*). Cronbach's alphas showed very good reliability across all three measurement points ($\alpha_{T1} = .83$, $\alpha_{T2} = .81$, and $\alpha_{T3} = .84$). The autoregressive coefficients for somatic signs of test anxiety ranged from $\beta = .46$ ($p < .001$) to $\beta = .57$ ($p < .001$) and thus indicated a medium stability. This is in line with the definition of trait test anxiety as a trait emotion with medium stability (Pekrun, 2006). To assess the average level of somatic signs of test anxiety, we aggregated students' data on class level based on the classroom composition at the beginning of Grade 7. *Worry* assessed the cognitive component of students' trait test anxiety in mathematics and it was measured with three items (see Appendix A). On a seven-point Likert scale students were asked about their cognitive concerns about tests in mathematics (e.g., "Before you take a test in math, how nervous do you get?"; 1 = *not at all* to 7 = *very much*). Cronbach's alphas were .87 (T1), .90 (T2), and .89 (T3), thus, showing high reliability across all three measurement points. The autoregressive coefficients for worry indicated a medium stability of this trait test anxiety measure: They ranged from $\beta = .43$ ($p < .001$) to $\beta = .57$ ($p < .001$). To assess class-average levels of worry, we aggregated students' data on class level based on the classroom composition at the beginning of Grade 7.

Control variables. Control variables included gender (dummy coded as 1 = female) and academic performance in mathematics. Results from a curriculum-based and standardized test called *Michigan Educational Assessment Program* (MEAP) that all students had to pass in Grade 5 were used as an indicator of prior mathematical competencies ($M = 24.18$, $SD = 4.26$, Range = 1–28).

Statistical Analyses

To investigate peer contagion effects of class-average trait test anxiety on individual trait test anxiety, a series of hierarchical multiple regression models were conducted in Mplus 7.4 (Muthén & Muthén, 1998–2010). The "type = twolevel" procedure was used to estimate adjusted standard errors that take into account the hierarchical structure of the data with

students nested in classrooms (Raudenbush & Bryk, 2002). Regression analyses were conducted separately for our two measures of trait test anxiety: somatic signs and worry. Despite the fact, that these two components are related as they both indicate the same emotion (see also correlation coefficients in Appendix B), confirmatory factor analyses showed that they are distinct: A two-factor model showed a significant better fit than a one-factor model (fit indices for the one-factor model: $\chi^2(9) = 459.679$; CFI = .891; TLI = .818; RMSEA = .166; fit indices for the two-factor model: $\chi^2(8) = 71.050$; CFI = .985; TLI = .971; RMSEA = .066).

Addressing the first research question, several multilevel regression analyses were conducted that examined the association of class-average test anxiety with the development of individual test anxiety. In the conducted models students' individual test anxiety was predicted by class-average test anxiety at the class level as a manifest predictor capturing the classroom composition of test anxiety. To investigate the development of individual test anxiety, prior students' test anxiety was entered as a predictor on the individual level. First, to test initial peer contagion effects within a new classroom right after the transition from elementary to middle school (research question 1a), the association of class-average test anxiety at the beginning of Grade 7 with students' development of individual test anxiety from the end of Grade 6 to the beginning of Grade 7 was investigated. Second, to test continuing peer contagion effects throughout the first year in middle school (research question 1b), we estimated multilevel regression models examining the association of class-average test anxiety at the beginning of Grade 7 with the development of individual test anxiety from the beginning to the end of Grade 7. In a set of more restrictive models, gender and prior achievement were included as control variables.

To answer research question 2, we examined whether the association of class-average test anxiety on the development of students' individual test anxiety varied by students' individual characteristics. To this end, we tested for moderation effects using the sets of multilevel regression models previously described. To test for differences in peer contagion effects based on initial levels of test anxiety (research question 2a), a cross-level interaction term between students' initial test anxiety and class-average test anxiety was added to the models. To test for potential gender differences (research question 2b), a cross-level interaction term between students' individual gender and class-average test anxiety was added to the models. Finally, to test the robustness of the interaction effects, a set of models was estimated which included both cross-level interaction terms.

To account for differences between classes, continuous covariates were grand-mean centered in all models. Furthermore, for ease of interpretation, all continuous variables were z-standardized ($M = 0$, $SD = 1$) at the individual level. The aggregation of trait test anxiety on the class level was also based on z-standardized individual data and again z-standardized on class level after aggregation. In all models, we controlled for prior trait test anxiety, gender, and prior math achievement. To address missing data which ranged from 1.8% to 4.7% for the test anxiety variables to 22.3% for math achievement, we used the Full Information Maximum Likelihood procedure, which is implemented in Mplus (FIML; Muthén & Muthén, 1998–2010). This procedure allows for the inclusion of all available data and estimates parameter values with the highest probability to represent the sample data (Baraldi & Enders, 2010; Graham, 2012).

To test the robustness of peer contagion effects, we replicated the estimated models in separate models controlling for several individual characteristics and characteristics of the learning environment that are assumed or known to affect the development of test anxiety. As individual characteristics, we added math self-concept and intrinsic value of math, because there is evidence, that self-concept is negatively related (Lauermann et al., 2017; Pekrun, 1984, 2006; Wigfield & Meece, 1988) and intrinsic value is positively related to test anxiety (Lauermann et al., 2017; Wigfield & Meece, 1988). As environmental determinants of test anxiety, student-perceived teacher fairness, student-perceived teacher ability to create a warm atmosphere, student-perceived competitive atmosphere in the classroom and class-average achievement are each included in separate models. These environmental characteristics are all assumed or known to increase test anxiety (Hembree, 1988; McDonald, 2001; Tanzer, 1990; Wigfield & Eccles, 1990; Zeidner, 1998). The results of the respective models are shown in Appendix C.

Results

Change in Trait Test Anxiety

Table 3.1 presents the unstandardized mean levels of somatic signs and worry at the three different time points for the entire sample and separated for boys and girls. In both components of trait test anxiety, students on average decreased statistically significant in their level of test anxiety from the end of Grade 6 to the end of Grade 7 (with the exception of the difference between somatic signs at the beginning and the end of Grade 7). In line with evidence on test anxiety in math, girls had statistically significantly higher levels of somatic signs and worry than boys.

Table 3.1
Descriptive statistics for somatic signs and worry for all measurement points

	Somatic signs			Worry		
	End of Grade 6 (<i>N</i> = 1,778)	Beginning of Grade 7 (<i>N</i> = 1,776)	End of Grade 7 (<i>N</i> = 1,726)	End of Grade 6 (<i>N</i> = 1,778)	Beginning of Grade 7 (<i>N</i> = 1,777)	End of Grade 7 (<i>N</i> = 1,729)
<i>M</i> (<i>SD</i>)	2.74 (1.67)	2.37 (1.49)	2.35 (1.53)	3.50 (1.74)	3.21 (1.70)	3.12 (1.70)
<i>M</i> _{girls} (<i>SD</i>)	2.91 (1.72)	2.53 (1.58)	2.46 (1.61)	3.85 (1.78)	3.51 (1.81)	3.31 (1.79)
<i>M</i> _{boys} (<i>SD</i>)	2.57 (1.61)	2.20 (1.36)	2.24 (1.44)	3.14 (1.63)	2.90 (1.53)	2.91 (1.57)
<i>D M</i> _{girls} - <i>M</i> _{boys}	-.20	-.22	-.14	-.42	-.36	-.24
ICC	.13	.09	.11	.11	.10	.10

Peer Contagion Effect of Class-Average Test Anxiety on Individual Test Anxiety

Peer contagion effect of class-average test anxiety on individual test anxiety right after the transition into a new classroom

To examine a peer contagion effect of class-average test anxiety on individual test anxiety right after the transition into a new classroom, multilevel regression models were conducted estimating how class-average test anxiety at the beginning of Grade 7 was associated with the development of individual test anxiety from the end of Grade 6 to the beginning of Grade 7 (see Table 3.2).

Table 3.2

Multilevel regression models predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7

	Somatic signs at the beginning of Grade 7		Worry at the beginning of Grade 7	
	M1a	M1b	M2a	M2b
	β (SE)	β (SE)	β (SE)	β (SE)
<i>Individual level</i>				
Test anxiety at the end of Grade 6	.50*** (.02)	.48* (.22)	.45*** (.04)	.43*** (.03)
Prior math achievement		-.04 (.25)		-.07* (.03)
Gender		.13 (.10)		.16*** (.04)
<i>Class level</i>				
Class-average test anxiety at the beginning of Grade 7	.18*** (.02) (95% CI: [.16, .21])	.18* (.08) (95% CI: [.05, .31])	.25*** (.02) (95% CI: [.22, .28])	.25*** (.02) (95% CI: [.22, .27])
R ² Individual level/Class level	.27/.99	.26/.99	.23/.99	.22/.99

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Predicting class-average test anxiety at the beginning of Grade 7, a statistically significant positive association with the change in individual test anxiety from the end of Grade 6 to the beginning of Grade 7 was found. This peer contagion effect was found for both, somatic signs ($\beta = .18$, $p < .001$, see Model 1a) and worry ($\beta = .25$, $p < .001$, see Model 2a). Given the decrease in test anxiety over time in our sample, students who attended a class with a higher class-average test anxiety level at the beginning of Grade 7 decreased less in their individual test anxiety from the end of Grade 6 to the beginning of Grade 7 than students in a class with a lower class-average test anxiety level at the beginning of Grade 7. When additionally controlling for prior achievement and gender, the associations remained stable and statistically significant (somatic signs: $\beta = .18$, $p < .05$, Model 1b; worry: $\beta = .25$, $p < .001$, Model 2b). Gender and prior achievement was only associated with changes in worry, but not with changes in somatic signs. Thus, worry decreased more from the end of Grade 6 to the beginning of Grade 7 for boys and when students had a higher level of prior achievement.

None of the additional models controlling for further individual characteristics and characteristics of the learning environment changed the patterns of the results, although math self-concept was negatively and class-average test achievement was positively associated to changes in test anxiety for both, somatic signs and worry, and competitive atmosphere was

negatively associated to changes in worry (see Table C.1 and C.2 in Appendix C). Thus, the contagious effect of class-average test anxiety at the beginning of Grade 7 was existent over and above further individual and environmental variables that could additionally predict individual test anxiety.

Peer contagion effect of initial class-average test anxiety on individual test anxiety at the end of the school year in a new classroom

In a next step, multilevel regression models were conducted estimating the association of class-average test anxiety at the beginning of Grade 7 with students' individual development of test anxiety from the beginning of Grade 7 to the end of Grade 7 (see Table 3.3). These models represent the continuing peer contagion effect of class-average test anxiety on individual test anxiety throughout the first school year in a new classroom.

Table 3.3

Multilevel regression models predicting development of test anxiety from the beginning of Grade 7 to the end of Grade 7

	Somatic signs at the end of Grade 7		Worry at the end of Grade 7	
	M3a <i>b</i> (<i>SE</i>)	M3b <i>b</i> (<i>SE</i>)	M4a <i>b</i> (<i>SE</i>)	M4b <i>b</i> (<i>SE</i>)
<i>Individual Level</i>				
Test anxiety at the beginning of Grade 7	.55*** (.03)	.54*** (.03)	.55*** (.03)	.54*** (.03)
Prior math achievement		-.09** (.03)		-.07** (.02)
Gender		.01 (.04)		.04 (.04)
<i>Class Level</i>				
Class-average test anxiety at the beginning of Grade 7	.08** (.03) (95% CI: [.04, .13])	.07** (.03) (95% CI: [.03, .11])	.05* (.02) (95% CI: [.02, .09])	.05* (.02) (95% CI: [.02, .09])
R ² Individual level/Class level	.32/.68	.32/.54	.31/.36	.31/.33

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

As Model 3a and 4a show, class-average test anxiety at the beginning of Grade 7 is positively associated with changes in students' individual level of test anxiety from the beginning to the end of Grade 7 (somatic signs: $\beta = .08$, $p < .01$, Model 3a; worry: $\beta = .05$,

$p < .05$, Model 4a). Thus, somatic signs and worry decreased more from the beginning of Grade 7 to the end of Grade 7 for students with a higher level of prior achievement. After additionally controlling for prior math achievement and gender, the associations remained stable and statistically significant (somatic signs: $\beta = .07$, $p < .01$, Model 3b; worry: $\beta = .05$, $p < .05$, Model 4b). Prior achievement was negatively related to changes in individual reports of test anxiety from the end of Grade 6 to the beginning of Grade 7. However, gender was not statistically significantly related to changes in individual's reports of somatic signs and worry from the beginning of Grade 7 to the end of Grade 7. None of the additional models controlling for further individual characteristics and characteristics of the learning environment changed the patterns of the results, although math self-concept was negatively associated to the development of test anxiety for both, somatic signs and worry (see Table C.3 and C.4 in Appendix C). Thus, the contagious effect of class-average test anxiety from the beginning to the end of Grade 7 was existent over and above further individual and environmental variables that could additionally predict individual test anxiety.

Individual Differences in the Peer Contagion Effect of Class-Average Test Anxiety on Individual Test Anxiety

For research question 2, we analyzed differences in the positive association of class-average test anxiety with the development of individual test anxiety due to individual initial test anxiety and gender. Therefore, we added a cross-level interaction term of class-average test anxiety and initial test anxiety and a cross-level interaction term of class-average test anxiety and gender to regression models M1b and M2b (see Table 3.4). With regards to changes on test anxiety from the beginning of Grade 7 to the end of Grade 7, only small associations with class-average test anxiety were found (see Table 3.3). As a consequence, the interaction effects with class-average test anxiety for that time period was not significant at the 95% level (see Table D.1 to D.3 in Appendix D). Thus, in the following, we only present the results of interaction effects for changes in text anxiety from the end of Grade 6 to the beginning of Grade 7, i.e., the initial period in a new classroom.

Table 3.4

Regression models predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7, including interaction effects

	Somatic signs at the beginning of Grade 7			Worry at the beginning of Grade 7		
	M5a	M5b	M5c	M6a	M6b	M6c
	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)
<i>Individual level</i>						
Test anxiety at the end of Grade 6	.48*** (.04)	.48*** (.02)	.48*** (.05)	.43*** (.03)	.43*** (.03)	.43*** (.03)
Prior math achievement	-.03 (.04)	-.04* (.02)	-.03 (.06)	-.06* (.03)	-.06* (.03)	-.06* (.03)
Gender	.13** (.05)	.13** (.05)	.13** (.05)	.16*** (.04)	.17*** (.04)	.16*** (.04)
Test anxiety at the end of Grade 6*Class-average test anxiety at the beginning of Grade 7	.06* (.03) (95% CI: [.02, .10])		.06† (.03) (95% CI: [.01, .10])	.03† (.02) (95% CI: [.00, .06])		.02 (.02) (95% CI: [-.01, .05])
Gender*Class-average test anxiety at the beginning of Grade 7			.08† (.04) (95% CI: [.01, .15])	.11** (.04) (95% CI: [.05, .17])		.10* (.04) (95% CI: [.03, .17])
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.17*** (.03)	.14*** (.03)	.14*** (.03)	.25*** (.02)	.19*** (.03)	.20*** (.03)
R ² Individual level/Class level	.26/.99	.26/.99	.26/.98	.22/.99	.23/.99	.23/.99

Note: Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Individual differences in the peer contagion effect due to initial individual test anxiety

Including the interaction term between initial individual test anxiety X class-average test anxiety interaction (see Table 3.4) resulted in positive interaction effects for both, somatic signs ($\beta = .06$, $p < .05$, 95% CI: [.02, .10], Model 5a) and worry ($\beta = .03$, $p < .10$, 95% CI: [.00, .06], Model 6a). Figures 3.1 and 3.2 illustrate the moderation effects. The y-axes of the graphs represent students' individual test anxiety at the beginning of Grade 7, while the x-axes represent the class-average test anxiety. The two lines on the graphs represent the slopes

of the association between students' individual test anxiety at the beginning of Grade 7 and class-average test anxiety at the beginning of Grade 7 for students one standard deviation above the mean of individual test anxiety at the end of Grade 6 and one standard deviation below the mean of individual test anxiety at the end of Grade 6. The slopes of the associations between students' individual test anxiety at the beginning of Grade 7 and class-average test anxiety at the beginning of Grade 7 were steeper for those students who were more test anxious at the end of Grade 6. Conversely, a lower test anxiety at the end of Grade 6 reduced the impact of class-average test anxiety at the beginning of Grade 7 on individual test anxiety at the beginning of Grade 7. Thus, students with high initial test anxiety were not only more likely to show higher levels of subsequent test anxiety than students with low initial test anxiety, but were also more strongly affected by the class-average level of test anxiety. None of the additional models controlling for further individual characteristics and characteristics of the learning environment changed the patterns of the results (see Table C.5 and C.6 in Appendix C).

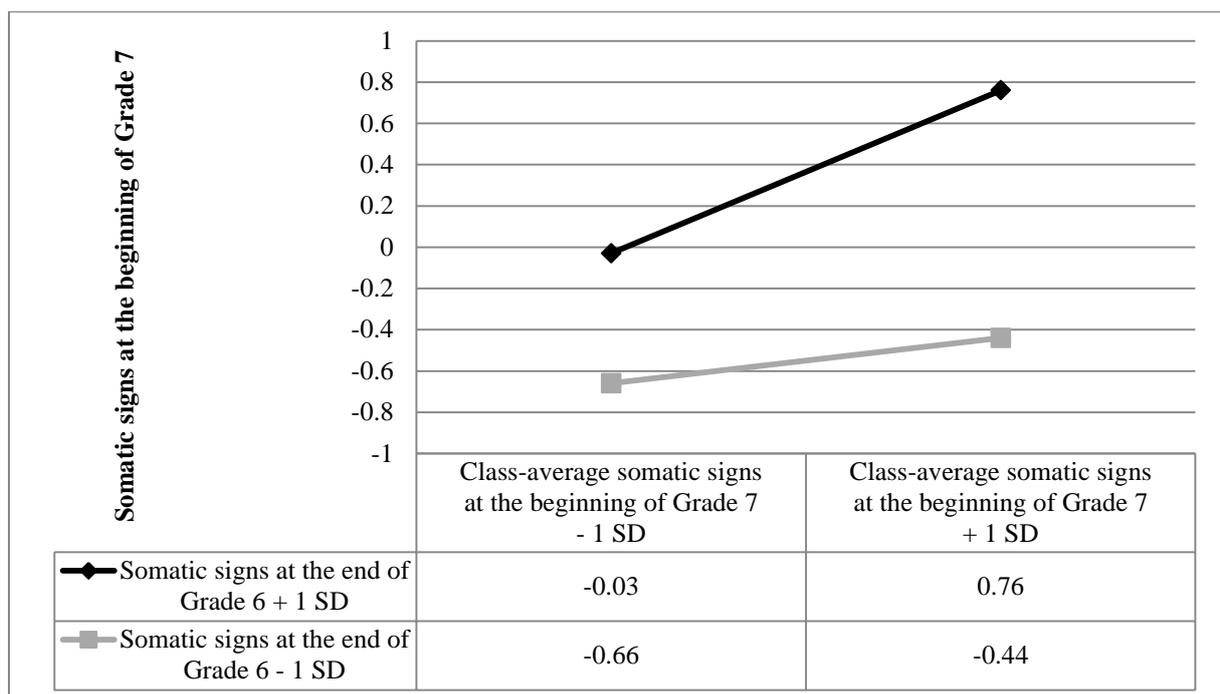


Figure 3.1. Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for 1 SD below and 1 SD above the mean of test anxiety at the end of Grade 6 (in regard to somatic signs).

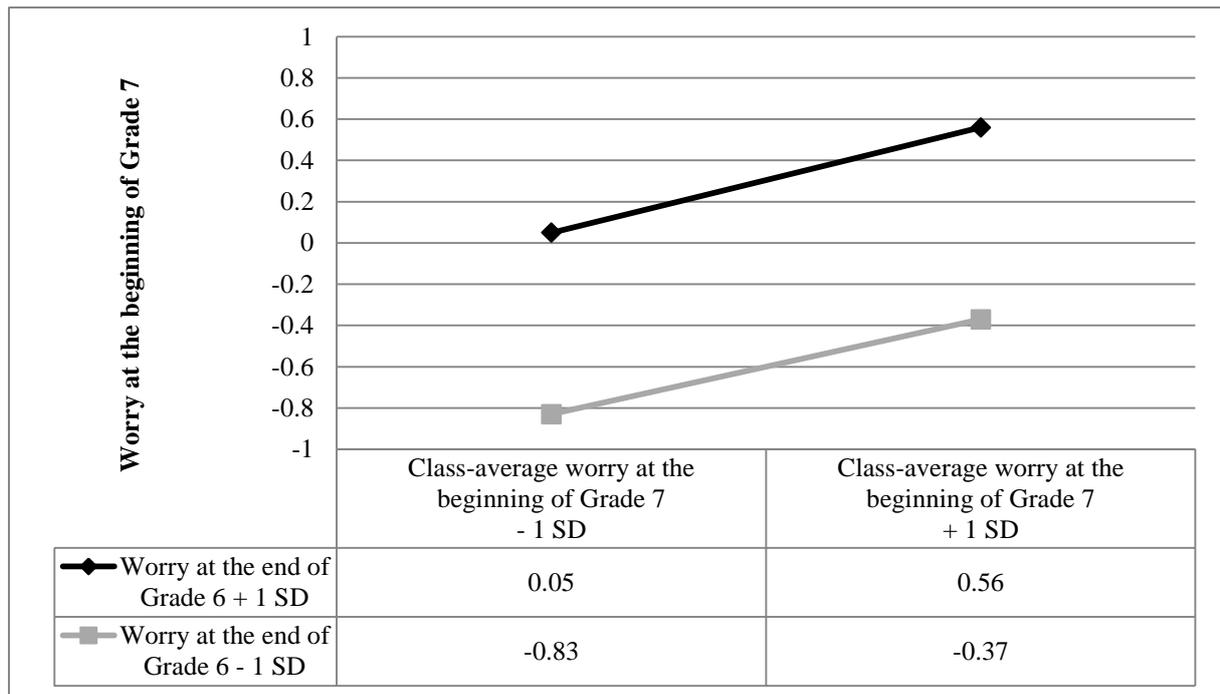


Figure 3.2. Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for 1 SD below and 1 SD above the mean of test anxiety at the end of Grade 6 (in regard to worry).

Individual differences in the peer contagion effect due to gender

Including the interaction term between gender and class-average test anxiety at the beginning of Grade 7 (see Models 5b and 6b) resulted in positive interaction effects (somatic signs: $\beta = .08$, $p < .10$, 95% CI: [.01, .15], see Model 5b; worry: $\beta = .11$, $p < .01$, see Model 6b). The positive interaction effects indicate that class-average test anxiety at the beginning of Grade 7 is more strongly associated with changes in test anxiety from the end of Grade 6 to the beginning of Grade 7 for girls than for boys. In other words, in a highly anxious class, girls decreased less in their test anxiety than boys. In low level anxiety classrooms, boys and girls show the same level of change. This was shown for both, somatic signs, and worry. Figures 3.3 and 3.4 illustrate the moderating effect of gender on the association between class-average and students' individual level of somatic signs and worry, respectively. None of the additional models controlling for further individual characteristics and characteristics of the learning environment changed the patterns of the results (see Table C.5 and C.6 in Appendix C).

As a test for robustness, both interaction terms were included simultaneously in Model 5c and 6c. As the interaction terms remained stable in size and statistically significant for somatic signs, prior test anxiety and gender moderated the association between class-average

somatic signs and changes in somatic signs independently from each other. In regard to worry, differences in a peer contagion effect due to gender were predominant.

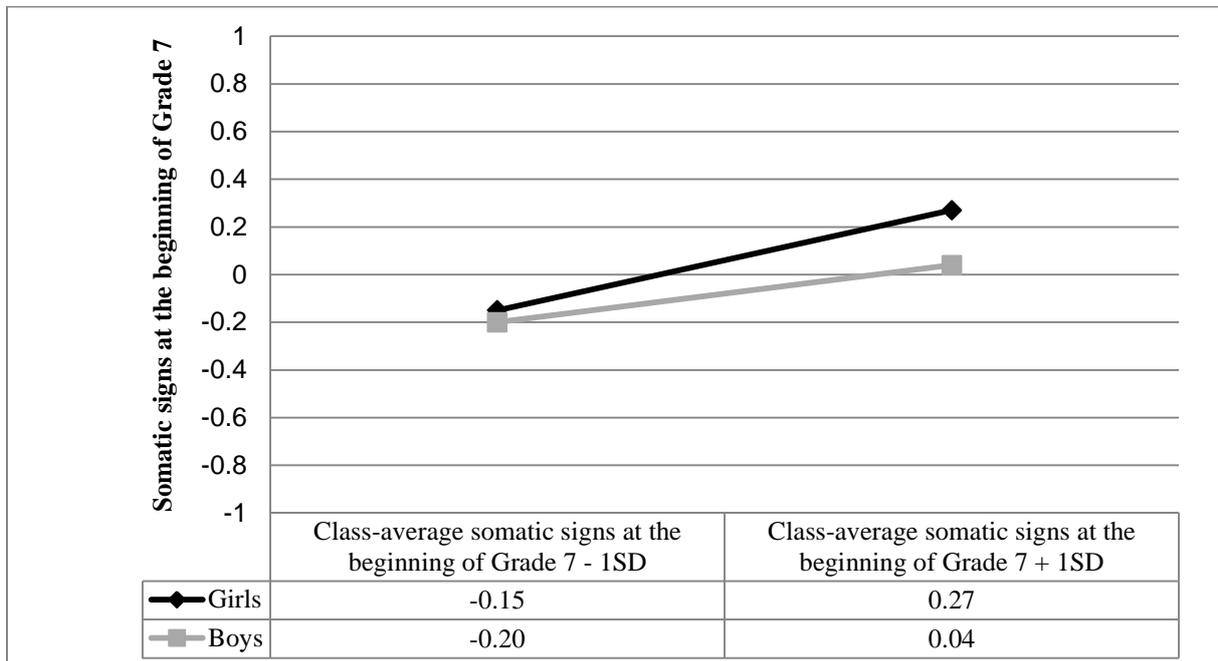


Figure 3.3. Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for girls and boys (in regard to somatic signs).

