

Gröhlich, Carola; Scharenberg, Katja; Bos, Wilfried
**Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen
Lernerfolg in der Sekundarstufe aus?**

Journal for educational research online 1 (2009) 1, S. 86-105



Quellenangabe/ Reference:

Gröhlich, Carola; Scharenberg, Katja; Bos, Wilfried: Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus? - In: Journal for educational research online 1 (2009) 1, S. 86-105 - URN: urn:nbn:de:0111-opus-45578 - DOI: 10.25656/01:4557

<https://nbn-resolving.org/urn:nbn:de:0111-opus-45578>

<https://doi.org/10.25656/01:4557>

in Kooperation mit / in cooperation with:



WAXMANN
www.waxmann.com

<http://www.waxmann.com>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Carola Gröhlich, Katja Scharenberg & Wilfried Bos

Wirkt sich Leistungsheterogenität in Schulklassen auf den individuellen Lernerfolg in der Sekundarstufe aus?

Zusammenfassung

Der Frage nach dem schulischen Umgang mit unterschiedlichen Leistungsvoraussetzungen wird nicht erst in den letzten Jahren verstärkte Aufmerksamkeit geschenkt. In der Diskussion geht es dabei auch darum, ob Schülerinnen und Schüler in leistungshomogenen oder -heterogenen Lerngruppen günstigere Voraussetzungen für ihre Lern- und Leistungsentwicklung vorfinden. Anhand der Hamburger Schulleistungsstudie KESS (Kompetenzen und Einstellungen von Schülerinnen und Schülern) kann der Zusammenhang zwischen der Leistungsstreuung in Schulklassen ($N = 514$) und den Lernerfolgen von Schülerinnen und Schülern ($N = 10\ 036$) in der 6. Jahrgangsstufe geprüft werden. Im Leseverständnis ist an den Haupt- und Realschulen und den Gymnasien die mittlere Streuung auf der Klassenebene deutlich geringer als an den Gesamtschulen. In der Kompetenzdomäne Mathematik ist das Gymnasium die leistungsheterogenste Schulform zu Beginn der Sekundarschulzeit. Mehrebenenanalysen ergeben, dass sich Leistungsheterogenität innerhalb von Lerngruppen nicht – d.h. weder positiv noch negativ – auf den Lernerfolg von Schülerinnen und Schülern auswirkt.

Schlagworte

Heterogenität, Homogenität, Leistungsgruppierung

Does Classroom Heterogeneity in Secondary Schools Affect Student Achievement?

Abstract

How to best accommodate the variety of preconditions that students bring to their learning is a frequently asked question. The study documented in this article sought to further illuminate answers to this question by focusing on whether the conditions for learning and for students' achievement are better and higher in homogeneous or heterogeneous ability groups. Data from the longitudinal study KESS (Competencies and Attitudes of Students) were used to analyze the relationship between heterogeneity at class level ($N = 514$) and student achievement ($N = 10\ 036$). Tracked classes were found to be more homogeneous than classes in comprehensive schools in terms of reading achievement. In relation to mathematics achievement, however, the highest track (Gymnasium) included the most heterogeneous classes. Multilevel analyses showed that achievement heterogeneity in classes had neither positive nor negative effects on individual learning.

Keywords

heterogeneity, homogeneity, ability grouping

1. Einleitung

Die verschiedenen Voraussetzungen von Schülerinnen und Schülern stellen eine der größten Herausforderungen an schulische Lern- und Vermittlungsprozesse dar. Unter dem Stichwort ‚Heterogenität‘ werden sowohl soziale oder kulturelle Unterschiede als auch die divergenten leistungsbezogenen Ausgangsbedingungen der Schülerschaft gefasst. Letztere stehen im Fokus des folgenden Beitrags. Der Umgang der Schule mit unterschiedlichen Leistungsvoraussetzungen sorgt für heftige Kontroversen sowohl unter den Lehrkräften als auch in der Bildungsforschung und Bildungspolitik. Dass die Schule auf die unterschiedlichen Bedingungen und Bedürfnisse der Schülerinnen und Schüler nicht gleichermaßen reagieren kann, ist dabei unstrittig. Fraglich ist jedoch, unter welchen strukturellen Rahmenbedingungen alle Kinder und Jugendlichen in der Schule bestmöglich gefördert werden können, ohne dass dies auf Kosten ihrer Mitschülerinnen und Mitschüler stattfindet. Die zunächst organisatorische Frage, ob Kinder und Jugendliche mit unterschiedlichen Voraussetzungen in Lerngruppen mit unterschiedlichem Anspruch und Lerntempo unterrichtet werden sollten, birgt durchaus pädagogischen Sprengstoff: Die Befürchtung lautet, dass leistungsschwächeren Schülerinnen und Schülern in leistungshomogenen Lerngruppen durch ein beschränktes Curriculum und mangelnde Aufstiegsmöglichkeiten in höhere Bildungsgänge Lern- und Lebenschancen verwehrt würden, während leistungsstärkere Schülerinnen und Schüler in leistungsheterogenen Lerngruppen möglicherweise keine angemessene Förderung ihrer Potenziale erhielten.

1.1 Homogenisierung im deutschen Bildungssystem

Dem deutschen Bildungssystem wird wegen seiner vielfältigen Homogenisierungsstrategien die Verstärkung von gesellschaftlicher Chancenungleichheit vorgeworfen. Eine externe Leistungsdifferenzierung von Lerngruppen, d.h. eine Verminderung der Leistungsbreite und damit eine Homogenisierung der Ausgangsbedingungen in Schulklassen, wird trotz der bestehenden Kritik in den Schulsystemen sehr vieler Staaten praktiziert. Sie wird entweder fachspezifisch (*tracking* bzw. *setting*) oder fächerübergreifend (*streaming*) vorgenommen. Für Deutschland kann die Homogenisierung von Lerngruppen als die „dominierende Strategie“ (Tillmann & Wischer, 2006, S. 45) des Schulsystems bezeichnet werden. Organisatorische Maßnahmen der externen Differenzierung zur Herstellung von möglichst gleichartigen Bedingungen im Klassenzimmer stellen darüber hinaus beispielsweise Zurückstellungen vom Schulbesuch, Klassenwiederholungen, Überweisungen in Sonderschulen oder der Wechsel auf andere Schulformen dar. Als eine bei-

spielhafte Debatte zur Gruppierung von Schülerinnen und Schülern ist hier sicherlich die Diskussion um das dreigliedrige Schulsystem zu nennen. Dieses stellt einen Grundpfeiler der leistungsbezogenen Homogenisierung im deutschen Bildungssystem dar, in dessen verschiedene Schulformen die Schülerinnen und Schüler zu Beginn der Sekundarstufe I überwiesen werden. Dadurch, dass diese jeweils ein bestimmtes Leistungsspektrum abdecken sollen, kann – so der Anspruch – eine gezielte Förderung des Einzelnen bestmöglich gewährleistet werden. Allerdings konnte für das deutsche Schulsystem gezeigt werden, dass die Zuweisung der Schülerinnen und Schüler in die unterschiedlichen Schulformen der Sekundarstufe I nicht allein nach Leistungsmaßstäben erfolgt, sondern erheblich mit Faktoren der sozialen Herkunft konfundiert ist (Arnold, Bos, Richert & Stubbe, 2007; Bos et al., 2004; Pietsch, 2007). Eine Folge dieser Zuweisungsmechanismen ist eine größere Leistungsheterogenität zwischen den Schulformen als eigentlich durch die Homogenisierungsmaßnahme angestrebt. Zwar sind deutsche Schulklassen im internationalen Vergleich die homogensten, allerdings ist aus allen Schulleistungsstudien der letzten Jahre bekannt, dass zwar das mittlere Leistungsniveau von Schülerschaften unterschiedlicher Schulformen deutlich voneinander abweicht, gleichzeitig aber breite Überlappungen in den Leistungsstreuungen zwischen den Schulformen vorhanden sind (Bos et al., 2007; Lehmann, Peek, Gänsfuß & Husfeldt, 2002; Prenzel et al., 2008).

