

## **Schule und Schulklasse als soziale Kontexte der Entwicklung im Jugendalter**

**Katja Scharenberg**

© Springer Fachmedien Wiesbaden 2014

**Zusammenfassung** Untersuchungen zur Bedeutung von sozialen Kontexten haben in der empirischen Sozial- und Bildungsforschung eine lange Tradition. Ein Anwendungsbereich der Kontextanalyse liegt im Mehrebenensystem Schule. Der Beitrag beschäftigt sich mit der Frage nach den relevanten Kontexten (Schulklasse und Schule) für die schulische Leistungsentwicklung als zentraler Aspekt der Entwicklung im Jugendalter. Dabei wird auch untersucht, ob die normative Kultur der Altersgruppe als Indikator für Vermittlungsprozesse zwischen Kontextmerkmalen und Schulleistung entwicklungsrelevant ist. Datengrundlage ist die Hamburger Schulleistungsstudie „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS), mit der die Leistungsentwicklung im Leseverständnis von 9681 Schülerinnen und Schülern in 467 Schulklassen und 142 Schulen in den Jahrgangsstufen 7 und 8 nachvollzogen werden kann. In schrittweisen Mehrebenenmodellen erweisen sich neben individuellen Schülermerkmalen die leistungsbezogene Klassenkomposition und die Schulformzugehörigkeit als bedeutsam. Die normative Kultur der Altersgruppe stellt hingegen nur im unmittelbaren Klassenkontext einen bedeutsamen Indikator für Vermittlungsprozesse von Kontexteffekten dar, wobei diese Prozesse geschlechtsspezifisch ausfallen.

**Schlüsselwörter** Schule · Schulklasse · Leistungsentwicklung · Jugendalter · Kontexteffekt · Vermittlungsprozesse

---

K. Scharenberg (✉)  
Institut für Soziologie, Universität Basel,  
Petersgraben 27,  
4051 Basel, Schweiz  
E-Mail: katja.scharenberg@unibas.ch

## Schools and Classrooms as Social Contexts for the Development in Adolescence

**Abstract** Studies dealing with the importance of social contexts already have a long tradition in empirical sociological and educational research. One possibility of applying context analysis is to consider schools as multi-level systems. The contribution focuses on the question of the relevant context factors on classroom and school level for the development of school achievement as a key aspect of development in adolescence. It is also examined whether social responsibility goals and norms are indicators of mediating processes between context variables on the one hand and school achievement on the other hand. Data base is the study “Competences and attitudes of students” (KESS) that examines the development of Reading achievement from grade 7 to 8 of a whole student cohort (9,682 students in 467 classes and 142 schools in Hamburg, Germany). Stepwise multi-level regression analyses reveal individual student characteristics, the average ability composition on the class level and the attended track on the school level as relevant predictors of the development of achievement. Social responsibility goals and norms are relevant indicators of mediating processes within the proximal context of the classroom, but not within the context of the schools. These processes are, however, gender-specific.

**Keywords** School · Classroom · Achievement development · Adolescence · Contextual effect · Mediation processes

### 1 Einleitung

Kennzeichnend für das deutsche Schulsystem in der Sekundarstufe ist die institutionelle Differenzierung in verschiedene Schulformen und Bildungsgänge. Diese Differenzierung soll nach Leistungskriterien erfolgen und beruht auf der Annahme, dass der Unterricht in leistungshomogeneren Lerngruppen einfacher an individuelle Lernvoraussetzungen anzupassen sei, damit die Schülerinnen<sup>1</sup> und Schüler gezielt und ihrem Leistungsvermögen angemessen gefördert werden können (z. B. Barr und Dreeben 1983; Hallinan 1994; Harker und Tymms 2004; Hattie 2002; Slavin 1990b). Anhand verschiedener empirischer Studien wurde jedoch gezeigt, dass diese Differenzierung nicht ausschließlich leistungsbasiert erfolgt, sondern erheblich mit sozialen Herkunftsmerkmalen zusammenhängt (z. B. Baumert und Schümer 2001; Baumert et al. 2006). Die Schulformgliederung führt somit nicht nur zur gewünschten leistungsbezogenen Stratifizierung, sondern auch zu einer damit einhergehenden sozialen Segregation (Baumert et al. 2003, 2006). Gerade in Deutschland ist der Zusammenhang zwischen Sozialschichtzugehörigkeit und Lesekompetenz besonders eng – auf schulischer noch stärker als auf individueller Ebene (Baumert et al. 2009; Baumert und Schümer 2001). Dies führt zu einer Konzentration von Hauptschulen in den leistungsschwächeren, sozial benachteiligten sowie von Gymnasien in den

<sup>1</sup> Im Weiteren wird aus Vereinfachungsgründen von Schülern gesprochen, wenn Schülerinnen *und* Schüler gemeint sind.

leistungsstärkeren, sozial privilegierteren Bereichen der Verteilung bei gleichzeitig starker Überlappung der schulformspezifischen Leistungsverteilungen im mittleren Bereich (Baumert et al. 2006). Da eine Reihe weiterer leistungsrelevanter Merkmale mit der sozialen Herkunft eng zusammenhängt, können die spezifischen Kompositionsprofile der Schulformen zu einer kumulativen Benachteiligung oder Privilegierung von Schulen aufgrund einer für die jeweilige Schülerschaft unterschiedlichen Verfügbarkeit an ökonomischem, kulturellem oder sozialem Kapital führen. Nach Dreeben und Barr (1988) ist dabei nicht allein die soziale Zusammensetzung der Schülerschaft ausschlaggebend. Vielmehr scheinen sich Lehrkräfte v. a. an das damit variierende Leistungsniveau der Schülerschaft didaktisch, curricular und organisatorisch anzupassen. Auch schulformspezifische Traditionen der Didaktik und Lehrerbildung fördern die Herausbildung selektionsbedingter „differenzielle[r] Lern- und Entwicklungsmilieus“ (Baumert und Schümer 2001, S. 462; Baumert et al. 2006, S. 98). Diese können dazu führen, dass die Entwicklung von Schülern unabhängig von und zusätzlich zu bereits existierenden individuellen Unterschieden je nach besuchter Schulform ungleich verläuft, erschwert oder sogar verhindert wird. Eine „faire“, ausschließlich leistungsbezogene externe Differenzierung der Schülerschaft gelingt somit nicht (vgl. z. B. auch Artelt et al. 2001; Bos und Scharenberg 2010; Scharenberg 2012, 2013). Dies wirft Fragen auf

- nach der Gerechtigkeit und Chancengleichheit des deutschen Bildungssystems – und zwar auf individueller wie aggregierter Ebene (siehe hierzu Berkemeyer et al. 2012; Berkemeyer und Manitius 2013);
- danach, inwiefern die Entwicklung im Jugendalter – oder, in der Terminologie der Kontextanalyse ausgedrückt, das Explanandum auf der Mikroebene des Individuums – durch verschiedene institutionelle Kontexte auf Makro- und ggf. Mesoebene beeinflusst wird; und
- welcher der verschiedenen Kontexte im schulischen Bereich konkret welchen Anteil zum Bildungserfolg der Schüler beisteuert (Ditton 2013).

## 2 Schule und Schulklasse als soziale Kontexte der Entwicklung – ein Vermittlungsmodell

Gründe für Schereneffekte bei der Leistungsentwicklung zwischen verschiedenen Schulformen, die bereits mit Beginn der Sekundarstufe einsetzen und deren Auswirkungen sich über einen längeren Zeitraum kumulieren können (Bonsen et al. 2009a; Gröhlich et al. 2010a), liegen nach Baumert et al. (2006) darin, dass

1. Schüler in Abhängigkeit ihrer individuellen Lernausgangslagen unterschiedlich große Lernfortschritte machen (*individueller Matthäus-Effekt*);
2. *institutionelle Unterschiede* hinsichtlich der Unterrichtskulturen, Lehrpläne und Stundentafeln zwischen den Schulformen bestehen; und
3. sich die Schulformen auch hinsichtlich ihrer sozialen, kulturellen und leistungsbezogenen Zusammensetzung ihrer Schülerschaft, also in ihrer *Komposition*, unterscheiden (*institutioneller Matthäus-Effekt*).

Diese Auswirkungen der Schülerzusammensetzung auf die erzielten Lernstände und die Lern- und Persönlichkeitsentwicklung von Schülern werden auch unter dem Stichwort *Kompositionseffekte* behandelt. Beispiele für solche Effekte aggregierter Individualmerkmale (*Kompositionsmerkmale*) sind z. B. die durchschnittliche soziale Zusammensetzung und der prozentuale Anteil von Schülern mit Migrationshintergrund innerhalb von Schulen, ferner das durchschnittliche Leistungsniveau und die Leistungsheterogenität innerhalb von Schulklassen. Als Kompositionsmerkmale haben sie eine neue, eigenständige Bedeutung für die Vorhersage einer Kriteriumsvariablen, die sich nicht nur als Wirkung individueller Merkmale herausstellt:

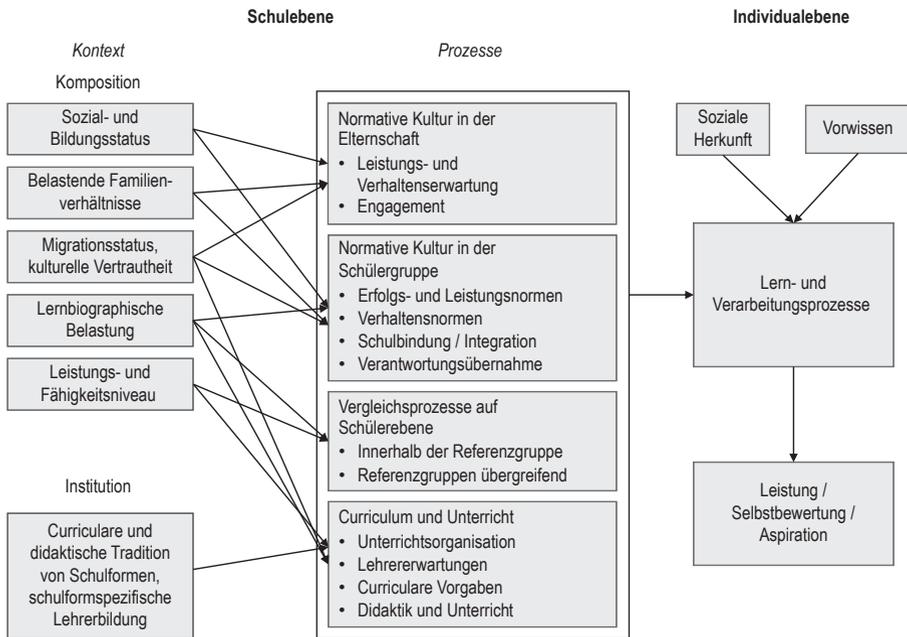
„In many school effects studies, achievement differences are found to be related to differences in the composition of the body of students. This is known as the ‚compositional‘ effect. Such an effect is often reported when a school-level aggregate of an individual-level variable makes an independent contribution to the explanation of outcome variance. In other words, the school-level aggregated variable makes a significant contribution to the explanatory model over and above the contribution of the same variable at the individual level“ (Harker und Tymms 2004, S. 177).

Bezogen auf den Bereich Schule kann dies bedeuten, dass Schüler auch bei gleichen Lernausgangslagen und individuellen Voraussetzungen je nach Komposition des Lernkontextes (z. B. in leistungsstarken Schulklassen oder sozial eher günstiger zusammengesetzten Schulen) unterschiedliche Lernerfolge erzielen: „[T]wo pupils with similar prior scores but attending different schools can be predicted to have different scores depending on the average score of the other pupils in their schools“ (ebd., S. 178). Erst wenn Institutions- oder Kompositionseffekte vorliegen, kann davon ausgegangen werden, dass verschiedene Schulkontexte differenzielle Lern- und Entwicklungsmilieus darstellen (Baumert et al. 2009, S. 37).

Von Effekten solcher Kompositionsmerkmale, die sich auf die Zusammensetzung der Schülerschaft beziehen, grenzen z. B. Harker und Tymms (2004) Effekte allgemeiner *Kontextmerkmale* ab, die nicht über die Aggregation individueller Merkmale operationalisiert werden.<sup>2</sup> Solche Kontextmerkmale sind Kennzeichen der weiteren Umgebung und z. B. von den Schulen selbst nicht direkt oder nur kaum zu beeinflussen, jedoch für die Entwicklung schulischer Leistungen relevant (Scheerens 1990). Zu den Kontextmerkmalen zählen demnach z. B. das schulische Einzugsgebiet und die Trägerschaft, regionale Unterschiede zwischen städtischen und ländlichen Gebieten, bundeslandspezifische Regelungen und Besonderheiten oder Unterschiede zwischen Schulsystemen verschiedener Staaten: „[T]he term, contextual effect‘ is now used in a way that includes other differences between schools [...]. [T]hese additional variables could be seen as other level-2 [...] variables to be examined for their mediating effects of a basic compositional effect“ (Harker und Tymms 2004, S. 183).

Zur systematischen Analyse der Auswirkungen von Schule und Schulklasse als soziale Kontexte der Entwicklung im Jugendalter ist die Skizzierung eines theore-

<sup>2</sup>Andere Autoren verwenden den Begriff *Kontexteffekt* hingegen als Oberbegriff und bezeichnen damit den Effekt jeglicher Merkmale des Kontextes, z. B. auch die Auswirkungen von Kompositionsmerkmalen (z. B. Lüdtke et al. 2002).



**Abb. 1** Vermittlungsmodell für Kontexteffekte. (Quelle: Baumert et al. (2006, S. 126))

tisch fundierten Vermittlungsmodells (Abb. 1) erforderlich, das von Baumert et al. (2006) vorgeschlagen wurde: In ihrem Vermittlungsmodell zählen kompositionelle und institutionelle Merkmale von Schulen zu den Kontextbedingungen für schulische Lernprozesse. Kontextmerkmale wirken jedoch nicht unmittelbar auf Leistungs- und Entwicklungsprozesse ein. Vielmehr sind Makro- und Mikroebene durch verschiedene Vermittlungsmechanismen als Brückenhypothesen miteinander verbunden. Dreeben und Barr (1988) unterscheiden vier Transmissionswege, wonach Kontextmerkmale erstens durch die Unterstützungsbereitschaft der Elternschaft, zweitens durch die normative Kultur der Altersgruppe (z. B. über Leistungs- und Verhaltensnormen), drittens durch soziale Vergleichsprozesse (siehe hierzu Köller 2004; Köller und Baumert 2001) auf Schülerebene (referenzgruppenintern und -übergreifend) und viertens durch didaktische, curriculare und organisatorische Rahmenbedingungen vermittelt werden (z. B. wenn Lehrkräfte und Schulleitungen auf eine unterschiedliche Zusammensetzung der Schülerschaft mit verschiedenen Formen der Unterrichtsorganisation und Differenzierung, jeweils spezifischen Erwartungen und curricularen Vorgaben oder didaktisch adaptiv reagieren). Letztere werden von Dreeben und Barr (1988) in ihrer Bedeutung besonders hervorgehoben.