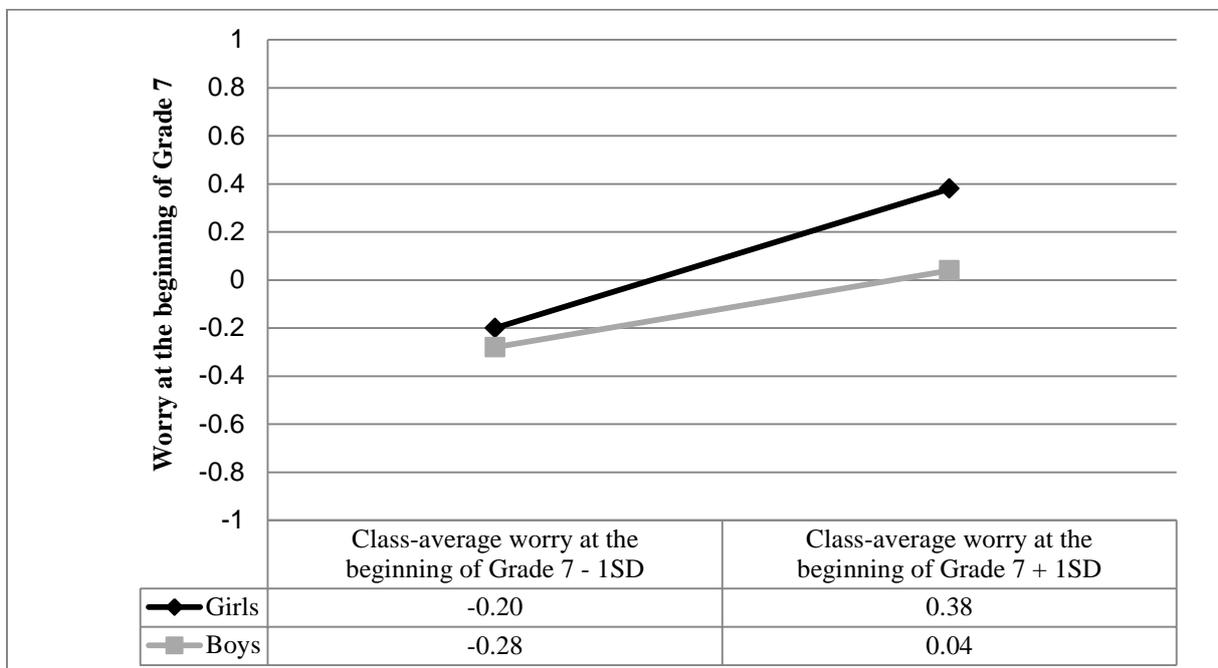


Figure 3.4. Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for girls and boys (in regard to worry).

Discussion

The purpose of the present study was to investigate the role of peers for students' development of individual trait test anxiety. In our first research question, we examined whether class-average test anxiety is associated with changes in students' reports of their individual test anxiety. Studying data from seventh graders, our results indicate a peer contagion effect from class-average test anxiety on changes in students' individual test anxiety, especially during the initial formation of a new classroom at the transition from elementary school to middle school. Thus, given the general decrease in test anxiety over time in this sample, students' individual test anxiety decreased less when they were surrounded by classmates with a higher level of test anxiety than in classes with a lower average level of test anxiety. In terms of contagion effects, the results indicate that classmates "infect" each other in their somatic signs of test anxiety and worry about tests. This is in line with research findings on the emotional contagion effect on general anxiety in out-of-school settings (Behnke et al., 1994; Bonaccio & Reeve, 2010; Donner & Schonfield, 1975; Gump & Kulik, 1997; Gurtman et al., 1990). Students may have perceived somatic symptoms of test anxiety of their classmates such as shaking hands in several test situations, whereby they might have orientated towards their classmates' test anxiety as a guide to evaluate the test situation. Furthermore, classmates may talk about their test anxiety and assimilate due to co-rumination. The peer contagion effect was slightly higher for the cognitive component of test anxiety than for the physiological component when considering the initial transition into the new classroom. One reason might be that students notice classmates' worry more than their somatic signs, because they rather talk about their worries than notice classmates' physiological reactions in written test situations due to their own preoccupation with the test. As two of the three items of the used somatic signs scale refer to written test situations, another pattern might occur with items referring to more public test situations in the class.

Throughout the first school year in the newly formed classroom, the impact of class-average test anxiety lost in importance, but class-average test anxiety at the beginning of Grade 7 still predicted changes in students' test anxiety from the beginning of Grade 7 to the end of Grade 7. In a newly formed reference group such as a new classroom, students may try to gauge and adapt to the new environment by focusing on their new classmates' characteristics and attitudes, and how they deal with test situations in the new school (Kulik & Mahler, 1993). After this initial adaptation and convergence, students might have assimilated to this new environment, including the test situations in the new school, thus, no longer being affected strongly by peers' attitudes.

In a second research question, we scrutinized differences between students in their proneness to a peer contagion effect with regard to trait test anxiety. The analyses revealed that students with high test anxiety at the end of Grade 6 were more prone to be affected by the level of class-average test anxiety at the beginning of Grade 7 in the development of their individual test anxiety from the end of Grade 6 to the beginning of Grade 7 than students with a lower test anxiety at the end of Grade 6. Similar to the theoretical assumption, that test anxious people are more prone to recognize indicators of evaluative situations (Zeidner, 2014), they may also be more sensitive for recognizing their classmates' test anxiety and respective symptoms. Thus, recognizing a higher level of test anxiety in their immediate environment, they are confronted with more opportunities to get "infected".

As a second differential effect, we found, that girls were more vulnerable to class-average test anxiety than boys in the change of their test anxiety. In line with stereotypical behavior adolescents may have incorporated due to socialization, it could be a general tendency of boys to communicate less than girls about emotions and boys might be less likely to confess to and openly express test anxiety (Feldner, Zvolensky, & Schmidt, 2004; Gürses, Kaya, Doğar, Günes, & Yolcu, 2010; Hall, 2005). It has to be noticed, that this differential effect was only found for the initial period in a new classroom, not after spending one school year together in a classroom.

Limitations of the Present Study

Some limitations of our study should be addressed. First, the class-average test anxiety was based on student responses in the first weeks of Grade 7. These responses might already be influenced by their new classmates' test anxiety as opportunities for observing classmates' test anxiety and for talking about test anxiety already arose. An optimal measurement point for student's test anxiety they experience right after transition into the new school and classroom would be the first days of Grade 7 before being exposed at length to the new classmates' attitudes.

Second, the regression models presented in the study cannot provide any causal claims (e.g. Imai, Keele, Tingley, & Yamamoto, 2011; Valeri & Vanderweele, 2013). To test for the robustness of the peer contagion effects we found in the present analyses, we replicated the shown models by adding math self-concept, intrinsic value of math, student-perceived teacher fairness, student-perceived teacher ability to create a warm atmosphere, and student-perceived competition in the class in separate models. None of the additional factors changed the patterns of the results, which points to the significance of our results. However, we do not presume to have controlled for all potentially confounding factors.

Third, the influence of test anxiety on performance is assumed to be curvilinear, i.e. moderate levels of test anxiety might enhance performance due to a higher motivation to put effort into the test preparation and the test (Dweck & Leggett, 1988; King, Ollendick, & Gullone, 1991; Osterhouse, 2001). The mean test anxiety in this sample was below the midpoint of the test anxiety scale indicating a relatively low test anxiety on average. Thus, we do not know if the assessed levels of test anxiety indeed had a debilitating effect for the students in our study or if they valued them as motivating. If the level of experienced test anxiety is still motivating, students in classrooms with a higher average test anxiety might still maintain a high motivation to pass math tests successfully. To provide conclusions about beneficial or detrimental consequences of the investigated peer contagion effect on subsequent motivation, detailed information about students' qualitative valuation of test anxiety levels and consequences for their performance would be needed.

Fourth, gender differences in the peer contagion effect might be biased due to gender stereotypes in stating test anxiety. Admitting to fear or anxiety is typically seen to be more acceptable for females than males (McLean & Anderson, 2009), which might have led boys to report lower levels of test anxiety than they actually experienced. Teacher-assessed measures of students' test anxiety might be one way to elude this problem.

Finally, it has to be taken into account that test anxiety has found to be domain-specific and we only investigated the domain of mathematics (Goetz, Frenzel, Pekrun, Hall, & Luedtke, 2007; Goetz et al., 2014). It needs to be investigated further whether the found peer contagion effects can be replicated across different domains.

Theoretical Significance of Our Findings

Taken together, the present study has shown that students' trait test anxiety was affected by class-average trait test anxiety. The average trait test anxiety in a class can be conceptualized as part of the learning culture in a class. In light of the lack of evidence on how the classroom environment affects students' test anxiety (Frenzel et al., 2007; Zeidner, 2014), the study provides new information in test anxiety research.

In terms of research on contagion effects, the positive association of class-average trait test anxiety on the development of individual trait test anxiety can be interpreted as a peer contagion effect in regard to trait test anxiety between classmates. Thus, the study contributes to the little evidence on peer contagion effects in the school setting that exist so far (Müller et al., 2016; Müller & Zurbriggen, 2016). The majority of studies in this vein investigate contagion effects between peer relations that are self-chosen such as dyadic friendships or

peer groups. The study reveals that the more formal relation between schoolmates (Brown, 1999) also serves as a frame for contagion effects. As socio-emotional student characteristics such as test anxiety can influence subsequent academic achievement (for meta-analysis, see Byron & Khazanchi, 2011) and academic trajectories (e.g. Betz & Hackett, 1983; Fennema, 1980; Wirtz et al., 2003), it is important to shed light on disparities in academic development between classes. Neither the development of school-related socio-emotional characteristics itself, nor achievement development and academic career choices should depend on group affiliation.

References

- Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48*(1), 5–37. doi:10.1016/j.jsp.2009.10.001
- Behnke, R. R., Sawyer, C. R., & King, P. E. (1994). Contagion theory and the communication of public speaking state anxiety. *Communication Education, 43*(3), 246–251. doi:10.1080/03634529409378981
- Betz, N. E., & Hackett, G. (1983). The relationship of mathematics self-efficacy expectations to the selection of science-based college majors. *Journal of Vocational Behavior, 23*(3), 329–345. doi:10.1016/0001-8791(83)90046-5
- Boekaerts, M. (2007). Understanding students' affective processes in the classroom. In P. A. Schutz & R. Pekrun (Eds.), *Educational psychology series. Emotion in education* (pp. 37–56). San Diego, CA: Elsevier Academic Press.
- Bonaccio, S., & Reeve, C. L. (2010). The nature and relative importance of students' perceptions of the sources of test anxiety. *Learning and Individual Differences, 20*(6), 617–625. doi:10.1016/j.lindif.2010.09.007
- Brown, B. B. (1999). Measuring the peer environment of American adolescents. In S. L. Friedman & T. D. Wachs (Eds.), *Measuring environment across the life span: Emerging methods and concepts* (pp. 59–90). Washington, DC: American Psychological Association.
- Byron, K., & Khazanchi, S. (2011). A meta-analytic investigation of the relationship of state and trait anxiety to performance on figural and verbal creative tasks. *Personality and Social Psychology Bulletin, 37*(2), 269–283. doi:10.1177/0146167210392788
- Cassady, J. C. (2004). The influence of cognitive test anxiety across the learning-testing cycle. *Learning and Instruction, 14*(6), 569–592. doi:10.1016/j.learninstruc.2004.09.002
- Cattell, R. B. (1950). *Personality: A systematic theoretical and factual study*. New York, NY: McGraw-Hill.
- Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268–274. doi:10.1037/0022-0663.97.2.268
- Connor, M. J. (2003). Pupil stress and standard assessment tasks (SATs): An Update. *Emotional and Behavioural Difficulties, 8*(2), 101–107. doi:10.1080/13632750300507010

- Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, *62*(1), 189–214.
doi:10.1146/annurev.psych.093008.100412
- Donner, L., & Schonfield, J. (1975). Affect contagion in beginning psychotherapists. *Journal of Clinical Psychology*, *31*(2), 332–339. doi:10.1002/1097-4679(197504)
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*(2), 256–273. doi:10.1037/0033-295X.95.2.256
- Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D. A., & Yee, D. (1989). Self-concepts, domain values, and self-esteem: relations and changes at early adolescence. *Journal of Personality*, *57*(2), 283–310. doi:10.1111/j.1467-6494.1989.tb00484.x
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353. doi:10.1037/1528-3542.7.2.336
- Feldner, M. T., Zvolensky, M. J., & Schmidt, N. B. (2004). Prevention of Anxiety Psychopathology: A Critical Review of the Empirical Literature. *Clinical Psychology: Science and Practice*, *11*(4), 405–424. doi:10.1093/clipsy.bph098
- Fennema, E. (1980). Teachers and sex bias in mathematics. *The Mathematics Teacher*, *73*(3), 169–173.
- 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*(5), 478–493. doi:10.1016/j.learninstruc.2007.09.001
- Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R., & Hall, N. C. (2013). Do girls really experience more anxiety in mathematics? *Psychological Science*, *24*(10), 2079–2087.
doi:10.1177/0956797613486989
- Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C., & Lüdtke, O. (2007). Between- and within-domain relations of students' academic emotions. *Journal of Educational Psychology*, *99*(4), 715–733. doi:10.1037/0022-0663.99.4.715
- Goetz, T., Haag, L., Lipnevich, A. A., Keller, M. M., Frenzel, A. C., & Collier, A. P. M. (2014). Between-domain relations of students' academic emotions and their judgments of school domain similarity. *Frontiers in Psychology*, *5*: 1153.
doi:10.3389/fpsyg.2014.01153
- Goetz, T., Pekrun, R., Zirngibl, A., Jullien, S., Kleine, M., Vom Hofe, R., & Blum, W. (2004). Leistung und emotionales Erleben im Fach Mathematik [Achievement and

- emotions in Mathematics]. *Zeitschrift für Pädagogische Psychologie//German Journal of Educational Psychology*, 18(3/4), 201–212. doi:10.1024/1010-0652.18.34.201
- Goetz, T., Zirngibl, A., & Pekrun, R. (2004). Lern- und Leistungsemotionen von Schülerinnen und Schülern [Students' learning and achievement emotions]. In T. Hascher (Ed.), *Schule positiv erleben. Erkenntnisse und Ergebnisse zum Wohlbefinden von Schülerinnen und Schülern* (pp. 49–66). Bern, Switzerland: Haupt AG.
- Graham, J. W. (Ed.) (2012). *Missing Data: Analysis and Design*. New York, NY: Springer New York.
- Gump, B. B., & Kulik, J. A. (1997). Stress, affiliation, and emotional contagion. *Journal of Personality and Social Psychology*, 72(2), 305–319. doi:10.1037/0022-3514.72.2.305
- Gürses, A., Kaya, Ö., Doğar, Ç., Günes, K., & Yolcu, H. H. (2010). Measurement of secondary school students' test-anxiety levels and investigation of their causes. *Procedia - Social and Behavioral Sciences*, 9, 1005–1008. doi:10.1016/j.sbspro.2010.12.276
- Gurtman, M. B., Martin, K. M., & Hintzman, N. M. (1990). Interpersonal reactions to displays of depression and anxiety. *Journal of Social and Clinical Psychology*, 9(2), 256–267. doi:10.1521/jscp.1990.9.2.256
- Hall, T. S. (2005). *Is test anxiety a form of specific social phobia? Paper presented at the Anxiety disorders Association of America Annual Convention*. (Masterthesis). University of Maryland, MD.
- Harker, R., & Tymms, P. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15(2), 177–199. doi:10.1076/sesi.15.2.177.30432
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion. Studies in emotion and social interaction*. Cambridge: Cambridge University Press.
- Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58(1), 47–77. doi:10.3102/00346543058001047
- Hill, K. T., & Sarason, S. B. (1966). The relation of test anxiety and defensiveness to test and school performance over the elementary-school years: A further longitudinal study. *Monographs of the Society for Research in Child Development*, 31(2), 1–76. doi:10.2307/1165770
- Hodapp, V. (1982). Causal inference from non-experimental research on anxiety and educational achievement. In H. W. Krohne & L. Laux (Eds.), *Achievement. stress and anxiety* (pp. 355–372). Washington, D.C.: Hemisphere.

- Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *The American Political Science Review*, *105*(4), 765–789.
- King, N. J., Ollendick, T. H., & Gullone, E. (1991). Test anxiety in children and adolescents. *Australian Psychologist*, *26*(1), 25–32. doi:10.1080/00050069108258829
- King, N. J., Ollier, K., Iacuone, R., Schuster, S., Bays, K., Gullone, E., & Ollendick, T. H. (1989). Fears of children and adolescents: A cross-sectional Australian study using the Revised-Fear Survey Schedule for Children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *30*(5), 775–784.
- Kulik, J.A., & Mahler, H.I.M. (2000). Social comparison, affiliation, and emotional contagion under threat. In J. Suls & L. Wheeler (Eds.), *Handbook of social comparison: Theory and research* (pp. 259–322). New York, NY: Kluwer Academic/Plenum Publishers.
- Lapointe, J. M., Legault, F., & Batiste, S. J. (2005). Teacher interpersonal behavior and adolescents' motivation in mathematics: A comparison of learning disabled, average, and talented students. *International Journal of Educational Research*, *43*(1–2), 39–54. doi:10.1016/j.ijer.2006.03.005
- Lauermann, F., Eccles, J. S., & Pekrun, R. (2017). Why do children worry about their academic achievement? An expectancy-value perspective on elementary students' worries about their mathematics and reading performance. *ZDM Mathematics Education*, Published online first. doi:10.1007/s11858-017-0832-1
- Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, *20*(3), 975–978. doi:10.2466/pr0.1967.20.3.975
- McDonald, A. S. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology*, *21*(1), 89–101. doi:10.1080/01443410020019867
- McLean, C. P., & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review*, *29*(6), 495–505. doi:10.1016/j.cpr.2009.05.003
- Müller, C. M., Hofmann, V., & Arm, S. (2016). Susceptibility to classmates' influence on delinquency during early adolescence. *The Journal of Early Adolescence*, Published online first. doi:10.1177/02724316166653475
- Müller, C. M., Hofmann, V., Fleischli, J., & Studer, F. (2015). „Sag' mir, was deine Klassenkameraden tun und ich sage dir, was du tun wirst“? Zum Einfluss der Klassenzusammensetzung auf die Entwicklung von schulischem Problemverhalten [„Tell

- me what your classmates do and I will tell you what you are going to do?" The influence of classroom composition on the development of problem behavior in school]. *Zeitschrift für Erziehungswissenschaft*, 18(3), 569–589. doi:10.1007/s11618-015-0629-3
- Müller, C. M., & Zurbriggen, C. (2016). An overview of classroom composition research on social-emotional outcomes: Introduction to the special issue. *Journal of Cognitive Education and Psychology*, 15(2), 163–184. doi:10.1891/1945-8959.15.2.163
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.
- Ollendick, T. H., King, N. J., & Frary, R. B. (1989). Fears in children and adolescents: Reliability and generalizability across gender, age and nationality. *Behaviour Research and Therapy*, 27(1), 19–26. doi:10.1016/0005-7967(89)90115-0
- Osterhouse, R. A. (2001). Classroom anxiety and the examination performance of test-anxious students. *The Journal of Educational Research*, 68(7), 247–250. doi:10.1080/00220671.1975.10884763
- Pekrun, R. (1984). An expectancy-value model of anxiety. In van der Ploeg, H. M., R. Schwarzer, & C. D. Spielberger (Eds.), *Advances in test anxiety research* (3rd ed., pp. 53–72). Lisse, Netherlands: Sweets & Zeitlinger.
- Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Advances in psychology. Motivational psychology of human development: Developing motivation and motivating development* (pp. 143–163). Amsterdam, Netherlands: Elsevier Science.
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18(4), 315–341. doi:10.1007/s10648-006-9029-9
- Pekrun, R., & Götz, T. (2006). Emotionsregulation: Vom Umgang mit Prüfungsangst [Emotional regulation. Dealing with test anxiety]. In H. Mandl & H. F. Friedrich (Eds.), *Handbuch Lernstrategien* (pp. 248–258). Göttingen, Germany: Hogrefe.
- Prinstein, M. J. (2007). Moderators of peer contagion: A longitudinal examination of depression socialization between adolescents and their best friends. *Journal of Clinical Child & Adolescent Psychology*, 36(2), 159–170. doi:10.1080/15374410701274934
- Putwain, D. W. (2007). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology*, 77(3), 579–593. doi:10.1348/000709906X161704

- Putwain, D. W., & Symes, W. (2011). Teachers' use of fear appeals in the Mathematics classroom: Worrying or motivating students? *British Journal of Educational Psychology*, 81(3), 456–474. doi:10.1348/2044-8279.002005
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (Advanced quantitative techniques in the social sciences)*. Thousand Oaks, CA: Sage Publications.
- Rheinberg, F. (1983). Achievement evaluation: A fundamental difference and its motivational consequences. *Studies in Educational Evaluation*, 9(2), 185–194. doi:10.1016/0191-491X(83)90026-3
- Rose, A. J. (2002). Co-rumination in the friendships of girls and boys. *Child Development*, 73(6), 1830–1843. doi:10.1111/1467-8624.00509
- Schwartz-Mette, R. A., & Rose, A. J. (2012). Co-rumination mediates contagion of internalizing symptoms within youths' friendships. *Developmental Psychology*, 48(5), 1355–1365. doi:10.1037/a0027484
- Snyder, J., Horsch, E., & Childs, J. (1997). Peer relationships of young children: Affiliative choices and the shaping of aggressive behavior. *Journal of Clinical Child Psychology*, 26(2), 145–156. doi:10.1207/s15374424jccp2602_3
- Spielberger, C. D. (1972). Conceptual and methodological issues in anxiety research. In C. D. Spielberger (Ed.), *Anxiety: Current trends in theory and research* (2nd ed., pp. 481–493). New York, NY: Academic Press.
- Spielberger, C. D., & Vagg, P. R. (1995). *Test anxiety: Theory, assessment, and treatment*. Washington, DC: Taylor and Francis.
- Tanzer, N. K. (1990). A dual-level approach to assessing social climates as perceived by students, teachers, and parents. *German Journal of Psychology*, 14(1), 13–20.
- Valeri, L., & Vanderweele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological methods*, 18(2), 137–150. doi:10.1037/a0031034
- Warr, M. (2002). *Companions in crime*. Cambridge, UK: Cambridge University Press.
- Wigfield, A., & Eccles, J. S. (1989). Test anxiety in elementary and secondary school students. *Educational Psychologist*, 24(2), 159–183. doi:10.1207/s15326985ep2402_3
- Wigfield, A., & Eccles, J. S. (1990). Test anxiety in the school setting. In M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 237–250). Boston, MA: Springer US. doi:10.1007/978-1-4615-7142-1_19

- Wigfield, A., & Meece, J. L. (1988). Math anxiety in elementary and secondary school students. *Journal of Educational Psychology, 80*(2), 210–216. doi:10.1037/0022-0663.80.2.210
- Wine, J. D. (1980). Cognitive-attentional theory of test-anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and applications* (pp. 349–385). Hillsdale, MI: Erlbaum.
- Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break? *Psychological Science, 14*(5), 520–524. doi:10.1111/1467-9280.03455
- Wong, S. S. (2008). The relations of cognitive triad, dysfunctional attitudes, automatic thoughts, and irrational beliefs with test anxiety. *Current Psychology, 27*(3), 177–191. doi:10.1007/s12144-008-9033-y
- Zeidner, M. (1997). Cross-cultural and individual differences in test anxiety. *World Psychology, 3*, 143–175.
- Zeidner, M. (1998). *Test anxiety: The state of the art. Perspectives on individual differences*. New York, NY: Plenum Press.
- Zeidner, M. (2010). Test anxiety. In I. B. Weiner & E. E. Craighead (Eds.), *Corsinis's encyclopaedia of psychology* (4th ed., pp. 1766–1768). New York, NY: Wiley.
- Zeidner, M. (2014). Anxiety in education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 265–288). New York, NY: Routledge.
- Zeidner, M., & Schleyer, E. J. (1998). The big-fish–little-pond effect for academic self-concept, test anxiety, and school grades in gifted children. *Contemporary Educational Psychology, 24*(4), 305–329. doi:10.1006/ceps.1998.0985

Appendices

Appendix A. Items used to measure text anxiety

Physiological component of trait test anxiety

- Does the hand you write with shake when you are taking a test?
 - Does your heart beat faster when you have to do a test?
 - When the teacher asks you to write on the blackboard, does the hand you write with sometimes shake?
-

Cognitive component of trait test anxiety

- Do math tests scare you?
 - While you are taking a test in math, how nervous do you get?
 - Before you take a test in math, how nervous do you get?
-

Appendix B. Intercorrelations of all variables considered in the analyses, on individual and class level

<i>Individual level</i>	1	2	3	4
1 Somatic signs at the beginning of Grade 7	1			
2 Worry at the beginning of Grade 7	.66***	1		
3 Prior math achievement	-.18***	-.15***	1	
4 Gender	.11***	.18***	.03	1
<hr/>				
<i>Class level</i>				
1 Class-average somatic signs at the beginning of Grade 7	1			
2 Class-average worry at the beginning of Grade 7	.67***	1		
3 Class-average prior math achievement	-.52 ***	-.20*	1	
4 Percentage of girls in class	-.05	0.17	.11	1

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Appendix C. Results of the replicated models controlling for further antecedents of test anxiety

Table C.1
Multilevel regression models predicting development of somatic signs from the end of Grade 6 to the beginning of Grade 7, controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the end of Grade 6	.48** (.17)	.48* (.21)	.48*** (.02)	.49*** (.02)	.46*** (.05)	.48*** (.08)
Prior math achievement	-.04 (.19)	-.04 (.24)	-.04* (.02)	-.08** (.03)	-.02 (.05)	-.04 (.08)
Gender	.13† (.08)	.13 (.09)	.13** (.05)	.12** (.04)	.12* (.05)	.13* (.06)
<i>Additional individual characteristic</i>					-.12*** (.03)	-.05 (.04)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.18** (.06)	.18* (.04)	.20*** (.02)	.22*** (.02)	.18*** (.02)	.18*** (.03)
<i>Additional environmental characteristic</i>	.01 (.05)	.01 (.03)	-.03† (.02)	.11*** (.02)		
R ² Individual level/Class level	.26/.99	.26/.99	.26/.99	.27/.99	.26/.99	.26/.99

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table C.2

Multilevel regression models predicting development of worry from the end of Grade 6 to the beginning of Grade 7, controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the end of Grade 6	.43*** (.03)	.43*** (.03)	.43*** (.03)	.44*** (.02)	.40*** (.03)	.43*** (.03)
Prior math achievement	-.07* (.03)	-.07* (.03)	-.07* (.03)	-.12*** (.02)	-.04† (.03)	-.07* (.02)
Gender	.16*** (.04)	.16*** (.04)	.16*** (.04)	.15*** (.04)	.15*** (.04)	.17*** (.04)
<i>Additional individual characteristic</i>					-.15*** (.02)	-.02 (.02)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.25*** (.02)	.25*** (.02)	.26*** (.02)	.26*** (.01)	.24*** (.02)	.25*** (.02)
<i>Additional environmental characteristic</i>	.01 (.01)	.01 (.02)	-.05** (.02)	.12*** (.02)		
R ² Individual level/Class level	.22/.99	.22/.99	.22/.99	.24/.99	.22/.99	.22/.99