Allerdings gilt es, die Betrachtung auf das deutsche Schulsystem zu differenzieren: Die Schulformgliederung stellt vom Beginn der Sekundarstufe an eine vollständige Aufteilung der Schülerinnen und Schüler in Hauptschul-, Realschul- und Gymnasialzweige dar. Neben diesen haben sich seit den 1970er Jahren integrierte Gesamtschulen (IGS) etabliert, in denen die Schülerinnen und Schüler unabhängig vom Leistungsstand in gemeinsamen Klassenverbänden unterrichtet werden (Köller, 2008). Ab dem Beginn der Jahrgangsstufe 7 wird in den Kernfächern (Deutsch, Mathematik, Englisch) ein leistungsdifferenziertes Kurssystem eingeführt, dieses kann auf andere Fächer ausgeweitet werden. An integrierten Gesamtschulen ist der Erwerb aller Schulabschlüsse möglich – die Entscheidung, welcher Abschluss letztlich erworben wird, kann jedoch deutlich länger als im gegliederten Schulsystem offen gehalten werden. Zudem ist eine Durchlässigkeit innerhalb des Kurssystems der Gesamtschulen eher gegeben als in den getrennten Bildungsgängen des gegliederten Schulsystems, für die nach wie vor eine hohe Haltekraft konstatiert wird (Cortina & Trommer, 2003).

1.2 Vor- und Nachteile von Homogenisierung

Slavin (1987) verweist auf Argumente, die für bzw. gegen – generelle und fachspezifische – Leistungsgruppierung sprechen: Prinzipiell wird angenommen, dass durch Homogenisierungsmaßnahmen die Schülerleistungen gesteigert werden können. Dies wird darauf zurückgeführt, dass der Lernstoff und die Lehrmethoden besser an die jeweiligen Schülervoraussetzungen angepasst werden können. In leis-

tungsstärkeren homogenen Lerngruppen soll eine gesteigerte Leistungserwartung die Schülerinnen und Schüler anspornen. Zudem könne der Lernstoff in solchen Gruppen schneller und tiefgehender bearbeitet werden. In leistungsschwächeren homogenen Lerngruppen hingegen stehe mehr Zeit zum Erklären von noch nicht verstandenen Sachverhalten und mehr Aufmerksamkeit für individuelle Förderung zur Verfügung. Gleichzeitig sei eine psychische Entlastung dadurch gegeben, dass der direkte Vergleich mit den leistungsstärkeren Mitschülerinnen und Mitschülern entfalle.

Geen Leistungsgruppierungen werden vor allem zwei Argumente angeführt, die sich vornehmlich auf die leistungsschwachen Schülerinnen und Schüler beziehen. So wird erstens darauf hingewiesen, dass eine Gegenüberstellung mit leistungsstärkeren Mitschülerinnen und Mitschülern nicht ausschließlich negativ wirken müsse, sondern in heterogenen Lerngruppen auch zu erhöhten Anreizen und damit zu Leistungssteigerungen vor allem bei leistungsschwächeren Schülerschaften führen könne. In leistungshomogenen schwachen Lerngruppen entfalle hingegen diese Vergleichsmöglichkeit ‚nach oben‘, so dass möglicherweise ein insgesamt niedrigerer Leistungsanspruch und zwar seitens der Lehrkräfte wie der Schülerschaft selbst entstehe. Zweitens wird auf eine mögliche stigmatisierende Wirkung der Zugehörigkeit zu leistungsschwächeren Lerngruppen hingewiesen, die mit negativen psychosozialen Folgen für Kinder und Jugendliche einhergehen könne.

In aktuellen Debatten zur Heterogenität in der Schule und im Unterricht wird zudem immer wieder darauf hingewiesen, dass diese nicht ausschließlich negativ als Belastung, sondern vielmehr als reale Bedingung des Zusammenlebens und damit als Chance für das Erlernen des Umgangs mit gesellschaftlicher Vielfalt und als eine Bereicherung für den Unterricht begriffen werden solle (Carle, 2005; Jennessen, Kastirke & Uhrlau, 2005; Prediger, 2003). Um dieser Herausforderung zu begegnen, sei jedoch ein anderer Unterricht und eine verbesserte Lehrerbildung als bisher vonnöten (Tillmann & Wischer, 2006). Schülerinnen und Schülern mit verschiedenen Ausgangsbedingungen innerhalb einer Lerngruppe müsse auch mit unterschiedlichen und individuell angemessenen pädagogischen Angeboten begegnet werden, so lautet der Anspruch der inneren Differenzierung des Unterrichts in leistungsheterogenen Lerngruppen (z. B. Weinert, 1997).

Die Debatte um Heterogenität in der Schule ist also durchaus als kontrovers zu bezeichnen. Einerseits wird diese Vielfalt als Chance und Herausforderung in der heutigen diversifizierten Gesellschaft verstanden. Dem steht andererseits die schulische Praxis der Lerngruppenhomogenisierung entgegen. Letztlich geht es in dieser Diskussion um nichts Geringeres als um die Verwirklichung des Bildungsauftrags der Schule, nämlich um den Anspruch einer *optimalen* Förderung *aller* Schülerinnen und Schüler (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland, 2005) und damit auch um die Frage, unter welchen organisatorischen Voraussetzungen dieses Ziel gewährleistet werden kann.

2. Forschungsstand

Es gibt zahlreiche Anhaltspunkte dafür, dass die leistungsbezogene Zusammensetzung der Lerngruppe im Zusammenhang mit dem Schulleistungserwerb steht. Solche Kompositionseffekte sind für das mittlere Leistungsniveau von Schulen oder Schulklassen international vielfach belegt (zusammenfassend Hattie, 2002; Thrupp, Lauder & Robinson, 2002). Für den deutschen Kontext konnten neben diesem kompositionellen Zusammenhang zusätzlich institutionelle Effekte der Schulformgliederung nachgewiesen werden (Baumert, Stanat & Watermann, 2006; Bos et al., 2007). Analysen, die neben dem mittleren Niveau von Intelligenz- oder Kompetenzindikatoren zusätzlich die Streuung dieser Merkmale als Kompositionsmerkmale der Lerngruppen untersuchen, sind bisher für den bundesdeutschen Kontext rar. Lediglich eine Arbeit von Lehmann (2006) greift diese Fragestellung auf. Die Analyse zeigt einen bemerkenswert hohen positiven Effekt der kognitiven Variation von Berliner Grundschulklassen auf die individuellen Leseleistungen am Ende der 4. Jahrgangsstufe. Da die querschnittlichen Befunde allerdings keine Kausalschlüsse zulassen, muss offen bleiben, ob dieser positive Zusammenhang tatsächlich einen Effekt abbildet, der allein durch die Leistungsstreuung innerhalb der Schulklassen zustande kommt. Möglicherweise handelt es sich hier auch um ein Auslesephänomen, insofern als leistungsheterogene Schulklassen auch vornehmlich aus leistungsstärkeren Schülerinnen und Schülern bestehen. Für die Leistungen in Mathematik zeigen die Berliner Analysen keinen bedeutsamen Effekt der Leistungsvariation auf der Klassenebene.