Soziologische Ansätze richten sich bei der Erklärung von Kontexteffekten hingegen eher auf die kollektiv geteilten und vorherrschenden Normen (vgl. bereits Blau 1960). Bezogen auf den schulischen Kontext wird dabei die normative Kultur der Altersgruppe als primärer Vermittlungsprozess und Bindeglied zwischen Kontextfaktoren einerseits und individuellem Verhalten andererseits angesehen. Demnach führen bestimmte Kompositionsmerkmale zu gemeinsam geteilten Erfolgs- und Leis-

tungsnormen auf kollektiver Ebene, die über subjektive Deutungsprozesse aufgrund ihrer hohen Verbindlichkeit handlungsleitend werden (Ditton 2013). Die individuelle Bereitschaft zur Einhaltung von Leistungsnormen auch bei Normabweichungen der Klassenkameraden stellt somit ein prosoziales Ziel im schulischen Kontext dar.<sup>3</sup> Weitere prosoziale Ziele sind die Unterstützung anderer bei Problemen oder unterrichtsbezogenen Aktivitäten, die Verlässlichkeit gegenüber Gleichaltrigen, wie z. B. Versprechen zu halten und Geheimnisse für sich zu bewahren (Wentzel 1991, 1994, 1999). Solche selbstgesetzten persönlichen Ziele beeinflussen wiederum, wie viel Anstrengungsbereitschaft und welche Verhaltensweisen im Unterricht gezeigt werden (Wentzel 1999). Ihnen wird daher als Vermittlungsprozess eine leistungsförderliche Wirkung zugeschrieben.

Baumert et al. (2006) modellieren in ihren Analysen die komplexen Vermittlungsprozesse auf der Ebene von *Schulen*. Dagegen bildet die *Schulklasse* einen sozialen „Kontext, in den man stark und längerfristig eingebunden ist, [der] eher Wirkungen zeigt als ein Kontext, zu dem eine nur lose kurzzeitige Beziehung besteht“ (Ditton 2013, S. 174). Auch aufgrund der Bedeutung, die die leistungsbezogene Schülerkomposition für soziale Vergleiche, normative Wertorientierungen sowie das Curriculum und den Unterricht hat (Dreeben und Barr 1988), liegt es nahe, diese Vermittlungsprozesse auch im unmittelbaren, nahe liegenden (*proximalen*) Lernkontext von Schulklassen zu vermuten, der direkt mit der Lehr-Lernsituation verbunden ist (Ditton 2000). Die Schulklassen stellen wiederum eine „normativ geregelte Verbindung bestimmter Formen des sozialen Handelns bestimmter Akteure“ (Meulemann 2006, S. 224) dar mit einer sehr engen sozialräumlichen Umgebung (Friedrichs und Nonnenmacher 2010, S. 471) und Beziehung zwischen Individuum und Kollektiv (Ditton 2013, S. 174). In der Schul- und Unterrichtsforschung wird die Schulklasse deshalb als wirksame Entwicklungsumwelt angesehen und als eigenständige Analyseeinheit berücksichtigt (z. B. Schwetz 2003). Nach Fend (1991) stellen Schulklassen diejenigen sozialen Interaktionsfelder dar, in denen sich jeweils spezifische, gemeinsam geteilte Normen zur Bewertung von Schule, schulischer Leistungen und Anstrengungsbereitschaft herausbilden. So führen z. B. eine hohe Leistungsorientierung in der Schulklasse und die gemeinsame Verständigung auf leistungsbezogene Kriterien für Beliebtheit und gegenseitige Anerkennung unter den Gleichaltrigen dazu, dass sie ihre Selbstbewertung an der klassenspezifischen Leistungshierarchie festmachen, was sich langfristig über eine hohe subjektive Bedeutung von Motivation und Anstrengungsbereitschaft auf die Leistungsentwicklung niederschlagen sollte. Das soziale Lernumfeld hat demnach eine normative Funktion, wobei die Schulklasse die normgebende Bezugsgruppe bildet, in der sich eine jeweils spezifische Gruppen- und Lernkultur entwickelt und Verhaltensmodelle und -rollen bereitgestellt werden (z. B. Dar und Resh 1986).

Ausgehend von diesen Überlegungen soll nachfolgend mit Blick auf die Schule als einem Anwendungsbereich der Kontextanalyse das Ziel verfolgt werden, einen weiteren Beitrag, der sich hier nur auf einzelne Aspekte des theoretischen Vermitt-

<sup>3</sup>Die Akzeptanz von Leistungsnormen wird üblicherweise als Individualmerkmal erhoben, kann aber durch Aggregation auf höherer Ebene (z. B. Schulklasse oder Schule) als analytisches Kontextmerkmal behandelt werden, das unmittelbar mit dem Lernprozess verknüpft ist (Baumert et al. 2001).

lungsmodells von Baumert et al. (2006) beschränken kann, zur Klärung der Frage zu leisten, was der entwicklungsrelevante Kontext im Mehrebenensystem Schule ist und über welche Vermittlungsmechanismen sich Kontexteffekte auf die Schulleistung als zentraler Aspekt der Entwicklung im Jugendalter auswirken.

### 3 Empirische Forschungsbefunde

International ist der Forschungsstand zu den Auswirkungen unterschiedlicher Schulkontexte und der Zusammensetzung der Schülerschaft breit gefächert, insbesondere mit Fokus auf die soziale Komposition von Schulen (z. B. Harker und Tymms 2004; Hattie 2002; Oakes 1990; Opdenakker und van Damme 2001; Thrupp et al. 2002). Nachfolgend wird auf einige ausgewählte Forschungsbefunde näher eingegangen (für eine vertiefende Forschungsübersicht siehe z. B. Dumont et al. 2013; Maaz et al. 2008).

Für das deutsche Schulsystem wurden Kompositions- und v. a. Institutionseffekte nachgewiesen: Anhand der nationalen PISA-Ergänzungsstudie (PISA-E 2000) zeigte sich, dass 15-Jährige an Gymnasien unter Kontrolle von Merkmalen des individuellen und schulischen Lernkontextes signifikant höhere Lesekompetenzwerte gegen Ende der Pflichtschulzeit erzielen als Gleichaltrige an Hauptschulen (Baumert et al. 2006). Darüber hinaus wurde eine starke Konfundierung von Kompositions- und Institutionsmerkmalen deutlich: Das Intelligenzniveau der Schülerschaft scheint das bedeutendste Kompositionsmerkmal zu sein. Die soziale Zusammensetzung übt dagegen nur einen kleinen spezifischen, aber signifikanten Effekt auf die Lesekompetenz aus. Auch für den Kompetenzbereich Mathematik liegt empirische Evidenz zu differenziellen Entwicklungsverläufen an verschiedenen Schulformen vor: Im Rahmen einer Reanalyse der nationalen längsschnittlichen Erweiterung der dritten internationalen Mathematik- und Naturwissenschaftsstudie (TIMSS) zeigt sich anhand von Mehrgruppen-Strukturgleichungsmodellen, dass die um Messfehler bereinigten, latenten Lernzuwächse in den Jahrgangsstufen 7 und 8 an Gymnasien und Realschulen mehr als doppelt so hoch ausfallen wie an Hauptschulen (Becker et al. 2006). Die Befunde sind dabei mit Effektstärken in mittlerer Höhe von praktischer Relevanz.

Für den Stadtstaat Hamburg liegen hingegen andere Befunde vor: Im Gegensatz zu anderen Befunden für die Lesekompetenz (Baumert et al. 2006) stellen Gröhlisch et al. (2010a) hinsichtlich der Entwicklung der Mathematikleistung in den Jahrgangsstufen 7 und 8 unter Kontrolle von Kompositions- und Institutionsmerkmalen keine höhere Lerneffektivität an Gymnasien fest. Allerdings fällt an Realschulen die Lernentwicklung im Vergleich zu Hauptschulen signifikant höher aus. Darüber hinaus scheinen sich institutionelle Effekte der Schulformgliederung nicht erst gegen Ende der Sekundarstufe, sondern bereits nach zwei Sekundarschuljahren zu manifestieren: Insbesondere an Gymnasien im Vergleich zu anderen Schulformen werden signifikant höhere Lernstände in Mathematik erzielt. Ihr Leistungsvorsprung kann nicht allein auf eine günstigere leistungsbezogene und soziale Komposition der Schülerschaft zurückgeführt werden (Bonsen et al. 2009a). Auch bei direkter Modellierung des Leistungszuwachses in Mathematik als abhängiger Variable treten im Drei-Ebenen-Modell schulformspezifische Schereneffekte zugunsten der Hamburger Gym-

nasien und Realschulen gegenüber Hauptschulen auf, wobei auch das schulische Leistungsniveau für unterschiedliche Wissenszuwächse ausschlaggebend ist (Guill und Gröhlich 2013).

Die Annahme von Schereneffekten wird auch anhand einer nordrhein-westfälischen Stichprobe gestützt: Demnach führen schulformspezifische Zuwachsraten zu einer ungünstigeren Entwicklung der Mathematikleistung an Gesamt- und Hauptschulen bereits in den ersten beiden Sekundarschuljahren. Dies kann u. a. daran liegen, dass mit der Schulformzugehörigkeit soziodemografische Kompositions- und auch Unterrichtsmerkmale (z. B. Überforderung, Disziplin und Schülerorientierung) variieren (van Ophuysen und Wendt 2009). Es gibt allerdings auch Studien, die für die Kompetenzentwicklung in Lesen und Mathematik keine Schereneffekte zwischen Hauptschulen und Gymnasien nachweisen können (Schneider und Stefanek 2004).

Auch für die Schweiz gibt es Studien zu institutionellen Effekten der Leistungs-differenzierung: So untersuchen z. B. Ramseier und Brühwiler (2003) in vertiefenden Analysen im Rahmen von PISA 2000 anhand von Pfadmodellen die Lese- und Mathematikleistungen von Schülern in den Kantonen Bern und St. Gallen. Unter Kontrolle der non-verbalen kognitiven Grundfähigkeit zeigt sich ein vergleichsweise hoher signifikanter Leistungsvorteil, wenn ein Schultyp mit erweiterten oder hohen Anforderungen besucht wird verglichen mit Schultypen mit Grundanforderungen. Der Schultyp erweist sich seinerseits als vermittelnde Variable zwischen personalen, kulturellen und sozialen Herkunftsmerkmalen einerseits sowie den Lese- und Mathematikleistungen andererseits, ist allerdings größtenteils mit diesen konfundiert. Neumann et al. (2007) zeigen anhand von Mehrebenenanalysen für Schüler der 8. Jahrgangsstufe, dass auch die Lernzuwächse im Fach Französisch als Fremdsprache unter Berücksichtigung zentraler Merkmale der Eingangsselektivität je nach Bildungsgang unterschiedlich ausfallen: Demnach werden an Schultypen mit erweiterten Anforderungen zwischen einem Drittel und einer halben Standardabweichung signifikant höhere Lernzuwächse erzielt als an Schultypen mit Grundanforderungen. Gleichzeitig erweisen sich Effekte von Bildungsgang und Leistungsniveau als konfundiert.

*Zusammenfassend* liegt also eine Reihe empirischer Forschungsbefunde zur Bedeutung unterschiedlicher Lernkontexte für die schulische Leistungsentwicklung in der Sekundarstufe vor, die jedoch über verschiedene Kompetenzbereiche hinweg nicht eindeutig sind (Baumert et al. 2009). In der Zusammenschau zeigen die Befunde jedoch, dass der mit der Schulformgliederung einhergehende Anspruch einer optimalen Förderung *aller* Schüler nicht erreicht werden kann. Die Frage, welche Rolle die normative Kultur der Altersgruppe als ein Aspekt von Vermittlungsmechanismen für Kontextmerkmale einnimmt und in welchem schulischen Kontext oder auf welcher Ebene diese wirksam werden, wurde zwar aufgeworfen, ist empirisch aber bislang weitgehend unbeantwortet geblieben. Es fehlt daher an Studien, die der Bedeutung von Vermittlungsmechanismen Rechnung tragen und die in der Lage sind, diese angemessen auf allen relevanten Kontextebenen zu berücksichtigen. In ihren Analysen zeigen Baumert et al. (2006) jedoch, dass normbezogene Vermittlungsprozesse auf schulischer Ebene nur in wenigen Schulformen leistungsförderlich sind. Die Autoren kommen zu dem Schluss, dass diese Prozesse offenbar nur auf individueller Ebene leistungsprädiktiv und schulformübergreifend auch weitgehend unabhängig

von schulischen Kontextmerkmalen sind. Das postulierte Vermittlungsmodell sei demzufolge nicht haltbar, wenn Zusammenhänge nur auf aggregierter Ebene berücksichtigt werden. Somit ist „die Frage der Vermittlung der nachgewiesenen Kompositionseffekte weiterhin offen“ (ebd., S. 149) und bislang „keineswegs eindeutig oder abschließend beantwortet“ (Ditton 2013, S. 173).

#### 4 Herleitung der Fragestellung und Hypothesen

Ausgehend von den theoretischen Überlegungen und den hierzu vorliegenden empirischen Forschungsbefunden lassen sich in dem für die Kontextanalyse typischen schrittweisen Vorgehen (ebd., S. 177) folgende Forschungsfragen mit Fokus auf die Kompetenzentwicklung als im Jugendalter zentraler Entwicklungsaspekt im schulischen Kontext konkretisieren:

1. In welcher Weise wirken sich Schule und Schulklasse als soziale Kontexte auf die Entwicklung im Jugendalter aus?

Zunächst wird gefragt, ob und inwieweit Schule und Schulklasse relevante Entwicklungskontexte darstellen, d. h., ob und in welcher Größenordnung signifikante Varianzanteile schulischer Kompetenzen auf Unterschiede zwischen Schulklassen und zwischen Schulen entfallen. In einem *ersten Schritt* wird also simultan die Varianzzerlegung der Kriteriumsvariable auf den Schul- und Klassenkontext bestimmt und gezeigt, welchem Kontext die größere Bedeutung zur Erklärung individueller Entwicklungsunterschiede zukommt. In einem *zweiten Schritt* wird regressionsanalytisch überprüft, welchen Beitrag bereits herkunfts- und personenbezogene Schülermerkmale zur Vorhersage unterschiedlicher Lernentwicklungen liefern. In einem *dritten Schritt* wird danach gefragt, ob und in welcher Richtung (positiv oder negativ) weitere schul- und klassenbezogene Merkmale über individuelle Schülermerkmale hinaus eine Rolle für die Leistungsentwicklung spielen: Lassen sich im Mehrebenenmodell Kompositionseffekte der leistungsbezogenen und sozialen Zusammensetzung von Schulen und Schulklassen sowie Institutionseffekte der Schulformgliederung in der Sekundarstufe nachweisen? Und ggf. auf welcher Ebene entfalten einzelne Kompositionsmerkmale ihre Wirkung?