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table C.3

Multilevel regression models predicting development of somatic signs from the beginning of Grade 7 to the end of Grade 7, controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the beginning of Grade 7	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)	.53*** (.03)	.54*** (.03)
Prior math achievement	-.08** (.03)	-.08** (.03)	-.08** (.03)	-.09** (.03)	-.08** (.03)	-.09** (.03)
Gender	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)
<i>Additional individual characteristic</i>					-.06* (.02)	-.01 (.02)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.07** (.03)	.07** (.03)	.07** (.03)	.08** (.03)	.07** (.03)	.07** (.03)
<i>Additional environmental characteristic</i>	-.04† (.02)	-.03 (.02)	-.00 (.02)	.02 (.03)		
R ² Individual level/Class level	.32/.72	.32/.62	.32/.54	.32/.55	.31/.55	.32/.55

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table C.4

Multilevel regression models predicting development of worry from the beginning of Grade 7 to the end of Grade 7, controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the beginning of Grade 7	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)	.52*** (.03)	.53*** (.03)
Prior math achievement	-.07** (.02)	-.07** (.02)	-.07** (.02)	-.07** (.03)	-.06** (.02)	-.07** (.02)
Gender	.04 (.04)	.04 (.04)	.04 (.04)	.04 (.04)	.03 (.04)	.04 (.04)
<i>Additional individual characteristic</i>					-.07** (.02)	-.02 (.02)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.05* (.02)	.05* (.02)	.06* (.02)	.05* (.02)	.05* (.02)	.05* (.02)
<i>Additional environmental characteristic</i>	-.02 (.02)	.01 (.03)	-.01 (.03)	.01 (.02)		
R ² Individual level/Class level	.31/.34	.31/.34	.31/.34	.31/.33	.30/.28	.30/.31

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table C.5
Multilevel regression models predicting development of somatic signs from the end of Grade 6 to the beginning of Grade 7, including interaction effects and controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the end of Grade 6	.48*** (.05)	.48*** (.05)	.48*** (.09)	.48*** (.02)	.46*** (.04)	.47*** (.04)
Prior math achievement	-.03 (.06)	-.03 (.06)	-.03 (.10)	-.07** (.03)	-.01 (.04)	-.03 (.05)
Gender	.13** (.05)	.13** (.05)	.12* (.05)	.12** (.04)	.11** (.04)	.13** (.05)
Test anxiety at the end of Grade 6*Class-average test anxiety at the beginning of Grade 7	.06† (.03)	.06† (.03)	.06 (.04)	.06** (.02)	.06* (.03)	.06* (.03)
Gender*Class-average test anxiety at the beginning of Grade 7	.08 (.05)	.08 (.05)	.07 (.06)	.07† (.04)	.07† (.04)	.08† (.05)
<i>Additional individual characteristic</i>					-.12*** (.03)	-.05* (.03)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.14*** (.03)	.13*** (.03)	.15*** (.03)	.19*** (.03)	.14*** (.03)	.14*** (.03)
<i>Additional environmental characteristic</i>	.01 (.02)	.01 (.02)	-.03 (.02)	.11*** (.02)		
R ² Individual level/Class level	.26/.98	.26/.98	.26/.98	.27/.99	.26/.98	.26/.98

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table C.6

Multilevel regression models predicting development of worry from the end of Grade 6 to the beginning of Grade 7, including interaction effects and controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the end of Grade 6	.43*** (.03)	.43*** (.03)	.44*** (.03)	.44*** (.02)	.43*** (.03)	.43*** (.03)
Prior math achievement	-.06* (.03)	-.06* (.03)	-.07* (.03)	-.12*** (.02)	-.04 (.02)	-.06** (.02)
Gender	.16*** (.04)	.16*** (.04)	.15*** (.04)	.15*** (.04)	.08*** (.02)	.16*** (.04)
Test anxiety at the end of Grade 6*Class-average test anxiety at the beginning of Grade 7	.02 (.02)	.02 (.02)	.02 (.02)	.03 (.02)	.03 (.02)	.02 (.02)
Gender*Class-average test anxiety at the beginning of Grade 7	.10* (.04)	.10* (.04)	.10* (.04)	.09* (.04)	.08* (.03)	.11* (.04)
<i>Additional individual characteristic</i>					-.16*** (.03)	-.02 (.02)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.20*** (.03)	.20*** (.03)	.21*** (.03)	.21*** (.03)	.20*** (.03)	.20*** (.03)
<i>Additional environmental characteristic</i>	.02 (.01)	.01 (.02)	-.06** (.02)	.12*** (.02)		
R ² Individual level/Class level	.23/.99	.23/.99	.23/.99	.25/.99	.23/.98	.23/.98

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

**Appendix D. Results of interaction effect models for the time period of one school year
in a new class**

Table D.1
Multilevel regression models predicting development of somatic signs and worry from the beginning of Grade 7 to the end of Grade 7, including interaction effects

	Somatic signs at the end of Grade 7			Worry at the end of Grade 7		
	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)	<i>b</i> (SE)
<i>Individual level</i>						
Test anxiety at the beginning of Grade 7	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)
Prior math achievement	-.08** (.03)	-.09** (.03)	-.08** (.03)	-.07** (.02)	-.07** (.02)	-.07** (.02)
Gender	.01 (.04)	.01 (.04)	.01 (.04)	.04 (.04)	.04 (.04)	.04 (.04)
Test anxiety at the beginning of Grade 7*Class-average test anxiety at the beginning of Grade 7	-.01 (.02)		-.01 (.02)	.00 (.03)		.01 (.04)
Gender*Class-average test anxiety at the beginning of Grade 7		.00 (.05)	.00 (.04)		.00 (.04)	.00 (.03)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.07** (.03)	.07 (.04)	.07 (.04)	.05* (.02)	.05† (.03)	.05 (.03)
R ² Individual level/Class level	.32/.54	.32/.54	.32/.53	.30/.32	.31/.31	.30/.30

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table D.2

Multilevel regression models predicting development of somatic signs from the beginning of Grade 7 to the end of Grade 7, including interaction effects and controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the beginning of Grade 7	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)	.53*** (.03)	.54*** (.03)
Prior math achievement	-.08** (.03)	-.08** (.03)	-.08** (.03)	-.09** (.03)	-.07* (.03)	-.08** (.03)
Gender	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.03)	.01 (.04)
Test anxiety at the beginning of Grade 7*Class-average test anxiety at the beginning of Grade 7	-.01 (.02)	-.01 (.02)	-.01 (.02)	-.01 (.02)	-.01 (.02)	-.01 (.02)
Gender*Class-average test anxiety at the beginning of Grade 7	.00 (.04)	.00 (.04)	.00 (.04)	.00 (.04)	.00 (.04)	.00 (.04)
<i>Additional individual characteristic</i>					-.06* (.02)	-.01 (.02)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.07 (.04)	.07 (.04)	.07 (.04)	.08† (.04)	.07† (.04)	.07 (.04)
<i>Additional environmental characteristic</i>	-.03 (.02)	-.03 (.02)	.00 (.02)	.02 (.02)		
R ² Individual level/Class level	.32/.68	.32/.60	.32/.53	.32/.56	.31/.56	.32/.55

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

Table D.3

Multilevel regression models predicting development of worry from the beginning of Grade 7 to the end of Grade 7, including interaction effects and controlling for additional antecedents of test anxiety

<i>Additional individual or environmental characteristic</i>	Teacher fairness	Teacher creates a warm atmosphere	Competitive atmosphere	Class-average math achievement	Math self-concept	Intrinsic value of math
	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>	<i>b (SE)</i>
<i>Individual level</i>						
Test anxiety at the beginning of Grade 7	.54*** (.03)	.54*** (.03)	.54*** (.03)	.54*** (.03)	.52*** (.03)	.53*** (.03)
Prior math achievement	-.07** (.02)	-.07** (.02)	-.07** (.02)	-.07** (.02)	-.06* (.02)	-.07** (.02)
Gender	.04 (.04)	.04 (.04)	.04 (.04)	.04 (.04)	.03 (.04)	.04 (.04)
Test anxiety at the beginning of Grade 7*Class-average test anxiety at the beginning of Grade 7	.00 (.03)	.0 (.03)	.00 (.03)	.00 (.03)	.01 (.03)	.00 (.03)
Gender*Class-average test anxiety at the beginning of Grade 7	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)	.01 (.04)
<i>Additional individual characteristic</i>					-.08** (.02)	-.02 (.02)
<i>Class level</i>						
Class-average test anxiety at the beginning of Grade 7	.05 (.03)	.05 (.03)	.05 (.03)	.05 (.03)	.05 (.03)	.05 (.03)
<i>Additional environmental characteristic</i>	-.02 (.02)	.01 (.02)	-.01 (.02)	.01 (.02)		
R ² Individual level/Class level	.30/.31	.30/.31	.30/.31	.30/.30	.29/.26	.30/.29

Note. Standardized coefficients are reported. Student gender was coded: 0 = male, 1 = female; *** $p < .001$, ** $p < .01$, * $p < .05$, † $p < .10$.

4

Teilstudie 3:

What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept

Note: The final publication is available via <http://dx.doi.org/10.1037/edu0000135>.

Stäbler, F., Dumont, H., Becker, M., & Baumert, J. (2017). What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept. *Journal of Educational Psychology*, *109*(2), 191–207. doi:10.1037/edu0000135

4 What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept

Abstract

Empirical studies have demonstrated that students who are taught in a group of students with higher average achievement benefit in terms of their achievement. However, there is also evidence showing that being surrounded by high-achieving students has a negative effect on students' academic self-concept, also known as the big-fish-little-pond effect. In view of the reciprocal relationship between achievement and academic self-concept, the present study aims to scrutinize how the average achievement of a class affects students' achievement and academic self-concept, and how that, in turn, affects subsequent achievement and academic self-concept. Using a sample of 6,463 seventh-graders from 285 classes in Germany, multilevel path models showed that the class-average achievement at the beginning of the school year positively affected individual achievement in the middle and at the end of the school year, while negative effects on academic self-concept occurred only at the beginning of Grade 7, but not later in the school year. In addition, mediation analyses revealed that the effects of class-average achievement on students' achievement and academic self-concept at the end of the school year were mediated by mid-term achievement, but not by mid-term academic self-concept. This pattern was found for mathematics, biology, physics, and English as a foreign language. The results of our study indicate that the consequences for students of belonging to a group of high-achieving students should be analyzed with respect to both academic self-concept and achievement.

Keywords: academic self-concept, BFLPE, compositional effects, reciprocal effects model

5

Teilstudie 4:

Ein Blick in die Klassenzimmer Deutschlands –
wie heterogen sind Klassen zusammengesetzt?

5 Ein Blick in die Klassenzimmer Deutschlands – wie heterogen sind Klassen zusammengesetzt?

Zusammenfassung

Das deutsche Bildungssystem ist aktuell von Umstrukturierungen gekennzeichnet, die eine größere Heterogenität in Schulen und Klassenzimmern hervorbringen. Beispielhaft lässt sich der Trend von der Mehrgliedrigkeit zur zweigliedrigen Sekundarstufe nennen (Neumann, Maaz & Becker, 2013) sowie die Umgestaltung allgemeinbildender Schulen zu inklusiven Schulen (Klemm, 2015). In der konkreten Unterrichtssituation sind Lehrerinnen und Lehrer durch diese Umstrukturierungen noch stärker mit der Herausforderung konfrontiert, gemeinsame Klassenziele bei gleichzeitiger Berücksichtigung von Heterogenität und Individualität zu erreichen. Trotz eines breiten Diskurses über Heterogenität ist wenig darüber bekannt, wie heterogen die Schülerschaft innerhalb deutscher Klassenzimmern tatsächlich ist und ob dies nach Schulform variiert.

In der vierten Teilstudie soll anhand einer repräsentativen Stichprobe von Neuntklässlerinnen und Neuntklässlern ein erster deskriptiver Einblick geboten werden, in welcher Form sich Heterogenität innerhalb deutscher Klassenzimmer bezüglich verschiedener Heterogenitätskategorien zeigt. Die Befunde weisen darauf hin, dass trotz der leistungsbezogenen Schulformgliederung nicht von homogenen Lerngruppen gesprochen werden kann. Die klassenspezifischen Standardabweichungen in den leistungs- und herkunftsbezogenen sowie psychosozialen Schülermerkmalen verweisen darauf, dass die Schülermerkmale an allen Schulformen streuen. Zudem ergaben sich aus der Kombination der klassenspezifischen Standardabweichungen und der mittleren Ausprägung eines Merkmals in der Klasse unterschiedliche Profile von heterogenen und homogenen Klassen an den Schulformen. Mit dieser ersten deskriptiven Annäherung an die Heterogenität innerhalb deutscher Schulklassen über die Leistungsheterogenität hinaus sollen Anstöße für den Diskurs über und die empirische Untersuchung von Heterogenität gegeben werden.

Schlagworte: Heterogenität, psychosoziale Schülermerkmale, schulische Leistung

Einleitung

Traditionell wurde im deutschen Schulsystem durch die Verteilung der Lernenden auf verschiedene Schulformen gemäß ihres Leistungs- und Fähigkeitsniveaus versucht, der Heterogenität von Lernenden durch die Bildung homogener Lerngruppen gerecht zu werden (Reh, 2005). Dem liegt die Annahme zugrunde, dass homogene Gruppen von Schülerinnen und Schülern besser zu unterrichten seien als heterogene Gruppen (LeTendre, Hofer & Shimizu, 2003). In aktuellen bildungspolitischen und -wissenschaftlichen Diskursen wird diese Annahme zunehmend kritisch hinterfragt. So wird im gegenwärtigen Diskurs über den Umgang mit Heterogenität häufig die Position vertreten, dass die Heterogenität einer Schülerschaft eine Bereicherung darstellt und produktiv für den Unterricht genutzt werden kann und sollte (siehe z. B. Höhmann, 2009; Klippert, 2014). Auf der anderen Seite wird von einigen Pädagoginnen und Pädagogen seit jeher angemerkt, dass es durchaus eine Herausforderung sei, in einer heterogenen Lerngruppe zu unterrichten (siehe im Überblick Budde, 2012). In der Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (KMK, 2015, S. 15) wurde der „Umgang mit Heterogenität“ zu einer der „Schlüselfragen der Schul- und Unterrichtsentwicklung“ erklärt. Dieser sei durch individuelle Förderung zu bewerkstelligen. Vor diesem Hintergrund ist eine Auseinandersetzung mit dem Thema Heterogenität erforderlich, die sich nicht nur auf der Ebene der Schule bewegt, sondern auch die Heterogenität innerhalb von Klassen – als Ebene, auf der der Unterricht stattfindet – einbezieht.

In der erziehungswissenschaftlichen Fachliteratur sind in den letzten Jahren eine Fülle von *theoretischen* und *konzeptuellen* Arbeiten entstanden, die sich mit der Heterogenität von Schülerinnen und Schülern auseinandersetzen (z. B. Bräu & Schwerdt, 2005; Decristan & Jude, 2017; Walgenbach, 2014; Werning, 2010). Im vorliegenden Beitrag soll der Versuch unternommen werden, die Heterogenität der Schülerschaft auf Klassenebene im Hinblick auf multiple Heterogenitätskategorien *empirisch* zu untersuchen. Ziel ist es, deskriptiv darzustellen, welcher Heterogenität Lehrerinnen und Lehrer alltäglich im Unterricht begegnen. Dabei wird die Annahme vertreten, dass nicht nur das Leistungs- und Fähigkeitsniveau sowie die familiäre Herkunft der Schülerinnen und Schüler, sondern auch deren psychosoziale Merkmale für das Unterrichtsgeschehen relevant sind.

Der Diskurs zur Heterogenität von Lerngruppen

Der gegenwärtige theoretische Diskurs zur Heterogenität von Lernenden in deutschen Klassenzimmern ist von der Auffassung geprägt, dass Verschiedenheit anzuerkennen und

wertzuschätzen sei (Höhmann, 2009; Miller, 2014; Prengel, 2005). Die Vorstellung eines „imaginären Durchschnittsschülers“ (Helmke, 2013; Tillmann, 2007) wird als unzeitgemäß betrachtet und weicht der Ansicht, dass Unterricht an die individuellen Bedürfnisse der Lernenden anzupassen sei (Feyerer, 2011; Fischer, 2014; Kunze, 2008). So wird Heterogenität als „Chance“ für den Unterricht, aber auch als Herausforderung für Lehrerinnen und Lehrer betrachtet (Budde, 2012).

In zahlreichen praktischen Ratgebern für Lehrerinnen und Lehrer (Boller, 2007; Klippert, 2014; Rebel, 2011) und Konzeptbeiträgen (Breidenstein, 2014; Budde, 2013; Trautmann & Wischer, 2011) wird diskutiert, welche Merkmale in ihrer Heterogenität eine Rolle für den schulischen Lernprozess spielen. Diskutiert werden dabei vor allem Unterschiede zwischen Lernenden hinsichtlich des Leistungs- und Fähigkeitsniveaus (Tillmann, 2008; Werning, 2010), des sozioökonomischen und akademischen Hintergrunds (Becker, 2014; Dräger, 2009), des kulturellen und ethnischen Hintergrunds (Gogolin, 2013; Neumann, 2008), des Geschlechts (Faulstich-Wieland, 2010; Koch-Priewe, 2009) sowie hinsichtlich der körperlichen und geistigen Gesundheit (Hinz, 2008; Schwerdt, 2005). Dem Leistungs- und Fähigkeitsniveau kommt dabei eine übergeordnete Bedeutung als „Leitdifferenz“ zu (Rabenstein, Idel & Ricken, 2015, S. 241). Im individualisierten oder adaptiven Unterricht, der für den Umgang mit Heterogenität vorgeschlagen wird, und in dem das Unterrichtsangebot an die individuellen Bedürfnisse der Schülerinnen und Schüler angepasst wird (Breidenstein, 2014; Wischer, 2014; Fischer, 2013), werden darüber hinaus auch psychosoziale Variablen wie Motivation oder Interesse als bedeutsam für den Lernprozess angesehen. Diese Ansicht kann durch Erkenntnisse der Pädagogischen Psychologie gestützt werden, dass psychosoziale Variablen das Lernverhalten und damit auch den Lernerfolg beeinflussen (siehe im Überblick Gutman & Schoon, 2013). Im theoretischen Diskurs zur Heterogenität von Lerngruppen spielen sie jedoch bisher nur eine marginale Rolle.

Empirische Beiträge zur Zusammensetzung von Lerngruppen

Unabhängig von theoretischen Beiträgen zur Heterogenität von Lerngruppen sind Einflüsse der Lerngruppenzusammensetzung auf die individuelle Leistungsentwicklung seit Langem ein bedeutender Gegenstand der empirischen Bildungsforschung (Coleman et al., 1966). In erster Linie wird die Lerngruppenzusammensetzung dazu über die mittlere Ausprägung bestimmter Individualmerkmale definiert, z. B. die mittlere Leistung einer Klasse (Danhier & Martin, 2014; Hanushek, Kain, Markman & Rivkin, 2003; Marks, 2010). Über die an der Heterogenität von Lernenden konzeptualisierte Lerngruppenzusammensetzung gibt es bislang erst wenige empirische Studien.

Für die mittlere Leistung und den mittleren sozioökonomischen Status einer Lerngruppe ist vielfach gezeigt worden, dass sich die Zugehörigkeit zu einer leistungsstarken Klasse sowie zu einer Klasse mit sozial begünstigter Schülerschaft positiv auf die individuelle Leistungsentwicklung auswirkt (siehe im Überblick Dumont et al., 2013). In Deutschland gibt es aufgrund der Schulformgliederung besonders große Unterschiede im mittleren Leistungsniveau und im mittleren familiären Hintergrund zwischen verschiedenen Schulformen (Baumert, Stanat & Watermann, 2006). Gymnasien stellen im Hinblick auf die Schülerschaft die günstigsten Lernumwelten dar. An Hauptschulen wurde dagegen eine Kumulation benachteiligender Ausprägungen der Zusammensetzung festgestellt. Auch für psychosoziale Merkmale der Schülerinnen und Schüler bestehen Hinweise, dass diese an den Schulformen unterschiedlich ausgeprägt sind. Baumert et al. (2006) fanden an Gymnasien eine höhere Schulzufriedenheit der Lernenden und geringere aggressive Verhaltensweisen als an anderen Schulformen, wobei die ungünstigsten Werte an Hauptschulen vorlagen.

In einigen neueren Untersuchungen wird die Heterogenität der Lerngruppe hinsichtlich des Leistungs- und Fähigkeitsniveaus betrachtet, indem die Standardabweichung oder Varianz der Leistung in einer Klasse oder Schule als Maß für die Heterogenität herangezogen wird (Kiss, 2013; Lehmann, 2006; Scharenberg, 2012). Über den Einfluss der Heterogenität der Schülerschaft einer Klasse oder Schule auf die individuelle Leistungsentwicklung kommen die genannten Studien bisher zu uneinheitlichen Ergebnissen. So zeigte die Heterogenität in Klassen teilweise positive Effekte auf die individuelle Leistungsentwicklung (Lehmann, 2006; Scharenberg, 2012), andere Untersuchungen ermittelten negative (Ding & Lehrer, 2007; Künsting, Post, Greb, Faust & Lipowsky, 2010) oder nicht signifikante Effekte (Kiss, 2013; Künsting et al., 2010).

Die Frage, von welchem Ausmaß der Leistungsheterogenität sowie weiterer Heterogenitätskategorien in deutschen *Schulklassen* eigentlich ausgegangen werden muss, ist jedoch noch nicht geklärt. Für die Heterogenität von *Schulen* zeigte sich, dass die Leistungen an deutschen Sekundarschulen im Vergleich zu anderen OECD-Staaten, die an PISA 2015 teilnahmen, tatsächlich überdurchschnittlich streuten (Reiss, Sälzer, Schiepe-Tiska, Klieme & Köller, 2016). Wie die Heterogenität hinsichtlich der Leistung oder gar weiterer Schülermerkmale innerhalb von *Klassen* ausgeprägt ist, ist ungeklärt. Es bestehen lediglich Hinweise, dass die Gruppe der Gymnasiastinnen und Gymnasiasten am homogensten ist; sowohl in verschiedenen Kompetenzbereichen (Prenzel, Sälzer, Klieme & Köller, 2013) als auch im akademischen Selbstkonzept und in der Lernmotivation (Watermann, Klingebiel & Kurtz, 2010).

Die Bedeutung der Heterogenität von Lerngruppen für die Tätigkeit der Lehrerinnen und Lehrer

Nicht nur für den schulischen Lernprozess der Schülerinnen und Schüler ist die Klassenzusammensetzung relevant, sondern auch für das Erleben und Handeln der Lehrerinnen und Lehrer. So gaben Lehrkräfte und Lehramtsanwärterinnen und -anwärter in verschiedenen Studien als größte Herausforderung der täglichen Arbeit eine hohe Leistungsheterogenität an (Baumert & Lehmann, 1997; Gebauer, McElvany & Klukas, 2013; Reh 2005). In einer Lehrerbefragung an integrativ orientierten Schulen (Reh, 2005) gab ein Großteil der Lehrerinnen und Lehrer bei einer hohen Leistungsheterogenität in Klassen das unbefriedigende Gefühl an, nicht alle Schülerinnen und Schüler genügend berücksichtigen zu können. Eine ältere Beobachtungsstudie zeigte, dass es Lehrkräften in leistungsheterogenen Klassen weniger gut gelang, die Instruktionen an die akademischen und affektiven Bedürfnisse der Lernenden anzupassen als in homogenen Klassen (Evertson, Sanford & Emmer, 1981). Spezifische Maßnahmen zum Umgang mit einer erhöhten Leistungsheterogenität scheinen bisher selten angewandt zu werden. So konnte die naheliegende Annahme nicht bestätigt werden, dass Lehrerinnen und Lehrer bei einer hohen Leistungsheterogenität in der Klasse eventuell häufiger binnendifferenzieren als in leistungshomogenen Klassen (Warwas, Hertel & Labuhn, 2011; Westphal, Gronostaj, Vock, Emmrich & Harych, 2016). In den aufgeführten Studien wurde die Heterogenität der untersuchten Klassen vor allem über die Standardabweichung der Leistung gemessen oder über die vom Schulpersonal wahrgenommene Leistungsheterogenität.

Die vorliegende Studie

Zusammenfassend ist herauszustellen, dass zwar Erkenntnisse zur *Leistungsheterogenität* an deutschen *Schulen* vorliegen, nicht aber für die Heterogenität hinsichtlich der Leistung und *weiterer* Schülermerkmale innerhalb von *Klassen*. Dies ist insbesondere angesichts der zahlreichen theoretischen Auseinandersetzungen mit der Heterogenität von Lernenden und empirischen Studien zur Lerngruppenzusammensetzung verwunderlich. Daher soll im Folgenden anhand einer für Deutschland repräsentativen Studie deskriptiv beleuchtet werden, in welcher Form sich die derzeit intensiv diskutierte Heterogenität innerhalb deutscher Klassenzimmer bezüglich verschiedener Heterogenitätskategorien zeigt. Dies soll einen ersten Einblick gewähren, mit welchem Ausmaß der Heterogenität Lehrerinnen und Lehrer an den verschiedenen Schulformen tatsächlich arbeiten. Dazu wird nicht nur die Heterogenität

hinsichtlich des Leistungs- und Fähigkeitsniveaus, sondern auch hinsichtlich des familiären Hintergrunds und ausgewählter psychosozialer Merkmale berücksichtigt.

Methode

Stichprobe

Eine für die Betrachtung der Heterogenität geeignete Untersuchung stellt die DESI-Studie (Klieme & Beck, 2007; Klieme, 2008) dar, da eine Reihe an Schülermerkmalen in *kompletten* Schulklassen erhoben wurde und da es sich um eine für Deutschland repräsentative Stichprobe handelt. Das Anliegen der DESI-Studie war die Erfassung der Sprachkompetenzentwicklung in Deutsch und Englisch. Im Schuljahr 2003/2004 wurden zu Beginn und am Ende der 9. Jahrgangsstufe Lernende an 219 Schulen getestet und befragt, wobei pro Schule jeweils zwei Klassen zufällig gezogen wurden.

Die im vorliegenden Beitrag betrachtete Stichprobe der Schülerinnen und Schüler, die am ersten Messzeitpunkt teilgenommen haben, umfasste 10 062 Schülerinnen und Schüler (50.8% weiblich) in 427 Klassen. Zu Beginn der 9. Jahrgangsstufe lag das Durchschnittsalter der Lernenden bei 14.8 Jahren ($SD = .73$). 76.5% der Lernenden wies keinen Zuwanderungshintergrund auf. Die Lernenden verteilten sich auf 80 Hauptschulklassen, 43 Mittel-, Sekundar- oder Regelschulklassen, 106 Realschulklassen, 35 Gesamtschulklassen und 163 Gymnasialklassen.