Eine Vielzahl internationaler Studien zeigt, dass durchaus Zweifel an der Annahme der leistungsförderlichen Wirkung von Lerngruppenhomogenisierung angebracht sind. Zudem werden in vielen Untersuchungen differenzielle Effekte berichtet: Befunde aus Meta-Analysen (Kulik & Kulik, 1982, 1984) weisen zwar leicht positive Effekte für leistungshomogene Lerngruppen aus, allerdings zeigen sich diese vor allem für leistungsstärkere Schülerinnen und Schüler. Tendenziell ähnlich fällt eine Untersuchung niederländischer Grundschulen von Luyten und van der Hoeven-van Doornum (1995) aus. Positive Effekte der Leistungshomogenisierung zeigen sich in ihrer Untersuchung vor allem für die leistungsstärkeren Kinder. Ihre leistungsschwächeren Mitschülerinnen und Mitschüler scheinen hingegen durch diese Maßnahme in ihrer Leistungsentwicklung eher gehemmt zu werden. Zudem können sie keine Bestätigung für einen negativen Effekt einer großen Leistungsheterogenität von Schulklassen auf die individuelle Leistungsentwicklung finden. Zu einem vergleichbaren Schluss kommen auch Dar und Resh (1986) und Resh und Dar (1992) für israelische Sekundarschulen. Auch sie können zeigen, dass eine Homogenisierung von Lerngruppen gerade für leistungsschwache Kinder eine Benachteiligung darstellt und dass durch diese Maßnahme Leistungsunterschiede im Vergleich zu leistungsstärkeren Kindern vergrößert werden. Dar und Resh (1986) berichten zudem differenzielle Effekte der Klassenzusammensetzung für die leistungsstarken und -schwachen Schülerinnen und Schüler: Der Leistungsgewinn, den

die leistungsschwachen Kinder aus heterogen zusammengesetzten Klassen ziehen können, überwiegt die Leistungseinbußen, die für die leistungsstarken Kinder in heterogenen Klassen entstehen, um ein Vielfaches. Verglichen mit dem Einfluss individueller Lernvoraussetzungen fallen die Effekte für die Klassenzusammensetzung in dieser Studie allerdings gering aus. Auch Slavin (1987) attestiert der Aufteilung der Schülerschaft nach ihren Leistungen negative Wirkungen. Insbesondere gingen soziale Benachteiligungen mit dieser Maßnahme einher. Von Interesse sind auch die Befunde von Linchevski und Kutscher (1998). Sie vergleichen die Lernentwicklung von Schülerinnen und Schülern in heterogenen und homogenen Lerngruppen in israelischen Schulen und kommen zu dem Schluss, dass Leistungserträge beider Gruppen vergleichbar ausfallen. Daneben können auch sie aus ihren Daten schließen, dass gerade in heterogenen Lerngruppen die leistungsschwächeren Kinder besonders große Lernzuwächse zeigen, während gleichzeitig ihre leistungsstärkeren Mitschülerinnen und Mitschüler nicht in ihren Lernerträgen beeinträchtigt werden.

Allerdings gibt es auch empirische Belege, die weder einen Effekt für unterschiedlich leistungshomogene Lerngruppen noch differenzielle Effekte für unterschiedliche Schülergruppen finden: So können beispielsweise Opdenakker und van Damme (2001) in ihren Analysen keinen bedeutsamen Effekt der kognitiven Heterogenität von Schulen feststellen. Slavin (1990) findet in einer Meta-Analyse kaum bedeutsame Effekte für bzw. gegen Leistungsgruppierung, auch nicht für unterschiedliche Leistungsvoraussetzungen von Schülerinnen und Schülern.

Auch für den bundesdeutschen Kontext liegen durchaus widersprüchliche Befunde zum Effekt von unterschiedlich heterogenen Lerngruppen vor. Diese Untersuchungen berücksichtigen zwar die hierarchische Struktur der Daten nicht in der heute üblichen Weise, sollen aber an dieser Stelle berichtet werden, da sie von Relevanz für die deutsche Debatte sind. Mehrere Untersuchungen konnten zeigen, dass ein Unterricht, der eine Verringerung von Leistungsunterschieden in Schulklassen anstrebt, mit Leistungseinbußen bei leistungsstärkeren Schülerinnen und Schülern einhergeht (Helmke, 1988; Treiber & Weinert, 1985). Allerdings fanden Treinies und Einsiedler (1996) Gegenteiliges: Ihnen zufolge sind Leistungsdiskrepanzen in Schulklassen für leistungsstärkere Schülerinnen und Schüler ohne Belang.

Fasst man diese teils widersprüchlichen empirischen Befunde zusammen, kann man annehmen, dass eine Leistungshomogenisierung vermutlich eher positive Auswirkungen auf die Lernentwicklung der leistungsstarken Schülerschaften hat. Den Preis zahlen hier unter Umständen allerdings die leistungsschwächeren Schülerinnen und Schüler, deren Lernerträge in leistungshomogenen Gruppen oft besonders gering ausfallen. Tillmann und Wischer (2006, S. 44) sprechen in diesem Zusammenhang sogar eine „Warnung vor Homogenisierung am ‚unteren Ende‘ des Schulsystems“ aus. In leistungsheterogenen Lerngruppen hingegen scheint ein Unterricht, der die Verringerung des klasseninternen Leistungsspektrums zum Ziel hat, insbesondere für die leistungsschwächeren Schülerinnen und Schüler von Vorteil zu sein. Für ihre leistungsstarken Klassenkameraden und -kameradin-

nen muss dies kein Nachteil sein, denn für sie ist es offenbar unter bestimmten Bedingungen möglich, das hohe Niveau zu halten.

3. Fragestellungen

Unsere Analysen zielen auf die Lernerfolge von Schülerinnen und Schülern in Lerngruppen mit unterschiedlich großer Leistungsstreuung. Daher wird zunächst das Ausmaß der Leistungsvariation innerhalb von Schulklassen zu Beginn der Sekundarschulzeit untersucht. In diesem Zusammenhang kann geprüft werden, ob die intendierte externe Leistungsdifferenzierung anhand der Zuweisung der Schülerschaft in die unterschiedlichen Schulformen der Sekundarstufe tatsächlich vollzogen wird. Für die Schulformen des gegliederten Schulsystems müsste entsprechend dem Anspruch der Aufteilung eine deutliche Leistungshomogenisierung innerhalb der Schulklassen – vor allem im Vergleich zu den integrierten Gesamtschulen – zu erwarten sein.

In einem zweiten Schritt wird untersucht, ob sich Unterschiede in der Leistungsvariation von Schulklassen für die Lernerfolge der Schülerinnen und Schüler am Ende der 6. Jahrgangsstufe geltend machen. Dementsprechend wird geprüft, ob die Leistungsvariation als ein Kompositionsmerkmal der Schulklassen neben ihrem Leistungsniveau einen eigenständigen Effekt über individuelle und schulformspezifische Effekte hinaus auf den Leistungserwerb ausübt. Ausgehend von den bisherigen Forschungsbefunden wird auch für Hamburger Schulklassen erwartet, dass eine große Leistungsstreuung in Lerngruppen keinen negativen Effekt auf die individuellen Leistungen von Schülerinnen und Schülern ausübt.

4. Datengrundlage und Methode

Die hier vorgestellten Analysen beruhen auf Daten, die im Rahmen der Schulleistungsstudie KESS (Kompetenzen und Einstellungen von Schülerinnen und Schülern) erhoben wurden (Bos et al., 2007; Bos & Pietsch, 2006). Mit KESS kann die Leistungsentwicklung einer gesamten Schülerkohorte der Freien und Hansestadt Hamburg nachgezeichnet werden. Zu drei Messzeitpunkten (am Ende der Jahrgangsstufe 4, zu Beginn der Jahrgangsstufe 7 und am Ende der Jahrgangsstufe 8) wurden die Lernstände der Schülerinnen und Schüler in den Fächern Deutsch, Englisch, Mathematik und Naturwissenschaften erhoben. Gleichzeitig wurden auch Eltern, Lehrkräfte und Schulleitungen befragt, um die Rahmenbedingungen des schulischen Lernens zu erfassen. Für die hier vorgenommenen Analysen werden die Daten der ersten beiden Messzeitpunkte verwendet. Dieser Zeitraum umfasst die Jahrgangsstufen 5 und 6, für die im gegliederten Schulsystem eine Beobachtungsstufe vorgesehen ist, die den Übergang am Ende der Grundschule entschärfen soll. Die Daten wurden am Ende der Jahrgangsstufe 4

(KESS 4) sowie direkt nach dem Übergang der Schülerinnen und Schüler in die Jahrgangsstufe 7 (KESS 7) erhoben. Die Werte aus KESS 4 bilden somit in den folgenden Analysen die Lernausgangslagen zum Beginn der 5. Jahrgangsstufe, die Werte aus KESS 7 die Lernstände der Schülerschaft zum Ende der 6. Jahrgangsstufe nach einer zweijährigen Beschulung ab. Für die Schülerinnen und Schüler liegen Angaben zur Schulklassenzugehörigkeit in der 6. Jahrgangsstufe vor, so dass mit den hier durchgeführten Analysen das schulische Lernen in den Jahrgangsstufen 5 und 6 nachgezeichnet werden kann. Insgesamt haben 11 398 Kinder an diesen beiden Messzeitpunkten teilgenommen. Davon werden Schülerinnen und Schüler ausgeschlossen, für die die Schulklassenzugehörigkeit in den Jahrgangsstufen 5 bzw. 6 nicht rekonstruiert werden kann oder die in einer dieser beiden Jahrgangsstufen eine Förderschule besucht haben. Zusätzlich werden Schulklassen ausgeschlossen, für die weniger als zehn Kinder zugeordnet werden können, so dass davon ausgegangen werden kann, dass die aggregierten Kompositionsmerkmale ein realistisches Abbild der jeweiligen Lerngruppen darstellen. Für die Analysen stehen somit Angaben von 10 036 Kindern in 514 Schulklassen zur Verfügung. Empirisch fundierte Aussagen sind nur mit vollständigen Datensätzen sinnvoll. Für die Auswertungen wurden daher einzelne fehlende Werte mittels der Software NORM einfach geschätzt (Schafer, 1997). Die Schätzung fehlender Werte beruht auf den Angaben der Schülerinnen und Schüler zu ihrem kulturellen, sozialen und ökonomischen Kapital sowie ihren schulbezogenen Einstellungen.