Es wird erwartet, dass in der Varianzzerlegung signifikante Varianzanteile der abhängigen Variable auf den Schul- und Klassenkontext entfallen. Schule und Schulklasse sollten also entwicklungsrelevante Lernkontexte im schulischen Bereich darstellen, sodass es gerechtfertigt ist, beide Ebenen in die Analysen einzubeziehen. Regressionsanalytisch sollten analog zu Befunden anderer Studien (z. B. Baumert et al. 2006; Bonsel et al. 2009a; Gröhlich et al. 2010a, b; Guill und Gröhlich 2013; Neumann et al. 2007; Nikolova 2011) individuelle Schülermerkmale (Geschlecht, Migrationshintergrund, soziale Lage, allgemeine kognitive Grundfähigkeit und Ausgangsleistung zu einem früheren Erhebungszeitpunkt) die individuelle Lernentwicklung bereits gut vorhersagen können. Darüber hinaus wird auf Grundlage dieser Forschungsbefunde und als *Kontexthypothese* erwartet, dass eine günstige leistungsbezogene und soziale Zusammensetzung sowohl des Klassen- als auch Schulkontextes entwicklungsförderlich ist. Es wird vermutet, dass sich klassen-

bezogene Kompositionsmerkmale stärker durchsetzen, da der Klassenkontext die unmittelbare Lern- und Entwicklungsumwelt darstellt (z. B. Bellin 2009; Scharenberg 2012), in der das Lehren und Lernen stattfindet, und der Schulkontext weiter vom Unterrichtsgeschehen entfernt ist. Darüber hinaus wird vermutet, dass sich im schulischen Kontext Institutionseffekte der Schulformgliederung nachweisen lassen, die analog zu anderen Forschungsbefunden (z. B. Baumert et al. 2006; Bonsen et al. 2009a) v. a. in einer deutlichen, positiven Abgrenzung der Gymnasien zu anderen Schulformen bestehen sollten.

## 2. Durch welche Mechanismen werden Merkmale des Schul- und Klassenkontextes vermittelt?

In einem *vierten Schritt* wird das bereits auf der Schul- und Klassenebene wie oben beschrieben spezifizierte Mehrebenenmodell sozusagen in einer sukzessiven *Wirkungskette* um Vermittlungsmechanismen ergänzt. Nach Baumert et al. (2006) ist davon auszugehen, dass sich über solche Prozesse Kontextmerkmale höherer Ebenen auf individuelles Leistungsverhalten auf unterster Ebene auswirken. Dies entspricht dem von Friedrichs und Nonnenmacher (2010, S. 490) vorgeschlagenen Vorgehen, zunächst Effekte auf der Makroebene zu analysieren, bevor die Mechanismen, durch welche sich diese Effekte auf die Individuen auswirken, zu prüfen sind. Konkret wird anhand der verfügbaren Indikatoren überprüft, ob die mittlere Akzeptanz von Leistungsnormen innerhalb der Schulen und Schulklassen als ein möglicher Indikator für die normative Kultur der Altersgruppe auf der jeweiligen Kontextebene einen eigenständigen Erklärungsbeitrag leistet und von zusätzlichem Erklärungsmehrwert ist.

Es wird erwartet, dass bei einem hohen durchschnittlichen Ausgangsniveau der Leistungen im Schul- und Klassenkontext auch die gemeinsame Akzeptanz von Leistungsnormen überdurchschnittlich hoch ausfällt und diese ihrerseits zu höheren Testleistungen zu einem späteren Beobachtungszeitpunkt führen sollte (*Vermittlungshypothese*). Dementsprechend sollten sich diese Vermittlungsmechanismen in Verbindung mit den Kontextmerkmalen leistungsförderlich auf die individuelle Lernentwicklung auswirken. Die Spezifikation dieser Vermittlungsmechanismen auf den beiden Kontextebenen sollte einen eigenständigen Erklärungsbeitrag leisten und zu einem zusätzlichen Mehrwert des Gesamtmodells führen.

## 3. Sind Effekte des Klassen- oder Schulkontextes gruppenspezifisch?

In einem *fünften Schritt* wird untersucht, ob diese Vermittlungsmechanismen für alle Schüler gleichermaßen leistungsrelevant sind oder ob sie je nach individuellen Voraussetzungen unterschiedliche Auswirkungen haben. Durch die Modellierung von *Interaktionseffekten* lässt sich also prüfen, ob die Vermittlungsmechanismen des Klassen- oder Schulkontextes auf den Zusammenhang zwischen verschiedenen herkunfts- und personenbezogenen Schülermerkmalen und Kriteriumsvariable auf individueller Ebene wirken.

Es wird erwartet, dass sich die Vermittlungsmechanismen für Kontexteffekte in Abhängigkeit von individuellen Voraussetzungen unterschiedlich auswirken und dass sie bei ungünstigeren individuellen Lernausgangslagen kompensatorische Fördereffekte besitzen. So sollten z. B. Leistungsschwächere stärker von einer höheren Akzeptanz von Leistungsnormen hinsichtlich ihrer Leistungsentwicklung profitieren. Die

Annahme differenzieller Effekte von Merkmalen auf Aggregatebene beruht auf Befunden aus Meta-Analysen, die z. B. hinsichtlich der Leistungsheterogenität in Lerngruppen eine differenzielle Förderwirksamkeit für leistungsstarke und -schwächere Schüler belegt haben (z. B. Kulik und Kulik 1982; Lou et al. 1996; Slavin 1990a).

## 5 Datengrundlage und methodisches Vorgehen

### 5.1 Die Hamburger KESS-Studie

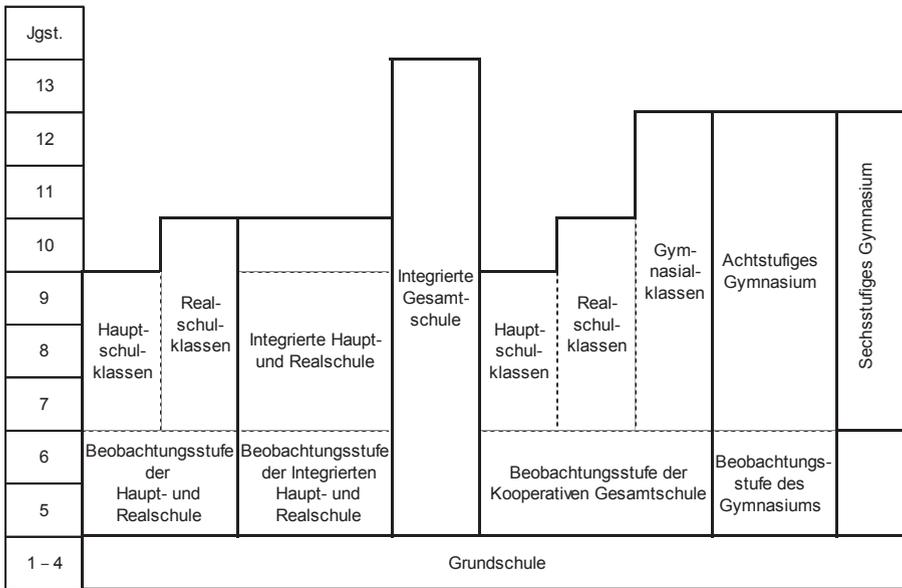
Datengrundlage ist die Hamburger Schulleistungsstudie „Kompetenzen und Einstellungen von Schülerinnen und Schülern“ (KESS). Im Rahmen von KESS wurden die Lernstände in den Kompetenzbereichen Deutsch (Leseverständnis und Orthografie), Mathematik, Englisch und Naturwissenschaften am Ende der Jahrgangsstufe 4 (KESS 4, 2003; Bos et al. 2006), unmittelbar zu Beginn der Jahrgangsstufe 7 (KESS 7, 2005; Bensen et al. 2009b) und am Ende der Jahrgangsstufe 8 (KESS 8, 2007; Bos et al. 2010a) erfasst.<sup>4</sup> Die Teilnahme an den Schulleistungstests war verpflichtend. Darüber hinaus wurden umfangreiche Hintergrundinformationen zu den individuellen und kontextspezifischen Rahmenbedingungen des schulischen Lehrens und Lernens mittels schriftlicher Befragungen der Schüler, ihrer Eltern, Lehrkräfte und Schulleitungen erhoben. Auftraggeberin der Studie war die Behörde für Schule und Berufsbildung der Freien und Hansestadt Hamburg. Die Organisation, Durchführung, Erfassung und Aufbereitung der Leistungsdaten erfolgten durch das Landesinstitut für Lehrerbildung und Schulentwicklung. Die Auswertung der Leistungsdaten in den Bereichen Leseverständnis und Mathematik fand am Institut für Schulentwicklungsforschung (IFS) der Technischen Universität Dortmund statt.

Das Hamburger Schulsystem zeichnete sich zum Zeitpunkt der KESS8-Studie im Jahr 2007 durch eine Reihe von Besonderheiten aus, die in Abb. 2 vereinfacht dargestellt sind (vgl. hierzu auch Bos et al. 2010a).<sup>5</sup>

In Hamburg erfolgt der Übergang in die weiterführende Schule nach dem Ende der vierjährigen Grundschulzeit. Mit dem Übergang in die Sekundarstufe findet dann eine zunehmende Differenzierung statt: An Haupt- und Realschulen, die in Hamburg eine organisatorische Einheit bilden, wird in den Jahrgangsstufen 5 und 6 im gemeinsamen Klassenverband ohne externe Leistungsdifferenzierung unterrichtet. Die Einrichtung dieser Beobachtungsstufe an Haupt- und Realschulen soll den Grundschulübergang entschärfen und eine Korrektur der Schulformentscheidung auch noch zu einem späteren Zeitpunkt ermöglichen. Ab der 7. Jahrgangsstufe wird innerhalb dieser Schulen in getrennten Hauptschul- oder Realschulklassen unterrichtet, in denen dann spä-

<sup>4</sup>Weitere Befragungen und Kompetenzmessungen des KESS-Jahrgangs erfolgten im Sommer 2009 am Ende der Sekundarstufe I (Jahrgangsstufe 10) bzw. zu Beginn der gymnasialen Oberstufe (Jahrgangsstufe 11), im Frühjahr 2011 am Ende der sogenannten Studienstufe an achtstufigen Gymnasien (Jahrgangsstufe 12) sowie ein Jahr später am Ende der Studienstufe der dreijährigen Oberstufe (Jahrgangsstufe 13) an Integrierten Gesamtschulen, Aufbaugymnasien und Beruflichen Gymnasien (Ivanov et al. 2011).

<sup>5</sup>Die Ausführungen beziehen sich auf die geltenden Regelungen zum Zeitpunkt der Durchführung der KESS-Studie. Auf die Besonderheiten, die sich durch die Hamburger Schulstrukturreform im Schuljahr 2010/11 ergeben haben, wird im Folgenden nicht näher eingegangen (vgl. hierzu z. B. Scharenberg 2012).



**Abb. 2** Das Hamburger Schulsystem (Stand 2007).

(Quelle: BSB/IFS/LI, Kompetenzen und Einstellungen von Schülerinnen und Schülern<sup>®</sup> KESS 8)

ter auch schulformspezifische Abschlüsse erworben werden können. Die Berechtigung für den Übergang in die Hauptschule oder Realschule wird auf Grundlage der Zeugnisnoten am Ende der Jahrgangsstufe 6 ausgesprochen. An Integrierten Haupt- und Realschulen (IHR-Schulen) erfolgt hingegen eine innere Differenzierung nach Hauptschul- und Realschulstatus (H- und R-Status) bei gemeinsamer Unterrichtung im Klassenverband in den Fächern Deutsch, Mathematik und Englisch.

An den grundständigen achtstufigen Gymnasien erfolgt nach der zweijährigen Beobachtungsstufe der Übergang in die Mittelstufe. Das Abitur kann nach insgesamt 12 Schuljahren (mit zweijähriger Oberstufe) erworben werden. Ab der Jahrgangsstufe 7 steht das sechsstufige Gymnasium auch den leistungsstärkeren Haupt- und Realschülern offen. An Integrierten Gesamtschulen (IGS) findet eine jahrgangsweise Fachleistungsdifferenzierung in ein leistungsstärkeres Kursniveau I und ein leistungsschwächeres Kursniveau II nur in den Fächern Mathematik und Englisch statt. Der Erwerb des Abiturs ist hier nach insgesamt 13 Schuljahren (mit dreijähriger Oberstufe) möglich. An Kooperativen Gesamtschulen werden ab der Jahrgangsstufe 7 Hauptschul-, Realschul- und Gymnasialzweige getrennt.

## 5.2 Methodisches Vorgehen

Für die Analyse von Kontexteffekten empfiehlt sich ein mehr Ebenenanalytisches Vorgehen, bei dem Prädiktoren simultan auf mehreren Ebenen, d. h. als individuelle und als auf höherer Ebene aggregierte Merkmale, berücksichtigt werden können. Gerade im Zusammenhang mit der Analyse von Kompositionseffekten ist deshalb die Mehrebenenanalyse das methodische Vorgehen der Wahl (Dreeben und Barr 1988,

S. 129), um die hierarchisch geschachtelte Struktur der Daten, wie sie gerade für die Schule als ein Anwendungsbereich der Kontextanalyse kennzeichnend ist, angemessen zu berücksichtigen. Die einzelnen Beobachtungseinheiten sind ineinander verschachtelt (Schüler in Schulklassen in Schulen) und nicht unabhängig voneinander.<sup>6</sup> So sind sich z. B. Schüler derselben Schulklasse einander ähnlicher als diejenigen, die verschiedene Schulklassen oder Schulen besuchen. Gründe hierfür können z. B. darin liegen, dass sie gemeinsam von denselben Lehrkräften unterrichtet werden, d. h. gemeinsamen Rahmenbedingungen unterliegen, und in Altersgruppen sind, d. h. gemeinsame Sozialbeziehungen pflegen.

Verallgemeinert ausgedrückt ermöglicht die Mehrebenenanalyse, Effekte auf der Mikro- und Makroebene simultan zu schätzen. Zunächst müssen die zu untersuchenden Kontexte und deren Effekte (direkt vs. indirekt) auf das individuelle Verhalten bestimmt werden. Darüber hinaus kann es erforderlich sein, zusätzlich zur Mikro- und Makroebene auch eine Mesoebene einzubeziehen. Bezogen auf den Schulbereich und mit Blick auf die zuvor abgeleiteten Fragestellungen soll die schulische Leistungsentwicklung als ein zentraler Aspekt der Entwicklung im Jugendalter durch individuelle Schülermerkmale und Merkmale des schulischen Kontexts vorhergesagt werden. Diese Kontextmerkmale können in verschiedene Ebenen eingebettet sein, wobei die Schulklasse als die unmittelbare Lernumwelt und Bezugsgruppe der Schüler angesehen wird (Baumert et al. 2004; Dar und Resh 1986). Diese wird wiederum von schulischen Rahmenbedingungen beeinflusst (Opdenakker und van Damme 2006).

Zur Beantwortung der Frage nach der Kontextabhängigkeit schulischer Leistungen, den Vermittlungsmechanismen, deren Gruppenspezifität und differenzieller Effektivität für verschiedene Personengruppen wird folgende Modellspezifikation gewählt: Die individuelle Entwicklung schulischer Leistungen stellt bei den Analysen die abhängige Variable dar. Dabei beschränken sich die Analysen auf die Entwicklung des Leseverständnisses, das als „Werkzeug“ für den selbstständigen Wissenserwerb auch in anderen Kompetenzbereichen eine grundlegende Voraussetzung ist (OECD 2001). Da schulische Kompetenzen durch eine Vielzahl von Determinanten bestimmt werden, welche die Rahmenbedingungen für individuelle Lernprozesse bilden (Helmke und Weinert 1997), werden zur Vorhersage der Kompetenzentwicklung in den Mehrebenenmodellen individuelle Schülermerkmale sowie weitere Merkmale des Klassen- und Schulkontextes hinzugezogen (zur Operationalisierung siehe Abschn. 5.3).