Instrumente

Kennwerte des Leistungs- und Fähigkeitsniveaus. Die *Deutsch- und Englischkompetenz* zu Beginn der 9. Jahrgangsstufe wurde in der DESI-Studie mit curriculumvalidierten Testinstrumenten erhoben (Klieme, 2008). Pro Sprache wurde eine Gesamtskala gebildet, die verschiedene Teilkompetenzen wie Leseverstehen und Sprachbewusstheit in Deutsch oder Hörverstehen und Textrekonstruktion in Englisch umfasst. Über ein probabilistisches Item-Response-Modell wurde jeweils ein *Weighted-Likelihood*-Schätzer (Warm, 1989) pro Sprache generiert, der die Kompetenzen der Lernenden und das Anspruchsniveau der Items einbezieht (Hartig, Jude & Wagner, 2008). Zusätzlich wurden die allgemeinen *kognitiven Grundfähigkeiten* mit dem Kognitiven Fähigkeitstest (KFT; Heller & Perleth, 2000) erfasst.

Herkunftsmerkmale. Der *sozioökonomische Hintergrund* wurde durch den International Socio-Economic Index of Occupational Status (ISEI; Ganzeboom, Graaf, Treiman & Leeuw, 1992) operationalisiert, wobei von beiden Eltern der jeweils höhere sozioökonomische Status (HISEI) herangezogen wurde. Als Index des *kulturellen Kapitals* (Bourdieu, 1997) wurde die Anzahl der Bücher herangezogen, die laut Schülerinnen und Schüler zum Haushaltsbestand

ihrer Eltern gehören. Auf einer siebenstufigen Skala konnten die Lernenden ankreuzen, ob in ihrem Elternhaus 0, 1–10, 11–50, 51–100, 101–250, 251–500 oder mehr als 500 Bücher vorhanden sind. Zur Erfassung der *ethnischen Herkunft* wurden die Lernenden unter anderem zu ihrer Nationalität befragt.

Psychosoziale Merkmale. Zusätzlich wurden exemplarisch zwei psychosoziale Merkmale betrachtet, die in der DESI-Studie für ganze Schulklassen erhoben wurden. Zum einen wurden die Lernenden hinsichtlich ihrer *Lesefreude* befragt (Beispielitem: „Wenn ich lese, vergesse ich alles um mich herum.“, zu beantworten auf einer vierstufigen Likert-Skala von 1 = *stimmt gar nicht* bis 4 = *stimmt ganz genau*). Die aus drei Items zusammengefasste Skala zur Freude am Lesen wies eine zufriedenstellende interne Konsistenz auf (Cronbachs $\alpha = .84$). Zum anderen wurde eine Skala mit sieben Items zum *sozialen Selbstkonzept* eingesetzt. Auf einer sechsstufigen Skala wurden die Lernenden unter anderem gefragt, wie gut sie in einer Gruppe arbeiten können (von 1 = *sehr gut* bis 6 = *ungenügend*). Die Skala wies ebenfalls eine zufriedenstellende interne Konsistenz auf (Cronbachs $\alpha = .81$).

Methodisches Vorgehen

In der vorliegenden Studie wurde eine deskriptive Betrachtung der Heterogenität innerhalb von Klassen mit der Statistiksoftware SPSS 22 vorgenommen. Die Heterogenität innerhalb von Schulklassen wurde als klassenspezifische Standardabweichung des jeweiligen Schülermerkmals operationalisiert. Dazu wurden die Schülermerkmale auf Individualebene z-standardisiert, um sie um den Mittelwert aller befragten bzw. getesteten Neuntklässlerinnen und Neuntklässler zu zentrieren und so die Heterogenität in den einzelnen Klassen auf die gesamte Bandbreite der Ausprägungen beziehen zu können.¹¹

In einem ersten deskriptiven Schritt wurde der Mittelwert der klassenspezifischen Standardabweichungen der betrachteten Schülermerkmale in den einzelnen Schulformen deskriptiv dargestellt. Zusätzlich wurden die jeweiligen Klassenmittelwerte der Schülermerkmale pro Schulform gemittelt.

In einem zweiten deskriptiven Schritt wurde eine differenziertere Betrachtung der klassenspezifischen Standardabweichung einzelner Schülermerkmale vorgenommen. Dazu wurden die Klassen in Klassen mit homogener, mit durchschnittlicher und mit heterogener Verteilung aufgeteilt, indem die klassenspezifische Standardabweichung mit der mittleren klassenspezifischen Standardabweichung aller Klassen in Relation gesetzt wurde: Eine Klasse wurde als homogen definiert, wenn die klassenspezifische Standardabweichung unter einer

¹¹ Die Standardabweichungen innerhalb der Klassen veränderten sich durch die z-Standardisierung jedoch nicht wesentlich.

Standardabweichung der mittleren klassenspezifischen Standardabweichung aller Klassen lag. Als heterogen wurde eine Klasse definiert, wenn die klassenspezifische Standardabweichung über einer Standardabweichung der mittleren klassenspezifischen Standardabweichung aller Klassen lag. Entsprechend wurden Klassen mit einer klassenspezifischen Standardabweichung, die weniger als eine Standardabweichung nach oben oder unten vom Mittelwert der mittleren klassenspezifischen Standardabweichung aller Klassen abwich, als durchschnittlich definiert (siehe Abb. 5.1). Diese Einteilung erfolgt exemplarisch für jeweils ein leistungsbezogenes, herkunftsbezogenes und psychosoziales Merkmal.

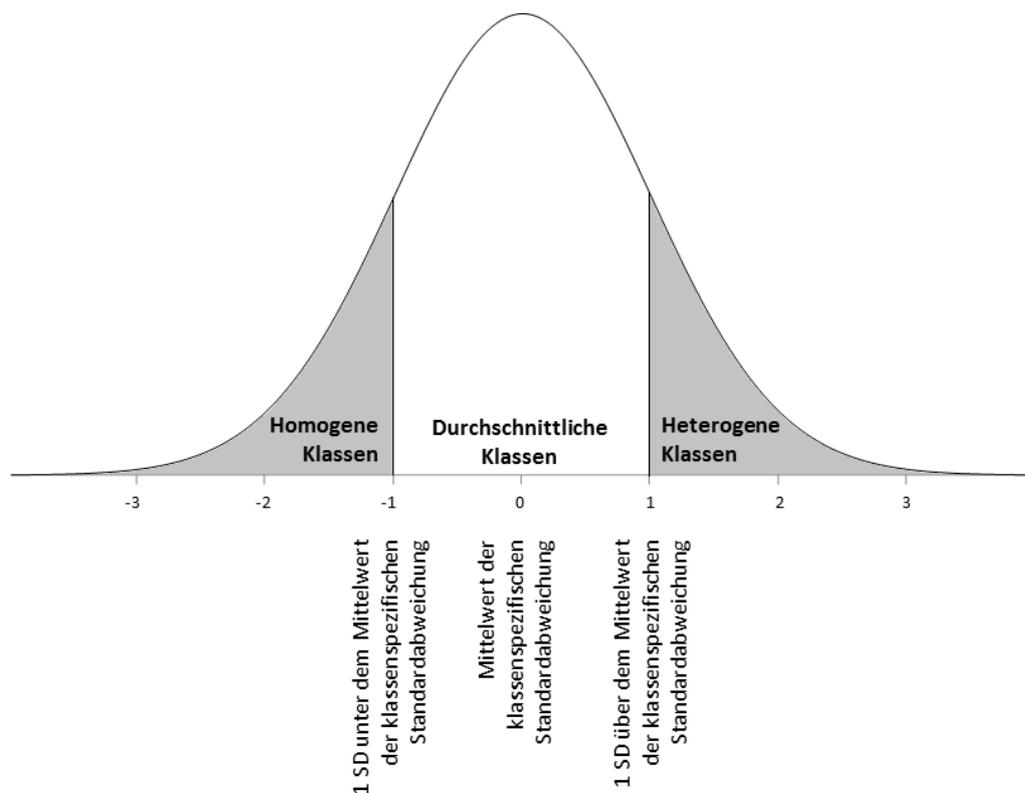


Abbildung 5.1. Schema zur Einteilung der Schulklassen in homogene, durchschnittliche und heterogene Klassen

Da sich die Zusammensetzung von Klassen nicht nur bezüglich der Verteilung, sondern auch hinsichtlich des mittleren Niveaus eines Merkmals unterscheidet, wurde zusätzlich zur dreistufigen Einteilung des Heterogenitätsgrads eine zweite Differenzierungsebene einbezogen. So wurden die oben gebildeten drei Gruppen von Klassen noch einmal dahingehend differenziert, ob der klassenspezifische Mittelwert eine Standardabweichung unter dem Mittelwert der klassenspezifischen Mittelwerte aller Klassen, zwischen einer

Standardabweichung unter und über dem Mittelwert oder eine Standardabweichung über dem Mittelwert aller klassenspezifischen Mittelwerte lag.

Die Heterogenität hinsichtlich der ethnischen Herkunft wurde über die mittlere Anzahl der verschiedenen Nationalitäten pro Klasse analysiert.

Da in der Stichprobe nur wenige fehlende Werte bei den betrachteten Schülermerkmalen bestanden (von 4.7% in der Deutschkompetenz bis maximal 9.0% bezüglich des kulturellen Kapitals), wurde von einem multiplen Imputationsverfahren (Rubin, 1987) abgesehen.¹² Für den HISEI wurden fehlende Werte bereits durch das Forschungsteam der DESI-Studie mit einer multiplen Imputation geschätzt, sodass für 92.0% der Lernenden Angaben generiert werden konnten (Hartig et al., 2008). In der deskriptiven Betrachtung der vorliegenden Studie wurden die fünf imputierten HISEI-Scores in SPSS gemittelt. Da in der geclusterten Stichprobenziehung aufgrund gewisser Untersuchungsziele von DESI ein Oversampling von bilingualen Schulen stattfand, wurde ein vom Forschungsteam der DESI-Studie generiertes Stichprobengewicht (Hartig et al., 2008) eingesetzt.

Befunde

Wie aus Abbildung 5.2 hervorgeht, unterschieden sich die Klassen verschiedener Schulformen entsprechend der leistungsbezogenen Schulformgliederung in ihrem mittleren Leistungs- und Fähigkeitsniveau stark voneinander. Dies zeigte sich auch bezüglich der mittleren Ausprägung familiärer Herkunftsmerkmale und des psychosozialen Merkmals Lesefreude. Wenn zusätzlich die klassenspezifischen Standardabweichungen betrachtet werden, lässt sich allerdings erkennen, dass die Schulformgliederung entlang der Leistung zwar zu den beabsichtigten Leistungsunterschieden zwischen den Schulformen geführt hat, nicht aber zur Ausbildung homogener Klassen. So lagen die mittleren klassenspezifischen Standardabweichungen der Deutschkompetenzen zwischen 0.64 bis 0.72 Punkten. Die klassenspezifischen Standardabweichungen des HISEI lagen zwischen 0.79 und 0.94 Punkten. Das soziale Selbstkonzept als psychosoziales Schülermerkmal streute innerhalb von Klassen mit klassenspezifischen Standardabweichungen zwischen 0.89 und 1.06 Punkten am stärksten. Insgesamt fielen die klassenspezifischen Standardabweichungen der psychosozialen Merkmale am höchsten aus, während jene der Leistungs- und Fähigkeitsmerkmale am geringsten ausfielen.

¹² Aufgrund der geringen Anzahl an fehlenden Werten wurde keine multiple Imputation vorgenommen. Mit Analysen im Statistikprogramm Mplus hätte immerhin die FIML-Methode angewandt werden können, wie es vom DESI-Konsortium vorgeschlagen wird (Hartig et al., 2008). Für die vorliegenden deskriptiven Analysen sind jedoch die Möglichkeiten des Mplus-Programms beschränkt, welches eher für komplexe statistische Verfahren als für detaillierte deskriptive Beschreibungen konzipiert wurde.

Hinzuzufügen ist, dass die Schulformen bezüglich der mittleren klassenspezifischen Standardabweichungen relativ ähnlich waren. Es lässt sich lediglich ein tendenzielles Muster finden: Die mittleren klassenspezifischen Standardabweichungen fielen an Gymnasien zumeist etwas niedriger aus.

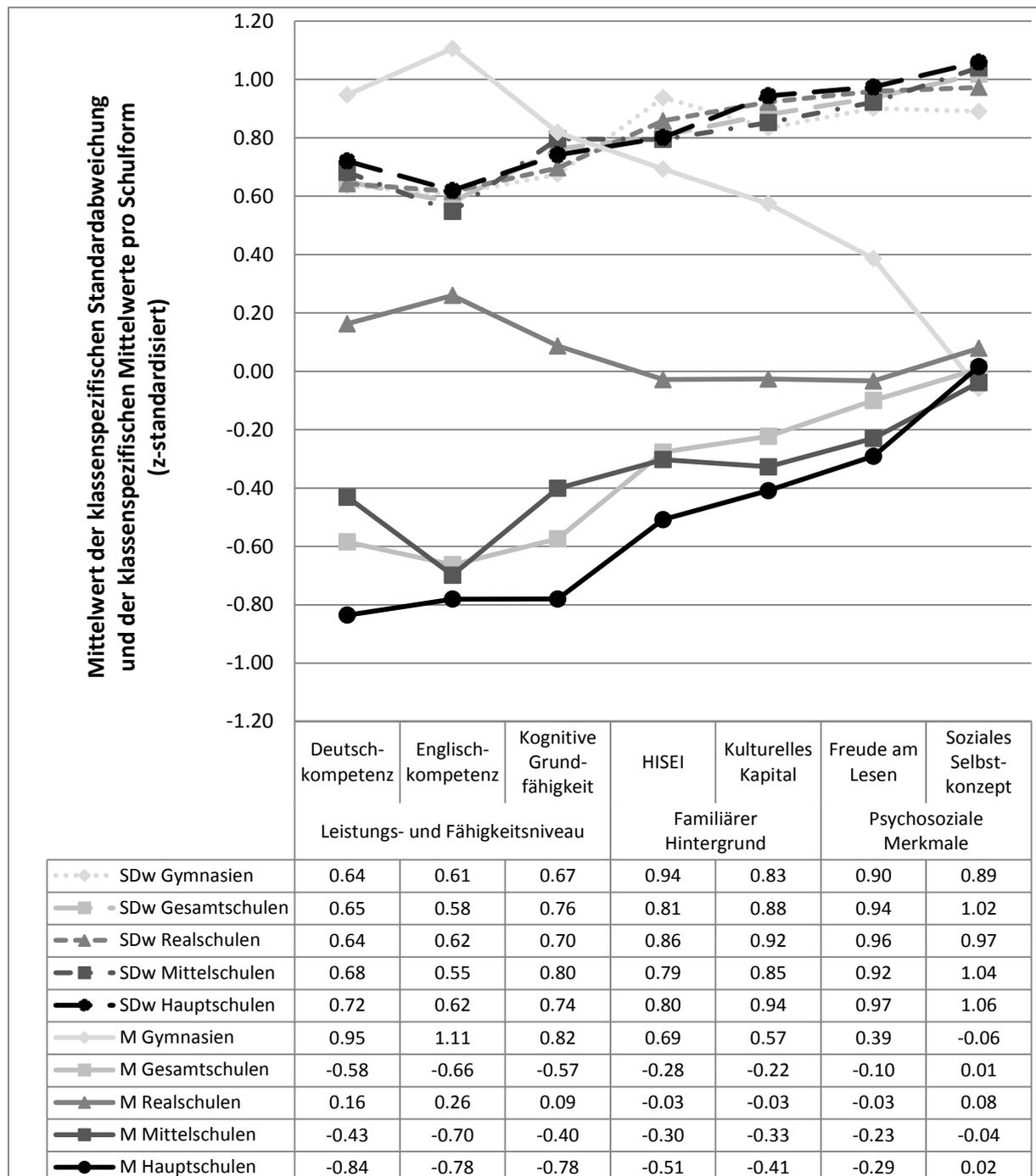


Abbildung 5.2. Mittelwerte der klassenspezifischen Standardabweichungen und der klassenspezifischen Mittelwerte von z-standardisierten Kompositionsmerkmalen pro Schulform

In einem zweiten Schritt wurde die klassenspezifische Verteilung exemplarisch für jeweils ein Schülermerkmal der Kategorie Leistungs- und Fähigkeitsniveau, sozioökonomische Herkunft, ethnische Herkunft und psychosoziale Merkmale differenzierter betrachtet. Diese sollen nun in den folgenden Abschnitten einzeln durchgegangen werden.

Deutschkompetenz

Die mittlere klassenspezifische Standardabweichung aller Klassen lag für die Deutschkompetenz bei $M_{SDk} = 0.66$ mit einer Standardabweichung von $SD_{M_{SDk}} = 0.14$. Eine durchschnittliche Klasse wies somit eine klassenspezifische Standardabweichung zwischen $SDk = 0.52$ und $SDk = 0.80$ auf, was auf 74.7% aller Klassen zutraf (siehe Abb. 5.3). Dies lässt sich an den fettgedruckten Prozentangaben in der Spalte „GESAMT“ ablesen, die – aufgeschlüsselt nach homogenen, durchschnittlichen und heterogenen Klassen – zusammen 100% der Klassen ergeben. Bei den Klassen mit durchschnittlicher klassenspezifischer Standardabweichung in der Deutschkompetenz handelte es sich bei 12.6% um Klassen mit unterdurchschnittlicher mittlerer Deutschkompetenz, bei 46.4% um Klassen mit durchschnittlicher und bei 15.7% um Klassen mit überdurchschnittlicher mittlerer Deutschkompetenz. Dies ist abzulesen an den *nicht* fettgedruckten Prozentangaben in der Spalte „GESAMT“, die – aufgeschlüsselt nach den beiden Differenzierungsebenen der klassenspezifischen Standardabweichung und der klassenspezifischen Mittelwerte – zusammen 100% der Klassen ergeben. Nach gleicher Aufschlüsselung wurden 12.8% aller Klassen mit einer klassenspezifischen Standardabweichung unter $SDk = 0.52$ als leistungshomogene Klassen definiert, wobei 0.9% eine unterdurchschnittliche, 9.6% eine durchschnittliche und 2.3% eine überdurchschnittliche mittlere Deutschkompetenz aufwiesen. Als leistungsheterogene Klassen wurden schließlich 12.4% aller Klassen mit einer klassenspezifischen Standardabweichung über $SDk = 0.80$ definiert, wobei 5.6% eine unterdurchschnittliche, 6.1% eine durchschnittliche und 0.7% eine überdurchschnittliche mittlere Deutschkompetenz aufwiesen.

Leistungshomogene Klassen waren hauptsächlich an Gymnasien zu finden (43.6% der homogenen Klassen, siehe fettgedruckte Prozentangaben in der Zeile „homogene Klassen bzgl. Deutschkompetenz“), gefolgt von der Realschule (32.7% der homogenen Klassen). Betrachtet man ausschließlich leistungshomogene Klassen mit unterdurchschnittlicher Deutschkompetenz, ist auffällig, dass es sich zu 75% um Hauptschulklassen handelte (siehe Prozentangaben in Klammern in der Zeile „homogene Klassen bzgl. Deutschkompetenz, mit unterdurchschnittlicher Deutschkompetenz“). Betrachtet man dagegen leistungsheterogene

Klassen mit überdurchschnittlicher Deutschkompetenz handelte es sich zu 100% um Gymnasialklassen.

Klassen mit durchschnittlicher klassenspezifischer Standardabweichung waren ebenfalls am häufigsten Gymnasialklassen (39.8% der Klassen mit durchschnittlichem *SDk*), gefolgt von Realschulklassen (25.4% der Klassen mit durchschnittlichem *SDk*). Auch hier waren jene Klassen mit unterdurchschnittlicher Deutschkompetenz mit 57.4% überwiegend Hauptschulklassen, wohingegen es sich bei jenen Klassen mit überdurchschnittlicher Deutschkompetenz zu 100% um Gymnasialklassen handelte.

Leistungsheterogene Klassen waren hauptsächlich an Hauptschulen zu finden (47.2% der heterogenen Klassen), gefolgt von Gymnasien (22.6% der heterogenen Klassen). Auch der überwiegende Teil der heterogenen Klassen mit unterdurchschnittlicher Deutschkompetenz fand sich an Hauptschulen (83.3%), während 100% der heterogenen Klassen mit überdurchschnittlicher Deutschkompetenz Gymnasialklassen waren.

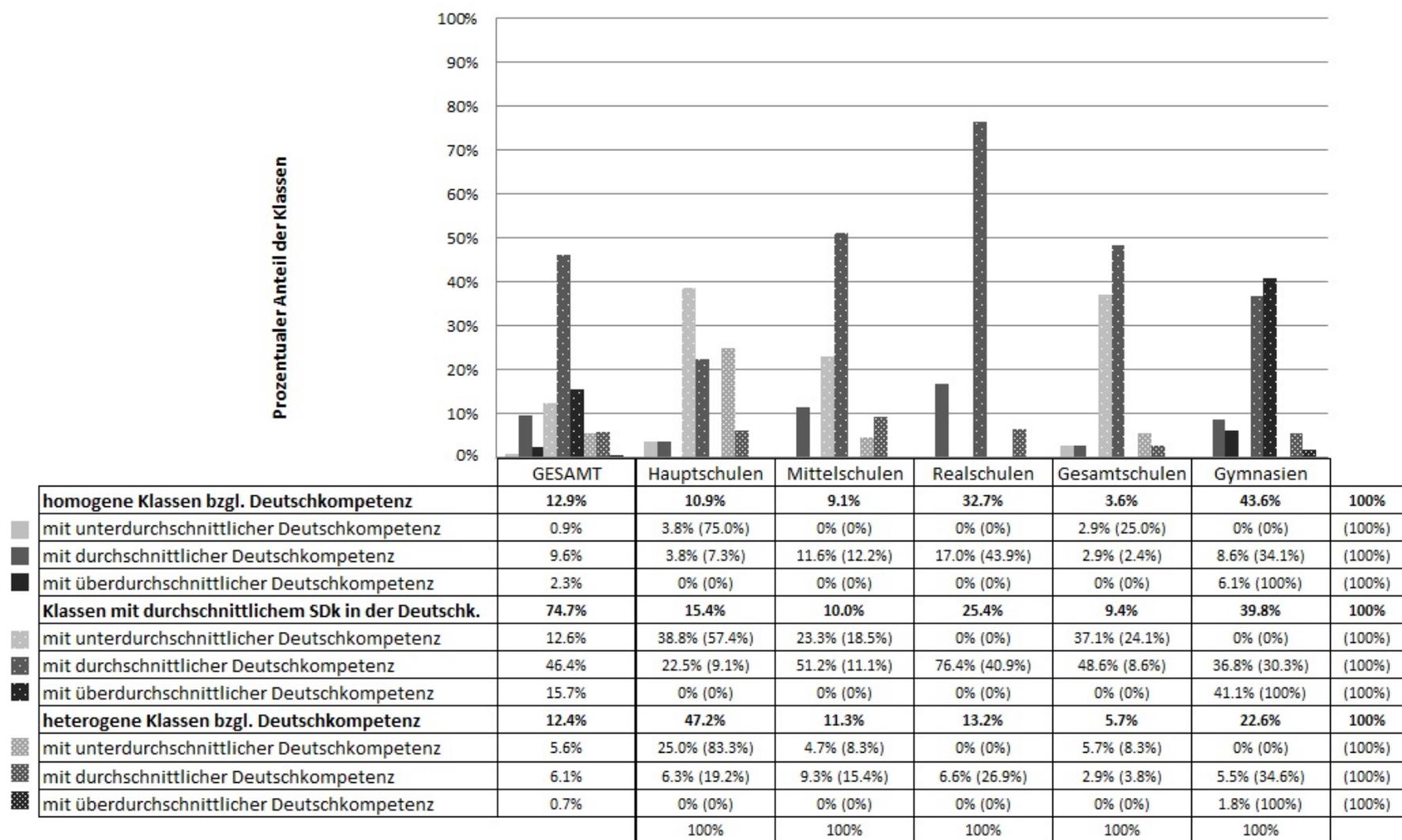


Abbildung 5.3. Die Verteilung homogener, durchschnittlicher und heterogener Klassen hinsichtlich der Deutschkompetenz, insgesamt und pro Schulform

Sozioökonomischer Hintergrund

Die mittlere klassenspezifische Standardabweichung aller Klassen bezüglich des HISEI lag bei $M_{SDk} = 0.87$ mit einer Standardabweichung von $SD_{M_{SDk}} = 0.17$. Eine durchschnittliche Klasse wies somit eine klassenspezifische Standardabweichung zwischen $SDk = 0.70$ und $SDk = 1.04$ auf. Dies traf auf 70.5% aller Klassen zu, wobei 10.5% einen unterdurchschnittlichen, 45.0% einen durchschnittlichen und 15.0% einen überdurchschnittlichen mittleren HISEI aufwiesen (siehe Abb. 5.4). Mit einer klassenspezifischen Standardabweichung unter $SDk = 0.70$ wurden 15.2% aller Klassen als homogene Klassen definiert, wobei 6.3% einen unterdurchschnittlichen, 7.5% einen durchschnittlichen und 1.4% einen überdurchschnittlichen mittleren HISEI aufwiesen. Die übrigen 14.3% der Klassen wurden mit einer klassenspezifischen Standardabweichung über $SDk = 1.04$ als heterogene Klassen definiert, wobei 0.9% einen unterdurchschnittlichen, 9.4% einen durchschnittlichen und 4.0% einen überdurchschnittlichen mittleren HISEI aufwiesen.

Homogene Klassen waren hauptsächlich an Hauptschulen zu finden (32.3%), gefolgt von den Mittel- und Gesamtschulen (jeweils 18.5% der homogenen Klassen). Betrachtet man ausschließlich homogene Klassen mit unterdurchschnittlichem HISEI sticht hervor, dass es sich zu 55.6% um Hauptschulklassen handelte. Dagegen fanden sich 100% der homogenen Klassen mit überdurchschnittlichem HISEI an Gymnasien.

Klassen mit durchschnittlicher klassenspezifischer Standardabweichung waren am häufigsten Gymnasialklassen, gefolgt von Realschulklassen (39.5% bzw. 27.9% der Klassen mit durchschnittlichem SDk). Auch hier waren jene Klassen mit unterdurchschnittlichem HISEI mit 66.7% überwiegend Hauptschulklassen, wohingegen es sich bei jenen Klassen mit überdurchschnittlichem HISEI zu 100% um Gymnasialklassen handelte.

Heterogene Klassen waren zu über der Hälfte an Gymnasien zu finden (57.4% der heterogenen Klassen). Bei den heterogenen Klassen mit unterdurchschnittlichem HISEI handelte es sich überwiegend um Hauptschulklassen (75.0%). Dagegen waren 100% der heterogenen Klassen mit überdurchschnittlichem HISEI Gymnasialklassen.