Die zu erklärenden Leistungen der Schülerinnen und Schüler werden durch die Kompetenzwerte am Ende der 6. Jahrgangsstufe operationalisiert. Diese wurden für das Leseverständnis und für Mathematik anhand einer eindimensionalen Raschskalierung ermittelt (Bos et al., 2007) und für die hier vorgenommenen Analysen auf der Individualebene z-standardisiert. Zur Erklärung der Lernstände in den beiden Kompetenzdomänen Leseverständnis und Mathematik werden Variablen herangezogen, die sich in verschiedenen nationalen und internationalen Schulleistungsstudien als zentrale exogene Faktoren für die Vorhersage schulischer Leistungen bewährt haben (z. B. Bos et al., 2003). Darüber hinaus kann mit den Leistungswerten aus KESS 4 auch das Vorwissen der Schülerinnen und Schüler am Ende der Grundschulzeit berücksichtigt werden. Zusätzlich zu den Leistungsindikatoren wird die in KESS 4 durch den kognitiven Fähigkeitstest (KFT) erfasste kognitive Grundfähigkeit der Schülerinnen und Schüler berücksichtigt (Heller & Perleth, 2000). Die soziale Lage der Schülerinnen und Schülern wird anhand des höchsten *International Socioeconomic Index* (HISEI) in der Familie dargestellt. Dieser Indikator wurde im Elternfragebogen erhoben. Die Prädiktoren werden innerhalb der Gesamtstichprobe auf der Individualebene z-standardisiert. Zusätzlich fließt das Geschlecht (Referenz: Mädchen, durch die Schülerteilnahmeliste erhoben) und der über die Sprache im Haushalt operationalisierte Migrationshintergrund (Referenz: nicht deutsch, mittels Schülerfragebogen erhoben) in die Analysen ein. Auf der Klassenebene werden das durchschnittliche Vorwissen, der durchschnittliche HISEI und die durchschnittlichen kognitiven Fähigkeiten als Mittelwerte der individuellen Schülermerkmale gebildet. Die Variation des Vorwissens und der

kognitiven Fähigkeiten der Lerngruppen zum Beginn der Jahrgangsstufe 5 wird über die Streuung der Merkmale innerhalb der einzelnen Schulklassen abgebildet. Zusätzlich werden die Angaben zur Schulform auf der Klassenebene modelliert. Referenzkategorie sind die in der Hamburger Beobachtungsstufe zu einer Schulform zusammengefassten Haupt- und Realschulen. Die mehr Ebenenanalytischen Auswertungen erfolgen mit der Software HLM 6.06 (Raudenbush, Bryk & Congdon, 2004). Allen Analysen wird ein Signifikanzniveau von $p < .05$ zugrunde gelegt.

5. Befunde

5.1 Leistungsvariation in Hamburger Schulklassen

Tabelle 1 beschreibt neben der mittleren Leistung der Schulklassen auch deren mittlere Leistungsstreuung sowie den Abstand zwischen den leistungsstärksten und -schwächsten Schülerinnen und Schülern für die beiden Kompetenzdomänen und die kognitive Grundfähigkeit zu Beginn der Jahrgangsstufe 5. Diese Werte stellen somit die Ausgangslage der Lerngruppen für die beiden folgenden Schuljahre bis zum Ende der 6. Jahrgangsstufe dar. Die Angaben werden getrennt nach Schulform ausgewiesen. In der linken Tabellenspalte ist die mittlere Leistung der Klassen innerhalb der jeweiligen Schulform sowie der Standardfehler abgetragen. Die mittlere Spalte gibt die durchschnittliche Leistungsstreuung der Klassen an. Die rechte Spalte zeigt, wie groß der durchschnittliche Abstand zwischen den leistungsstärksten und -schwächsten Schülerinnen und Schülern innerhalb der Klassen der jeweiligen Schulform ausfällt. Für die mittlere Streuung auf der Klassenebene und den mittleren Range wurde zudem anhand einfaktorieller Varianzanalysen überprüft, ob sich Klassen unterschiedlicher Schulformen bedeutsam voneinander unterscheiden. Unterscheiden sich die Klassen der unterschiedlichen Schulformen in ihrer mittleren Streuung und im mittleren Range in den drei Testdomänen mit $p < .05$ signifikant voneinander, sind sie durch unterschiedliche Buchstaben (a, b, c) gekennzeichnet. Gleiche Buchstaben zwischen den Schulformen innerhalb der drei Testdomänen bedeuten dementsprechend jeweils, dass sich deren Klassen in der mittleren Streuung und im mittleren Range nicht bedeutsam voneinander unterscheiden.

Die durchschnittlichen Leistungen der Klassen unterscheiden sich zwischen den Schulformen signifikant voneinander. Die höchsten Lernstände werden in den Gymnasialklassen erreicht, die niedrigsten in den Haupt- und Realschulklassen. In den Gesamtschulklassen werden zwar signifikant höhere Lernstände erreicht als an den Haupt- und Realschulen, die Lernstände verbleiben hier dennoch unterhalb des Stichprobendurchschnitts. Dieses Befundmuster gilt für beide Kompetenzdomänen und die kognitive Grundfähigkeit.

Für das Leseverständnis zeigt sich, dass die Differenzen innerhalb der Schulklassen an den Haupt- und Realschulen sowie an den Gymnasien rund drei

Viertel einer Standardabweichung ausmachen. Die mittlere Leistungsstreuung an diesen beiden Schulformen fällt vergleichbar aus ($SD = 0,78$ bzw. $SD = 0,75$). Signifikant höher ist dagegen die mittlere Leistungsstreuung in den Gesamtschulklassen ($SD = 0,84$). An den Haupt- und Realschulen sowie den Gymnasien sind also homogenere Klassen als an den Gesamtschulen anzutreffen. Entsprechend fällt der mittlere Range aus: An den Haupt- und Realschulen sowie an den Gymnasien ist der Abstand zwischen den leistungsstärksten und -schwächsten Schülerinnen vergleichbar ($R = 2,95$ bzw. $R = 2,97$). An den Gesamtschulen ist der durchschnittliche Range dagegen signifikant größer ($R = 3,17$).

Tabelle 1: Mittelwerte und Dispersionsmaße im Leseverständnis und in Mathematik sowie der kognitiven Grundfähigkeit auf der Klassenebene nach Schulform zu Beginn der Jahrgangsstufe 5

	Mittelwert (SE)	mittlere Streuung ¹	mittlerer Range ^{1,2} innerhalb der Schulklassen
Leseverständnis			
Haupt- und Realschule ^a	-0.63 (0.02)	0.78	2.95
Gesamtschule ^b	-0.36 (0.03)	0.84	3.17
Gymnasium ^a	0.60 (0.02)	0.75	2.97
Mathematik			
Haupt- und Realschule ^a	-0.58 (0.02)	0.73	2.74
Gesamtschule ^b	-0.38 (0.03)	0.78	2.99
Gymnasium ^c	0.58 (0.02)	0.84	3.36
Kognitive Grundfähigkeit			
Haupt- und Realschule ^a	-0.47 (0.03)	0.96	2.92
Gesamtschule ^a	-0.25 (0.02)	0.97	2.94
Gymnasium ^b	0.42 (0.01)	0.79	2.69

¹ Unterscheiden sich die Klassen der unterschiedlichen Schulformen in ihrer mittleren Streuung und im mittleren Range in den drei Testdomänen mit $p < .05$ signifikant voneinander, sind sie durch unterschiedliche Buchstaben (a, b, c) gekennzeichnet.