Die Modellgenese orientiert sich an dem von Hox (1995) vorgeschlagenen Vorgehen. Zunächst wird ein Modell ohne Prädiktoren (*Nullmodell*) spezifiziert, um die Varianzanteile der Kriteriumsvariablen zu bestimmen, die auf jeder der hier interessierenden Ebenen (Schüler, Schulklasse, Schule) maximal erklärt werden können sowie um entwicklungsrelevante Kontexte identifizieren und voneinander abgrenzen zu können. Dann werden zunächst auf unterster Ebene die Individualmerkmale Geschlecht, Migrationshintergrund, soziale Lage, kognitive Grundfähigkeit und

<sup>6</sup> Eine weitere Anwendungsmöglichkeit der Mehrebenenanalyse im schulischen Kontext ist z. B. die direkte Modellierung von Lernzuwächsen, wobei die Messwiederholungsdaten auf unterster Ebene verschachtelt sind und durch individuelle (zweite Ebene) und kontextuelle Merkmale von Schulen oder Schulklassen (dritte Ebene) vorhergesagt werden können (Anwendungsbeispiele finden sich u. a. in Bellin 2009; Guill und Gröhlich 2013).

Ausgangsleistung im Leseverständnis zu Beginn der Jahrgangsstufe 7 berücksichtigt (Variante I). Im nächsten Schritt werden als Kompositionsmerkmale des Klassenkontextes die soziale und leistungsbezogene Zusammensetzung der Schulklassen eingeführt, auf Schulebene wird zusätzlich die Schulformgliederung als Institutionsmerkmal berücksichtigt (Variante II). Anschließend wird die normative Kultur der Altersgruppe, welche nach Baumert et al. (2006) ein vermittelndes Prozessmerkmal zwischen Kontextfaktoren und individuellen Lernergebnissen darstellt, als zusätzlicher Prädiktor auf den beiden Aggregatebenen aufgenommen (Variante III).<sup>7</sup> Variante IV schätzt Effekte des Schul- und Klassenkontextes auf die Zusammenhänge auf individueller Ebene, um zu überprüfen, ob die angenommenen Vermittlungsmechanismen für alle Schüler gleichermaßen wirksam sind oder von individuellen Voraussetzungen moderiert werden.

Die Datenauswertung erfolgt mit der Software HLM 6.08 (Raudenbush et al. 2009). Kontinuierliche Prädiktoren werden auf individueller Ebene z-standardisiert und dann auf den höheren Ebenen aggregiert, sodass sich die Kompositionseffekte direkt ablesen lassen (Ditton 1998, S. 90 ff.). In den Varianten II und III werden ausschließlich direkte Effekte von Aggregatmerkmalen auf die abhängige Variable geschätzt. In Variante IV werden Interaktionseffekte (*Cross-level*-Interaktionen) zwischen einem Aggregatmerkmal und dem Zusammenhang zwischen verschiedenen individuellen Merkmalen und der abhängigen Variable zugelassen (für eine anschauliche Darstellung zur Schätzung dieser Effektarten siehe auch Snijders und Bosker 1999). Um abschließend beurteilen zu können, welche dieser Modellspezifikationen die empirischen Daten am besten repräsentiert, werden die verschiedenen Modelle anhand von Likelihood-Ratio- $\chi^2$ -Tests miteinander verglichen (Langer 2008, S. 119).

Fehlende Werte werden zunächst, sofern vorhanden, durch Angaben der Schüler zu früheren Messzeitpunkten ersetzt. Darüber hinausgehende fehlende Werte werden mittels eines einfachen Imputationsverfahrens auf Basis eines komplexen Modells an Hintergrundvariablen geschätzt (Software NORM 2.03; Schafer 1999). Hierfür werden die Variablen verwendet, die inhaltliche Zusammenhänge und Korrelationen mit den analyserelevanten Variablen in mindestens mittlerer Höhe aufweisen ( $|r| \geq 0,30$ ).<sup>8</sup>

Die Analysen beschränken sich auf die Entwicklung des Leseverständnisses in den Jahrgangsstufen 7 und 8, da Aspekte der normativen Kultur der Altersgruppe nur zum zweiten Messzeitpunkt der KESS-Studie (KESS 7) erfasst wurden und somit als Ausgangsbasis für die Lernentwicklung bis zum Ende der Jahrgangsstufe 8 verwendet werden können. Die Analysen erfolgen auf einem Signifikanzniveau von  $p < 0,05$ .

<sup>7</sup>In Anlehnung an Baumert et al. (2006, S. 145) wird die Akzeptanz von Leistungsnormen ausschließlich als Prädiktor auf aggregierter Ebene spezifiziert, um eine Überkontrolle des Modells auf Individualebene zu vermeiden, da davon ausgegangen werden kann, dass die individuelle Akzeptanz dieser Leistungsnormen wiederum von schulischen Qualitätsmerkmalen abhängt.

<sup>8</sup>Für die Imputation wurden Variablen aus den Schülerfragebögen verwendet, wie z. B. Anzahl der Bücher im Haushalt, soziales Kapital, allgemeines oder fachspezifisches Selbstkonzept, Zufriedenheit mit der Schule, Leistungsangst, Schüler-Lehrer-Verhältnis, Interesse der Eltern an den Freunden und Aktivitäten des Kindes, Verhältnis der Schüler zu ihren Eltern (zur Dokumentation der Erhebungsinstrumente vgl. Bos et al. 2009, 2010b).

### 5.3 Operationalisierung der verwendeten Variablen

#### *Prädiktoren auf Individualebene*

Abhängige Variable ist die individuelle Testleistung im Kompetenzbereich „Leseverständnis“ am Ende der Jahrgangsstufe 8. Das Leseverständnis wurde anhand verschiedener Testaufgaben erfasst, die sich bereits in anderen Schulleistungsstudien (z. B. PISA, IGLU, LAU) als valide und reliable Testinstrumente bewährt haben. Durch ein *Anker-Item-Design* wurden mehrere Messzeitpunkte miteinander verbunden (*fixed-parameters-Methode*; von Davier und von Davier 2007). Die Leistungsdaten wurden anhand des *einparametrischen Raschmodells* skaliert. Dadurch können die Testleistungen zu allen Messzeitpunkten auf einer gemeinsamen Metrik abgebildet werden (Bos et al. 2010a). Die Lernstände wurden für die KESS7-Population auf einen Mittelwert ( $M$ ) von 500 und eine Standardabweichung ( $SD$ ) von 100 Skalenpunkten standardisiert. Die Differenz der Testwerte zwischen zwei Messzeitpunkten bildet den Lernzuwachs ab.

Das *Geschlecht* (Referenz: Junge) wurde über die Schülerteilnahmeliste erfasst, die von den Testleitern ausgefüllt wurde. Der *Migrationshintergrund* wird über Angaben aus dem Elternfragebogen zu deren Geburtsland abgebildet. Bei diesem dichotomen Merkmal bilden die Schüler, deren Eltern beide in Deutschland geboren wurden, die Referenzkategorie gegenüber jenen mit mindestens einem im Ausland geborenen Elternteil. Die *soziale Lage* wird über den höchsten sozioökonomischen Status in der Familie (HISEI; Ganzeboom und Treiman 1996) erfasst, der auf den Elternangaben zu ihrer Berufstätigkeit basiert. Als ein Indikator für die *Intelligenz* wird die non-verbale kognitive Grundfähigkeit verwendet, die sprachfrei anhand von figuralen Analogien aus dem kognitiven Fähigkeitstest (KFT; Heller und Perleth 2000) erfasst wurde. Auf die Notwendigkeit korrekt spezifizierter Schätzmodelle, die ausleserelevante Leistungsmerkmale auf individueller und schulischer Ebene berücksichtigen, haben bereits Baumert und Schümer (2001) hingewiesen. Deshalb wird nachfolgend auch die *Ausgangsleistung* im Leseverständnis zu Beginn der Jahrgangsstufe 7 kontrolliert. Dadurch kann die vorhergesagte Variable indirekt als *Lernentwicklung* in den Jahrgangsstufen 7 und 8 interpretiert werden.<sup>9</sup>

#### *Prädiktoren auf Ebene des Klassenkontextes*

Die *soziale und leistungsbezogene Zusammensetzung* als Kompositionsmerkmale des Klassenkontextes werden anhand des arithmetischen Mittels der individuellen Merkmalsausprägungen über die Schulklassenzugehörigkeit in der Jahrgangsstufe 7 aggregiert. Im Unterschied zu globalen Merkmalen ohne weiteren Rückgriff auf individuelle Informationen (z. B. Schulträgerschaft) stellen solche Kompositions-

<sup>9</sup>Die *direkte* Modellierung des individuellen Lernzuwachses wäre auch anhand eines Drei-Ebenen-Modells mit der Schachtelung von Messwiederholungsdaten innerhalb von Individualeinheiten (Schüler) auf unterster Ebene möglich. Da in diesem Beitrag jedoch – neben individuellen Merkmalen – v.a. die gemeinsamen Auswirkungen des Schul- und Klassenkontextes auf zwei höheren Analyseebenen im Vordergrund stehen, die *direkte* Modellierung des Lernzuwachses aber nur die Berücksichtigung *einer* weiteren Aggregatebene erlaubt (Ditton 1998, S. 148), wird die am Ende der Jahrgangsstufe 8 erzielte Testleistung unter Kontrolle der Ausgangsleistung zu Beginn der Jahrgangsstufe 7 vorhergesagt und als *Leistungsentwicklung* in diesen beiden Schuljahren gedeutet.

merkmale nach Ditton (2013, S. 175) *analytische Merkmale* dar, die nur über die Aggregation individueller Merkmalsausprägungen innerhalb des Kontextes gewonnen werden.

Als ein Aspekt der normativen Kultur in der Altersgruppe, der im Rahmen der KESS-Studie als Indikator für Vermittlungsmechanismen von Kontexteffekten (Baumert et al. 2006) verfügbar ist, wird die *Akzeptanz von Leistungsnormen* auch bei Normabweichungen der anderen berücksichtigt, von der eine leistungsförderliche Wirkung vermutet wird. In KESS 7 wurde im allgemeinen Schülerfragebogen danach gefragt, wie sehr die Schüler darauf achten, im Unterricht nicht zu stören und Rücksicht auf andere zu nehmen. Die Akzeptanz von Leistungsnormen lässt sich über ihre Zustimmung zu den Fragen abbilden, wie oft sie versuchen,

- das zu tun, wozu die Lehrkraft sie auffordert;
- leise zu sein, wenn andere versuchen, im Unterricht zu lernen;
- weiterzuarbeiten, obwohl sie müde sind; und
- weiterzuarbeiten, obwohl die anderen nicht mehr ernsthaft bei der Sache sind.

Die Zustimmung zu diesen Aussagen konnte über ein vierstufiges Antwortformat (1 *immer*, 2 *oft*, 3 *selten*, 4 *nie*) ausgedrückt werden. Für die Analysen wurden die Antwortkategorien aller Items rekodiert, sodass höhere Skalenwerte für eine stärkere Akzeptanz von Leistungsnormen im Klassenkontext stehen. Die Skalenbildung erfolgte über den Mittelwert aller vier Items (Cronbachs  $\alpha=0,92$ , Trennschärfe  $r_{ii}$  bei allen Items  $\geq 0,79$ ). Zur Dokumentation der Erhebungsinstrumente für die gesamte KESS7-Population sei auf Bos et al. (2009) verwiesen.

#### *Prädiktoren auf Ebene des Schulkontextes*

Analog zur Operationalisierung auf Klassenebene wird die *soziale und leistungsbezogene Zusammensetzung* der Schulen über die mittleren Ausprägungen der entsprechenden Merkmale über die Schulzugehörigkeit aggregiert. Als institutionelles Merkmal wird die *Schulformgliederung* in der Sekundarstufe berücksichtigt. Referenzkategorie bilden die Hauptschulklassen an Haupt- und Realschulen. Da die Leistungsdifferenzierung an Integrierten Gesamtschulen nur in den Fächern Mathematik und Englisch erfolgt, werden sie hier als organisatorische Einheit betrachtet (zu den für den KESS-Jahrgang geltenden Besonderheiten des Hamburger Schulsystems vgl. Abschnitt 5.1). Als Prozessmerkmal zwischen Kontextfaktoren der Schulen und den Lernergebnissen der Schüler fließt auch auf der Schulebene die *mittlere Akzeptanz von Leistungsnormen* mit in die Analysen ein, deren Operationalisierung analog zum Vorgehen auf der Klassenebene erfolgt.

Um inhaltliche Zusammenhänge zwischen den Analysevariablen identifizieren zu können, sind in Tab. 1 die Korrelationen zwischen Prädiktor- und Kriteriumsvariablen dargestellt. In der linken Spalte sind die Korrelationen (Pearsons  $r$ ) der Merkmale Geschlecht, Migrationshintergrund, soziale Lage, kognitive Grundfähigkeit und Ausgangsleistung im Leseverständnis zu Beginn der Jahrgangsstufe 7 mit der Leseleistung am Ende der Jahrgangsstufe 8 auf individueller Ebene dargestellt. Da diese Einfachkorrelationen jedoch noch keine Rückschlüsse über Ursachen der Zusammenhänge erlauben und auch auf eine differenzielle Eingangsselektivität von Schulen und Schulformen zurückzuführen sein können, sind in der rechten Tabellen-

**Tab. 1** Korrelationen der unabhängigen Variablen mit der Leseleistung am Ende der Jahrgangsstufe 8

	<i>r</i>	Partialkorrelation
Geschlecht <sup>a</sup>	<b>0,18</b>	<b>0,16</b>
Migrationshintergrund <sup>b</sup>	<b>-0,23</b>	<b>-0,09</b>
Soziale Lage	<b>0,37</b>	<b>0,18</b>
Kognitive Grundfähigkeit	<b>0,48</b>	<b>0,22</b>
Ausgangsleistung zu Beginn der Jgst. 7 (T1)	<b>0,69</b>	

<sup>a</sup>Referenz: Junge

<sup>b</sup>Referenz: beide Eltern in Deutschland geboren

Signifikante Korrelationen fettgedruckt ( $p < 0,05$ ).  $n = 9.681$ .

Linke Spalte: Korrelationen auf individueller Ebene mit der Leseleistung am Ende der Jgst. 8 (T2)

Rechte Spalte: Korrelationen auf individueller Ebene mit der Leseleistung am Ende der Jgst. 8 (T2) unter Auspartialisierung der Ausgangsleistung zu Beginn der Jgst. 7 (T1)

spalte zusätzlich die Partialkorrelationen unter Kontrolle der Ausgangsleistung zu Beginn der Jahrgangsstufe 7 aufgeführt.