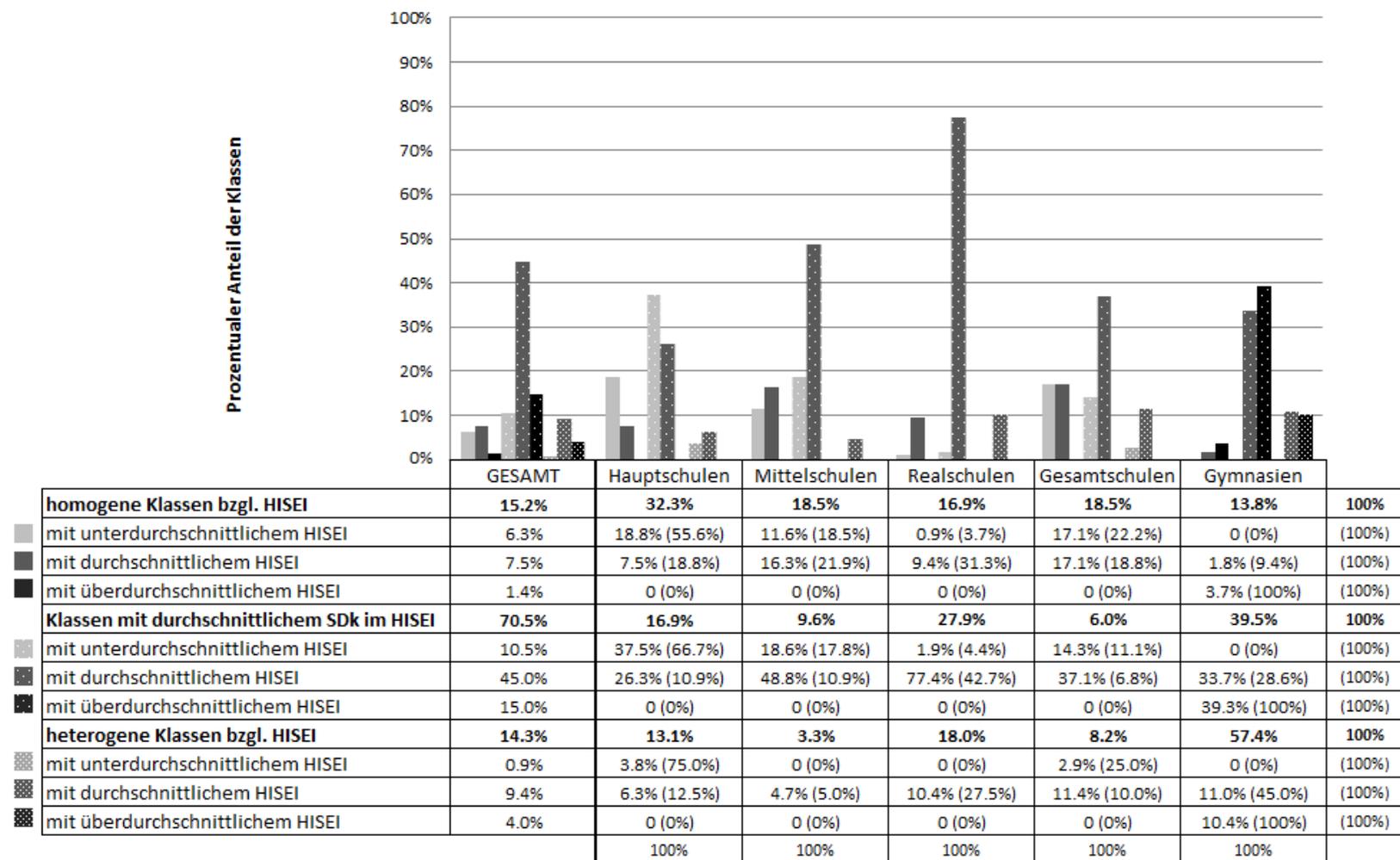


Abbildung 5.4. Die Verteilung homogener, durchschnittlicher und heterogener Klassen hinsichtlich des HISEI, insgesamt und pro Schulform

Ethnische Herkunft

Abbildung 5.5 bildet die Heterogenität der Klassen hinsichtlich der ethnischen Herkunft ihrer Schülerschaft an den verschiedenen Schulformen ab. Die Anzahl der Nationalitäten, die die Schülerinnen und Schüler pro Klasse aufwiesen, war mit im Durchschnitt 4.19 verschiedenen Nationalitäten in Hauptschulklassen am höchsten und mit durchschnittlich 2.75 verschiedenen Nationalitäten in Gymnasialklassen bzw. 1.72 in Mittelschulklassen deutlich geringer. Dieser Wert streute tendenziell zwischen Hauptschulklassen am stärksten.

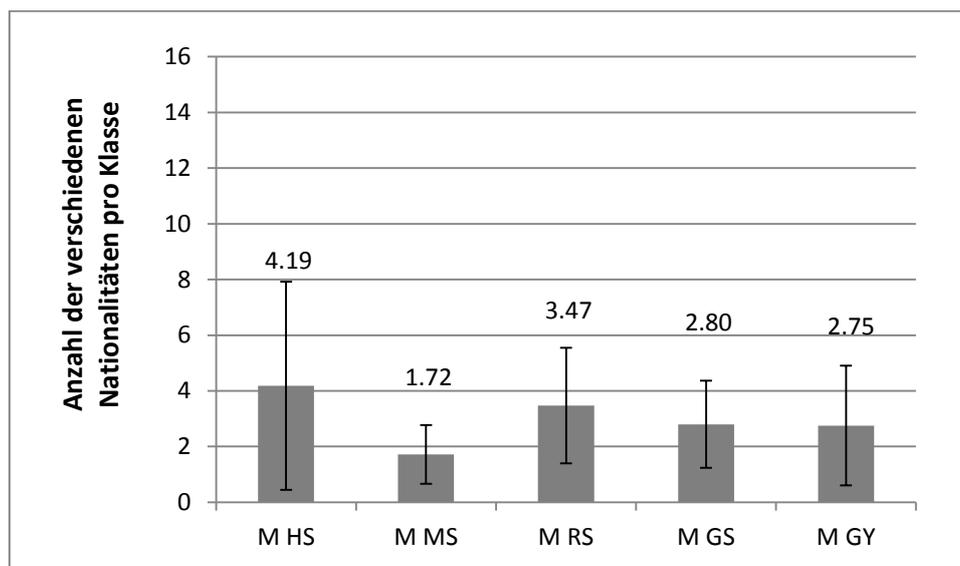


Abbildung 5.5. Mittlere Anzahl der verschiedenen Nationalitäten pro Klasse (inkl. deutsch), getrennt nach Schulform; HS = Hauptschulen, MS = Mittelschulen, RS = Realschulen, GS = Gesamtschulen, GY = Gymnasien

Lesefreude

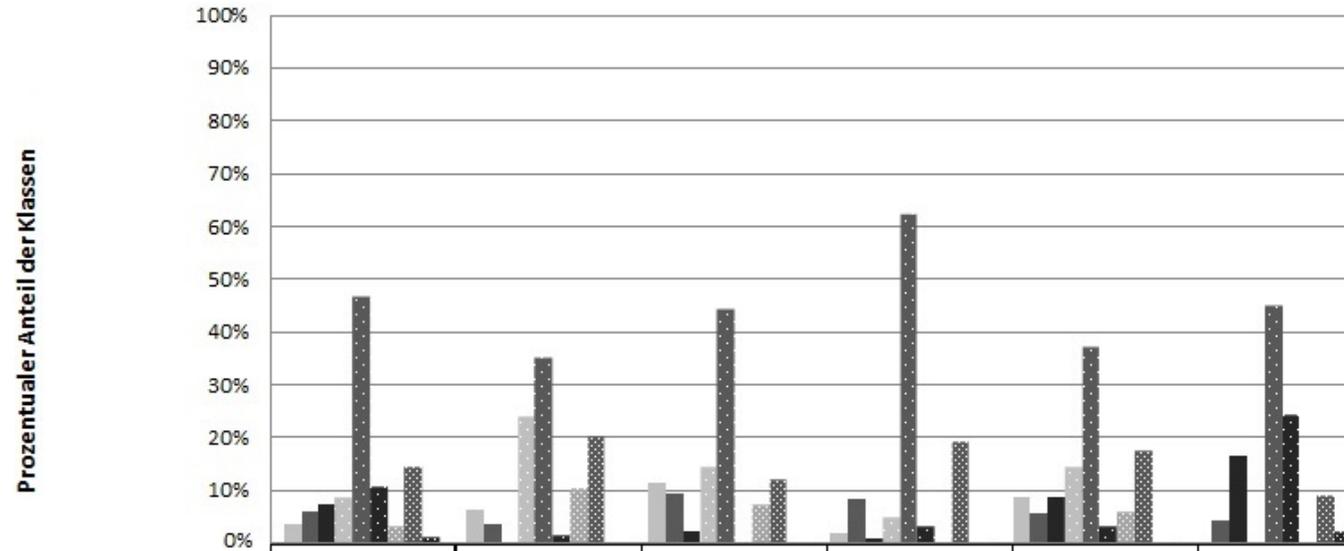
Die mittlere klassenspezifische Standardabweichung aller Klassen lag für die Lesefreude bei $M_{SDk} = 0.93$ mit einer Standardabweichung von $SD_{M_{SDk}} = 0.56$. Eine durchschnittliche Klasse wies somit eine klassenspezifische Standardabweichung zwischen $SDk = 0.79$ und $SDk = 1.07$ auf. Dies traf auf 65.1% der Klassen zu, wobei 8.2% eine unterdurchschnittliche, 46.6% eine durchschnittliche und 10.3% eine überdurchschnittliche mittlere Lesefreude aufwiesen (siehe Abb. 5.6). Mit einer klassenspezifischen Standardabweichung unter $SDk = 0.79$ wurden 16.9% aller Klassen als homogene Klassen definiert. Davon zeigten 3.5% eine unterdurchschnittliche, 5.9% eine durchschnittliche und 7.5% eine überdurchschnittliche mittlere Lesefreude. Die übrigen 18.0% der Klassen wurden mit einer klassenspezifischen Standardabweichung über $SDk = 1.07$ als heterogene Klassen definiert, wovon 3.0% eine

unterdurchschnittliche, 14.3% eine durchschnittliche und 0.7% eine überdurchschnittliche mittlere Lesefreude aufwiesen.

Homogene Klassen fanden sich zu fast der Hälfte an Gymnasien (47.2% der homogenen Klassen). Homogene Klassen mit unterdurchschnittlicher Lesefreude waren zu je 33.3% Hauptschul- und Mittelschulklassen, wogegen es sich bei den homogenen Klassen mit überdurchschnittlicher Lesefreude zu 84.4% um Gymnasialklassen handelte.

Klassen mit durchschnittlicher klassenspezifischer Standardabweichung waren am häufigsten Gymnasialklassen, gefolgt von Realschulklassen (40.3% bzw. 26.6% der Klassen mit durchschnittlichem *SDk*). Auch hier waren jene Klassen mit unterdurchschnittlicher Lesefreude mit 54.3% überwiegend Hauptschulklassen, wohingegen es sich bei jenen Klassen mit überdurchschnittlicher Lesefreude zu 88.6% um Gymnasialklassen handelte.

Heterogene Klassen waren am häufigsten an Hauptschulen vorfindbar, gefolgt von Realschulen (31.2% bzw. 26.0% der heterogenen Klassen). Aber auch an Gymnasien fanden sich 22.1% der heterogenen Klassen. Der überwiegende Teil der heterogenen Klassen mit unterdurchschnittlicher Lesefreude waren an Hauptschulen (61.5%), während 100% der heterogenen Klassen mit überdurchschnittlicher Lesefreude Gymnasialklassen waren.



	GESAMT	Hauptschulen	Mittelschulen	Realschulen	Gesamtschulen	Gymnasien	
homogene Klassen bzgl. Lesefreude	16.9%	11.1%	13.9%	16.7%	11.1%	47.2%	100%
mit unterdurchschnittlicher Lesefreude	3.5%	6.3% (33.3%)	11.6% (33.3%)	1.9% (13.3%)	8.6% (20.0%)	0 (0%)	(100%)
mit durchschnittlicher Lesefreude	5.9%	3.8% (12.0%)	9.3% (16.0%)	8.5% (36.0%)	5.7% (8.0%)	4.3% (28.0%)	(100%)
mit überdurchschnittlicher Lesefreude	7.5%	0 (0%)	2.3% (3.1%)	0.9% (3.1%)	8.6% (9.4%)	16.6% (84.4%)	(100%)
Klassen mit durchschnittlichem SDk der Lesefreude	65.1%	17.3%	9.0%	26.6%	6.8%	40.3%	100%
mit unterdurchschnittlicher Lesefreude	8.2%	23.8% (54.3%)	14.0% (17.1%)	4.7% (14.3%)	14.3% (14.3%)	0 (0%)	(100%)
mit durchschnittlicher Lesefreude	46.6%	35.0% (14.1%)	44.2% (9.5%)	62.3% (33.2%)	37.1% (6.5%)	44.8% (36.7%)	(100%)
mit überdurchschnittlicher Lesefreude	10.3%	1.3% (2.3%)	0 (0%)	2.8% (6.8%)	2.9% (2.3%)	23.9% (88.6%)	(100%)
heterogene Klassen bzgl. Lesefreude	18.0%	31.2%	10.4%	26.0%	10.4%	22.1%	100%
mit unterdurchschnittlicher Lesefreude	3.0%	10.0% (61.5%)	7.0% (23.1%)	0 (0%)	5.7% (15.4%)	0 (0%)	(100%)
mit durchschnittlicher Lesefreude	14.3%	20.0% (26.2%)	11.6% (8.2%)	18.9% (32.8%)	17.1% (9.8%)	8.6% (23%)	(100%)
mit überdurchschnittlicher Lesefreude	0.7%	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1.8% (100%)	(100%)
		100%	100%	100%	100%	100%	

Abbildung 5.6. Die Verteilung homogener, durchschnittlicher und heterogener Klassen hinsichtlich der Lesefreude, insgesamt und pro Schulform

Diskussion

Zusammenfassung der Ergebnisse

Im vorliegenden Beitrag wurde ein erster Einblick gegeben, wie heterogen die Schülerschaft *innerhalb von Schulklassen* in Deutschland zusammengesetzt ist. Dargestellt wurde auch, ob dies nach Schulform variiert. Aus der deskriptiven Betrachtung der klassenspezifischen Standardabweichungen ging hervor, dass leistungsbezogene, herkunftsbezogene und psychosoziale Schülermerkmale innerhalb von Klassen streuten. So kann trotz der leistungsbezogenen Schulformgliederung, die das mittlere Leistungsniveau der Klassen an den Schulformen beeinflusst, nicht von „homogenen“ Lerngruppen gesprochen werden. Die Aussage Tillmanns (2007, S. 8), dass die „homogene Lerngruppe eine Fiktion sei“, die er aufgrund der in der PISA-Studie gezeigten Leistungsüberschneidungen zwischen Lernenden verschiedener Schulformen traf (Baumert & Schümer, 2001), wird damit bestätigt. In der vorliegenden Studie hatten Lehrerinnen und Lehrer an allen Schulformen mit Klassen zu tun, die nicht nur heterogen hinsichtlich der üblicherweise einbezogenen Heterogenitätskategorien Leistung und familiärer Hintergrund waren, sondern vor allem auch hinsichtlich psychosozialer Merkmale. Im Vergleich der verschiedenen Merkmale zeigten sich die von uns betrachteten psychosozialen Merkmale Lesefreude und soziales Selbstkonzept innerhalb der Klassen heterogener als leistungs- und herkunftsbezogene Merkmale.

Die differenziertere Betrachtung von leistungsbezogenen, herkunftsbezogenen und psychosozialen Schülermerkmalen erweiterte die Sicht auf Heterogenität durch die gleichzeitige Betrachtung von Heterogenität und mittlerer Ausprägung. Es zeigte sich, dass Heterogenität an den verschiedenen Schulformen eine unterschiedliche „Gestalt“ annehmen kann. So waren heterogene Klassen an Hauptschulen gleichzeitig von einem ungünstigen mittleren Niveau hinsichtlich der leistungsbezogenen, herkunftsbezogenen und psychosozialen Schülermerkmale geprägt. Dagegen wiesen heterogene Klassen an Gymnasien günstige mittlere Ausprägungen hinsichtlich ebendieser Merkmale auf. Während dieses Muster für die leistungsbezogenen und herkunftsbezogenen Merkmale aufgrund der Kenntnisse über ihre generell ungünstigere Ausprägung an Hauptschulen (Baumert et al., 2006) nicht sehr überraschend ist, ist dies besonders für das hier betrachtete psychosoziale Merkmal Lesefreude neu.

Limitationen

Grundsätzlich ist anzumerken, dass die präsentierten Darstellungen rein deskriptiver Natur sind und die klassenspezifische Standardabweichung nur als eine Annäherung an das

theoretisch komplexe Konzept der Heterogenität gelten kann. Darüber hinaus wurden die Cut-Off-Werte für die Einteilung in homogene, durchschnittliche und heterogene Klassen ausschließlich auf der Basis statistischer Kennwerte gewählt. Inwieweit Lehrerinnen und Lehrer die auf dieser Grundlage als heterogen, homogen und durchschnittlich definierten Klassen tatsächlich als solche wahrnehmen, bleibt eine offene Frage. Auch wurden die Merkmale in der vorliegenden Studie nur einzeln betrachtet. Dies spiegelt nur bedingt die Realität wider, in der die Merkmale gemeinsam auftreten und zahlreiche Wechselwirkungen zwischen einzelnen Merkmalen existieren (unter dem Begriff der *Intersektionalität* bekannt; Walgenbach, 2014).

Weiterhin muss für die Interpretation der deskriptiven Darstellungen beachtet werden, dass in den vergangenen Jahren viele Bundesländer auf eine Verringerung der Mehrgliedrigkeit umgestiegen sind (Neumann et al., 2013). In vielen Bundesländern wurden die nicht akademisch orientierten Schulformen wie Haupt- und Realschulen zu einer Sekundarstufe zusammengelegt, was in den vorliegenden Analysen nicht berücksichtigt werden konnte.

Implikationen für Forschung und Praxis

Die Befunde sind für Studien zu Effekten der Heterogenität in Lerngruppen auf die individuelle Leistungsentwicklung aufschlussreich. Angesichts der geringen Varianz zwischen den Schulformen in den klassenspezifischen Standardabweichungen ist nicht verwunderlich, dass Studien zu Effekten der Heterogenität bisher keine eindeutigen Befunde ergaben. Da statistische Effekte auf Varianzunterschieden basieren und mittlere Ausprägungen von Schülermerkmalen stark zwischen den Schulformen variieren (Baumert et al., 2006), zeigen Studien zu Effekten der mittleren Ausprägung von Schülermerkmalen auf die individuelle Leistungsentwicklung ein viel klareres Bild. Untersuchungen zu Effekten der Heterogenität von Lerngruppen stehen somit vor der Aufgabe, eine geeignete methodische Umsetzung zu finden. Jenseits der Aggregation von Schülermerkmalen sind beispielweise Effekte der durch die Lehrerinnen und Lehrer wahrgenommenen Zusammensetzung der Schülerschaft denkbar oder durch Klassenprofile, die aus der Kombination von klassenspezifischer Standardabweichung und Mittelwert verschiedener Schülermerkmale gebildet werden.

Aus den Befunden der vorliegenden Untersuchung lässt sich für zukünftige theoretische Beiträge und Studien zur Heterogenität ableiten, dass ein stärkerer Fokus auf die Heterogenität psychosozialer Schülermerkmale erforderlich ist. Die Heterogenität der

Schülerschaft zeigte sich nicht nur in den üblicherweise betrachteten leistungs- und herkunftsbezogenen Heterogenitätskategorien, sondern insbesondere auch hinsichtlich psychosozialer Merkmale. Letztere waren innerhalb der Klassen an allen Schulformen heterogener verteilt als leistungs- und herkunftsbezogene Schülermerkmale.

Für die Unterrichtspraxis und den von der KMK als Schlüsselfrage der Schul- und Unterrichtsentwicklung postulierten Umgang mit Heterogenität (KMK, 2015) bedeutet dies, dass diesbezüglich nicht nur die Leistung einbezogen werden sollte, sondern es ebenfalls auf einen förderlichen Umgang mit der Heterogenität der Lernenden hinsichtlich psychosozialer Merkmale ankommt. Die vorliegende Untersuchung leistet somit einen Beitrag zur Sensibilisierung für die Heterogenität der Schülerschaft in ihren psychosozialen Merkmalen, die für einen erfolgreichen Umgang mit Heterogenität mit dem Ziel der optimalen Förderung jedes Einzelnen zu berücksichtigen sind (Wischer, 2014).

Literaturverzeichnis

- Baumert, J. & Lehmann, R. (1997). *TIMSS. Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde*. Opladen: Leske + Budrich.
- Baumert, J. & Schümer, G. (2001). Schulformen als selektionsbedingte Lernmilieus. 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. 454–467). 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.
- Becker, R. (2014). Heterogenität als Problem oder als Potenzial? In BMBF (Hrsg.), *Bildungsforschung 2020 -Herausforderungen und Perspektiven. Dokumentation der Tagung des Bundesministeriums für Bildung und Forschung vom 29.-30. März 2012* (S. 44–48). Berlin: BMBF.
- Boller, S. (Hrsg.) (2007). *Heterogenität in Schule und Unterricht. Handlungsansätze zum pädagogischen Umgang mit Vielfalt*. Weinheim: Beltz.
- Bourdieu, P. (1997). Ökonomisches Kapital-Kulturelles Kapital-Soziales Kapital. In F. Baumgart (Hrsg.), *Theorien der Sozialisation* (S. 217–231). Bad Heilbrunn: Klinkhardt.
- Bräu, K. & Schwerdt, U. (Hrsg.) (2005). *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule*. Münster: LIT-Verlag.
- Breidenstein, G. (2014). Die Individualisierung des Lernens unter den Bedingungen der Institution Schule. In B. Kopp, S. Martschinke, M. Munser-Kiefer, M. Haider, E.-M. Kirschhock, G. Ranger & G. Renner (Hrsg.), *Individuelle Förderung und Lernen in der Gemeinschaft. Jahrbuch Grundschulforschung* (S. 35–50). Wiesbaden: Springer VS.
- Budde, J. (Hrsg.) (2013). *Unscharfe Einsätze: (Re-)Produktion von Heterogenität im schulischen Feld*. Wiesbaden: Springer Fachmedien.
- Budde, J. (2012). *Problematisierende Perspektiven auf Heterogenität als ambivalentes Thema der Schul- und Unterrichtsforschung*. Weinheim: Beltz Juventa.

- 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, DC: US Government Printing Office.
- Danhier, J. & Martin, É. (2014). Comparing Compositional Effects in Two Education Systems: The Case of the Belgian Communities. *British Journal of Educational Studies*, 62(2), 171–189.
- Decristan, J. & Jude, N. (2017). Heterogenitätskategorie Schulleistung. In T. Bohl, J. Budde & M. Rieger-Ladich (Hrsg.), *Umgang mit Heterogenität in Schule und Unterricht* (S. 109–122). Stuttgart: UTB.
- Ding, W. & Lehrer, S. (2007). Do peers affect student achievement in China's secondary schools? *Review of Economics and Statistics*, 89(2), 300–312.
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60(3), 163–183.
- Dräger, J. (2009). Individuelle Förderung für ein faires und leistungsstarkes Schulsystem. In Bertelsmann Stiftung (Hrsg.), *Heterogenität und Bildung. Individuelle Förderung in Deutschland - Hindernisse und Herausforderungen* (S. 4–8). Gütersloh: Bertelsmann-Stiftung.
- Evertson, C. M., Sanford, J. P. & Emmer, E. T. (1981). Effects of class heterogeneity in junior-high school. *American Educational Research Journal*, 18(2), 219–232.
- Faulstich-Wieland, H. (2010). Mädchen und Jungen im Unterricht. In A. Buholzer & A. Kummer Wyss (Hrsg.), *Alle gleich - alle unterschiedlich! Zum Umgang mit Heterogenität in Schule und Unterricht* (S. 16–27). Zug: Klett.
- Feyerer, E. (2011). Inklusion meint mehr als Integration. *Journal für LehrerInnenbildung*, 11(4), 9–18.
- Fischer, C. (2013). Individuelle Förderung. Umgang mit Vielfalt als Herausforderung für die Schule und Lehrerbildung. *Engagement*, 4, 281–290.
- Fischer, C. (2014). *Individuelle Förderung als schulische Herausforderung*. Berlin: Friedrich-Ebert-Stiftung.
- Ganzeboom, H. B. G., de Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.

- Gebauer, M. M., McElvany, N. & Klukas, S. (2013). Einstellungen von Lehramtsanwärterinnen und Lehramtsanwärttern zum Umgang mit heterogenen Schülergruppen in Schule und Unterricht. In N. McElvany (Hrsg.), *Jahrbuch der Schulentwicklung. Bd. 17. Sprachliche, kulturelle und soziale Heterogenität in der Schule als Herausforderung und Chance der Schulentwicklung* (S. 191–216). Weinheim: Beltz Juventa.
- Gogolin, I. (2013). Unterricht in sprachlich und kulturell heterogenen Schulen - Herausforderungen an die Lehrfähigkeit. In N. McElvany (Hrsg.), *Jahrbuch der Schulentwicklung. Bd. 17. Sprachliche, kulturelle und soziale Heterogenität in der Schule als Herausforderung und Chance der Schulentwicklung* (S. 181–190). Weinheim: Beltz Juventa.
- Gutman, L. & Schoon, I. (2013). *The impact of non-cognitive skills on outcomes for young people. Literature review*. London: Education Endowment Fund.
- Hanushek, E. A., Kain, J. F., Markman, J. M. & Rivkin, S. G. (2003). Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18 (5), 527–544.
- Hartig, J., Jude, N. & Wagner, W. (2008). Methodische Grundlagen der Messung und Erklärung sprachlicher Kompetenzen. In E. Klieme (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 34–54). Weinheim: Beltz Pädagogik.
- Heller, K.A. & Perleth Ch. (2000). *Kognitiver Fähigkeits-Test (Rev.) für 5.-12. Klassen (KFT 5-12+ R)*. Göttingen: Beltz-Testgesellschaft.
- Helmke, A. (2013). Individualisierung: Hintergrund, Missverständnisse, Perspektiven. *Pädagogik*, 65(2), 34–37.
- Hinz, A. (2008). Integration von Schülern mit (und ohne zugeschriebenen) Behinderungen. In R. Lehberger & U. Sandfuchs (Hrsg.), *Schüler fallen auf. Heterogene Lerngruppen in Schule und Unterricht* (S. 127-136). Bad Heilbrunn: Klinkhardt.
- Höhmann, K. (2009). Heterogenität: eine begriffliche Klärung. In K. Höhmann, R. Kopp, H. Schäfers & M. Demmer (Hrsg.), *Lernen über Grenzen. Auf dem Weg zu einer Lernkultur, die vom Individuum ausgeht* (S. 27–36). Opladen: Budrich.
- 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.
- Klemm, K. (2015). *Inklusion in Deutschland - Daten und Fakten*. Gutachten im Auftrag der Bertelsmann Stiftung. Gütersloh.
- Klieme, E. (2008). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie*. Weinheim: Beltz Pädagogik.