² Differenz zwischen den leistungsstärksten und -schwächsten Schülerinnen und Schülern auf der Klassenebene nach Schulform.

$N = 135$ (Haupt- und Realschule), 169 (Gesamtschule), 210 (Gymnasium).

Leistungsindikatoren und kognitive Grundfähigkeit auf der Individualebene z-standardisiert.

Hinsichtlich der mittleren Leistungsstreuung in Mathematik unterscheiden sich Klassen aller Schulformen signifikant voneinander. In den Haupt- und Realschulklassen beträgt die mittlere Streuung etwa drei Viertel einer Standardabweichung ($SD = 0,73$). Für Gesamtschulklassen ($SD = 0,78$) fällt die durchschnittliche Streuung der Mathematikleistungen im Vergleich signifikant höher aus als in Haupt- und Realschulen. Die Gymnasialklassen weichen mit $SD = 0,84$ nochmals signifikant nach oben von den Gesamtschulklassen ab. Anders als beim Leseverständnis sind die Gymnasialklassen in Mathematik also im Schnitt am Heterogensten. Dies zeigt sich auch am mittleren Range, der an den Haupt- und

Realschulen ($R = 2,74$) und den Gesamtschulen ($R = 2,99$) deutlich geringer ausfällt als an den Gymnasien ($R = 3,36$).

Eine Abgrenzung der Gymnasialklassen von Klassen der anderen beiden Schulformen zeigt sich auch – allerdings in gegenläufiger Tendenz – bei der Streuung der kognitiven Grundfähigkeit zu Beginn der Jahrgangsstufe 5. An den Gymnasien ist die Schülerschaft diesbezüglich am Homogensten ($SD = 0,79$). An den anderen beiden Schulformen fällt die mittlere klasseninterne Streuung dagegen signifikant höher aus ($SD = 0,96$ bzw. $SD = 0,97$). Dies spiegelt sich auch im durchschnittlichen Abstand zwischen den Schülerinnen und Schülern mit eher hoher bzw. eher niedriger kognitiver Grundfähigkeit wider, der an den Haupt- und Realschulen sowie den Gesamtschulen knapp drei Standardabweichungen ausmacht und an den Gymnasien etwas geringer ausfällt.

5.2 Individueller Lernerfolg und Leistungsvariation in Schulklassen

Will man Kompositionseffekte von Lerngruppen auf den individuellen Lernerfolg modellieren, so ist bei der Datenauswertung ein mehr Ebenenanalytisches Vorgehen erforderlich. Zur Vorhersage der Lese- und Mathematikleistung am Ende der Jahrgangsstufe 6 wird ein Modell auf zwei Ebenen (Schülerinnen und Schüler in Klassen) gewählt. Hierbei wird zunächst für beide Kompetenzdomänen überprüft, wie viel Varianz auf Unterschiede zwischen den Schülerinnen und Schülern bzw. zwischen einzelnen Klassen entfällt. Beim Leseverständnis entfallen 34.6% der Varianz auf die Klassenebene. Für die Mathematikleistung fällt die Variabilität zwischen den Klassen mit 44.0% sogar noch höher aus als für das Leseverständnis. Im Folgenden werden die Ergebnisse der Mehrebenenanalysen beschrieben. Dabei wird zunächst auf das Leseverständnis und anschließend auf die mathematischen Kompetenzen eingegangen. Auf der Individualebene werden in allen Modellen für beide Kompetenzdomänen die gleichen Prädiktoren eingeführt. Einbezogen werden das Geschlecht, der Migrationshintergrund, die kognitive Grundfähigkeit, die soziale Lage und das Vorwissen der Schülerinnen und Schüler. Da sich mit dem vorliegenden Datensatz die Möglichkeit ergibt, die Befunde von Lehmann (2006) auch für Hamburger Schülerinnen und Schüler zu prüfen, wird in Modell 1 zunächst der Effekt der kognitiven Grundfähigkeit (Mittelwert und Streuung der Schulklassen) auf die individuelle Lese- bzw. Mathematikleistung geschätzt. Den Effekt der Kompositionsmerkmale Leistungsniveau und -streuung stellt Modell 2 dar. Modell 3 prüft schließlich, wie sich die kognitive Grundfähigkeit und die Leistungsindikatoren in ihrer gemeinsamen Wirkungsweise – gerade in Bezug auf die Streuung innerhalb der Schulklassen – darstellen.

Tabelle 2: Vorhersage der individuellen Leseleistung am Ende der Jahrgangsstufe 6

	Modell 1	Modell 2	Modell 3
Konstante	-0.34 (0.06)	-0.27 (0.05)	-0.35 (0.07)
<i>Individualebene</i>			
Geschlecht ¹	-0.09 (0.01)	-0.09 (0.01)	-0.09 (0.01)
Sprache im Haushalt ²	0.09 (0.02)	0.09 (0.02)	0.09 (0.02)
Kognitive Grundfähigkeit ³	0.10 (0.01)	0.10 (0.01)	0.10 (0.01)
Soziale Lage (HISEI) ³	0.04 (0.01)	0.04 (0.01)	0.04 (0.01)
Vorwissen ³	0.50 (0.01)	0.49 (0.01)	0.49 (0.01)
R ² innerhalb der Klassen	0.327	0.327	0.327
<i>Klassenebene</i>			
Mittlere kognitive Grundfähigkeit	0.10 (0.04)		0.03 (0.04)
Streuung kognitive Grundfähigkeit	0.13 (0.06)		0.10 (0.06)
Mittleres Vorwissen		0.15 (0.02)	0.14 (0.04)
Streuung Vorwissen		0.10 (0.05)	0.09 (0.06)
Mittlere soziale Lage (HISEI)	0.07 (0.03)	0.01 (0.03)	0.01 (0.03)
<i>Schulform⁴</i>			
Gesamtschule	0.09 (0.02)	0.08 (0.02)	0.08 (0.02)
Gymnasium	0.36 (0.02)	0.30 (0.04)	0.31 (0.04)
R ² zwischen den Klassen	0.902	0.910	0.907

¹ Referenzkategorie: Mädchen.

² Referenzkategorie: nicht deutsch.

³ Prädiktoren auf der Individualebene z-standardisiert.

⁴ Referenzkategorie: Haupt- und Realschule.

Signifikante Parameter sind fettgedruckt ($p < .05$), Standardfehler in Klammern.

Zur Vorhersage des Leseverständnisses am Ende der 6. Jahrgangsstufe werden in das Modell 1 auf der Klassenebene die mittlere soziale Lage sowie das Niveau und die Variation der kognitiven Grundfähigkeit der Schülerschaft aufgenommen (vgl. Tabelle 2). Um der leistungsbezogenen Stratifizierung des deutschen Schulsystems gerecht zu werden, wird außerdem die Schulform berücksichtigt. Für die Hamburger Schülerinnen und Schüler lässt sich auf der Individualebene die Bedeutung der kognitiven Grundfähigkeit und insbesondere des Vorwissens herausstellen. Erwartungsgemäß weisen die Mädchen einen leichten Vorsprung gegenüber den Jungen auf. Unter Berücksichtigung des Vorwissens fällt der mittlerweile vielfach empirisch nachgewiesene Zusammenhang zwischen sozialer Lage und Leseleistung für die Hamburger Schülerinnen und Schüler hier noch signifikant, aber nur sehr gering aus. Ebenfalls klein fällt der Zusammenhang für die im Haushalt der Schülerfamilien gesprochene Sprache aus. Auf der Klassenebene übertreffen die Institutionseffekte alle anderen Merkmale deutlich: Schülerinnen und Schüler an den Gymnasien schneiden bei ansonsten gleichen Voraussetzungen auf der Individual- und der Klassenebene im Leseverständnis am Ende der Jahrgangsstufe 6 deutlich besser ab als Schülerinnen und Schüler an den Haupt- und Realschulen. Für die Gesamtschulen fällt der Effekt im Vergleich zur Referenzgruppe hingegen sehr viel geringer aus. Auch für die mittlere soziale Lage fällt der Effekt signifi-

kant, wenngleich jedoch relativ klein, aus. Des Weiteren lässt sich ein eigenständiger Einfluss der kognitiven Variation von Schulklassen auf die Leseleistung ihrer Schülerinnen und Schüler beobachten: Je größer die Streuung der kognitiven Grundfähigkeit innerhalb der Schulklassen zu Beginn der Jahrgangsstufe 5 ist, desto besser schneiden die Schülerinnen und Schüler im Leseverständnis am Ende der 6. Jahrgangsstufe ab. Die mittlere kognitive Grundfähigkeit auf der Klassenebene weist einen ähnlich hohen Zusammenhang mit der individuellen Leseleistung auf.