Bereits bivariat zeigen sich signifikante Zusammenhänge zwischen individuellen Schülermerkmalen und der Leseleistung: Mädchen erreichen am Ende der Jahrgangsstufe 8 eine höhere Leseleistung als Jungen. Schüler mit Migrationshintergrund erzielen dagegen niedrigere Lernstände als solche, deren Eltern beide in Deutschland geboren wurden. Mit einer höheren sozialen Lage oder einer höheren kognitiven Grundfähigkeit gehen ebenfalls höhere Lernstände einher. Der stärkste Zusammenhang besteht jedoch zwischen den Testleistungswerten zu den beiden Messzeitpunkten. Anders ausgedrückt: Fast die Hälfte der Varianz des Leseverständnisses am Ende der Jahrgangsstufe 8 wird bereits durch die Ausgangsleistung zu Beginn der Jahrgangsstufe 7 erklärt ( $R^2 = 0,48$ ). Unter Berücksichtigung der Ausgangsleistung fallen die Zusammenhänge der Schülermerkmale mit der Leseleistung am Ende der Jahrgangsstufe 8 mit Ausnahme des Geschlechts geringer, aber immer noch signifikant aus. Ein Teil der Zusammenhänge lässt sich also bereits durch die differenzielle Eingangsselektivität der Schulen und Schulformen erklären.

#### 5.4 Beschreibung der Analytestichprobe

Ausgangsbasis für die Analysen sind die Schüler, die an KESS 7 und KESS 8 teilgenommen haben ( $n = 11.420$ ). In einem mehrschrittigen Ausschlussverfahren werden davon diejenigen Schüler ausgeschlossen,

- die eine Integrierte Haupt- und Realschule ( $n = 447$ ) oder Kooperative Gesamtschule ( $n = 356$ ) besuchen, da diese jeweils eine für Hamburg besondere, aber zahlenmäßig geringe Schülerschaft umfassen,
- die in Schulklassen oder Schulen gehen, für die weniger als zehn Schüler zugeordnet werden können ( $n = 187$ ), um für die über die Klassen- oder Schulzugehörigkeit aggregierten Kompositionsmerkmale realistische Schätzer der jeweiligen Lernumwelt zu erhalten,
- deren Schulen einzülig sind ( $n = 138$ ), um Klassen- und Schulkontext analytisch eindeutig voneinander trennen zu können,

**Tab. 2** Mittlere Zusammensetzung der Schülerschaft in den verschiedenen Schulformen zu Beginn der Jahrgangsstufe 7

	Leseverständnis		Soziale Lage		Akzeptanz von Leistungsnormen		Anzahl Schüler
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Hauptschule	406,56 <sup>a</sup>	77,73	38,94 <sup>a</sup>	13,48	2,76 <sup>a</sup>	1,02	805
Realschule	478,79 <sup>b</sup>	73,40	44,50 <sup>b</sup>	14,10	3,09 <sup>b</sup>	0,86	1203
Gesamtschule	468,20 <sup>c</sup>	89,94	46,69 <sup>c</sup>	15,85	2,78 <sup>a</sup>	1,00	3074
Gymnasium	564,93 <sup>d</sup>	81,36	57,55 <sup>d</sup>	16,25	2,91 <sup>c</sup>	0,89	4599
Gesamt	510,34	99,41	50,93	17,00	2,88	0,94	9681

Schulformen mit ungleichen Suffixen (a, b, c, d) unterscheiden sich hinsichtlich der jeweiligen Zusammensetzung signifikant ( $p < 0,05$ ).

- die in Jahrgangsstufe 7 oder 8 die Schulklasse, Schule oder Schulform gewechselt haben ( $n=611$ ), um die Klassen- und Schulkomposition über diese zwei Schuljahre hinweg konstant zu halten und um ein allzu „komplexes Ineinandergreifen unterschiedlicher Kontexte, an denen das Individuum teil hatte und aktuell (noch) teilhat“ (Ditton 2013, S. 186), ausschließen zu können.

Somit verbleiben für die Analysen 9681 Schüler in 467 Schulklassen und 142 Schulen. Tabelle 2 beschreibt die Zusammensetzung dieser Schülerschaft hinsichtlich der mittleren Lernstände im Leseverständnis, der durchschnittlichen sozialen Zusammensetzung und der mittleren Akzeptanz von Leistungsnormen getrennt nach Schulform zu Beginn der Jahrgangsstufe 7.

An Hauptschulen liegt die durchschnittlich erzielte Testleistung rund eine Standardabweichung unter dem Gesamtdurchschnitt. An Integrierten Gesamtschulen und Realschulen fallen die mittleren Lernstände um mehr als eine halbe Standardabweichung höher aus. An Gymnasien bestehen die günstigsten Lernausgangslagen. Die Mittelwertsunterschiede fallen zwischen allen Schulformen signifikant aus. Integrierte Gesamtschulen umfassen eine insgesamt leistungsheterogenere Schülerschaft, die Leistungsstreuung fällt hier entsprechend etwas höher aus als an den anderen Schulformen. Insgesamt besteht jedoch auch innerhalb der verschiedenen Schulformen eine beträchtliche Leistungsheterogenität, die dem Anspruch des gegliederten Schulsystems nach Homogenisierung der Leistungen entgegensteht (Scharenberg 2012, 2013). Analog zu den leistungsbezogenen Lernvoraussetzungen zeigt sich auch hinsichtlich der sozialen Komposition, dass lediglich an Gymnasien überdurchschnittliche Werte erzielt werden. Während an Hauptschulen erneut die ungünstigsten Lernvoraussetzungen bestehen, ist die Schülerschaft an Gymnasien am privilegiertesten zusammengesetzt, wobei auch hier die Unterschiede zwischen allen Schulformen signifikant sind. Die Effektstärken für die Schulformunterschiede in der leistungsbezogenen ( $0,12 \leq |d| \leq 1,96$ ) und sozialen Komposition ( $0,40 \leq |d| \leq 1,17$ ) der Schülerschaft variieren jedoch je nach Wahl der Vergleichsgruppe beträchtlich, deuten aber auf bedeutsame und z. T. substantielle Unterschiede zwischen den Schulformen hin.

Schulformspezifische Unterschiede zeigen sich auch für die normative Kultur der Altersgruppe: Die Akzeptanz von Leistungsnormen ist an Realschulen am stärksten

ausgeprägt, gefolgt von Gymnasien. An Hauptschulen und Integrierten Gesamtschulen, die sich im Vergleich zu allen anderen Schulformen nicht signifikant voneinander unterscheiden, ist die mittlere Skalenausprägung hingegen am niedrigsten. Bei den signifikanten Schulformunterschieden handelt es sich um kleine Effekte ( $0,14 \leq |d| \leq 0,36$ ). Die vergleichsweise großen Standardabweichungen an Hauptschulen und Integrierten Gesamtschulen indizieren eine erhebliche Variabilität selbst zwischen Schulen derselben Schulform.

## 6 Ergebnisse

Nachfolgend werden die Ergebnisse der multivariaten und mehrbenenanalytischen Auswertungen dargestellt. Dabei wird zunächst die Bedeutung individueller Herkunfts- und Personenmerkmale für die Entwicklung schulischer Kompetenzen am Beispiel des Leseverständnisses (Abschn. 6.1) beschrieben, bevor in einem Mehrebenenmodell Merkmale des Klassen- und Schulkontextes (Abschn. 6.2) ergänzt werden. Abschließend wird auf gruppenspezifische Effekte (Abschn. 6.3) eingegangen.

### 6.1 Zur Bedeutung individueller Herkunfts- und Personenmerkmale

Zunächst wird in einem „leeren“, d. h. vollständig unspezifizierten Modell die Varianz der abhängigen Variable in drei Komponenten zerlegt, und zwar in Varianzanteile, die auf

- Unterschiede zwischen Schülern (Ebene 1) innerhalb von Schulklassen,
- Unterschiede zwischen Schulklassen (Ebene 2) innerhalb von Schulen,
- Unterschiede zwischen Schulen (Ebene 3) entfallen.

57,2% der Gesamtvarianz des Leseverständnisses am Ende der Jahrgangsstufe 8 entfallen auf individuelle Unterschiede. 34,1% der Varianz können durch Schulmerkmale erklärt werden. Auf die Klassenebene entfallen im Drei-Ebenen-Modell hingegen nur 8,7% der Varianz.<sup>10</sup> Diese Varianzzerlegung deutet bereits auf eine unterschiedliche Bedeutsamkeit der verschiedenen sozialen Kontexte für die Entwicklung schulischer Kompetenzen hin.

Tabelle 3 berücksichtigt zur Vorhersage des Leseverständnisses ausschließlich individuelle Schülermerkmale. Auf die Notwendigkeit eines korrekt spezifizierten Schätzmodells auf der Individualebene haben bereits Baumert et al. (2006) hingewiesen. Im Mehrebenenmodell gibt der Intercept den Wert im Leseverständnis am Ende

<sup>10</sup>Bei einer Spezifikation als Zwei-Ebenen-Modell (Scharenberg 2012) fällt der Varianzanteil des Leseverständnisses am Ende der Jahrgangsstufe 8, der durch die Klassenzugehörigkeit erklärt wird, mit 44,6% deutlich höher aus und ergibt sich aus der Differenz der Gesamtvarianz (100%) und dem Varianzanteil, der auf individuelle Schülermerkmale zurückgeht (55,4%). Im Drei-Ebenen-Modell zeigt sich hingegen, dass diese Varianz zwischen den Schulklassen innerhalb einer Schule in erster Linie auf eine deutlich höhere Varianz zwischen den Schulen zurückzuführen ist. In beiden Fällen, d. h. im Zwei- wie im Drei-Ebenen-Modell, sind jedoch die Varianzanteile auf der Klassenebene signifikant. Die leicht unterschiedlichen Varianzanteile, die jeweils auf die Individualebene entfallen, sind durch die Wahl einer anderen Analysestichprobe begründet.

**Tab. 3** Vorhersage des Leseverständnisses am Ende der Jahrgangsstufe 8 durch individuelle Schülermerkmale

	Koeffizient	SE
Intercept	<b>-0,07</b>	0,03
Individualebene		
Geschlecht <sup>a</sup>	<b>0,25</b>	0,02
Migrationshintergrund <sup>b</sup>	<b>-0,07</b>	0,02
Soziale Lage	<b>0,06</b>	0,01
Kognitive Grundfähigkeit	<b>0,13</b>	0,01
Ausgangsleistung zu Beginn der Jgst. 7 (T1)	<b>0,45</b>	0,01
Erklärte Varianz innerhalb der Schulklassen (in %) <sup>c</sup>	35,2	

<sup>a</sup>Referenz: Junge

<sup>b</sup>Referenz: beide Eltern in Deutschland geboren

<sup>c</sup>Varianzaufklärung im Verhältnis zum Nullmodell (vollständig unspezifiziert)

Varianzzerlegung im Nullmodell:

Ebene 1 (innerhalb der Schulklassen): 57,2 %

Ebene 2 (zwischen den Schulklassen): 8,7 %

Ebene 3 (zwischen den Schulen): 34,1 %

Kontinuierliche Prädiktoren auf der Individualebene z-standardisiert

Signifikante Koeffizienten fett ( $p < 0,05$ ).

der Jahrgangsstufe 8 an, den ein Junge ohne Migrationshintergrund mit durchschnittlicher sozialer Lage, kognitiver Grundfähigkeit und durchschnittlicher Ausgangsleistung zu Beginn der Jahrgangsstufe 7 erreichen würde. Handelt es sich hingegen um ein Mädchen oder einen Schüler mit einer überdurchschnittlichen Ausgangsleistung zum vorangegangenen Messzeitpunkt, so verändert sich die erreichte Testleistung am Ende der Jahrgangsstufe 8 um den jeweils angegebenen Parameter. Da neben dem Geschlecht, dem Migrationshintergrund, der sozialen Lage und der kognitiven Grundfähigkeit auch die Ausgangsleistung im Leseverständnis zum vorherigen Messzeitpunkt als ein Maß der differenziellen Eingangsselektivität kontrolliert wird, lassen sich die Koeffizienten auch als Prädiktoren der *Lernentwicklung* in den Jahrgangsstufen 7 und 8 interpretieren.

Die wichtigsten Prädiktoren auf individueller Ebene, allesamt signifikant, sind die Ausgangsleistung im Leseverständnis zu Beginn der Jahrgangsstufe 7 und das Geschlecht: Unter Kontrolle der jeweils anderen Merkmale erzielen leistungsstarke Schüler mit einer Ausgangsleistung, die zu Beginn der Jahrgangsstufe 7 eine Standardabweichung über dem Durchschnitt liegt, zwei Jahre später um rund eine halbe Standardabweichung höhere Testleistungen im Leseverständnis. Der Entwicklungsvorsprung der Mädchen fällt gegenüber Jungen um rund ein Viertel einer Standardabweichung höher aus. Die restlichen Prädiktoren folgen mit deutlichem Abstand und deuten auf individuelle Leistungsvorteile bei höherer kognitiver Grundfähigkeit und höherer sozialer Lage sowie auf Leistungseinbußen bei Schülern mit Migrationshintergrund hin, welche jedoch unter ansonsten gleichen Voraussetzungen nur sehr gering ausfallen.

Die Intraklassenkorrelation, die im Nullmodell bei  $\rho = 0,59$  lag, sinkt bei Spezifikation des Individualmodells auf  $\rho_{\text{res}} = 0,38$ . 35,2 % der Varianz innerhalb der Schul-

klassen – das sind rund 20% der Gesamtvarianz des Leseverständnisses am Ende der Jahrgangsstufe 8 – können bereits durch individuelle Herkunftsmerkmale der Schüler erklärt werden, die zu einer differenziellen Eingangsselektivität von Schulen und Schulformen führen.

## 6.2 Zur Bedeutung des Klassen- und Schulkontextes

Tabelle 4 ergänzt das zuvor bereits auf der Individualebene spezifizierte Modell um Merkmale des Klassen- und Schulkontextes. Auf der Klassenebene werden nun in

**Tab. 4** Vorhersage des Leseverständnisses am Ende der Jahrgangsstufe 8 durch Individual-, Klassen- und Schulmerkmale

	Modell 1		Modell 2	
	Koeffizient	SE	Koeffizient	SE
Intercept	<b>-0,32</b>	0,13	<b>-0,33</b>	0,13
Individualebene				
Geschlecht <sup>a</sup>	<b>0,25</b>	0,02	<b>0,25</b>	0,02
Migrationshintergrund <sup>b</sup>	<b>-0,06</b>	0,02	<b>-0,06</b>	0,02
Soziale Lage	<b>0,06</b>	0,01	<b>0,06</b>	0,01
Kognitive Grundfähigkeit	<b>0,13</b>	0,01	<b>0,13</b>	0,01
Ausgangsleistung zu Beginn der Jgst. 7 (T1)	<b>0,42</b>	0,01	<b>0,43</b>	0,01
Klassenebene				
Mittlere Ausgangsleistung zu Beginn der Jgst. 7 (T1)	<b>0,21</b>	0,05	<b>0,19</b>	0,05
Mittlere soziale Lage	0,07	0,06	0,05	0,06
Mittlere Akzeptanz von Leistungsnormen			<b>0,10</b>	0,04
Schulebene				
Mittlere Ausgangsleistung zu Beginn der Jgst. 7 (T1)	0,05	0,09	0,05	0,09
Mittlere soziale Lage	-0,09	0,07	-0,07	0,07
Mittlere Akzeptanz von Leistungsnormen			-0,04	0,05
Schulform <sup>c</sup>				
Realschule	0,22	0,18	0,22	0,19
Integrierte Gesamtschule	0,12	0,12	0,13	0,12
Gymnasium	<b>0,32</b>	0,14	<b>0,33</b>	0,14
Erklärte Varianz (in %) <sup>d</sup>				
Innerhalb der Schulklassen (Individualebene)		35,4		35,4
Zwischen den Schulklassen (Klassenebene)		48,0		51,7
Zwischen den Schulen (Schulebene)		98,2		98,2

<sup>a</sup>Referenz: Junge

<sup>b</sup>Referenz: beide Eltern in Deutschland geboren

<sup>c</sup>Referenz: Hauptschule

<sup>d</sup>Varianzaufklärung im Verhältnis zum Nullmodell (vollständig unspezifiziert)

Varianzzerlegung im Nullmodell:

Ebene 1 (innerhalb der Schulklassen): 57,2%

Ebene 2 (zwischen den Schulklassen): 8,7%

Ebene 3 (zwischen den Schulen): 34,1%

Kontinuierliche Prädiktoren auf der Individualebene z-standardisiert

Signifikante Koeffizienten fett ( $p < 0,05$ ).