- Klieme, E. & Beck, B. (2007). *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Ergebnisse Band 1*. Weinheim: Beltz Pädagogik.
- Klippert, H. (2014). *Heterogenität im Klassenzimmer: Wie Lehrkräfte effektiv und zeitsparend damit umgehen können*. Weinheim: Beltz Pädagogik.
- Koch-Priewe, B. (2009). Reflexive Koedukation. In R. Hinz & R. Walthes (Hrsg.), *Heterogenität in der Grundschule. Den pädagogischen Alltag erfolgreich bewältigen* (S. 150–157). Weinheim: Beltz Pädagogik.
- Künsting, J., Post, S., Greb, K., Faust, G. & Lipowsky, F. (2010). Leistungsheterogenität im mathematischen Anfangsunterricht. Ein Risiko für die Leistungsentwicklung? *Zeitschrift für Grundschulforschung*, 3(1), 46–64.
- Kunze, I. (2008). Begründungen und Problembereiche individueller Förderung in der Schule - Vorüberlegungen zu einer empirischen Untersuchung. In I. Kunze & C. Solzbacher (Hrsg.), *Individuelle Förderung in der Sekundarstufe I und II* (S. 13–26). Baltmannsweiler: Schneider Verlag Hohengehren.
- Lehmann, R. (2006). Zur Bedeutung der kognitiven Heterogenität von Schulklassen für den Lernstand am Ende der Klassenstufe 4. In A. Schröder-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung: Migration, Leistungsangst und Schulübergang* (S. 109–121). Wiesbaden: VS Verlag für Sozialwissenschaften.
- LeTendre, G. K., Hofer, B. K. & Shimizu, H. (2003). What is tracking? Cultural expectations in the United States, Germany, and Japan. *American Educational Research Journal*, 40(1), 43–89.
- Marks, G. N. (2010). What aspects of schooling are important? School effects on tertiary entrance performance. *School Effectiveness and School Improvement*, 21(3), 267–287.
- Miller, S. (2014). Umgang mit Heterogenität. Stärkung der Selbst- und Sozialkompetenz von Kindern in Risikolagen. In C. Rohlf, M. Harring & C. Palentien (Hrsg.), *Kompetenz-Bildung. Soziale, emotionale und kommunikative Kompetenzen von Kindern und Jugendlichen* (2. Aufl.) (S. 243–260). Wiesbaden: Springer VS.
- Neumann, M., Maaz, K. & Becker, M. (2013). Die Abkehr von der traditionellen Dreigliedrigkeit im Sekundarschulsystem: Auf unterschiedlichen Wegen zum gleichen Ziel? *Recht der Jugend und des Bildungswesens*, 61(3), 274–292.

- Neumann, U. (2008). Schülerinnen und Schüler mit Migrationshintergrund in der Sekundarstufe I. In R. Lehberger & U. Sandfuchs (Hrsg.), *Schüler fallen auf. Heterogene Lerngruppen in Schule und Unterricht* (S. 249–263). Bad Heilbrunn: Klinkhardt.
- Prenzel, A. (2005). Heterogenität in der Bildung. Rückblick und Ausblick. In K. Bräu & U. Schwerdt (Hrsg.), *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule* (S. 19–35). Münster: LIT-Verlag.
- Prenzel, M., Sälzer, C., Klieme, E. & Köller, O. (2013). *PISA 2012. Fortschritte und Herausforderungen in Deutschland*. Münster: Waxmann.
- Rabenstein, K., Idel, T.-S. & Ricken, N. (2015). Zur Verschiebung von Leistung im individualisierten Unterricht: Empirische und theoretische Befunde zur schulischen Leitdifferenz. In J. Budde, N. Blasse, A. Bossen & G. Rißler (Hrsg.), *Heterogenitätsforschung: Empirische und theoretische Perspektiven* (S. 241–258). Weinheim: Beltz Juventa.
- Rebel, K. (Hrsg.) (2011). *Heterogenität als Chance nutzen lernen*. Bad Heilbrunn: Klinkhardt.
- Reh, S. (2005). Warum fällt es Lehrerinnen und Lehrern so schwer, mit Heterogenität umzugehen? Historische und empirische Deutungen. *Die deutsche Schule*, 97(1), 76–86.
- Rubin, D.E. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.
- Scharenberg, K. (2012). *Leistungsheterogenität und Kompetenzentwicklung: Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie*. Münster: Waxmann.
- Schwerdt, U. (2005). Impulse der Integrationspädagogik für die Heterogenitätsdebatte. In K. Bräu & U. Schwerdt (Hrsg.), *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule*. (S. 95–112). Münster: LIT-Verlag.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2015). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring* (Beschluss der 350. Kultusministerkonferenz vom 11.06.2015).
- Tillmann, K.-J. (2007). Heterogenität- Homogenität. Was wir darüber wissen. *Schul-Management*, 38(5), 8–10.
- Tillmann, K.-J. (2008). Viel Selektion - wenig Leistung: Erfolg und Scheitern in deutschen Schulen. In R. Lehberger & U. Sandfuchs (Hrsg.), *Schüler fallen auf. Heterogene Lerngruppen in Schule und Unterricht* (S. 62–78). Bad Heilbrunn: Klinkhardt.
- Trautmann, M. & Wischer, B. (2011). *Heterogenität in der Schule. Eine kritische Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Walgenbach, K. (2014). *Heterogenität - Intersektionalität - Diversity in der Erziehungswissenschaft*. Stuttgart: UTB/ Budrich.
- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.
- Warwas, J., Hertel, S. & Labuhn, A. S. (2011). Bedingungsfaktoren des Einsatzes von adaptiven Unterrichtsformen im Grundschulunterricht. *Zeitschrift für Pädagogik*, 57(6), 854–867.
- Watermann, R., Klingebiel, F. & Kurtz, T. (2010). Die motivationale Bewältigung des Grundschulübergangs aus Schüler- und Elternsicht. In K. Maaz, J. Baumert, C. Gresch & M. McElvany (Hrsg.), *Der Übergang von der Grundschule in die weiterführende Schule - Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten* (S. 355–384). Bonn: BMBF.
- Werning, R. (2010). Mehr Vielfalt für die Schule - Welche Chancen bieten heterogene Lerngruppen. *Schule heute*, 50(6), 4–10.
- Westphal, A., Gronostaj, A., Vock, M., Emmrich, R. & Harych, P. (2016). Differenzierung im gymnasialen Mathematik und Deutschunterricht - vor allem bei guten Diagnostiker/innen und in heterogenen Klassen? *Zeitschrift für Pädagogik*, 62 (1), 131–148.
- Wischer, B. (2014). Individuelle Förderung als neue Leitidee? Kritische Anmerkungen zu einer aktuellen Reformstrategie. In V. Eisenbraun & S. Uhl (Hrsg.), *Geschlecht und Vielfalt in Schule und Lehrerbildung* (S. 163–176). Münster: Waxmann.

6

Gesamtdiskussion

6 Gesamtdiskussion

In der vorliegenden Arbeit wurde das Ziel verfolgt, die bislang schmale empirische Evidenz zu Effekten der Lerngruppenzusammensetzung hinsichtlich psychosozialer Schülermerkmale zu erweitern. Dazu wurde der Fokus auf drei verschiedene Aspekte gelegt: (1) Psychosoziale Schülermerkmale werden als Merkmale betrachtet, deren Entwicklung im Schulkontext von der Komposition der Lerngruppe bedingt werden kann. (2) Psychosoziale Schülermerkmale werden in ihrer mittleren Ausprägung und in ihrer Verteilung innerhalb der Lerngruppe als Kompositionsmerkmale betrachtet. (3) Es wird das Zusammenspiel verschiedener Kompositionseffekte betrachtet. Im Folgenden werden die zentralen Befunde der vier Teilstudien der vorliegenden Dissertation zunächst entlang dieser drei Aspekte zusammengefasst, integrativ diskutiert und in die Forschung zu Kompositionseffekten eingeordnet (Kapitel 6.1). Anschließend werden die Grenzen der Arbeit bzw. ihrer Teilstudien aufgezeigt (Kapitel 6.2), bevor aus den zentralen Befunden Implikationen für die pädagogische Praxis und für die administrative Ebene des Schulsystems in ihrem gemeinsamen Wirken sowie für die zukünftige Forschung abgeleitet werden (Kapitel 6.3) sowie eine abschließende Konklusion gezogen wird (Kapitel 6.4).

6.1 Zusammenfassung und Einordnung der Befunde

6.1.1 Zentrale Befunde zur Rolle psychosozialer Schülermerkmale als Kriteriumsvariablen von Kompositionseffekten

Der erste thematische Schwerpunkt der vorliegenden Dissertation betraf die Rolle psychosozialer Schülermerkmale als Merkmale, deren Entwicklung im Schulkontext von der Komposition der Lerngruppe bedingt werden kann. Während in Teilstudie 1 Schulzufriedenheit und normverletzendes Verhalten als psychosoziale Kriteriumsvariablen von Kompositionseffekten untersucht wurden, wurden in Teilstudie 2 die Testangst und in Teilstudie 3 das akademische Selbstkonzept als Kriteriumsvariablen analysiert. Bezüglich der Entwicklung aller vier psychosozialen Merkmale zeigte sich, dass die Zusammensetzung der Schülerschaft, die Lernende angehören, mitentscheidend war.

Hinsichtlich der Entwicklung der individuellen Schulzufriedenheit zeigte sich in Teilstudie 1, dass die Schülerschaft der Schule dahingehend mitentscheidend war, dass ihre mittlere Leistung und ihre mittlere Schulzufriedenheit einen positiven Effekt ausübten. So waren Neuntklässlerinnen und Neuntklässler mit gleicher Mathematikleistung und Schulzufriedenheit nach einem Schuljahr desto zufriedener mit der Schule, je leistungsstärker die Schülerschaft war.

Auch waren sie nach einem Schuljahr umso zufriedener mit der Schule, je zufriedener die Mitschülerinnen und Mitschüler im Durchschnitt mit ihr waren. Diese Effekte blieben auch unter Kontrolle der Schulform bestehen, was darauf verweist, dass sie unabhängig von Einflüssen institutioneller Merkmale der Schulform wirkten. Zudem ergaben sich Hinweise, dass die individuelle Schulzufriedenheit negativ vom Anteil von Lernenden mit Zuwanderungshintergrund an der Schule beeinflusst wurde. Vom mittleren sozioökonomischen Status an der Schule ging kein statistisch signifikanter Effekt auf die individuelle Schulzufriedenheit aus. Während die Befunde zu Kompositionseffekten durch die mittlere Leistung und die mittlere Schulzufriedenheit auf die individuelle Schulzufriedenheit jeweils ein Novum darstellen, wurden Kompositionseffekte durch den Anteil von Lernenden mit Zuwanderungshintergrund und durch den mittleren sozioökonomischen Status bereits von Scharenberg (2016) sowie Peetsma, van der Veen, Koopman und van Schooten (2006) untersucht, dies jedoch auf Klassenebene. Im Gegensatz zum hier gefundenen negativen Effekt des Anteils von Lernenden mit Zuwanderungshintergrund auf die Schulzufriedenheit auf Schulebene, verliefen die Effekte dabei in positiver Richtung (Scharenberg, 2016) oder waren statistisch nicht signifikant (Peetsma et al., 2006). Der mittlere sozioökonomische Status der Klasse zeigte bei Scharenberg (2016) im Gegensatz zum nichtsignifikanten Effekt in der vorliegenden Studie einen positiven Effekt auf die individuelle Schulzufriedenheit. Die unterschiedlichen Befunde könnten an den verschiedenen betrachteten Ebenen liegen. Zu diskutieren wäre, welche Ebene hier die geeignetere ist. Die Klassenumwelt ist die proximale Ebene, in denen Peereinflüsse im Unterricht stattfinden können. Wird Schulzufriedenheit jedoch dahingehend gesehen, dass es konstituierend für das Schulklima ist, so könnte auch die mittlere Schulzufriedenheit in der Schule durch Interaktionen zwischen Lernenden außerhalb des Unterrichts, z. B. in den Pausen, in gemischten Arbeitsgemeinschaften oder bei Schulveranstaltungen übertragen werden.

Die Untersuchung der Entwicklung des individuellen normverletzenden Verhaltens in Teilstudie 1 ergab Hinweise eines negativen Effekts des mittleren sozioökonomischen Status an der Schule und der mittleren Schulzufriedenheit an der Schule sowie auf einen positiven Effekt des mittleren Ausmaßes des normverletzenden Verhaltens. Diese Effekte zeigten sich jedoch nicht mehr unter Kontrolle der Schulform. Bezüglich der individuellen Testangst zeigte sich in Teilstudie 2, dass deren Entwicklung im Schulkontext von der Komposition der Lerngruppe bedingt wurde: Je höher das mittlere Niveau der Testangst in der Klasse am Anfang eines

Schuljahrs war, desto höher war die individuelle Testangst am Ende des Schuljahrs. Nähere Erläuterungen hierzu folgen aufgrund der inhaltlichen Überschneidung im nächsten Kapitel.

Hinsichtlich des individuellen akademischen Selbstkonzepts wurde in Teilstudie 3 der bekannte positive BFLPE durch die mittlere Leistungskomposition repliziert. Als Nebenbefund ergab Teilstudie 3 jedoch durch die längsschnittliche Analyse über drei Messzeitpunkte und die Kontrolle des jeweils vorherigen akademischen Selbstkonzepts einen Befund, der selten diskutiert wird. Das mittlere Leistungsniveau der Klasse an einem bestimmten Zeitpunkt verlor nach einiger Zeit an Relevanz für das akademische Selbstkonzept. Stattdessen schien das eigene akademische Selbstkonzept immer wieder neu an die aktuelle mittlere Leistung adaptiert worden zu sein. Dieser Befund deckt sich mit einer aktuellen Studie, die den langfristigen Effekt der mittleren Leistung einer Lerngruppe auf das individuelle akademische Selbstkonzept nach einem Wechsel in eine neue Referenzgruppe analysierte und zu dem Ergebnis kam, dass die mittlere Leistung der vorherigen Referenzgruppe keinen Effekt mehr auf das aktuelle individuelle akademische Selbstkonzept in der neuen Referenzgruppe ausübte (Becker & Neumann, 2016).

Zusammenfassend lieferten Teilstudie 1, 2 und 3 Erkenntnisse darüber, inwieweit die Lernenden auch in der Entwicklung ihrer psychosozialen Merkmale von den Merkmalen der sie umgebenden Schülergruppe mitbedingt werden können. Dies ist insofern relevant, da mit Schule multiple Ziele verfolgt werden – so auch die psychosoziale Entwicklung von Schülerinnen und Schülern (Eccles & Roeser, 2011; Fend, 2009; Kunter, 2005). Neben dem akademischen Selbstkonzept wurden bisher jedoch nur wenige psychosoziale Schülermerkmale als Zielkriterien von Kompositionseffekten betrachtet, sodass keine eindeutige Befundlage zu Kompositionseffekten auf weitere psychosoziale Schülermerkmale besteht. An dieser Stelle kann die vorliegende Dissertation mit den Befunden beitragen, dass auch die Schulzufriedenheit und die Testangst durch verschiedene Schülermerkmale in ihrer Komposition beeinflusst werden. Zudem wurden die Befunde zum bereits vielfach untersuchten akademischen Selbstkonzept als Zielvariable von Kompositionseffekten durch eine Betrachtung des Effekts im Zeitverlauf erweitert.

6.1.2 Zentrale Befunde zur Rolle psychosozialer Schülermerkmale als Kompositionsmerkmale

Der zweite Schwerpunkt der vorliegenden Dissertation fokussierte die Rolle psychosozialer Schülermerkmale als Kompositionsmerkmale. Bisher wurde der Einfluss der Schülerkomposition

hauptsächlich bezüglich der mittleren Zusammensetzung hinsichtlich der Leistung sowie sozioökonomischer und ethnischer Hintergrundmerkmale betrachtet. Dies kann vor dem Hintergrund der Entstehungsgeschichte des Forschungszweigs zu Kompositionseffekten begründet werden, der entstand, als Forscherinnen und Forscher den Folgen der sozialen Segregation von Schulen nachgingen (Blau, 1960; Coleman et al., 1966; Wilson, 1959). Obwohl psychosoziale Merkmale zentrale Bestandteile der Wirkmechanismen sind, die zur Erklärung dieser Kompositionseffekte herangezogen werden, wird die Komposition der Lernenden hinsichtlich ihrer psychosozialen Merkmale selten betrachtet. Daher wurden in Teilstudie 1, 2 und 4 verschiedene psychosoziale Merkmale in ihrer *mittleren Ausprägung* sowie in ihrer *Verteilung* innerhalb von Lerngruppen betrachtet. Während in Teilstudie 1 Kompositionseffekte durch Schulzufriedenheit und normverletzendes Verhalten in ihrem mittleren Ausmaß innerhalb der Schule als Kompositionsmerkmale analysiert wurden, standen im Fokus der Teilstudie 2 Kompositionseffekte durch die mittlere Testangst innerhalb der Klasse. In Teilstudie 4 wurden Lesefreude und soziales Selbstkonzept in ihrer Verteilung innerhalb der Klasse deskriptiv als Kompositionsmerkmale betrachtet.

Die bereits in Kapitel 6.1.1 dargelegten Befunde aus Teilstudie 1, dass die mittlere Schulzufriedenheit positiv auf die individuelle Schulzufriedenheit und das mittlere Ausmaß des normverletzenden Verhaltens tendenziell positiv auf das individuelle normverletzende Verhalten wirkte, reihen sich in Befunde zu Peer-Contagion-Effekten als besondere Form von Kompositionseffekten ein. Ein in der Peergruppe stark auftretendes Verhalten wird demnach selbst stärker gezeigt. Zu solchen Angleichungen an die Gruppennorm kommt es besonders in der Pubertät. Die hier betrachtete Stichprobe umfasste Neuntklässlerinnen und Neuntklässler – eine Alterskohorte inmitten der Pubertät. Der Befund, dass sich die in dieser Studie untersuchten Schülerinnen und Schüler in ihrem normverletzenden Verhalten tendenziell an das Niveau des normverletzenden Verhaltens der Gruppe angleichen, ist konsistent zu Befunden von Peer-Contagion-Effekten bezüglich Aggressivität, die in zahlreichen Studien für den außerschulischen Bereich (siehe im Überblick Dishion & Tipsord, 2011) und in ersten Studien auch für den schulischen Bereich bezüglich aggressivem und unterrichtsstörendem Verhalten gefunden wurden (z. B. Müller, Hofmann, Fleischli & Studer, 2015; Thomas, Bierman & Powers, 2011; Warren, Schoppelrey, Moberg & McDonald, 2005). Über unerwünschte psychosoziale Merkmale wie Aggressivität hinaus, liegen kaum Befunde für Peer-Contagion-Effekte vor, somit auch nicht bezüglich wünschenswerter psychosozialer Merkmale. An diesem Punkt erweitert der Befund

zum positiven Effekt der mittleren Schulzufriedenheit auf die individuelle Schulzufriedenheit das Forschungsfeld.

Neben solchen als Peer-Contagion-Effekte definierbaren Kompositionseffekte, bei denen die Komposition eines psychosozialen Schülermerkmals auf die individuelle Entwicklung des gleichen psychosozialen Schülermerkmals wirkt, ergaben sich in Teilstudie 1 auch Hinweise für Einflüsse auf andere psychosoziale Schülermerkmale. So verhielten sich Lernende tendenziell nach einem Schuljahr umso weniger normverletzend, je höher die mittlere Schulzufriedenheit in der Schule war. Gemäß Theorien zum Schulklima ist die Schulzufriedenheit der Lernenden konstituierend für das Schulklima (siehe im Überblick Cohen, McCabe, Michelli & Pickeral, 2009). Ein Schulklima, welches sich durch eine hohe Schulzufriedenheit unter den Lernenden kennzeichnet, könnte so weniger Anlass für normverletzendes Verhalten geben. Auch wenn dieser Effekt unter Kontrolle der Schulform keine statistische Signifikanz mehr zeigte, könnte er für pädagogische Maßnahmen relevant sein. So zeigten sich an der akademisch anspruchsvollsten Schulform – der Gymnasien – höhere Ausprägungen der mittleren Schulzufriedenheit als an den akademisch weniger anspruchsvollen Schulformen, sodass von einer Konfundierung der mittleren Schulzufriedenheit und Merkmalen der Schulform auszugehen ist.

Als weiteres psychosoziales Kompositionsmerkmal stellte sich die Testangst heraus, die in Teilstudie 2 analysiert wurde. So ging von der mittleren Testangst innerhalb einer neu formierten Schulklasse ein positiver Kompositionseffekt auf die individuelle Testangst der Lernenden aus. Dies bestätigt Vermutungen vonseiten der Forschung zu Testangst, dass die individuelle Entwicklung von Leistungsemotionen wie der Testangst nicht nur von individuellen Merkmalen bedingt wird, sondern auch von sozialen Umweltfaktoren (Zeidner, 2014). Zugleich deckt es sich mit der Annahme der Theorien zu Peer-Contagion-Effekten (Brown, Bakken, Ameringer & Mahon, 2008) und zur Bezugsgruppentheorie (Kelley, 1968), dass Bezugsgruppen durch ihre normative Funktion eine Orientierung für eigene Einstellungen, eigenes Verhalten und weiterer psychosozialer Merkmale sein können. In einer neuen sozialen Umwelt wie einer neuen Schulklasse nach dem Übergang von der Elementarschule in die Sekundarschule, wie es in Teilstudie 2 der Fall war, bestehen mehr Unsicherheiten als in einer gewohnten sozialen Umgebung (Brown, 2013; Wigfield & Eccles, 1990). Auch für die Entwicklung der eigenen Testangst könnten dann die Mitschülerinnen und Mitschüler als Orientierung herangezogen werden, um einzuschätzen, wie viel Angst z. B. vor den Leistungstests der neuen Lehrerinnen und Lehrer „vonnöten“ ist. Sie können beobachten, wie viel Testangst die anderen haben und so

einen gewissen Normwert von Testangst in der Klasse wahrnehmen. Weiterhin ist denkbar, dass eine bestimmte Klassenkultur hinsichtlich der Testangst herrscht und damit auch eine Norm in der Klasse indirekt oder direkt vorgegeben wird, wieviel Testangst angemessen ist. In der Einschätzung der eigenen Testangst könnte sich dann entsprechend an die Norm der Klasse angepasst werden. Anzumerken ist, dass die mittlere Testangst der Klasse nach einer gewissen Zeit, die in dem neuen Umfeld verbracht wurde, keinen Effekt mehr hatte. Neben dem Befund, dass Testangst als psychosoziales Kompositionsmerkmal eine Rolle für die individuelle schulische Entwicklung spielen kann, konnten differenzielle Unterschiede für den untersuchten Kompositionseffekt auf die individuelle Testangst herausgestellt werden. Bei Lernenden mit einer hohen individuellen Testangst sowie bei Mädchen zeigte sich ein stärkerer Kompositionseffekt der mittleren Testangst auf die individuelle Testangst. Dies bestätigt Annahmen in der Theorie zu Peergruppen, dass der Einfluss durch Peers unter anderem von individuellen Merkmalen abhängt (Prinstein, 2007).

In den benannten Analysen von Teilstudie 1 und 2 wurden die Kompositionsmerkmale jeweils in ihrer mittleren Ausprägung innerhalb der Lerngruppe betrachtet. Da diskutiert wird, ob auch die Heterogenität von Schülermerkmalen in Lerngruppen eine Rolle für die individuelle Entwicklung spielt (Hattie, 2002; Hong, Corter, Hong & Pelletier, 2012; Lou et al., 1996; Roßbach & Wellenreuther, 2002), wurden Kompositionsmerkmale in der vorliegenden Dissertation weiterhin im Sinne der Verteilung von Schülermerkmalen in Klassen konzeptualisiert. Während für die Leistungsentwicklung inkonsistente Befunde für Effekte durch die Verteilung der Leistungsmerkmale bestehen, kann bei der Analyse psychosozialer Merkmale auf keine Befunde gebaut werden. Generell ist bislang erst wenig bekannt darüber, wie Schülermerkmale sich innerhalb von Schulklassen verteilen. Teilstudie 4 stellt hierzu deskriptiv Befunde über die Verteilung der Leistung sowie herkunftsbezogener und psychosozialer Merkmale in Schulklassen im deutschen Schulsystem dar. Es zeigte sich, dass Schülerinnen und Schüler durch die Schulformgliederung zwar leistungsdifferenten Schulformen zugewiesen wurden, dadurch jedoch keine homogenen Klassen gebildet wurden. Dies bestätigt die Aussage Tillmanns (2007), dass die „homogene Lerngruppe eine Fiktion sei“ (S. 8). Vor allem die beiden hierbei betrachteten psychosozialen Merkmale Lesefreude und soziales Selbstkonzept streuten innerhalb von Schulklassen stark. Inwieweit solch eine Verteilung psychosozialer Merkmale auf die schulische Entwicklung wirkt, muss in weiteren Studien geklärt werden. Die Teilstudie machte jedoch deutlich, dass Heterogenität auch bezüglich psychosozialer Merkmale besteht,

welche in bildungspolitischen und -wissenschaftlichen Debatten zur Heterogenität nur selten mitdiskutiert werden, obwohl sie offenbar einen relevanten Aspekt der Schülerkomposition darstellen.

Resümierend ergibt sich, dass die Komposition einer Lerngruppe multidimensional ist und auch die psychosozialen Merkmale umfassen bzw. umfassen sollte. Ausgehend von der Theorie zu Bezugsgruppen, Peer-Effekten im Jugendalter sowie zum Schulklima ließ sich die Bedeutung der psychosozialen Komposition für die individuelle Entwicklung bestätigen. In der vorliegenden Dissertation zeigten sich Effekte auf die individuelle schulische Entwicklung durch die psychosoziale Komposition der Lerngruppe am Beispiel der mittleren Testangst innerhalb der Klasse sowie der mittleren Schulzufriedenheit und des mittleren Ausmaßes des normverletzenden Verhaltens innerhalb der Schule.