Im Rahmen von KESS bietet sich nicht nur die Möglichkeit, die Wirkungsweise der kognitiven Heterogenität, sondern auch die Wirkungsweise der Leistungsheterogenität auf die Schülerleistungen zu prüfen. Modell 2 berücksichtigt deshalb anstelle der kognitiven Grundfähigkeit das Niveau und die Streuung des Vorwissens auf der Klassenebene. Für das mittlere Vorwissen ergibt sich ein positiver Zusammenhang mit der Leseleistung. Der leichte Effekt der Leistungsvariation von Schulklassen auf die Leistungen ihrer Schülerinnen und Schüler fällt hier nicht signifikant aus, verfehlt aber mit $p = .065$ nur knapp das Signifikanzniveau. Der schwache Effekt für die mittlere soziale Lage der Schulklassen, der für das Modell 1 noch nachgewiesen werden konnte, nivelliert sich hier. Die institutionellen Effekte fallen im Vergleich zu Modell 1 etwas geringer, aber noch immer sehr viel günstiger für die Gymnasien aus.

In Modell 3 werden sowohl die kognitiven Grundfähigkeit als auch das Vorwissen auf der Klassenebene eingeführt, um zu prüfen, inwieweit sich für diese Faktoren eigenständige Effekte nachweisen lassen. Der stärkste Prädiktor auf der Klassenebene ist hier ebenfalls die Schulform. Auch in diesem Modell geht ein höheres Vorwissen innerhalb von Schulklassen zu Beginn der Jahrgangsstufe 5 mit höheren Lernständen ihrer Schülerinnen und Schüler im Leseverständnis am Ende der 6. Jahrgangsstufe einher. Für die Variation des Vorwissens und der kognitiven Grundfähigkeit innerhalb der Schulklassen ergeben sich zwar ebenfalls nominell leicht positive, aber keine signifikanten Effekte.

Alle drei Modelle zeichnen sich insgesamt durch eine hohe Modellgüte aus. Auf der Individualebene kann jeweils knapp ein Drittel der Varianz der Lesekompetenz durch die hier berücksichtigten Merkmale erklärt werden. Auf der Klassenebene erklären die Prädiktoren mehr als 90 Prozent der Varianz.

Auch in der Kompetenzdomäne Mathematik ist das Vorwissen in Modell 1 erwartungsgemäß der stärkste Prädiktor auf der Individualebene, wenngleich der Effekt geringfügig schwächer ausfällt als bei der Leseleistung (vgl. Tabelle 3). Der Effekt für die kognitive Grundfähigkeit fällt ähnlich stark wie beim Leseverständnis ins Gewicht. Erwartungsgemäß schneiden die Jungen in der Kompetenzdomäne Mathematik auch unter Kontrolle weiterer Variablen besser ab als die Mädchen. Die geschlechtsspezifischen Unterschiede fallen hier etwas stärker aus als beim Leseverständnis. Für die soziale Lage und die Sprache im Haushalt lassen sich zwar signifikante, aber nur relativ geringe Effekte nachweisen. Auf der Klassenebene sind wiederum die Institutionseffekte am stärksten. Die Gymnasiastinnen und Gymnasiasten weisen bei gleichen individuellen Voraussetzungen und vergleich-

barer Eingangsselektivität der Schulklassen am Ende der Jahrgangsstufe 6 um eine halbe Standardabweichung höhere Mathematikleistungen auf als Schülerinnen und Schüler der Haupt- und Realschulen. Der Vorsprung der Gymnasiastinnen und Gymnasiasten fällt für Mathematik damit noch größer aus als für das Leseverständnis. Der Effekt der Streuung der kognitiven Grundfähigkeit innerhalb der Schulklassen auf die individuelle Mathematikleistung fällt zwar positiv, aber schwächer aus als für das Leseverständnis und lässt sich nicht mehr zufallskritisch absichern. Für die durchschnittlichen kognitiven Fähigkeiten von Schulklassen lässt sich ein signifikanter – allerdings leichter – Zusammenhang hingegen bestätigen. Ohne den Einbezug des Vorwissens zeigt sich auch für die mittlere soziale Lage (HISEI) ein signifikanter Effekt auf der Ebene der Schulklassen.

In Modell 2 wird anstelle der kognitiven Grundfähigkeit das Vorwissen auf der Klassenebene zur Vorhersage der individuellen Mathematikleistung herangezogen. Der Effekt für die mittlere soziale Lage fällt in diesem Modell nicht mehr signifikant aus. Für das mittlere Vorwissen lässt sich ein schwacher positiver Effekt ermitteln. Die Streuung des Vorwissens der Schulklassen erweist sich hingegen als bedeutungslos für die individuellen Mathematikleistungen am Ende der Jahrgangsstufe 6. Die institutionellen Effekte verändern sich im Vergleich zum Modell 1 nur marginal.

Modell 3 berücksichtigt auf der Klassenebene nun sowohl die kognitive Grundfähigkeit als auch das Vorwissen. Neben dem hohen Effekt der Schulform steht einzig das mittlere Vorwissen in einem signifikanten Zusammenhang mit der Mathematikleistung. Eine höhere Variation der kognitiven Grundfähigkeit in Schulklassen geht nominell zwar mit höheren Lernständen der Schülerinnen und Schüler einher, dieser Effekt fällt jedoch nicht signifikant aus. Für die Streuung des Vorwissens lässt sich ein solcher Effekt nicht beobachten.

Auch für die Mathematikleistung wird in allen Modellen mit über 93 Prozent eine hohe Varianzaufklärung auf der Klassenebene erreicht. Auf der Individualebene kann etwa ein Drittel der Varianz durch die hier berücksichtigten Prädiktoren aufgeklärt werden.

Tabelle 3: Vorhersage der individuellen Mathematikleistung am Ende der Jahrgangsstufe 6

	Modell 1	Modell 2	Modell 3
Konstante	-0.42 (0.07)	-0.34 (0.05)	-0.41 (0.08)
<i>Individualebene</i>			
Geschlecht ¹	0.11 (0.01)	0.11 (0,01)	0.11 (0.01)
Sprache im Haushalt ²	0.05 (0.02)	0.05 (0,02)	0.05 (0.02)
Kognitive Grundfähigkeit ³	0.14 (0.01)	0.14 (0,01)	0.14 (0.01)
Soziale Lage (HISEI) ³	0.05 (0.01)	0.05 (0,01)	0.05 (0.01)
Vorwissen ³	0.42 (0.01)	0.41 (0,01)	0.41 (0.01)
R ² innerhalb der Klassen	0.331	0.331	0.331
<i>Klassenebene</i>			
Mittlere kognitive Grundfähigkeit	0.08 (0.04)		0.04 (0.04)
Streuung kognitive Grundfähigkeit	0.09 (0.07)		0.09 (0.07)
Mittleres Vorwissen		0.11 (0,04)	0.10 (0.04)
Streuung Vorwissen		0.02 (0,06)	0.00 (0.06)
Mittlere soziale Lage (HISEI)	0.10 (0.03)	0.06 (0,04)	0.06 (0.04)
<i>Schulform⁴</i>			
Gesamtschule	0.05 (0.02)	0.06 (0,02)	0.05 (0.02)
Gymnasium	0.51 (0.04)	0.48 (0,04)	0.48 (0.04)
R ² zwischen den Klassen	0.937	0.935	0.935

¹ Referenzkategorie: Mädchen.