Modellvariante 1 die durchschnittliche Testleistung innerhalb der Schulklassen zu Beginn der Jahrgangsstufe 7 und die mittlere soziale Lage (durchschnittlicher HISEI innerhalb der Schulklassen) aufgenommen. In Modellvariante 2 wird zusätzlich die mittlere Akzeptanz von Leistungsnormen als im Rahmen der gewählten Datenbasis verfügbarer Indikator für die normative Kultur der Altersgruppe berücksichtigt. Auf der Schulebene werden dieselben Kompositionsmerkmale (Modellvariante 1) sowie Vermittlungsmechanismen (Modellvariante 2) berücksichtigt, die dort jedoch über die Schulzugehörigkeit aggregiert werden. Außerdem fließt die Schulformzugehörigkeit in beiden Modellvarianten mit ein, um institutionelle Effekte der Leistungsdifferenzierung kontrollieren zu können. Der Intercept gibt nun den Wert an, den ein Junge ohne Migrationshintergrund mit durchschnittlicher sozialer Lage, kognitiver Grundfähigkeit und Ausgangsleistung im Leseverständnis am Ende der Jahrgangsstufe 8 erreicht, der eine im Hinblick auf die Merkmale auf den beiden höheren Ebenen durchschnittlich zusammengesetzte Schulklasse an einer durchschnittlich zusammengesetzten Hauptschule (Referenzkategorie) besucht. Ein solch durchschnittlicher Hauptschüler würde eine Leseleistung erreichen, die mit einem Drittel einer Standardabweichung signifikant unterhalb des Durchschnitts liegt.

Auf der Individualebene bleiben die Koeffizienten auch bei der Spezifikation von Aggregatmerkmalen im Vergleich zu Tab. 3 konstant: Signifikante und deutliche Leistungsvorteile am Ende der Jahrgangsstufe 8 ergeben sich bei günstigeren Lernausgangslagen und für die Mädchen. Auf der Klassenebene lassen sich Leistungsvorteile für leistungsstärkere Lernkontexte beobachten: Schüler aus Schulklassen, in denen die mittlere Ausgangsleistung eine Standardabweichung über dem Durchschnitt liegt, erzielen unter ansonsten gleichen Bedingungen knapp zwei Jahre später eine um rund ein Fünftel einer Standardabweichung signifikant höhere Testleistung im Leseverständnis im Vergleich zum Besuch von Schulklassen mit durchschnittlichem Ausgangsniveau. Unter Berücksichtigung der Leistungszusammensetzung lässt sich für die soziale Komposition der Schulklassen kein eigenständiger Effekt mehr nachweisen. Darüber hinaus erweist sich die normative Kultur innerhalb der Schulklassen (Modellvariante 2) als leistungsrelevanter Prädiktor: In Schulklassen, in denen die mittlere Akzeptanz von Leistungsnormen überdurchschnittlich hoch ist, werden um ein Zehntel einer Standardabweichung signifikant höhere Lernstände erzielt. Die Berücksichtigung dieses Prädiktors liefert einen zusätzlichen Erklärungswert von 3,7 Prozentpunkten gegenüber Modellvariante 1.

Auf der Schulebene ist in beiden Modellvarianten nur der institutionelle Lernkontext relevant für die Leistungsentwicklung: Schüler an Gymnasien erzielen unter ansonsten gleichen individuellen, klassenbezogenen und schulischen Voraussetzungen um ein Drittel einer Standardabweichung signifikant höhere Lernstände im Leseverständnis am Ende der Jahrgangsstufe 8 als jene an Hauptschulen. Auch an Realschulen und Integrierten Gesamtschulen zeichnet sich ein Leistungsvorsprung gegenüber Hauptschulen ab, der sich jedoch nicht zufallskritisch absichern lässt. Gleiches gilt für die Effekte der sozialen und leistungsbezogenen Zusammensetzung der Schulen. Hier deutet sich an, dass Kompositions- und Institutionseffekte auf der Schulebene stark miteinander konfundiert sind (vgl. z. B. auch Baumert et al. 2006; Scharenberg 2012). Insgesamt bestätigt sich hier der bereits bekannte Befund, wonach die verschiedenen Schulformen differenzielle Lern- und Entwicklungsmi-

lieus darstellen, die den Schülern unabhängig von und zusätzlich zu individuellen Lernvoraussetzungen unterschiedliche Entwicklungschancen bieten – v. a. an Gymnasien (siehe z. B. auch Baumert et al. 2006; Gröhlich et al. 2010a, b; Guill und Gröhlich 2013). Die Akzeptanz von Leistungsnormen übt auf institutioneller Ebene der Schulen ebenfalls keinen signifikanten Effekt auf die erzielte Testleistung im Leseverständnis aus und ist ausschließlich im Klassenkontext entwicklungsrelevant.

Durch die in Tab. 4 dargestellte Modellspezifikation können auf individueller Ebene 35,4% der Leistungsvarianz – das entspricht 20,3% der Gesamtvarianz<sup>11</sup> –, 51,7% der Varianz auf der Klassenebene – das entspricht 4,5% der Gesamtvarianz<sup>12</sup> – und 98,2% der Leistungsvarianz zwischen den Schulen – das entspricht 33,5% der Gesamtvarianz<sup>13</sup> – erklärt werden. In der Summe liefert das Drei-Ebenen-Modell somit einen Erklärungsbeitrag von 58,2% der Gesamtvarianz in den am Ende der Jahrgangsstufe 8 erzielten Leseleistungen. Somit bleibt noch ein Erklärungsspielraum von insgesamt 41,8% der Gesamtvarianz. Herunter gerechnet verbleiben damit für Merkmale des Klassenkontextes noch 4,6% und für Merkmale des Schulkontextes noch 0,6% der Gesamtvarianz, die nicht bereits durch die hier berücksichtigten Merkmale erklärt werden. Der Löwenanteil der bislang noch unerklärten Varianz verbleibt jedoch mit 36,9% für individuelle Schüler- und familiäre Herkunftsmerkmale.

### 6.3 Gruppenspezifische Effekte

Die bisherigen Analysen haben gezeigt, dass nicht davon ausgegangen werden kann, dass die verschiedenen Dimensionen der Zusammensetzung der Schülerschaft gleichermaßen leistungsrelevant sind. Vielmehr scheint v. a. die Leistungszusammensetzung des Klassenkontextes für die Entwicklung des Leseverständnisses in den Jahrgangsstufen 7 und 8 bedeutsam zu sein. Darüber hinaus konnte gezeigt werden, dass auch eine hohe Akzeptanz von Leistungsnormen, welche hier als Indikator für Vermittlungsprozesse zwischen der Schülerzusammensetzung und individuellen Lernergebnissen herangezogen wird, leistungsförderlich ist. Bislang unbeantwortet geblieben ist jedoch die Frage, ob diese Vermittlungsprozesse für alle Schüler die gleichen Auswirkungen haben oder ob sie auch in Wechselwirkung mit individuellen Schülermerkmalen auftreten, ob also von differenziellen Effekten in Abhängigkeit individueller Voraussetzungen auszugehen ist. In einem nächsten Schritt werden deshalb im Mehrebenenmodell Interaktionseffekte zwischen der mittleren Akzeptanz von Leistungsnormen im Klassenkontext und den Individualmerkmalen zugelassen.

Tabelle 5 zeigt zunächst, dass bei der Modellierung von Interaktionseffekten im Vergleich zu einfacher spezifizierten Modellen die Auswirkungen von Individualmerkmalen, Kompositionsmerkmalen und der mittleren Akzeptanz von Leistungsnormen im Klassenkontext konstant bleiben. Darüber hinaus lässt sich ein zusätzlicher, negativer Interaktionseffekt zwischen dem Geschlecht und der mittleren Akzeptanz von Leistungsnormen auf die Lernentwicklung in den Jahrgangsstufen 7 und 8 beobachten: Je stärker die Akzeptanz von Leistungsnormen im Klassenkontext ist, desto

<sup>11</sup>  $R^2 = 0,354 \cdot 0,572 = 0,203$ .

<sup>12</sup>  $R^2 = 0,517 \cdot 0,087 = 0,045$ .

<sup>13</sup>  $R^2 = 0,982 \cdot 0,341 = 0,335$ .

**Tab. 5** Interaktionseffekte zwischen individuellen Schülermerkmalen und mittlerer Akzeptanz von Leistungsnormen innerhalb der Schulklassen bei der Vorhersage des Leseverständnisses am Ende der Jahrgangsstufe 8

	Koeffizient	SE
Intercept	<b>-0,34</b>	0,13
Individualebene		
Geschlecht <sup>a</sup>	<b>0,25</b>	0,02
Migrationshintergrund <sup>b</sup>	<b>-0,06</b>	0,02
Soziale Lage	<b>0,06</b>	0,01
Kognitive Grundfähigkeit	<b>0,13</b>	0,01
Ausgangsleistung zu Beginn der Jgst. 7 (T1)	<b>0,43</b>	0,01
Klassenebene		
Mittlere Ausgangsleistung zu Beginn der Jgst. 7 (T1)	<b>0,19</b>	0,05
Mittlere soziale Lage	0,04	0,06
Mittlere Akzeptanz von Leistungsnormen	<b>0,12</b>	0,05
Interaktionseffekte auf der Klassenebene		
Mittlere Akzeptanz von Leistungsnormen * Geschlecht	<b>-0,08</b>	0,04
Mittlere Akzeptanz von Leistungsnormen * Migrationshintergrund	0,05	0,04
Mittlere Akzeptanz von Leistungsnormen * soziale Lage	0,00	0,02
Mittlere Akzeptanz von Leistungsnormen * kognitive Grundfähigkeit	0,02	0,02
Mittlere Akzeptanz von Leistungsnormen * Ausgangsleistung T1	-0,01	0,02
Schulebene		
Mittlere Ausgangsleistung zu Beginn der Jgst. 7 (T1)	0,05	0,09
Mittlere soziale Lage	-0,05	0,07
Mittlere Akzeptanz von Leistungsnormen	-0,04	0,05
Schulform <sup>c</sup>		
Realschule	0,23	0,18
Integrierte Gesamtschule	0,13	0,12
Gymnasium	<b>0,34</b>	0,14
Erklärte Varianz (in %) <sup>d</sup>		
Innerhalb der Schulklassen (Individualebene)		35,6
Zwischen den Schulklassen (Klassenebene)		57,8
Zwischen den Schulen (Schulebene)		97,7

<sup>a</sup>Referenz: Junge<sup>b</sup>Referenz: beide Eltern in Deutschland geboren<sup>c</sup>Referenz: Hauptschule<sup>d</sup>Varianzaufklärung im Verhältnis zum Nullmodell (vollständig unspezifiziert)

Varianzzerlegung im Nullmodell:

Ebene 1 (innerhalb der Schulklassen): 57,2%

Ebene 2 (zwischen den Schulklassen): 8,7%

Ebene 3 (zwischen den Schulen): 34,1%

Kontinuierliche Prädiktoren auf der Individualebene z-standardisiert

Signifikante Koeffizienten fett ( $p < 0,05$ ).

geringer ist der Effekt, den das Geschlecht auf die Lernentwicklung ausübt. Da die Entwicklung des Leseverständnisses von Jungen unter ansonsten identischen individuellen, klassenbezogenen und schulischen Voraussetzungen an Hauptschulen (Referenzkategorie) ungünstiger verläuft als die der Mädchen, scheint eine hohe Akzeptanz

**Tab. 6** Vergleich der verschiedenen Modellspezifikationen

Variante	I	II	III	IV
	Tabelle 3	Tabelle 4, Modell 1	Tabelle 4, Modell 2	Tabelle 5
Spezifikation	Nur individuelle Schülermerkmale	+ Kompositions- und Institutionsmerkmale	+ Kompositions- und Institutionsmerkmale + Vermittlungsmechanismen	+ Interaktionseffekte
Kontextmodellierung		Klassen und Schulen	Klassen und Schulen	Effekte des Klassen- und Schulkontextes auf individuelle Leistungsentwicklung
Mittlere Devianz	19.103,54	18.941,65	18.933,39	18.906,28
df	141	136	135	130
Referenzvariante		(I)	(II)	(III)
$\Delta$ Devianz		161,89	8,26	27,11
$\Delta$ df		5	1	5
Prüfgröße		11,07	3,84	11,07
Veränderung des Modellfits <sup>a</sup>		+	+	+

<sup>a</sup>Nicht signifikante (0) bzw. signifikante Verbesserung (+) des Modellfits gegenüber dem Referenzmodell ( $p < 0,05$ )

von Leistungsnormen einen leichten kompensatorischen Effekt zu besitzen, der diesen Nachteil der Jungen geringfügig, aber signifikant, verringern kann. Dieser geschlechtsspezifische Interaktionseffekt ist jedoch kein Effekt einer unterschiedlichen Geschlechterzusammensetzung innerhalb der Schulklassen, da er sich auch dann noch nachweisen lässt, wenn der Mädchenanteil innerhalb der Schulklassen zusätzlich kontrolliert wird (hier nicht dargestellt). Darüber hinaus lässt sich jedoch für die mittlere Akzeptanz von Leistungsnormen kein zusätzlicher Interaktionseffekt mit weiteren Schülermerkmalen beobachten. Es liegt also keine Mehrfachinteraktion oder eine Kombination mehrerer Benachteiligungsfaktoren auf individueller Ebene vor. Die Berücksichtigung von Interaktionseffekten mit allen Individualmerkmalen erhöht den Anteil der erklärten Varianz um rund 6 Prozentpunkte auf der Klassen-ebene auf nunmehr 57,8% – das entspricht 5,0% der Gesamtvarianz.<sup>14</sup>

#### 6.4 Modellvergleichendes Resümee

Tabelle 6 gibt abschließend Auskunft über die Anpassungsgüte der verschiedenen Modellvarianten. In der ersten Modellvariante (Tab. 3) wurden nur individuelle Schülermerkmale spezifiziert. Die zweite Modellvariante (Tab. 4, Modell 1) berücksichtigte zusätzlich den Schul- und Klassenkontext und schätzte kompositionelle Effekte der leistungsbezogenen und sozialen Schülerzusammensetzung sowie institutionelle Effekte der Schulformzugehörigkeit. In einem dritten Schritt wurde ein Indikator für Vermittlungsmechanismen für Kontexteffekte ergänzt (Tab. 4, Modell 2). In der vierten Modellvariante (Tab. 5) wurden *Cross-Level*-Interaktionen zwischen Merkmalen

<sup>14</sup> $R^2 = 0,578 \cdot 0,087 = 0,050$ .

des Schul- und Klassenkontextes und dem Zusammenhang von Herkunfts- und Personenmerkmalen und der abhängigen Variable auf individueller Ebene zugelassen.