6.1.3 Zentrale Befunde zum Zusammenspiel verschiedener Kompositionseffekte

Den dritten Schwerpunkt der vorliegenden Dissertation bildete das Zusammenspiel verschiedener Kompositionseffekte. Aufgrund der Komplexität interpersonalen sozialer Einflüsse innerhalb des Ökosystems Schule (Eccles & Roeser, 2011) und intrapersonalen Wechselspielen zwischen verschiedenen Schülermerkmalen (Pekrun & Helmke, 1991) können einzelne Kompositionseffekte nur einen begrenzten Ausschnitt der Realität darstellen. Ein Kompositionsmerkmal kann jedoch für ein Schülermerkmal einen erwünschten Effekt ausüben und gleichzeitig für ein anderes Schülermerkmal einen unerwünschten. Betrachtet man die Ergebnisse zur mittleren Leistung auf die Schulzufriedenheit aus Teilstudie 1 und jene auf das akademische Selbstkonzept in Teilstudie 3, so fällt auf, dass eine hohe mittlere Leistung gemäß des BFLPE nachteilig für das individuelle akademische Selbstkonzept war, jedoch förderlich für die individuelle Schulzufriedenheit. Ein BFLPE muss sich somit nicht für jedes psychosoziale Merkmal ergeben.

Auch im Vergleich der Wirkung des Kompositionseffekts der mittleren Leistung auf die individuelle Leistung und auf das individuelle akademische Selbstkonzept übt dasselbe Kompositionsmerkmal einen entwicklungsförderlichen und einen unerwünschten Effekt aus. Da zu beiden Effekte eine breite Befundlage besteht, wurde in Teilstudie 3 das Zusammenspiel verschiedener Kompositionseffekte anhand dieses Beispiels betrachtet und beide Effekte wurden simultan analysiert. Als „Nettoeffekt“ aus den gegenläufigen Effekten der mittleren Leistung ergab sich, dass der positive Effekt auf die individuelle Leistung im Verlauf eines Schuljahrs

stärker als der negative Effekt auf das akademische Selbstkonzept war. Dies kann dahingehend interpretiert werden, dass sich in leistungsstarken Klassen keine „Spirale“ aus geschwächtem akademischen Selbstkonzept und daran anschließender weiterführender Schwächung der Leistung ergibt. Stattdessen blieb der positive Effekt der mittleren Leistung auf die individuelle Leistungsentwicklung über das ganze Schuljahr bestehen.

Zusammenfassend zeigten die Befunde zum simultanen Effekt der mittleren Leistung auf die individuelle Leistung und auf das individuelle akademische Selbstkonzept, wie wichtig es ist, das Zusammenspiel verschiedener Kompositionseffekte zu betrachten und den Rahmen der Forschung zu Kompositionseffekten, der sich ursprünglich aus der Forschung zu sozialen Segregationseffekten ergab, zu erweitern.

6.2 Methodische Bewertung und Grenzen der Dissertation

Bevor Implikationen für Praxis und Forschung aus den Befunden der vorliegenden Dissertation geschlossen werden, werden zur Einschätzung und Interpretation der Befunde die Grenzen der durchgeführten Teilstudien benannt. Zunächst wird die Erfüllung der Bedingungen für die Spezifizierung von Kompositionseffekten betrachtet. Den in Kapitel 1.4.1 beschriebenen methodischen Bedingungen für die Schätzung von Kompositionseffekten wurde soweit wie möglich nachgekommen. Bei allen Analysen zu Kompositionseffekten (siehe Teilstudie 1 bis 3) handelte es sich um hierarchisch lineare Modellierungen im Längsschnitt-Design. In den Teilstudien 2 und 3 wurde die Schulklasse als Aggregatebene für die Generierung der berücksichtigten Kompositionsmerkmale gewählt, da sie nach van Ewijk und Slegers (2010a) die unmittelbare Ebene für soziale Interaktionen zwischen den beteiligten Akteuren darstellt. Bestimmte Schülermerkmale könnten jedoch in ihrer mittleren Zusammensetzung auf Schulebene eine größere Rolle spielen als auf Klassenebene, wie Gommans, Stevens, ter Bogt und Cillessen (2016) für Kompositionseffekte hinsichtlich Verhaltensnormen bezüglich aggressiven Verhaltens fanden. Für Teilstudie 1 wurde daher die Ebene der Schule gewählt, da davon auszugehen ist, dass die untersuchten Merkmale Schulfriedenheit und normverletzendes Verhalten für das Schulklima der gesamten Schule konstituierend sind (siehe im Überblick Cohen et al., 2009). Es können jedoch keine Aussagen darüber getroffen werden, inwieweit die Befunde auch für Effekte auf Klassenebene übertragen werden können.

Weiterhin wurden die Besonderheiten des jeweiligen Schulsystems berücksichtigt. Dies trifft auf Teilstudie 1 und 3 zu, da es sich um deutsche Datensätze handelte und somit das gegliederte

Bildungssystem zu beachten ist. Für Teilstudie 2, die auf einem amerikanischen Datensatz basierte, entfiel dies, da hier kein explizites tracking vorgenommen wird. Im amerikanischen Schulsystem überwiegt ein implizites tracking, sodass sich die Schulen vornehmlich hinsichtlich ihrer sozioökonomischen Komposition unterscheiden. Daher wäre es sinnvoll gewesen, in Teilstudie 2 zusätzlich die sozioökonomische Komposition zu kontrollieren. Dies war mit der verwendeten Stichprobe jedoch nicht möglich, da die Stichprobe Schulen umfasste, in denen die Lernenden aufgrund der Charakteristiken der dortigen Nachbarschaften in übermäßiger Zahl in gleichen sozioökonomischen Verhältnissen aufwuchsen und darüber hinaus den gleichen ethnischen Hintergrund aufwiesen. In Zusatzanalysen zu Teilstudie 2 mit gleichzeitiger Spezifikation von mittlerer Leistung und mittlerer Testangst der Klasse ging von beiden Kompositionsmerkmalen ein positiver, statistisch signifikanter Kompositionseffekt aus (siehe Tabelle C.2 im Appendix von Teilstudie 2).

Zur Berücksichtigung verschiedener gleichzeitig ablaufender Kompositionseffekte, könnte auch in Teilstudie 1 die mittlere Leistungskomposition sowie die sozioökonomische und ethnische Komposition hinzugezogen werden. Da der Fokus in den hier vorgenommenen Analysen zu Kompositionseffekten auf den psychosozialen Schülermerkmalen in ihrer Komposition lag und zu diesen bisher wenig Evidenz besteht, wurde in der vorliegenden Dissertation kleinschrittiger vorgegangen. Zu komplexe Modelle werfen die ungeklärte Frage nach Unter- und Überkontrolle auf. Einerseits müssen Überschätzungen von Effekten vermieden werden, die durch eine Unterkontrolle entstehen. Andererseits waren die Kompositionsmerkmale in Teilstudie 1 hoch miteinander korreliert, sodass die spezifischen Effekte der leistungs- und herkunftsbezogenen Komposition sowie der Komposition der Schulzufriedenheit und des normverletzenden Verhaltens schwer voneinander abgegrenzt werden können und eine Überkontrolle möglich ist. In Teilstudie 1 kam die Frage der Unter- und Überkontrolle bereits durch die Berücksichtigung der Schulformen auf. Wurde für die Schulform kontrolliert, so zeigten die untersuchten Effekte teilweise keine statistische Signifikanz mehr. Es ist jedoch möglich, dass die leistungsbedingte Allokation auf verschiedene Schulformen gleichzeitig mit einer Selektion von Lernenden nach ihren Einstellungen zu Schule einherging und die Schulform so mit der Kriteriumsvariable Schulzufriedenheit korreliert. Wäre dies der Fall, könnte bei Berücksichtigung der Schulform eine Überkontrolle vorliegen und die Gefahr der Unterschätzung bestehen (mediation bias, siehe Baumert, Stanat & Watermann, 2006). Vor dem Hintergrund der bislang ungeklärten Frage nach Unter- und Überkontrolle stellen die Befunde der vorliegenden

Dissertation, auch wenn nicht alle relevanten Kompositionsmerkmale simultan spezifiziert wurden, eine Basis für weitergehende Analysen zu Effekten durch die psychosoziale Komposition bereit.

Weiterhin ist anzumerken, dass die Größe der Regressionskoeffizienten für Einflüsse auf psychosoziale Merkmale oder von psychosozialen Kompositionsmerkmalen aus relativ gering ausfiel. Dies ist insofern nicht erstaunlich, als dass Unterschiede zwischen Lerngruppen in psychosozialen Merkmalen geringer ausfallen als in Bezug auf Leistung (Bonell et al., 2013; Sellström & Bremberg, 2006). Bei Metaanalysen internationaler Studien ergaben sich für den Effekt der leistungs- und hintergrundbezogenen Komposition auf individuelle Leistungsmerkmale Regressionskoeffizienten zwischen $-.26$ und $.07$ (van Ewijk & Sleegers, 2010b; Mickelson, Bottia & Lambert, 2013). Es wird zudem angenommen, dass auch kleine Effekte im Zeitverlauf relevant für die schulische Entwicklung werden können, wenn sie sich akkumulieren (van Ewijk & Sleegers, 2010b; Johnson, Crosnoe & Elder 2001). Möglicherweise hat in der untersuchten Stichprobe der Neuntklässlerinnen und Neuntklässler in Teilstudie 1 der größte Teil des Peereinflusses bereits stattgefunden, da die Lernenden schon mehrere Jahre miteinander in einer Klasse verbrachten. Ein anderer Grund für die geringen Effekte könnte in der Operationalisierung des normverletzenden Verhaltens und der Testangst liegen. Während das akademische Selbstkonzept und die Schulzufriedenheit eine subjektive Einschätzung implizieren, könnte die Angabe des normverletzenden Verhaltens und der Testangst bereits durch Peer-Effekte beeinflusst worden sein und nicht das tatsächlich gezeigte normverletzende Verhalten bzw. die eigentlich gefühlte Testangst darstellen. Besteht die Norm in einer Gruppe beispielsweise darin, dass Testangst als „weichlich“ verpönt wird, so könnten die individuellen Einschätzungen der Testangst niedriger ausfallen. Um einem objektiveren Wert näherzukommen, könnten zusätzlich zu den subjektiven Maßen Angaben durch Lehrerinnen und Lehrer hinzugezogen werden. Bei den kognitiven Komponenten der Testangst wie sorgenvolle Gedanken ist dies kaum möglich. Inwieweit die Schülerinnen und Schüler Angstsymptome vor oder während Testsituationen zeigen oder normverletzendes Verhalten, könnte jedoch auch von Lehrerinnen und Lehrern angegeben werden.

Weitere Einschränkungen betreffen die Messinvarianz. Schüler- und Kompositionsmerkmale wurden in der vorliegenden Dissertation nicht latent modelliert. Stattdessen wurden Schülermerkmale als manifeste Konstrukte ins Modell aufgenommen sowie Aggregate aus den manifesten Konstrukten auf Klassen- oder Schulebene gebildet. Die Reliabilität der Konstrukte

wurde so auf individueller Ebene nicht getestet, wodurch es zu einem Aggregationsbias mit Überschätzung von Kompositionseffekten kommen kann (Baumert et al., 2006). Auch die Kompositionsmerkmale wurden als manifeste Werte in die Modelle eingeführt. Dies ist in Teilstudie 2 und 3 als legitim zu bewerten, da Lüdtke et al. (2008) für *formative* Level 2-Konstrukte mit einem hohen Samplingrate, wie es in diesen Teilstudien der Fall war, den *multilevel manifest covariate* (MMC)-Ansatz, d.h. die manifeste Aggregation, empfehlen. Auch in Teilstudie 1 wurden die mittlere Schulzufriedenheit und das mittlere Ausmaß an normverletzendem Verhalten in der Klasse als formative Level 2-Konstrukte konzeptualisiert, da Level 1- und Level 2-Werte hier inhaltlich als verschiedene Konstrukte angesehen werden; im Gegensatz zu *reflective* Level 2-Konstrukten, wobei Level 1- und Level 2-Werte im Idealfall, d.h. bei hoher Übereinstimmung der Werte auf Level 1, als austauschbar gelten (Lüdtke et al., 2008). Studienbedingt lag in Teilstudie 1 jedoch ein höherer Stichprobenausfall vor, da nicht die gesamte Schülerschaft einer Schule befragt und getestet wurde. Ist zusätzlich zum höheren Stichprobenausfall der Intraklassenkorrelationskoeffizient (ICC) gering, wie es bei psychosozialen Schülermerkmalen häufig der Fall ist, wird von Lüdtke et al. (2008) empfohlen, sowohl den MMC-Ansatz als auch den *multilevel latent covariate* (MLC)-Ansatz zu deklinieren und zu prüfen, ob beide Ansätze ähnliche Befunde liefern. Da die auf Schulebene aggregierten Werte für normverletzendes Verhalten und Schulzufriedenheit in den Analysen zu Teilstudie 1 nur manifest spezifiziert wurden, könnten die Kompositionseffekte unterschätzt worden sein. Bei vielen Studien zum BFLPE wurde ebenfalls nur der manifeste Aggregatwert der Leistung modelliert (z. B. Marsh & Hau, 2003; Marsh, Köller & Baumert, 2001), auch wenn nicht die gesamte Schülerschaft gesampelt wurde, wobei Nagengast und Marsh (2011) feststellen konnten, dass die BFLP-Effekte in latent modellierten Analysen höher ausfielen.

In einigen Studien zu Kompositionseffekten auf psychosoziale Merkmale wird zudem ein weiterer Aspekt bei der Bildung des Klassenaggregats berücksichtigt, indem der Individualwert herausgerechnet wird und jedem Lernenden ein individueller Klassenwert zugeordnet wird, der als *individual surrounding peer environment* bezeichnet wird (Barth, Dunlap, Dane, Lochman & Wells, 2004; Henry et al., 2000; Høglund & Leadbeater, 2004). Mit diesem Vorgehen wird die Kollinearität zwischen Individualmerkmal und Kompositionsmerkmal berücksichtigt. Es stellt sich zudem die Frage, ob die Schülerkomposition durch eine andere Generierung der Klassen- oder Schulaggregate besser abgebildet werden könnte, z. B. durch den Median oder Minimal- und Maximalwerte.

In Studien zu Peer-Contagion-Effekten wird weiterhin proklamiert, den sozialen Status der Mitglieder innerhalb der Peergruppe zu berücksichtigen (Müller & Zurbriggen, 2016). Es wird davon ausgegangen, dass das Verhalten von Peers mit einem hohen sozialen Status in der Gruppe besonders relevant für das Auftreten und die Stärke von Peereinflüssen ist (Asher & McDonald, 2009). So ist auch im Schulbereich davon auszugehen, dass von populären Schülerinnen und Schülern ein besonderer Einfluss auf die schulische Entwicklung der Lernenden ausgeht. Dies kann durch soziale Netzwerkanalysen berücksichtigt werden sowie durch die Generierung angepasster Klassenmittelwerte, die das Verhalten oder die Eigenschaften der populären Lernenden stärker gewichten (Jackson, Cappella & Neal, 2015; Müller & Zurbriggen, 2016).

Fehlende Daten wurden in den Teilstudien 1 bis 3 durch die zwei derzeitigen State-of-the-Art-Methoden berücksichtigt. Um unverzerrte Parameterschätzer zu erhalten, ohne durch die fehlenden Werte an Teststärke einzubüßen (Graham, 2012), wurde in Teilstudie 1 das Verfahren der multiplen Imputation mit dem Ansatz der Multivariate Imputation by Chained Equations (MICE; van Buuren & Oudshoorn, 1999) vorgenommen. In den Teilstudien 2 und 3 wurde das in Mplus implementierte *Full-Information-Maximum-Likelihood-Verfahren* (FIML; Muthén & Muthén, 1998–2010) angewandt. In Teilstudie 4 gestaltete sich dies schwieriger. Das Verfahren der multiplen Imputation würde die Anknüpfung der Ergebnisse mit Ergebnissen anderer Untersuchungen an Daten der DESI-Studie beeinträchtigen. Die meisten Untersuchungen basieren auf den vom DESI-Konsortium imputierten Datensätzen, wobei bewusst auf die Imputation fehlender Daten der psychosozialen Merkmale verzichtet und diesbezüglich die FIML-Prozedur vorgeschlagen wurde (Hartig, Klieme & Leutner, 2008). Die in Mplus integrierte FIML-Funktion konnte jedoch nicht genutzt werden, da Mplus wenig Möglichkeiten für eine umfassende deskriptive Untersuchung zur Verteilung der Schülermerkmale innerhalb der Klassen bietet. Daher musste auf das Statistikprogramm SPSS zurückgegriffen werden, wobei die fehlenden Werte für die psychosozialen Merkmale und die Leistung (< 5% fehlende Werte) über den fallweisen Ausschluss behandelt wurden. Für die fehlenden Werte bezüglich der sozioökonomischen Herkunft wurden die vom DESI-Konsortium fast vollständig imputierten Daten gemittelt und die restlichen fehlenden Werte (9%) ebenfalls durch fallweisen Ausschluss behandelt.

Schließlich muss berücksichtigt werden, dass mit den vorliegenden Untersuchungen, die auf Surveydaten basierten, keine Kausalität nachgewiesen wird. Der Königsweg hierfür wäre ein Experimentaldesign (Wilkinson et al., 2000). Bei der Durchführung von Surveys im Schulkontext

kann die Zuweisung der Lernenden zu Experimental- und Kontrollgruppen nicht randomisiert erfolgen. Die Lernenden sind bereits auf Schulformen, Schulen, Klassen und/oder Kurse aufgeteilt, wodurch es zum Endogenitätsproblem bzw. Selection Bias kommt (Baumert et al., 2006; sowie Duncan, Magnuson & Ludwig, 2004; Heckman, 1979). Das heißt, Effekte der Schülerkomposition auf die schulische Entwicklung können nicht mehr von Selbstselektionseffekten unterschieden werden, die eine Lernende oder ein Lernender in einer Schülerschaft mit einer anderen Komposition vollzogen hätte (Imai, Keele, Tingley & Yamamoto, 2011). Während dieses Problem bei der randomisierten Zuteilung zu Experimental- und Kontrollbedingungen berücksichtigt wird, indem die Korrelation zwischen Gruppenzugehörigkeit und Outcome aufgehoben wird und sich die Eingangswerte der Gruppen so nicht mehr systematisch unterscheiden, bleibt die kausalanalytische Interpretierbarkeit der geschätzten Effekte in Schuluntersuchungen eingeschränkt.

6.3 Implikationen für die pädagogische Praxis und zukünftige Forschung

6.3.1 Implikationen für die pädagogische Praxis

Die Untersuchungen der vorliegenden Arbeit setzen an einem Aspekt an, der aufgrund des vorrangigen Ziels von Schule zur Förderung des Leistungs- und Fähigkeitsniveaus meist im Hintergrund steht: der Einfluss von Schule auf die Entwicklung der Lernenden hinsichtlich psychosozialer Merkmale (Eccles & Roeser, 2011; Pekrun & Helmke, 1991). Die Entwicklung der Lernenden in psychosozialen Merkmalen wird als eines der multiplen Ziele von Schule gesehen (Eccles & Roeser, 2011; Fend, 2009; Kunter, 2005). Auch Effizienz von Schule wird nicht nur am Kriterium der Leistungsentwicklung gemessen, sondern ebenfalls an Aspekten wie dem Schul- und Klassenklima (Schnabel, 1998). Da Schulen im Zuge der *Neuen Steuerung* wettbewerbsfähig sein müssen und Eltern heutzutage häufig mehr Möglichkeiten bei der Einzelschulwahl offenstehen (Altrichter & Merki, 2010), müssen Schulen zudem auch für „Kundenzufriedenheit“ sorgen. Hier spielt eine Rolle, ob neben der Leistungsentwicklung auch die psychosoziale Entwicklung gefördert wird, da Eltern bei der Einzelschulwahl neben der leistungsbezogenen Unterrichtsqualität beispielsweise auch einbeziehen, welche Werte durch die Schule vermittelt werden oder ob die Schule für das Wohlbefinden der Schülerinnen und Schüler sorgen kann (Klinge, 2016). Daher sollte in der Forschung zu Kompositionseffekten die Untersuchung von Effekten der Schülerzusammensetzung hinsichtlich psychosozialer Schülermerkmale aufgenommen werden.

Befunde zum Effekt der leistungs- und herkunftsbezogenen Komposition auf die individuelle Entwicklung psychosozialer Merkmale wie jene der vorliegenden Arbeit auf die individuelle Schulzufriedenheit und das normverletzende Verhalten halten für die Steuerung des Bildungssystems die Information bereit, dass die leistungs- und herkunftsbezogene Komposition nicht nur die Leistungsentwicklung bedingt, sondern auch weitere multiple Ziele von Beschulung tangiert. Zudem geht aus den Untersuchungen hervor, dass Effekte von verschiedenen Kompositionsmerkmalen zeitgleich ablaufen und dass von einem Kompositionsmerkmal gleichzeitig erwünschte wie unerwünschte Effekte auf die verschiedenen Ziele von Schule ausgehen. So muss bei der Auswahl struktureller Maßnahmen um bestimmte Kompositionseffekte zu hemmen, beachtet werden, dass ein spezifisches Schülermerkmal gefördert werden kann, sich aber Nachteile für ein anderes ergeben könnten. Daher ist abzuschätzen, welche Kompositionseffekte im jeweiligen Zusammenspiel eine prominentere Rolle einnehmen und welcher Nettoeffekt sich aus verschiedenen Kompositionseffekten für die schulische Entwicklung ergibt.

In der praktischen Umsetzung der strukturellen Maßnahmen des Schulsystems und der multiplen Ziele von Unterricht, müssen weitere Kompositionsmerkmale beachtet werden, auf die Lehrkräfte direkt Einfluss nehmen können, wie z. B. die Komposition der Lernenden hinsichtlich der Testangst und der Schulzufriedenheit. Die psychosozialen Merkmale der Schülerschaft geben einen direkteren Anknüpfungspunkt für pädagogische Maßnahmen im Vergleich zu Herkunftsmerkmalen, deren Veränderung nicht im Einflussbereich der Schule oder der Lehrerinnen und Lehrer liegt. Da Lehrkräfte mit der gegebenen Komposition hinsichtlich des sozioökonomischen Hintergrunds oder des Bildungshintergrunds arbeiten müssen, ist es von zentraler Bedeutung, an welchen Kompositionsmerkmalen sie pädagogisch ansetzen können. Die vorliegende Dissertation lieferte neben Erkenntnissen zum Einfluss der Leistungskomposition Informationen über weitere Kompositionsmerkmale wie Testangst und Schulzufriedenheit, welche die individuelle schulische Entwicklung ebenfalls mitbedingten. Beispielsweise könnten durch die gezielte Bildung von Kleingruppen aus Lernenden mit einer bestimmten Konstellation psychosozialer Merkmale, unerwünschte Kompositionseffekte eingeschränkt oder förderliche Kompositionseffekte ausgelöst werden. Auch wenn die Mechanismen der Peereinflüsse in weitergehenden Untersuchungen geklärt werden müssen, verweisen die Befunde der vorliegenden Dissertation auf normative und komparative Prozesse zwischen Lernenden. Sind

Lehrerinnen und Lehrer für solche Bezugsgruppenprozesse sensibilisiert, können sie beispielweise ungünstigen Normen, die sich in der Gruppe etabliert haben, entgegenwirken.

6.3.2 Implikationen für die Forschung zu Kompositionseffekten und Ausblick auf zukünftige Forschungsfragen

Die vorliegende Dissertation erweiterte die Befundlage zu Kompositionseffekten hinsichtlich mehrerer Aspekte. Zum einen wurde der zentrale Effekt der Schülerkomposition auf die psychosoziale Entwicklung nachgewiesen. Am Beispiel der individuellen Entwicklung der Schulzufriedenheit, des normverletzenden Verhaltens und der Testangst wurde gezeigt, dass auch psychosoziale Schülermerkmale von unterschiedlichen Kompositionsmerkmalen beeinflusst werden können. Damit leistet die vorliegende Dissertation einen Beitrag zur Entstehung eines neuen Forschungsbereichs, wie er auch von Müller und Zurbriggen (2016, S. 163f.) vorgeschlagen wurde: „The aim is to systematically explore how characteristics of the classroom composition, meaning the characteristics of students attending a specific classroom, are related to the social-emotional outcomes of individual children and adolescents. Social-emotional outcomes here are considered as an umbrella term that encompasses inter- and intrapersonal competences and processes that are essential for functioning in social interactions and relationships“.

Zum anderen liefert die vorliegende Arbeit bedeutsame Hinweise für die in der Forschung zu Kompositionseffekten bestehende Annahme, dass die Zusammensetzung der Schülerschaft über Peergruppenphänomene auf die individuelle schulische Entwicklung der Lernenden wirkt. Besonders sind hier die Befunde zu nennen, dass die individuelle Entwicklung psychosozialer Merkmale wie Schulzufriedenheit oder Testangst von der Zusammensetzung der Klasse oder Schule hinsichtlich der jeweiligen Merkmale mitbedingt ist, wobei eine Angleichung an die Normen in der Lerngruppe stattfindet. Hiermit konnte den Peer-Effekten, die als Wirkmechanismen von Effekten der leistungs- und herkunftsbezogenen Komposition auf die Leistungsentwicklung angenommen werden, etwas nähergekommen werden. In den vorliegenden Teilstudien wurden keine Wirkmechanismen untersucht. Jedoch lassen sich aus den Befunden, dass sich Lernende hinsichtlich der Schulzufriedenheit und der Testangst von der Komposition der Lerngruppe in den jeweiligen Merkmalen beeinflussen lassen sowie auch aus den tendenziellen Kompositionseffekten der mittleren Schulzufriedenheit auf das normverletzende Verhalten konkretere Vermutungen anstellen, was in der Black Box der Wirkmechanismen von Kompositionseffekten stattfindet. Zur Erhellung dieser Frage müssten in weiteren Studien die

Analysen der üblicherweise betrachteten Kompositionseffekte mit Analysen zu den hier gezeigten Peer-Effekten verknüpft werden. Es ist jedoch andererseits auch möglich, dass psychosoziale Kompositionsmerkmale nicht nur die Rolle von Mediatoren einnehmen, die sich aus leistungs- und herkunftsbezogenen Kompositionsmerkmalen ergeben, sondern als eigenständiges Kompositionsmerkmal konzeptualisiert werden könnten. Insofern kann in Frage gestellt werden, ob sich tatsächlich alle psychosozialen Merkmale aus der leistungs- und herkunftsbezogenen Komposition ergeben. Dies wird durch den Befund von Baumert et al. (2006) gestützt, dass der Anteil von aggressiven Lernenden an der Schülerschaft weder von der leistungs- noch von der herkunftsbezogenen Komposition der Schule abhing. Zudem ist zu hinterfragen, ob der Rückgriff auf diese leistungs- und herkunftsbezogenen Kompositionsmerkmale in jedem Fall die relevantere Information für bildungspolitische und pädagogische Maßnahme bereithält, vor allem, wenn im Unterrichtsalltag mit den Umständen der gegebenen Komposition gearbeitet werden muss.