² Referenzkategorie: nicht deutsch.

³ Prädiktoren auf der Individualebene z-standardisiert.

⁴ Referenzkategorie: Haupt- und Realschule.

Signifikante Parameter sind fettgedruckt ($p < .05$), Standardfehler in Klammern.

6. Diskussion und Ausblick

Die Analysen zeigen, dass das Leistungsspektrum in Hamburger Schulklassen zu Beginn der 5. Jahrgangsstufe im Leseverständnis und in Mathematik sehr unterschiedlich ausfällt: Im Leseverständnis ist an den Haupt- und Realschulen und den Gymnasien die mittlere Streuung auf der Klassenebene deutlich geringer als an den Gesamtschulen. In der Kompetenzdomäne Mathematik zeichnen sich die Haupt- und Realschulklassen im Durchschnitt ebenfalls durch die geringste Streuung, die Gymnasialklassen jedoch durch die größte Streuung aus. Die Leistungsstreuung fällt demnach für das Leseverständnis erwartungsgemäß im Sinne der intendierten Homogenisierung des gegliederten Schulsystems aus. Für die Ausgangslage der Schulklassen zu Beginn der Jahrgangsstufe 5 in Mathematik fallen die gefundenen Differenzen jedoch erwartungswidrig aus. Bemerkenswerterweise ist das Gymnasium die leistungsheterogenste Schulform zu Beginn der Sekundarschulzeit. Eventuell ist dies darauf zurückzuführen, dass sich die Leistungseinschätzungen der Grundschullehrkräfte und damit auch die Übergangsempfehlungen eher an den verbalen Fähigkeiten orientieren. Zwar findet in Hamburg der Übergang von der

Grund- in die Sekundarschule anhand des Elternwunsches statt, allerdings ist bekannt, dass ein Großteil der Eltern sich an der Grundschulempfehlung orientiert.

Die bereits bekannten differenziellen Fördereffekte der verschiedenen Schulformen können auch in den hier vorgenommenen Analysen bestätigt werden. Die Schülerinnen und Schüler, die ein Gymnasium besuchen, schneiden – unter Kontrolle individueller und kompositioneller Merkmale – bereits nach dem zweijährigen Besuch der Sekundarschule deutlich besser ab als Schülerinnen und Schüler an Haupt- und Realschulen oder an Gesamtschulen. Ebenfalls bestätigt werden die bekannten Effekte der Prädiktoren auf der Individualebene, insbesondere der hohe Effekt des Vorwissens. Dieses spielt eine – wenn nicht gar *die* – entscheidende Rolle für schulische Lernerfolge. Es ist davon auszugehen, dass das Niveau des Vorwissens eines der bestimmenden Kompositionsmerkmale ist, das wiederum eine entscheidende Bedingung für das Entstehen differenzieller Lernmilieus an den unterschiedlichen Schulformen darstellen dürfte. Wechselwirkungen zwischen Schülerkomposition und schulformspezifischen didaktisch-curricularen Traditionen sind hierbei anzunehmen. Beunruhigen muss in diesem Zusammenhang vor allem die Situation an den Haupt- und Realschulen, die in beiden Kompetenzdomänen nicht nur die niedrigste Streuung, sondern auch die geringsten mittleren Lernstände aufweisen. Da die Leistungen der Haupt- und Realschülerinnen und -schüler bereits zu Beginn der Jahrgangsstufe 5 im unteren Leistungsbereich im Vergleich zu anderen Schulformen liegen (Bos et al., 2007), würde es bei einer weiteren Verringerung des klasseninternen Leistungsspektrums im weiteren Verlauf der Sekundarschule zu einer ‚Homogenisierung am unteren Ende‘ kommen (Tillmann & Wischer 2006, S. 46).

Die Breite des klasseninternen Leistungsspektrums scheint unseren Analysen zufolge für die Kompetenzen im Leseverständnis und in Mathematik am Ende der Jahrgangsstufe 6 von nachgeordneter Bedeutung im Vergleich zu anderen Prädiktoren zu sein. Ob Schülerinnen und Schüler in ihren Schulklassen eher homogene oder eher heterogene Leistungen vorfinden, hat nach den hier vorgestellten Analysen unter Einbezug aller Prädiktoren nominell schwach positive, wenngleich keine signifikanten Effekte auf ihre Leistungen. Immerhin lassen sich zumindest für das Leseverständnis ohne Berücksichtigung des Vorwissens (Modell 1) signifikant positive Effekte einer höheren Streuung der kognitiven Grundfähigkeit innerhalb der Schulklassen finden. Dies bestätigt Effekte, die ebenfalls für das Leseverständnis bereits von Lehmann (2006) für Grundschulklassen gefunden wurden – allerdings fallen die Effekte dort wesentlich höher aus als in den hier vorgelegten Analysen.

Zudem finden sich keine negativen Effekte eines größeren Leistungsspektrums in Schulklassen. Damit könnte anhand der vorliegenden Analysen davon ausgegangen werden, dass in Schulklassen, deren kognitive und leistungsbezogene Zusammensetzung heterogen ausfällt, kein Nachteil für die Lernerfolge der einzelnen Schülerinnen und Schüler besteht, sondern sich zumindest tendenziell eher ein Vorteil abzeichnet. Möglicherweise könnte ein Unterricht in der Sekundarstufe,

der diese Heterogenität nutzt, wie z. B. im Lesepatentmodell (Voss, Blatt, Gebauer, Müller & Masanek, 2008), diesen Vorteil noch verstärken.

Die hier dargestellten Analysen geben zunächst einen Überblick über generelle Effekte der Leistungsvariation auf zukünftige Lernerfolge innerhalb der ersten beiden Schuljahre der Sekundarstufe. Um einen differenzierten Blick auf die Bedingungen des Lernens in leistungshomogenen und -heterogenen Schulklassen für diesen Zeitraum zu gewinnen, sind weitere Untersuchungen notwendig. In einem nächsten Schritt könnten daher schulformspezifische Analysen vorgenommen werden, da davon auszugehen ist, dass in den drei Schulformen der Umgang mit Leistungsheterogenität in unterschiedlichem Maße verankert ist. Weiterhin ist zu vermuten, dass innerhalb der Lerngruppen durchaus unterschiedliche Bedingungen eine Rolle für den produktiven Umgang mit Heterogenität spielen und die Wirkungen der Leistungsgruppierung von der Unterrichtsqualität und dem Umgang der Lehrkräfte mit den Ausgangsbedingungen der Schülerinnen und Schüler abhängen. Ein Fokus weiterer Analysen wird sich daher vor allem auf die Unterrichtsgestaltung richten müssen. Welche Gründe letztlich dafür verantwortlich sind, dass Schülerinnen und Schüler in einigen Kontexten besser lernen als in anderen, soll und kann nicht die Frage dieser Arbeit sein und ist bis heute auch nicht gänzlich geklärt. Vermutlich muss die Antwort hierauf sowohl in der Qualität des Unterrichts als auch in Interaktionsstrukturen im Schüler-Schüler-Verhältnis gesucht werden. Einen direkten Zusammenhang mit den Schülerleistungen kann man zwischen dem unterrichtlichen Lehrangebot, der Art und Weise der Vermittlung sowie im zur Verfügung gestellten Zeitrahmen annehmen (Helmke, 2003). Da individuelles Lernen aber immer auch innerhalb von interpersonellen Interaktionen stattfindet (z. B. Bandura, 1982), kann angenommen werden, dass sich auch soziale Dynamiken in Schulklassen (beispielsweise Bezugsgruppeneffekte, Motivation, Angst) nachhaltig auf den Schulleistungserwerb auswirken.