Diese Modellvarianten können nun über ihre Anpassungsgüte an die empirische Datengrundlage anhand eines Likelihood-Ratio- $\chi^2$ -Tests miteinander verglichen werden. Dieser Test bestimmt die Devianzwerte und die Anzahl der Freiheitsgrade eines Alternativmodells im Vergleich zu einem Referenzmodell. Die daraus resultierenden Differenzwerte werden anhand eines  $\chi^2$ -Tests daraufhin getestet, ob das jeweilige Alternativmodell zu einer signifikanten Verbesserung gegenüber dem Referenzmodell führt (siehe hierzu Langer 2008, S. 119).

Der Modellvergleich zeigt, dass die Berücksichtigung der Drei-Ebenen-Struktur (Variante II) zu einer signifikant besseren Modellanpassung gegenüber jenem Modell führt, das ausschließlich individuelle Schülermerkmale berücksichtigt (Variante I). Berücksichtigt man zusätzlich die Vermittlungsmechanismen (Variante III), von denen angenommen wird, dass Kontexteffekte über diese vermittelt werden, ergibt sich im Vergleich zu Variante II ebenfalls eine signifikante Modellverbesserung. Die Variante IV, bei der zusätzlich Interaktionseffekte zwischen Vermittlungsmechanismen auf aggregierter Ebene und Schülermerkmalen auf individueller Ebene zugelassen werden, weist schließlich die beste Modellanpassung auf.

## 7 Zusammenfassung und Diskussion

Ausgangspunkt des Beitrags war die Frage, welcher der verschiedenen Kontexte im Mehrebenensystem Schule entwicklungsrelevant ist und über welche Vermittlungsmechanismen sich Kontexteffekte auf die Schulleistung als zentraler Aspekt der Entwicklung im Jugendalter entfalten. Die Ergebnisse zeigen anhand von Daten der Hamburger KESS-Studie zunächst die unterschiedliche Bedeutung verschiedener sozialer Kontexte für die schulische Kompetenzentwicklung im Leseverständnis: Im vollständig unspezifizierten, also „leeren“, Drei-Ebenen-Modell entfällt rund ein Drittel der Varianz der Kriteriumsvariable auf den Schulkontext, hingegen nur weniger als 10% auf den Klassenkontext. Die Variabilität zwischen den Schulklassen ist also schon von vornherein durch den Schulkontext eingeschränkt, der wiederum mit der Schulform und der sozio-kulturellen Zusammensetzung des Einzugsgebiets konfundiert ist. Zwar scheint dem Schulkontext eine höhere Bedeutung für die Leistungsentwicklung zuzukommen, dennoch ist der Varianzanteil, der auf den Klassenkontext entfällt, signifikant. Die Schulklasse ist somit als eigenständiger Entwicklungskontext im Jugendalter keineswegs zu vernachlässigen, stellt sie doch die unmittelbare Lernumwelt im schulischen Bereich dar, in der die Schüler über viele Jahre hinweg gemeinsam mit und von anderen lernen.

Mehr als die Hälfte der Leistungsvarianz entfällt im unspezifizierten Modell hingegen auf individuelle Unterschiede zwischen den Schülern. Diese werden zwar bereits v. a. durch unterschiedliche kognitive und leistungsbezogene Voraussetzungen, aber auch durch geschlechtsspezifische Disparitäten zugunsten der Mädchen erklärt. Allerdings lassen sich auf individueller Ebene auch migrationsbedingte und soziale Disparitäten nachweisen. Dies deutet darauf hin, dass auch der familiäre Kontext einen eigenständigen Beitrag zu ungleichen Entwicklungsverläufen leistet.

Insgesamt lässt sich in dieser *ersten Modellvariante* unter Berücksichtigung individueller Herkunftsmerkmale mehr als ein Drittel der Entwicklungsunterschiede auf individueller Ebene oder rund ein Fünftel der Gesamtvarianz erklären. Diese differenzielle Eingangsselektivität der Schulkontexte hatte sich zuvor bereits deskriptiv angedeutet.

Anschließend wurde anhand von Mehrebenenanalysen schrittweise überprüft, in welcher Weise sich Schule und Schulklasse als soziale Kontexte auf die Entwicklung im Jugendalter auswirken, durch welche Mechanismen verschiedene Kontextmerkmale vermittelt werden und ob solche Vermittlungsprozesse gruppenspezifisch auftreten. In einer *zweiten Modellvariante* zeigen sich dabei Vorteile eines leistungsstarken Klassenkontextes. Außerdem lässt sich eine höhere Lerneffektivität an Gymnasien, v. a. in Abgrenzung zu Hauptschulen, beobachten. Dies bestätigt die Befunde früherer Arbeiten (z. B. Baumert et al. 2006; Bensen et al. 2009a; Gröhlich et al. 2010a, b; Neumann et al. 2007; Nikolova 2011). Es ist somit einerseits bedeutsam, *wo*, also in welchem Schulkontext die Schüler lernen. Andererseits macht es aber auch einen Unterschied, *mit wem* sie zusammen lernen. Schulklassen und Schulen stellen also gemeinsam differenzielle Lern- und Entwicklungsmilieus dar, die den Schülern, ungeachtet ihrer individuellen Unterschiede, je nach Kontextzugehörigkeit unterschiedliche Entwicklungsmöglichkeiten bieten. Es bestehen also neben *individuellen* auch erhebliche *institutionelle Matthäus-Effekte*. Überraschend erscheinen jedoch die ansonsten vielfach nachgewiesenen, hier in der Drei-Ebenen-Spezifikation aber ausbleibenden Kompositionseffekte des Schulkontextes. Bei gleichzeitiger Modellierung von Schul- und Klassenkontext schlugen sich Merkmale der sozialen und leistungsbezogenen Zusammensetzung von Schulen offenbar ausschließlich in Kompositionseffekten des Klassenkontextes *innerhalb* der jeweiligen Schulformen nieder. Dies legt vertiefende schulformspezifische Analysen nahe. Zudem wäre es interessant zu überprüfen, ob die hier ausbleibenden Kompositionseffekte der sozialen und leistungsbezogenen Schulzusammensetzung möglicherweise auf nicht-lineare Effekte zurückzuführen sind oder sich nur ab bestimmten Schwellenwerten nachweisen lassen (Friedrichs und Nonnenmacher 2010, S. 478).

In einer *dritten Modellvariante* wurden *Vermittlungsprozesse* auf den beiden Kontextebenen ergänzt, für die hier als Indikator die normative Kultur der Altersgruppe berücksichtigt wurde und von denen theoretisch anzunehmen ist, dass Effekte von Kontextmerkmalen über diese vermittelt werden (Baumert et al. 2006, S. 126). Als Ergebnis lässt sich festhalten, dass diese Vermittlungsprozesse auf schulischer Ebene unter Kontrolle kompositioneller und institutioneller Merkmale offenbar nur eine untergeordnete Rolle spielen. Im Klassenkontext zeigt sich hingegen ein direkter Effekt der geteilten Akzeptanz von Leistungsnormen als Vermittlungsprozess, der als Merkmal der unmittelbaren Lernumwelt der Schüler für die konkrete Lehr-Lernsituation im Unterricht kennzeichnend ist (Ditton 2000). Eine hohe Akzeptanz von Leistungsnormen scheint also durchaus leistungsförderlich zu sein und liefert einen zusätzlichen Erklärungswert bei der Vorhersage der Leistungsentwicklung. Der Nachweis von Vermittlungsprozessen innerhalb von Schulklassen als unmittelbare Entwicklungskontexte im Jugendalter ist auch theoretisch zu erwarten, da „*Vermittlungsprozesse über soziale Interaktionen* stattfinden und erst dadurch ein Kollektiv oder Kontext überhaupt für ein Individuum handlungsrelevant wird“ (Dit-

ton 2013, S. 178; Hervorhebung im Original). Die Überprüfung gruppenspezifischer Effekte dieser Vermittlungsprozesse ergab in einer *vierten Modellvariante*, dass eine hohe Akzeptanz von Leistungsnormen im Klassenkontext nur für männliche Hauptschüler (Referenzkategorie) einen leichten kompensatorischen Fördereffekt hat.

*Zusammenfassend* zeigte sich im Vergleich, dass mit jeder Modellvariante ein zusätzlicher Erklärungsbeitrag zur Vorhersage der Kompetenzentwicklung im Leseverständnis gewonnen werden konnte, sodass das vollständig spezifiziertere Gesamtmodell die höchste Anpassung an die empirischen Daten lieferte. Zur Vorhersage der schulischen Leistungsentwicklung im Jugendalter muss also davon ausgegangen werden, dass diese nicht nur von individuellen Lernvoraussetzungen, sondern gleichzeitig auch von Bedingungen verschiedener Kontexte und von Vermittlungsprozessen abhängt, wobei letztere gruppenspezifisch auftreten. Bezogen auf den schulischen Bereich als ein Anwendungsgebiet der Kontextanalyse haben sich die Mehrebenenanalysen hier als geeignetes methodisches Verfahren erwiesen, das der hierarchischen Struktur der vorliegenden Daten gerecht wird und es ermöglicht, simultan Effekte der Makro- und Mesoebene auf die Mikroebene zu schätzen.

In methodischer Hinsicht eröffnen die hier präsentierten Befunde gleichzeitig vielfältige Anknüpfungspunkte für zukünftige Analysen: Einerseits legen sie die Überprüfung eines solchen Vermittlungsmodells für Kontexteffekte anhand einer weiteren Schülerpopulation, für einen anderen Kompetenzbereich sowie die Ausdehnung auf einen längeren Untersuchungszeitraum nahe. Andererseits wurde hier als im Rahmen der vorliegenden Datenbasis verfügbarer Indikator für Vermittlungsprozesse von Kontexteffekten lediglich die Akzeptanz von Leistungsnormen als ein Aspekt der normativen Kultur der Altersgruppe herausgegriffen. Neben diesen führen Baumert et al. (2006, S. 126) auch deviante Verhaltensnormen, Schulbindung und Verantwortungsübernahme sowie Vergleichsprozesse auf Schülerebene als bezugsgruppenbezogene Prozesse auf. Unberücksichtigt geblieben sind hier auch Engagement, Leistungs- und Verhaltensnormen der Elternschaft sowie curriculare und didaktische Aspekte. Eine Gesamtüberprüfung eines solch mehrfach gestuften Vermittlungsmodells steht also nach wie vor aus. Bereits eine Übertragung ihres komplexen Vermittlungsmodells auf den Klassenkontext ohne sonstige Modifikation wäre aufschlussreich, da sich ihr Vermittlungsmodell auf der Schulebene als nicht tragfähig erweist (ebd., S. 149).

Darüber hinaus wäre es wünschenswert, gemeinsame Effekte verschiedener Lern- und Entwicklungskontexte direkt auf individuelle Lernzuwächse zu modellieren. Die Mehrebenenanalyse wäre hierfür ein geeignetes Verfahren, da sie die Schachtelung von Messwiederholungsdaten innerhalb von Individualeinheiten ermöglicht. Vielversprechend wäre es auch – analog zu Baumert et al. (2006) – im Klassen- und Schulkontext Pfadmodelle zu konzipieren, bei denen die Kontextmerkmale die exogenen Variablen bilden und die Vermittlungsmechanismen als Mediatorvariablen fungieren. Der vorliegende Beitrag verdeutlicht also einmal mehr, „dass die Ermittlung von *Kontexteffekten* nicht trivial ist“ (Ditton 2013, S. 175, Hervorhebung im Original).

## Literatur

- Artelt, Cordula, Petra Stanat, Wolfgang Schneider und Ulrich Schiefele. 2001. Lesekompetenz: Testkonzeption und Ergebnisse. In *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*, Hrsg. Jürgen Baumert, Eckhard Klieme, Michael Neubrand, Manfred Prenzel, Ulrich Schiefele, Wolfgang Schneider, Petra Stanat, Klaus-Jürgen Tillmann und Manfred Weiß, 69–137. Opladen: Leske + Budrich.
- Barr, Rebecca, und Robert Dreeben. 1983. *How schools work*. Chicago: University of Chicago Press.
- Baumert, Jürgen, und Gundel Schümer. 2001. Schulformen als selektionsbedingte Lernmilieus. In *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*, Hrsg. Jürgen Baumert, Eckhard Klieme, Michael Neubrand, Manfred Prenzel, Ulrich Schiefele, Wolfgang Schneider, Petra Stanat, Klaus-Jürgen Tillmann und Manfred Weiß, 454–467. Opladen: Leske + Budrich.
- Baumert, Jürgen, Eckhard Klieme, Michael Neubrand, Manfred Prenzel, Ulrich Schiefele, Wolfgang Schneider, Klaus-Jürgen Tillmann und Manfred Weiß. 2001. *Soziale Bedingungen von Schulleistungen. Zur Erfassung von Kontextmerkmalen durch Schüler-, Schul- und Elternfragebögen*. <http://www.mpib-berlin.mpg.de/pisa/Kontextmerkmale.pdf> (Zugegriffen: 28. Okt. 2013).
- Baumert, Jürgen, Ulrich Trautwein und Cordula Artelt. 2003. Schulumwelten und institutionelle Bedingungen des Lehrens und Lernens. In *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*, Hrsg. Jürgen Baumert, Cordula Artelt, Eckhard Klieme, Michael Neubrand, Manfred Prenzel, Ulrich Schiefele, Wolfgang Schneider, Klaus-Jürgen Tillmann und Manfred Weiß, 261–331. Opladen: Leske + Budrich.
- Baumert, Jürgen, Werner Blum und Michael Neubrand. 2004. Drawing the lessons from PISA 2000 – Long-term research implications: Gaining a better understanding of the relationship between system inputs and learning outcomes by assessing instructional and learning processes as mediating factors. *Zeitschrift für Erziehungswissenschaft, Beiheft 3*:143–157.
- Baumert, Jürgen, Petra Stanat und Rainer Watermann. 2006. Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit. Vertiefende Analysen im Rahmen von PISA 2000*, Hrsg. Jürgen Baumert, Petra Stanat und Rainer Watermann, 95–188. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, Jürgen, Kai Maaz, Petra Stanat und Rainer Watermann. 2009. Schulkomposition oder Institution – was zählt? Schulstrukturen und die Entstehung schulformspezifischer Entwicklungsverläufe. *Die Deutsche Schule* 101:33–46.
- Becker, Michael, Oliver Lüdtke, Ulrich Trautwein und Jürgen Baumert. 2006. Leistungszuwachs in Mathematik. Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem? *Zeitschrift für Pädagogische Psychologie* 20:233–242.
- Bellin, Nicole. 2009. *Klassenkomposition, Migrationshintergrund und Leistung. Mehrebenenanalysen zum Sprach- und Leseverständnis von Grundschulern*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Berkemeyer, Nils, und Veronika Manitus. 2013. Gerechtigkeit als Kategorie der Analyse von Schulsystemen – das Beispiel Chancenspiegel. In *Schul- und Bildungsforschung: Diskussionen, Befunde und Perspektiven. Festschrift für Wilfried Bos*, Hrsg. Knut Schwippert, Martin Bosen und Nils Berkemeyer, 223–240. Münster: Waxmann.
- Berkemeyer, Nils, Wilfried Bos und Veronika Manitus. 2012. Chancenspiegel. Zur Leistungsfähigkeit und Chancengerechtigkeit der deutschen Schulsysteme. In *Chancenspiegel. Zur Leistungsfähigkeit und Chancengerechtigkeit der deutschen Schulsysteme*, Hrsg. Bertelsmann Stiftung und Institut für Schulentwicklungsforschung, 1–192. Gütersloh: Verlag Bertelsmann Stiftung.
- Blau, Peter M. 1960. Structural effects. *American Sociological Review* 25:178–193.
- Bosen, Martin, Carola Gröhlich und Wilfried Bos. 2009a. Differentielle Lern- und Entwicklungsmilieus in der Hamburger Beobachtungsstufe? In *KESS 7 – Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*, Hrsg. Wilfried Bos, Martin Bosen und Carola Gröhlich, 113–122. Münster: Waxmann.
- Bosen, Martin, Wilfried Bos, Carola Gröhlich und Anna Rau. 2009b. Ziele der Untersuchung KESS 7. In *KESS 7 – Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen zu Beginn der Jahrgangsstufe 7*, Hrsg. Wilfried Bos, Martin Bosen und Carola Gröhlich, 13–21. Münster: Waxmann.