In Anbetracht der erweiterten Multidimensionalität von Kompositions- und Zielmerkmalen von Kompositionseffekten wird die Relevanz der gemeinsamen Betrachtung unterschiedlicher Kompositionseffekte umso deutlicher. Einzelne Kompositionseffekte können nur einen begrenzten Ausschnitt der Realität darstellen. Soziale Interaktionen und Beziehungen innerhalb des Lern- und Entwicklungsortes Schule sind jedoch von enormer Komplexität. So konnte gezeigt werden, dass von der mittleren Leistung einer Lerngruppe nicht nur unerwünschte Effekte auf psychosoziale Merkmale wie beim BFLPE ausgehen, sondern auch erwünschte wie der positive Effekt auf die Schulzufriedenheit. Da zusätzlich auch von intrapersonalen Wechselspielen zwischen verschiedenen Schülermerkmalen auszugehen ist (Pekrun & Helmke, 1991), stellt sich die Frage, wie gleichzeitig ablaufende Kompositionseffekte miteinander im Zusammenspiel stehen. So kann die vonseiten der Forschung zu Kompositionseffekten gegebene Empfehlung, dass verschiedene Kompositionsmerkmale gleichzeitig berücksichtigt werden sollten (Baumert et al., 2006; Dumont, Neumann, Maaz & Trautwein, 2013; Thrupp, Lauder & Robinson, 2002), zusätzlich dahingehend erweitert werden, dass auch die gleichzeitige Wirkung eines Kompositionsmerkmals auf verschiedene Schülermerkmale berücksichtigt werden sollte. Die in dieser Arbeit dargelegten Befunde der gleichzeitigen Betrachtung des Kompositionseffekts der mittleren Leistung einer Lerngruppe auf die Individualleistung und das akademische Selbstkonzept untermauern dies. Hierbei zeigte sich, dass verschiedene gleichzeitig bestehende Kompositionseffekte unterschiedlich gewichtige Konsequenzen für die schulische Entwicklung

haben können. So wird der untersuchte negative Effekt der mittleren Leistung einer Lerngruppe auf das akademische Selbstkonzept in der Forschung zum BFLPE zumeist als ungünstig für den Lernprozess interpretiert (Marsh, 2005). Aus dem vorliegenden Befund, dass das vom BFLPE beeinflusste akademische Selbstkonzept selbst nicht weiter negativ auf die Leistungsentwicklung wirkt, könnte dagegen gefolgert werden, dass das schwächere akademische Selbstkonzept in einer leistungsstarken Lernumwelt möglicherweise gar nicht so einschneidend zu bewerten ist.

Weiterhin ergaben sich in der vorliegenden Dissertation Belege für die Annahme, dass Peergruppeneinflüsse verschiedenen individuellen Merkmalen unterliegen (Dishion & Tipsord, 2011; Prinstein, 2007). Der Effekt der mittleren Testangst in der Klasse auf die individuelle Testangst war stärker bei Mädchen sowie bei Lernenden ausgeprägt, die bereits vorher eine höhere Testangst angaben. Differenzielle Effekte durch verschiedene Schülermerkmale könnten in zukünftigen Studien berücksichtigt werden, um den dahinterliegenden Prozessen näherzukommen.

Die vorliegende Arbeit verweist zudem auf die Relevanz der Betrachtung von Schülermerkmalen in ihrer *Verteilung* innerhalb der Lerngruppe. Der Begriff der Komposition wird in der empirischen Bildungsforschung vorwiegend über die mittlere Zusammensetzung von Lerngruppen hinsichtlich eines Merkmals konzeptualisiert. Die Definition der Komposition über die Verteilung von Schülermerkmalen erhielt bisher weniger Aufmerksamkeit. Dies könnte mitunter daran liegen, dass zur Rolle der *Verteilung* von leistungs- und herkunftsbezogenen sowie psychosozialen Schülermerkmalen für die schulische Entwicklung noch ein weites Forschungsdesiderat besteht. Angesichts des aktuell prominenten Begriffs der Heterogenität und der damit verbundenen *Schlüsselfrage der Schul- und Unterrichtsentwicklung* (KMK, 2015) nach einem optimalen Umgang mit einer heterogenen Schülerschaft, scheint es von Relevanz, diese Forschungslücke zu füllen. Das gilt besonders hinsichtlich der multikategorialen Heterogenität auf Klassenebene als der Ebene, auf der Unterricht stattfindet. Die vorliegende Dissertation liefert diesbezüglich den dafür grundlegenden Befund, dass trotz der leistungsbezogenen Schulformgliederung in Deutschland nicht von homogenen Klassen gesprochen werden kann. Besonders heterogen zeigten sich die Klassen in den betrachteten psychosozialen Merkmalen. Dies verdeutlicht, dass Heterogenität von Schülerschaften nicht nur hinsichtlich der Leistung besteht, sondern multiple Heterogenitätskategorien berücksichtigt werden müssen. Weiterhin zeigte sich, dass sich die Verteilung der Schülermerkmale innerhalb der Klassen zwischen den Schulformen kaum unterschied. Methodisch gesehen könnte daraus abgeleitet werden, dass

möglicherweise ein geeigneteres Maß für die Verteilung von Schülermerkmalen existiert als die Varianz oder Standardabweichung, wenn Effekte der Verteilung von Schülermerkmalen untersucht werden. Statistische Effekte basieren jedoch auf Varianzunterschieden. Somit könnte die viel klarere Befundlage zu Effekten auf die individuelle Leistungsentwicklung durch die mittlere Leistung – verglichen mit Effekten durch die Leistungsverteilung – damit zusammenhängen, dass sich Schulformen in der Verteilung der Leistung innerhalb der Klassen kaum unterscheiden, während die mittlere Leistung zwischen den Schulformen stark variiert (Baumert et al., 2006). Es könnte aber auch angezweifelt werden, dass solch ein Maß der Heterogenität überhaupt generierbar ist, da Heterogenität zum einen relativ ist und es zahlreiche Schülermerkmale gibt, bezüglich derer die Schülerschaft heterogen sein kann. Es fragt sich dann, bei welchem Merkmal Heterogenität entscheidend für Schülerinnen und Schüler sowie für Lehrkräfte ist. Zum anderen unterliegen Schülermerkmale nicht nur interindividuellen Unterschieden, sondern zeigen sich auch intraindividuell unterschiedlich (Prenzel, 1995). Die Heterogenität einer Schulklasse kann daher von Tag zu Tag und von Lernsituation zu Lernsituation eine ganz andere sein.

Ein weiterer wichtiger Aspekt für zukünftige Forschungsfragen besteht in der Ausweitung der Untersuchungen von Effekten der Komposition psychosozialer Schülermerkmale auf andere Jahrgangskohorten. Die hier betrachteten Datensätze umfassen die 6., 7., 9. und 10. Jahrgangsstufe, sodass die Altersgruppen der Sekundarschule I gut abgedeckt sind. Weitergehende Forschung über Effekte der Komposition psychosozialer Schülermerkmale wäre für den Elementarbereich sowie für die Sekundarstufe II bzw. für die Berufsausbildung interessant. Im frühen Elementarbereich wäre einerseits von geringeren Peereinflüssen auszugehen, da diese ihren Höhepunkt bei Jugendlichen zwischen 11 und 16 Jahren erreichen (siehe im Überblick Müller & Minger, 2013). Andererseits könnten Analysen hier auch stärkere Effekte ergeben, da sich die Schülerinnen und Schüler noch nicht so lange gegenseitig beeinflusst haben können wie in höheren Klassenstufen. In der Sekundarstufe II bzw. in der Berufsausbildung befinden sich die Jugendlichen anfangs noch im kritischen Altersbereich. Gerade da sich Jugendliche in der Berufsausbildung neuen Rollen und Herausforderungen stellen müssen, ist eine erneute verstärkte Orientierung an den sie umgebenden Peers wie den neuen Mitschülerinnen und Mitschülern in der Berufsausbildung anzunehmen. Auch Gymnasiastinnen und Gymnasiasten in der Sekundarstufe II kommen dem Berufsleben näher und müssen sich berufsrelevante Fragen stellen, wobei der Peergruppe der Mitschülerinnen und Mitschüler, die

sich mit gleichen Fragen beschäftigt, ein relevanter normativer und komparativer Einfluss zukommen dürfte. Da Jugendliche mit höherem Alter weniger anfällig für Peereinflüsse sind, wäre es weiterhin von Interesse, wie sich Peereinflüsse im späteren Verlauf der Berufsausbildung bzw. in der Sekundarstufe II auswirken. Neben diesem altersbezogenen Aspekt von Peereinflüssen innerhalb von Klassen könnte weiterhin das Zusammenspiel zwischen den Kompositionseffekten durch die psychosoziale Komposition der Klasse oder Schule und ähnlichen sozialen Einflüssen durch andere Peergruppen außerhalb der Schule untersucht werden. Auch abgegrenzte Freundeskreise innerhalb der Klasse könnten den Effekt, der von der gesamten Klassenpeergruppe ausgeht, beeinflussen, wie auch Müller und Zurbriggen (2016) vermuten. Einen Aspekt, der in den vorliegenden Untersuchungen dieser Dissertation nur mit Blick auf den BFLPE untersucht wurde, aber in zukünftigen Studien näher beleuchtet werden könnte, stellen komparative Einflüsse der Schülerschaft hinsichtlich psychosozialer Merkmale dar. Soziale Vergleichsprozesse könnten auch hinsichtlich der sozialen Stellung angenommen werden. Es wird vermutet, dass Lernende an leistungsschwachen Schulen ihr Grundbedürfnis nach Kompetenz und sozialer Anerkennung in andere Bereiche wie soziale Anerkennung durch aggressives Verhalten verlagern, wenn sie das Bedürfnis nicht über Leistungskompetenzen befriedigen können (Cillessen & Mayeux, 2004; Eccles & Roeser, 2011; Werner & Hill, 2010). Würde an leistungsschwachen Schulen dadurch insgesamt ein aggressives Klima vorliegen, so dürften sich Schülerinnen und Schüler besonders aggressive Mitschülerinnen und Mitschüler als Referenzmaß heranziehen.

6.4 Schlussfolgerung

Die vorliegende Dissertation hat einen Beitrag zur Erforschung der Schülerschaft als Mitgestalter des Lern- und Entwicklungsorts Schule geleistet. Sie konnte zeigen, dass Schülerinnen und Schüler in ihrer Zusammensetzung als Schülerschaft einer Lerngruppe nicht nur eine Rolle für das primäre schulische Ziel der Leistungsentwicklung spielen, sondern auch für die Entwicklung psychosozialer Merkmale. Zudem verweisen Effekte der psychosozialen Komposition auf die individuelle Entwicklung psychosozialer Merkmale auf soziale Interaktionsprozesse zwischen Lernenden, auf die pädagogisch reagiert werden kann. Diesbezügliche pädagogische Maßnahmen können eine Ergänzung von strukturellen Maßnahmen darstellen, über die intendiert und nicht intendiert die leistungs- und herkunftsbezogene Schülerkomposition gesteuert wird. Dabei sollten sie Hand in Hand mit solchen strukturellen

Maßnahmen verlaufen, um die gesellschaftlichen Ziele von Beschulung umsetzen zu können (Fend, 2009). Die Komposition von Lerngruppen hinsichtlich psychosozialer Merkmale kann im Gegensatz zur herkunftsbezogenen Komposition als justierbare Stellschraube für die pädagogische Praxis gesehen werden, um den Einflüssen einer ungünstigen leistungs- und herkunftsbezogenen Komposition der Schülerschaft einer Schule zu begegnen.

“[...] the holistic conceptualization of both educational contexts and outcomes in terms of multiple domains of development (e.g., the intellectual, the social, the moral, and the emotional) is an important valued-added perspective that developmental science can add to the study of schooling and its influences on human development.”

Eccles & Roeser, 2011, S. 580

Literaturverzeichnis II

- Altrichter, H. & Merki, K. M. (2010). *Handbuch Neue Steuerung Im Schulsystem*: Springer VS.
- Asher, S. R. & McDonald, K. L. (2009). The behavioral basis of acceptance, rejection, and perceived popularity. In K.H. Rubin, W. Bukowski & B. Laursen (Eds.), *Handbook of peer interactions, relationships, and groups* (pp. 232–248), New York, NY: Guilford.
- 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*(2), 115–133.
- Baumert, J., Stanat, P. & Watermann, R. (Hrsg.) (2006). *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Becker, M. & Neumann, M. (2016). Context-related changes in academic self-concept development. On the long-term persistence of big-fish-little-pond effects. *Learning and Instruction, 45*, 31–39.
- Blau, P. M. (1960). Structural effects. *American Sociological Review, 178–193*.
- Bonell, C., Parry, W., Wells, H., Jamal, F., Fletcher, A., Harden, A et al. (2013). The effects of the school environment on student health: A systematic review of multi-level studies. *Health Place, 21*, 180–191.
- Brown, B. B. (2013). Adolescents' relationships with peers. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology: Contextual influences on adolescent development* (Vol. 2, 2nd ed., pp. 363–394). Hoboken, NJ: Wiley & Sons.
- Brown, B. B., Bakken, J. P., Ameringer, S. W. & Mahon, S. D. (2008). A comprehensive conceptualization of the peer influence process in adolescence. In M. J. Prinstein & K. A. Dodge (Eds.), *Understanding peer influence in children and adolescents* (pp. 17–44). New York, NY: The Guilford Press.
- Cillessen, A. H. N. & Mayeux, L. (2004). From censure to reinforcement: Developmental changes in the association between aggression and social status. *Child Development, 75*(1), 147–163.
- Cohen, J., McCabe, E. M., Michelli, N. M. & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record, 111*(1), 180–213.

- 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, DC: US Government Printing Office.
- Dishion, T. J. & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, 62, 189–214.
- Dumont, H., Neumann, M., Maaz, K. & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60(3), 163–183.
- Duncan, G. J., Magnuson, K. A. & Ludwig, J. (2004). The endogeneity problem in developmental studies. *Research in Human Development*, 1(1/2), 59–80.
- Eccles, J. S. & Roeser, R. W. (2011). School and community influences on human development. In M. H. Bornstein & M. E. Lamb (Eds.), *Developmental science. An advanced textbook* (6th ed., pp. 571–643). New York, NY: Psychology Press.
- Fend, H. (2009). *Neue Theorie der Schule: Einführung in das Verstehen von Bildungssystemen* (2. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gommans, R., Stevens, G. W. J. M., ter Bogt, T. F. M. & Cillessen, A. H. N. (2016). Adolescent substance use, aggressive behaviors, and peer context behavioral norms. *Journal of Cognitive Education and Psychology*, 15(2), 185–205.
- Graham, J. W. (Ed.). (2012). *Missing data: Analysis and design*. New York, NY: Springer New York.
- Hartig, J., Klieme, E. & Leutner, D. (Eds.). (2008). *Assessment of competencies in educational contexts: State of the art and future prospects*. Göttingen: Hogrefe.
- Hattie, J. A. C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37(5), 449–481.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.
- Henry, D., Guerra, N., Huesmann, R., Tolan, P., VanAcker, R. & Eron, L. (2000). Normative influences on aggression in urban elementary school classrooms. *American Journal of Community Psychology*, 28(1), 59–81.
- Hoglund, W. L. & Leadbeater, B. J. (2004). The effects of family, school, and classroom ecologies on changes in children's social competence and emotional and behavioral problems in first Grade. *Developmental Psychology*, 40(4), 533–544.

- Hong, G. L., Corter, C., Hong, Y. H. & Pelletier, J. (2012). Differential effects of literacy instruction time and homogeneous ability grouping in kindergarten classrooms. Who will benefit? Who will suffer? *Educational Evaluation and Policy Analysis*, 34(1), 69–88.
- Imai, K., Keele, L., Tingley, D. & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, 105(4), 765–789.
- Jackson, D. R., Cappella, E. & Neal, J. W. (2015). Aggression norms in the classroom social network: contexts of aggressive behavior and social preference in middle childhood. *American Journal of Community Psychology*, 56(3–4), 293–306.
- Johnson, M. K., Crosnoe, R. & Elder Jr, G. H. (2001). Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education*, 74(4), 318–340.
- Kelley, H. H. (1968). Two functions of reference groups. In H.H. Hyman & E. Singer (Eds.), *Readings in reference group theory and research* (pp. 77–83). New York, NY: The Free Press.
- Klinge, D. (2016). *Die elterliche Übergangsentscheidung nach der Grundschule: Werte, Erwartungen und Orientierungen*. Wiesbaden: Springer VS.
- Kunter, M. (2005). *Multiple Ziele im Mathematikunterricht*. Münster: Waxmann.
- Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B. & d'Apollonia, S. (1996). Within-class grouping. A meta-analysis. *Review of Educational Research*, 66(4), 423–458.
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T. & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13(3), 203–229.
- Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift für Pädagogische Psychologie*, 19(3), 119–127.
- Marsh, H. W. & Hau, K. T. (2003). Big-fish-little-pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58(5), 364–376.
- Marsh, H. W., Köller, O. & Baumert, J. (2001). Reunification of East and West German school systems: Longitudinal multilevel modeling study of the big-fish-little-pond effect on academic self-concept. *American Educational Research Journal*, 38(2), 321–350.

- 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(1), 121–158.
- Müller, C. M., Hofmann, V., Fleischli, J. & Studer, F. (2015). „Sag‘ mir, was deine Klassenkameraden tun und ich sage dir, was du tun wirst“? Zum Einfluss der Klassenzusammensetzung auf die Entwicklung von schulischem Problemverhalten. *Zeitschrift für Erziehungswissenschaft*, 18(3), 569–589.
- Müller, C. M. & Minger, M. (2013). Welche Kinder und Jugendliche werden am stärksten durch die Peers beeinflusst? Eine systematische Übersicht für den Bereich dissozialen Verhaltens. *Empirische Sonderpädagogik*, 5(2), 107–129.
- Müller, C. M. & Zurbriggen, C. L. A. (2016). An overview of classroom composition research on social–emotional outcomes: Introduction to the special issue. *Journal of Cognitive Education and Psychology*, 15(2), 163–184.
- Muthén, B. O. & Muthén, L. K. (1998–2010). *Mplus user’s guide*. Los Angeles, CA: Author.
- Nagengast, B. & Marsh, H. W. (2011). The negative effect of school–average ability on science self–concept in the UK, the UK countries and the world: the big–fish–little–pond effect for PISA 2006. *Educational Psychology*, 31(5), 629–656.
- 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(3), 275–302.
- Pekrun, R. & Helmke, A. (1991). Schule und Persönlichkeitsentwicklung: Theoretische Perspektiven und Forschungsstand. In R. Pekrun & H. Fend (Hrsg.), *Schule und Persönlichkeitsentwicklung: Ein Resümee der Längsschnittforschung (Der Mensch als soziales und personales Wesen)* (Bd. 11, S. 33–56). Stuttgart: Enke.
- Prengel, A. (Hrsg.). (1995). *Pädagogik der Vielfalt. Verschiedenheit und Gleichberechtigung in Interkultureller, Feministischer und Integrativer Pädagogik* (2. Aufl.). Opladen: Leske und Budrich.
- Prinstein, M. J. (2007). Moderators of peer contagion: A longitudinal examination of depression socialization between adolescents and their best friends. *Journal of Clinical Child & Adolescent Psychology*, 36(2), 159–170.
- Roßbach, H.–G. & Wellenreuther, M. (2002). Empirische Forschungen zur Wirksamkeit von Methoden der Leistungsdifferenzierung in der Grundschule. In F. Heinzel & A. Prengel

- (Hrsg.), *Heterogenität, Integration und Differenzierung in der Primarstufe* (Jahrbuch Grundschulforschung, Bd. 6, S. 44–57). Opladen: Leske und Budrich.
- Scharenberg, K. (2016). The interplay of social and ethnic classroom composition, tracking, and gender on students' school satisfaction. *Journal of Cognitive Education and Psychology*, 15(2), 320–346.
- Schnabel, K. (1998). Schuleffekte. In D. H. Rost (Hrsg.), *Handwörterbuch für pädagogische Psychologie* (S. 431–435). Weinheim: Beltz.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK) (Hrsg.) (2015): *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring* (Beschluss der 350. Kultusministerkonferenz vom 11.06.2015).
- Sellström, E. & Bremberg, S. (2006). Is there a “school effect” on pupil outcomes? A review of multilevel studies. *Journal of Epidemiology and Community Health*, 60(2), 149–155.
- Thomas, D. E., Bierman, K. L. & Powers, C. J. (2011). The influence of classroom aggression and classroom climate on aggressive-disruptive behavior. *Child Development*, 82(3), 751–757.
- Thrupp, M., Lauder, H. & Robinson, T. (2002). School composition and peer effects. *International Journal of Educational Research*, 37(5), 483–504.
- Tillmann, K.-J. (2007). Heterogenität– Homogenität. Was wir darüber wissen. *Schul-Management*, 38(5), 8–10.
- van Buuren, S. & Oudshoorn, K. (1999). *Flexible multivariate imputation by MICE*. Leiden: TNO Prevention and Health.
- van Ewijk, R. & Slegers, P. (2010a). The effect of peer socioeconomic status on student achievement. A meta-analysis. *Educational Research Review*, 5(2), 134–150.
- van Ewijk, R. & Slegers, P. (2010b). Peer ethnicity and achievement: A meta-analysis into the compositional effect. *School Effectiveness and School Improvement*, 21(3), 237–265.
- Warren, K., Schoppelrey, S., Moberg, D. P. & McDonald, M. (2005). A model of contagion through competition in the aggressive behaviors of elementary school students. *Journal of Abnormal Child Psychology*, 33(3), 283–292.
- Werner, N. E. & Hill, L. G. (2010). Individual and peer group normative beliefs about relational aggression. *Child Development*, 81(3), 826–836.
- Wigfield, A. & Eccles, J. S. (1990). Test Anxiety in the School Setting. In M. Lewis & S. M. Miller (Eds.), *Handbook of Developmental Psychopathology* (pp. 237–250). Boston, MA: Springer US.

Wilkinson, I. A., Hattie, J., Parr, J. M., Townsend, Michael A. R., Fung, I. Y., Ussher, C. et al. (2000). *Influence of peer effects on learning outcomes. A review of the literature*. Auckland: Auckland UniServices Limited.

Wilson, A. B. (1959). Residential segregation of social classes and aspirations of high school boys. *American Sociological Review*, 24, 836–845.

Zeidner, M. (2014). Anxiety in education. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 265–288). New York, NY: Routledge.

Abbildungsverzeichnis

Abbildung 1.1	Modell des Schulkontextes nach Eccles und Roeser (2011).....	23
Abbildung 1.2	Klassische pfadanalytische Darstellung des leistungsbezogenen Kompositionseffekts.....	39
Abbildung 3.1	Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for 1 SD below and 1 SD above the mean of test anxiety at the end of Grade 6 (in regard to somatic signs).....	124
Abbildung 3.2	Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for 1 SD below and 1 SD above the mean of test anxiety at the end of Grade 6 (in regard to worry).....	125
Abbildung 3.3	Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for girls and boys (in regard to somatic signs).	126
Abbildung 3.4	Simple slopes of class-average test anxiety at the beginning of Grade 7 predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7 for girls and boys (in regard to worry).	126
Abbildung 4.1	Theoretical path model of the compositional effect of class-average achievement on individual achievement.	153
Abbildung 4.2	Theoretical path model of the big-fish-little-pond effect.	155
Abbildung 4.3	Theoretical path model of the reciprocal effects model.	156
Abbildung 4.4	Path model of the present study.	159
Abbildung 4.5	Path model of the relationship between the compositional effect on achievement, the BFLPE, and the REM (Model 13).....	172
Abbildung 5.1	Schema zur Einteilung der Schulklassen in homogene, durchschnittliche und heterogene Klassen.....	199
Abbildung 5.2	Mittelwerte der klassenspezifischen Standardabweichungen und der klassenspezifischen Mittelwerte von z-standardisierten Kompositionsmerkmalen pro Schulform.....	201

Abbildung 5.3	Die Verteilung homogener, durchschnittlicher und heterogener Klassen hinsichtlich der Deutschkompetenz, insgesamt und pro Schulform	204
Abbildung 5.4	Die Verteilung homogener, durchschnittlicher und heterogener Klassen hinsichtlich des HISEI, insgesamt und pro Schulform	206
Abbildung 5.5	Mittlere Anzahl der verschiedenen Nationalitäten pro Klasse (inkl. deutsch), getrennt nach Schulform	207
Abbildung 5.6	Die Verteilung homogener, durchschnittlicher und heterogener Klassen hinsichtlich der Lesefreude, insgesamt und pro Schulform	209

Tabellenverzeichnis

Tabelle 2.1 Korrelationsmatrix der betrachteten Variablen auf Schulebene.....	88
Tabelle 2.2 Mehrebenenmodelle verschiedener Kompositionseffekte auf die individuelle Schul- zufriedenheit in der 10. Jahrgangsstufe	91
Tabelle 2.3 Mehrebenenmodelle verschiedener Kompositionseffekte auf das individuelle normverletzende Verhalten in der 10. Jahrgangsstufe	93
Tabelle 2.4 Mehrebenenmodelle verschiedener Kompositionseffekte auf die individuelle Mathematiknote in der 10. Jahrgangsstufe.....	95
Tabelle 3.1 Descriptive statistics for somatic signs and worry for all measurement points	119
Tabelle 3.2 Multilevel regression models predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7	120
Tabelle 3.3 Multilevel regression models predicting development of test anxiety from the beginning of Grade 7 to the end of Grade 7	121
Tabelle 3.4 Regression models predicting development of test anxiety from the end of Grade 6 to the beginning of Grade 7, including interaction effects	123
Tabelle 4.1 The effect of class-average mathematics achievement on students' individual mathematics achievement over the course of Grade 7	165
Tabelle 4.2 The effect of class-average mathematics achievement on students' mathematics self- concept over the course of Grade 7	167
Tabelle 4.3 Cross-lagged associations between students' achievement and academic self-concept over the course of Grade 7	169
Tabelle 4.4 The effect of class-average achievement on students' individual achievement and academic self-concept over the course of Grade 7	171

LEBENS LAUF

Der Lebenslauf ist aus Gründen des Datenschutzes nicht enthalten.

ERKLÄRUNG

Hiermit erkläre ich an Eides statt, dass ich die Dissertation »Die Zusammensetzung der Lerngruppe und ihre Effekte auf psychosoziale Merkmale und Leistung von Schülerinnen und Schülern« selbstständig und ohne unzulässige Hilfe Dritter verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese Arbeit hat keiner anderen Prüfungsbehörde vorgelegen.

Franziska Stäbler

Berlin, den 29. Mai 2017