Literatur

- Arnold, K.-H., Bos, W., Richert, P. & Stubbe, T.C. (2007). Schullaufbahnpräferenzen am Ende der vierten Klassenstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 271–297). Münster: Waxmann.
- Bandura, A. (1982). Self efficacy mechanism in human agency. *American Psychologist*, 37(2), 122–147.
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen. Vertiefende Analysen im Rahmen von PISA 2000*. (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Bos, W., Bensen, M., Gröhlich, C., Guill, K., May, P., Rau, A., Stubbe, T.C., Vieluf, U. & Wocken, H. (2007). *KESS 7. Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 7*. Hamburg: Behörde für Bildung und Sport.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Walther, G. & Valtin, R. (Hrsg.). (2003). *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich*. Münster: Waxmann.
- Bos, W. & Pietsch, M. (Hrsg.). (2006). *KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen*. Münster: Waxmann. (Bd. 1, HANSE – Hamburger Schriften zur Qualität im Bildungswesen).
- Bos, W., Voss, A., Lankes, E.-M., Schwippert, K., Thiel, O. & Valtin, R. (2004). Schullaufbahneempfehlungen von Lehrkräften für Kinder am Ende der vierten Jahrgangsstufe. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 191–228). Münster: Waxmann.
- Carle, U. (2005). Leistungsvielfalt im Unterricht. In K. Bräu & U. Schwerdt (Hrsg.), *Heterogenität als Chance. Vom produktiven Umgang mit Gleichheit und Differenz in der Schule*. (S. 55–70). Paderborn: Lit.
- Cortina, K.S. & Trommer, L. (2003). Bildungswege und Bildungsbiographien in der Sekundarstufe I. In K.S. Cortina, J. Baumert, A. Leschinsky, K.U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 342–391). Reinbek: Rowohlt.
- Dar, Y. & Resh, N. (1986). Classroom intellectual composition and academic achievement. *American Educational Research Journal*, 23(3), 357–374.
- Hattie, J.A.C. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37, 449–481.
- Heller, K.A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen (Revision) KFT 4-12+ R, Manual*. (3. Aufl.). Göttingen: Beltz Test.
- Helmke, A. (1988). Leistungssteigerung und Ausgleich von Leistungsunterschieden in Schulklassen: unvereinbare Ziele? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 20(1), 45–76.
- Helmke, A. (2003). *Unterrichtsqualität. Erfassen, bewerten, verbessern*. Seelze: Kallmeyer.
- Jennessen, S., Kastirke, N. & Uhrlau, K. (2005). Heterogenität als Herausforderung in der Grundschule – Ausgewählte Aspekte schulischer Handlungsmöglichkeiten. In F. Hellmich (Hrsg.), *Lehren und Lernen nach IGLU – Grundschulunterricht heute* (S. 87–101). Oldenburg: Didaktisches Zentrum.
- Köller, O. (2008). Gesamtschule – Erweiterung statt Alternative. In K.S. Cortina, J. Baumert, A. Leschinsky, K.U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland. Strukturen und Entwicklungen im Überblick* (S. 437–465). Reinbek: Rowohlt.
- Kulik, C.-L.C. & Kulik, J.A. (1982). Effects of ability grouping on secondary school students: A meta-analysis of evaluation findings. *American Educational Research Journal*, 19(5), 415–428.
- Kulik, C.-L.C. & Kulik, J.A. (1984). Effects of ability grouping on elementary school pupils: A meta-analysis. Paper presented at the Annual Meeting of the American Psychological Association. Toronto, Ontario, Canada.
- Lehmann, R.H. (2006). Zur Bedeutung der kognitiven Heterogenität von Schulklassen für den Lernstand am Ende der Klassenstufe 4. In A. Schründer-Lenzen (Hrsg.), *Risikofaktoren kindlicher Entwicklung. Migration, Leistungsangst und Schulübergang*. (S. 109–121). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Lehmann, R.H., Peek, R., Gänsfuß, R. & Husfeldt, V. (2002). *LAU 9. Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9. Ergebnisse einer Längsschnittuntersuchung in Hamburg*. Hamburg: Behörde für Bildung und Sport.
- Linchevski, L. & Kutscher, B. (1998). Tell me with whom you're learning and I'll tell you how much you've learned: Mixed-ability versus same-ability grouping in Mathematics. *Journal for Research in Mathematics Education*, 29(5), 533–554.
- Luyten, H. & van der Hoeven-van Doornum, A. (1995). Classroom composition and individual achievement. Effects of classroom composition and teacher goals in Dutch elementary education. *Tijdschrift voor Onderwijsresearch*, 20(1), 42–62.
- Opdenakker, M.-C. & van Damme, J. (2001). Relationship between school composition and characteristics of school process and their effect on mathematics achievement. *British Educational Research Journal*, 27(4), 407–432.
- Pietsch, M. (2007). Schulformwahl in Hamburger Schülerfamilien und die Konsequenzen für die Sekundarstufe I. In W. Bos, C. Gröhlich & M. Pietsch (Hrsg.), *KESS 4 – Lehr- und Lernbedingungen in Hamburger Grundschulen* (S. 127–166). Münster: Waxmann.
- Prediger, S. (2003). Heterogenität macht Schule – Herausforderungen und Chancen. In W. Sailer (Hrsg.), *Schulbegleitforschung Bremen Jahrbuch 2004* (S. 90–97). Bremen: Landesinstitut für Schule.
- Prenzel, M., Schütte, K., Rönnebeck, S., Senkbeil, M., Schöps, K. & Carstensen, C. (2008). Der Blick in die Länder. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006 in Deutschland – Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 149–263). Münster: Waxmann.
- Raudenbush, S.W., Bryk, A.S. & Congdon, R. (2004). *HLM 6 for Windows [Computer software]*. Lincolnwood, IL: Scientific Software International, Inc.
- Resh, N. & Dar, Y. (1992). Learning segregation in junior high-schools in Israel: Causes and consequences. *Educational Effectiveness and School Improvement*, 3(4), 272–292.
- Schafer, J.L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards der Kultusministerkonferenz*. Zugriff am 24.09.2008 unter: <http://www.kmk.org/schul/Bildungsstandards/Argumentationspapier308KMK.pdf>
- Slavin, R.E. (1987). Ability grouping and student achievement in elementary schools: A best-evidence synthesis. *Review of Educational Research*, 57(3), 293–336.
- Slavin, R.E. (1990). Achievement effects of ability grouping in secondary schools: A best evidence synthesis. *Review of educational research*, 60(3), 471–499.
- Thrupp, M., Lauder, H. & Robinson, T. (2002). School composition and peer effects. *International journal of educational research*, 37(5), 483–504.
- Tillmann, K.-J. & Wischer, B. (2006). Heterogenität in der Schule. Forschungsstand und Konsequenzen. *Pädagogik*, 59(3), 44–48.
- Treiber, B. & Weinert, F.E. (1985). *Gute Schulleistungen für alle? Psychologische Studien zu einer pädagogischen Hoffnung*. Münster: Aschendorff.
- Treinius, G. & Einsiedler, W. (1996). Zur Vereinbarkeit von Steigerung des Leistungsniveaus und Verringerung von Leistungsunterschieden in Grundschulklassen. *Unterrichtswissenschaft*, 24, 290–311.

Voss, A., Blatt, I., Gebauer, M., Müller, A. & Masanek, N. (2008). Unterrichtsentwicklung als integrierte Schulentwicklung. Das Hamburger Leseförderprojekt (HeLp). In W. Bos, H.G. Holtappels, H. Pfeiffer, H.-G. Rolff & R. Schulz-Zander (Hrsg.), *Jahrbuch der Schulentwicklung, Band 15. Daten, Beispiele und Perspektiven* (S. 93–122). Weinheim: Juventa.

Weinert, F.E. (1997). Notwendige Methodenvielfalt. Unterschiedliche Lernfähigkeiten erfordern variable Unterrichtsmethoden. *Friedrich Jahresheft, 15*, 50–52.

Carola Gröhlich, Dipl.-Päd., Wissenschaftliche Mitarbeiterin im Institut für Schulentwicklungsforschung, Technische Universität Dortmund
E-Mail: groehlich@ifs.tu-dortmund.de

Katja Scharenberg, M.A., Wissenschaftliche Mitarbeiterin im Institut für Schulentwicklungsforschung, Technische Universität Dortmund
E-Mail: scharenberg@ifs.tu-dortmund.de

Wilfried Bos, Prof. Dr., Direktor des Instituts für Schulentwicklungsforschung, Technische Universität Dortmund
E-Mail: officebos@ifs.uni-dortmund.de