- Bos, Wilfried, und Katja Scharenberg. 2010. Lernentwicklung in leistungshomogenen und -heterogenen Schulklassen. In *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert*, Hrsg. Wilfried Bos, Eckhard Klieme und Olaf Köller, 173–194. Münster: Waxmann.
- Bos, Wilfried, Uta Brose, Svenja Bundt, Carola Gröhlich, Nina Hugk, Nike Janke, Peter May, Marcus Pietsch, Tobias C. Stubbe und Andreas Voss. 2006. Anlage und Durchführung der Studie „Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 4 (KESS 4)“. In *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Schulen*, Hrsg. Wilfried Bos und Marcus Pietsch, 9–32. Münster: Waxmann.
- Bos, Wilfried, Martin Bensen, Carola Gröhlich, Karin Guill und Katja Scharenberg. 2009. *KESS 7 – Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Bos, Wilfried, Carola Gröhlich, Karin Guill, Katja Scharenberg und Heike Wendt. 2010a. Ziele und Anlage der Studie KESS 8. In *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8*, Hrsg. Wilfried Bos und Carola Gröhlich, 9–20. Münster: Waxmann.
- Bos, Wilfried, Carola Gröhlich, Denisa-Felicia Dudas, Karin Guill und Katja Scharenberg. 2010b. *KESS 8 – Skalenhandbuch zur Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.
- Dar, Yehezkel, und Nura Resh. 1986. Classroom intellectual composition and academic achievement. A study of the effects of ability-based classes. *American Educational Research Journal* 23:357–374.
- Ditton, Hartmut. 1998. *Mehrebenenanalyse. Grundlagen und Anwendungen des hierarchisch linearen Modells*. Weinheim: Juventa.
- Ditton, Hartmut. 2000. Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der Empirischen Forschung. *Zeitschrift für Pädagogik* 41:73–92.
- Ditton, Hartmut. 2013. Kontexteffekte und Bildungsungleichheit: Mechanismen und Erklärungsmuster. In *Bildungskontexte. Strukturelle Voraussetzungen und Ursachen ungleicher Bildungschancen*, Hrsg. Rolf Becker und Alexander Schulze, 173–206. Wiesbaden: Springer VS.
- Dreeben, Robert, und Rebecca Barr. 1988. Classroom composition and the design of instruction. *Sociology of Education* 61:129–142.
- Dumont, Hanna, Marko Neumann, Kai Maaz und Ulrich Trautwein. 2013. Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen. *Psychologie in Erziehung und Unterricht* 60:163–183.
- Fend, Helmut. 1991. Schule und Persönlichkeit: Eine Bilanz der Konstanzer Forschungen zur „Sozialisation in Bildungsinstitutionen“. In *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung*, Hrsg. Reinhard Pekrun und Helmut Fend, 9–32. Stuttgart: Enke.
- Friedrichs, Jürgen, und Alexandra Nonnenmacher. 2010. Welche Mechanismen erklären Kontexteffekte? In *Komparative empirische Sozialforschung*, Hrsg. Tilo Beckers, Klaus Birkelbach, Jörg Hagenah und Ulrich Rosar, 469–497. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ganzeboom, Harry B.G., und Donald J. Treiman. 1996. Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research* 25:201–239.
- Gröhlich, Carola, Karin Guill, Katja Scharenberg und Wilfried Bos. 2010a. Kumulative Effekte differentieller Lern- und Entwicklungsmilieus innerhalb der Sekundarstufe I am Beispiel der Mathematikleistung. In *Erziehungswissenschaftliche Forschung – nachhaltige Bildung. Beiträge zur 5. DGfE-Sektionstagung „Empirische Bildungsforschung“/AEPF-KBBB im Frühjahr 2009*, Hrsg. Bernd Schwarz, Peter Nenniger und Reinhold S. Jäger, 473–479. Landau: Verlag Empirische Pädagogik.
- Gröhlich, Carola, Karin Guill, Katja Scharenberg und Wilfried Bos. 2010b. Differenzielle Lern- und Entwicklungsmilieus beim Erwerb der Lesekompetenz in den Jahrgangsstufen 7 und 8. In *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8*, Hrsg. Wilfried Bos und Carola Gröhlich, 100–106. Münster: Waxmann.
- Guill, Karin, und Carola Gröhlich. 2013. Individuelle Lernentwicklung im gegliederten Schulsystem der Bundesrepublik Deutschland. Fragen an die Sekundarstufe I. In *Schul- und Bildungsforschung: Diskussionen, Befunde und Perspektiven. Festschrift für Wilfried Bos*, Hrsg. Knut Schwippert, Martin Bensen und Nils Berkemeyer, 51–69. Münster: Waxmann.
- Hallinan, Maureen T. 1994. Further thoughts on tracking. *Sociology of Education* 67:89–91.
- Harker, Richard, und Peter Tymms. 2004. The effects of student composition on school outcomes. *School Effectiveness and School Improvement* 15:177–199.
- Hattie, John A.C. 2002. Classroom composition and peer effects. *International Journal of Educational Research* 37:449–481.
- Heller, Kurt A., und Christoph Perleth. 2000. *KFT 4–12 + R. Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz.

- Helmke, Andreas, und Franz E. Weinert. 1997. Bedingungsfaktoren schulischer Leistungen. In *Psychologie des Unterrichts und der Schule*, Hrsg. Franz E. Weinert, 71–176. Göttingen: Hogrefe.
- Hox, Joop J. 1995. *Applied multilevel analysis*. Amsterdam: TT-Publikaties.
- Ivanov, Stanislav, Roumiana Nikolova und Ulrich Vieluf. 2011. Einführung. In *KESS 10/11. Kompetenzen und Einstellungen von Schülerinnen und Schülern an Hamburger Schulen am Ende der Sekundarstufe I und zu Beginn der gymnasialen Oberstufe*, Hrsg. Ulrich Vieluf, Stanislav Ivanov und Roumiana Nikolova, 9–13. Münster: Waxmann.
- Köller, Olaf. 2004. *Konsequenzen von Leistungsgruppierungen*. Münster: Waxmann.
- Köller, Olaf, und Jürgen Baumert. 2001. Leistungsgruppierungen in der Sekundarstufe I. Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung. *Zeitschrift für Pädagogische Psychologie* 15:99–110.
- Kulik, Chen-Lin C., und James A. Kulik. 1982. Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal* 19:415–428.
- Langer, Wolfgang. 2008. *Mehrebenenanalyse. Eine Einführung für Forschung und Praxis*. 2. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lou, Yiping, Philip C. Abrami, John C. Spence, Catherine Poulsen, Bette Chambers und Sylvia d'Apollonia. 1996. Within-class grouping: A meta-analysis. *Review of Educational Research* 66:423–458.
- Lüttke, Oliver, Alexander Robitzsch und Olaf Köller. 2002. Statistische Artefakte bei Kontexteffekten in der pädagogisch-psychologischen Forschung. *Zeitschrift für Pädagogische Psychologie* 16:217–231.
- Maaz, Kai, Ulrich Trautwein, Oliver Lüttke und Jürgen Baumert. 2008. Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives* 2:99–106.
- Meulemann, Heiner. 2006. *Soziologie von Anfang an*. 2. überarb. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Neumann, Marko, Inge Schnyder, Ulrich Trautwein, Alois Niggli, Oliver Lüttke und Rico Cathomas. 2007. Schulformen als differenzielle Lernmilieus. Institutionelle und kompositionelle Effekte auf die Leistungsentwicklung im Fach Französisch. *Zeitschrift für Erziehungswissenschaft* 10:399–420.
- Nikolova, Roumiana. 2011. *Grundschulen als differenzielle Entwicklungsmilieus. Objektive Kontextmerkmale der Schülerzusammensetzung und deren Auswirkung auf die Mathematik- und Leseleistungen*. Münster: Waxmann.
- Oakes, Jeannie. 1990. *Multiplying inequalities. The effects of race, social class, and tracking on opportunities to learn mathematics and science*. Santa Monica: RAND Corporation.
- Opendakker, Marie-Christine, und Jan van Damme. 2001. Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal* 27:407–432.
- Opendakker, Marie-Christine, und Jan van Damme. 2006. Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools. *School Effectiveness and School Improvement* 17:87–117.
- Organisation for Economic Co-Operation and Development (OECD). 2001. *Lernen für das Leben. Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000. Ausbildung und Kompetenzen*. Paris: OECD.
- Ramseier, Erich, und Christian Brühwiler. 2003. Herkunft, Leistung und Bildungschancen im gegliederten Bildungssystem: Vertiefte PISA-Analyse unter Einbezug der kognitiven Grundfähigkeiten. *Schweizerische Zeitschrift für Bildungswissenschaften* 25:23–58.
- Raudenbush, Stephen W., Anthony S. Bryk und R. Congdon. 2009. *HLM 6.08 for Windows [Computer software]*. Lincolnwood: Scientific Software International, Inc.
- Schafer, Joseph L. 1999. *NORM: Multiple imputation of incomplete multivariate data under a normal model, version 2*. Software for Windows 95/98/NT. <http://www.stat.psu.edu/~jls/misoftwa.html> (Zugegriffen: 01. Okt. 2013).
- Scharenberg, Katja. 2012. *Leistungsheterogenität und Kompetenzentwicklung. Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie*. Münster: Waxmann.
- Scharenberg, Katja. 2013. Heterogenität in der Schule – Definitionen, Forschungsbefunde, Konzeptionen und Perspektiven für die empirische Bildungsforschung. In *Jahrbuch der Schulentwicklung, Bd. 17. Sprachliche, kulturelle und soziale Heterogenität in der Schule als Herausforderung und Chance der Schulentwicklung*, Hrsg. Nele McElvany, Miriam M. Gebauer, Wilfried Bos und Heinz Günter Holtappels, 10–49. Weinheim: Beltz Juventa.

- Scheerens, Jaap. 1990. School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement* 1:61–80.
- Schneider, Wolfgang, und Jan Stefanek. 2004. Entwicklungsveränderungen allgemeiner kognitiver Fähigkeiten und schulbezogener Fertigkeiten im Kindes- und Jugendalter. Evidenz für einen Schereneffekt? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie* 36:147–159.
- Schwetz, Herbert. 2003. *Die Klasse macht den Unterschied. Mehrebenenanalytische Untersuchung der Effekte von Unterricht*. Landau: Verlag Empirische Pädagogik.
- Slavin, Robert E. 1990a. Ability grouping in secondary schools: A response to Hallinan. *Review of Educational Research* 60:505–507.
- Slavin, Robert E. 1990b. Achievement effects of ability grouping in secondary schools: A best-evidence synthesis. *Review of Educational Research* 60:471–499.
- Snijders, Tom A. B., und Roel J. Bosker. 1999. *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. Thousand Oaks: Sage Publications.
- Thrupp, Martin, Hugh Lauder und Tony Robinson. 2002. School composition and peer effects. *International Journal of Educational Research* 37:483–504.
- van Ophuysen, Stefanie, und Heike Wendt. 2009. Zur Veränderung der Mathematikleistung von Klasse 4 bis 6. Welchen Einfluss haben Kompositions- und Unterrichtsmerkmale? *Zeitschrift für Erziehungswissenschaft, Sonderheft* 12:302–327.
- von Davier, Matthias, und Alina A. von Davier. 2007. A unified approach to IRT scale linking and scale transformations. *Methodology* 3:115–124.
- Wentzel, Kathryn R. 1991. Relations between social competence and academic achievement in early adolescence. *Child Development* 62:1066–1078.
- Wentzel, Kathryn R. 1994. Relations of social goal pursuit to social acceptance, classroom behavior, and perceived social support. *Journal of Educational Psychology* 86:173–182.
- Wentzel, Kathryn R. 1999. Social-motivational processes and interpersonal relationships: Implications for understanding motivation at school. *Journal of Educational Psychology* 91:76–97.

**Katja Scharenberg**, 1980, Dr. phil., Senior Researcher an der Universität Basel am Lehrstuhl für Sozialforschung und Methodologie. Forschungsgebiete: Schulleistungs- und Schuleffektivitätsforschung, Bildungsverläufe und Übergänge im Schul- und Ausbildungssystem sowie ins Erwerbsleben, Leistungsgruppierung und -heterogenität, differenzielle Lern- und Entwicklungsmilieus, Belastung und Beanspruchung bei Lehrkräften. Veröffentlichungen: Schulische Belastung und Beanspruchung von Lehrkräften – eine Überprüfung des Erfurter Modells im Rahmen der Hamburger KESS-Studie. In: Schul- und Bildungsforschung: Diskussionen, Befunde und Perspektiven – Festschrift für Wilfried Bos (hrsg. von K. Schwippert, M. Bonsen und N. Berkemeyer), Münster 2013 (mit W. Rollett); Leistungsheterogenität und Kompetenzentwicklung – Zur Relevanz klassenbezogener Kompositionsmerkmale im Rahmen der KESS-Studie, Münster 2012; Lernentwicklung in leistungshomogenen und -heterogenen Schulklassen. In: Schulische Lerngelegenheiten und Kompetenzentwicklung – Festschrift für Jürgen Baumert (hrsg. von W. Bos, E. Klieme und O. Köller), Münster 2010 (mit W. Bos).