

Becker, Michael; McElvany, Nele; Lüdtke, Oliver; Trautwein, Ulrich  
**Lesekompetenzen und schulische Lernumwelten: Besondere Fördereffekte des Frühübergangs in Gymnasien?**

*formal und inhaltlich überarbeitete Version der Originalveröffentlichung in:*

*formally and content revised edition of the original source in:*

*Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie 46 (2014) 1, S. 35-50*



Bitte verwenden Sie beim Zitieren folgende URN /

Please use the following URN for citation:

urn:nbn:de:0111-pedocs-146089 - <http://nbn-resolving.org/urn:nbn:de:0111-pedocs-146089>

DOI: 10.1026/0049-8637/a000104 - <http://dx.doi.org/10.1026/0049-8637/a000104>

#### Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen und das Werk bzw. den Inhalt nicht für kommerzielle Zwecke verwenden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

#### Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by-nc/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work, provided that the work or its contents are not used for commercial purposes.

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](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Mitglied der

  
Leibniz-Gemeinschaft

Akzeptierte Manuskriptfassung (nach peer review) des folgenden Artikels:

[Becker, M., McElvany, N., Lüdtke, O. & Trautwein, U. \(2014\). Lesekompetenzen und schulische Lernumwelten. Besondere Fördereffekte des Frühübergangs in Gymnasien? Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 46 \(1\). doi: 10.1026/0049-8637/a000104](#)

© Hogrefe Verlag, Göttingen 2014

Diese Artikelfassung entspricht nicht vollständig dem in der Zeitschrift veröffentlichten Artikel. Dies ist nicht die Originalversion des Artikels und kann daher nicht zur Zitierung herangezogen werden.

Die akzeptierte Manuskriptfassung unterliegt der Creative Commons License CC-BY-NC.

# Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie

## Lesekompetenzen und schulische Lernumwelten: Besondere Fördereffekte des Frühübergangs in Gymnasien? Reading literacy and schools as learning environments: Differential effects of early transition into the academic track of secondary schools?

--Manuskript-Entwurf--

<b>Manuskriptnummer:</b>	ZEPP-D-12-00008R2
<b>Vollständiger Titel:</b>	Lesekompetenzen und schulische Lernumwelten: Besondere Fördereffekte des Frühübergangs in Gymnasien? Reading literacy and schools as learning environments: Differential effects of early transition into the academic track of secondary schools?
<b>Artikeltyp:</b>	Originalarbeit
<b>Schlüsselwörter:</b>	Lesekompetenz; Wortschatz; Dekodierfähigkeit; Schereneffekt; Leistungsgruppierung; reading comprehension; vocabulary; word decoding ability; fan-spread effect; school tracking
<b>Korrespond. Autor:</b>	Michael Becker, Dr. FB Erziehungswissenschaft und Bildungssysteme Berlin, GERMANY
<b>Korrespondierender Autor, zweite Information:</b>	
<b>Korrespondierender Autor, Institution:</b>	FB Erziehungswissenschaft und Bildungssysteme
<b>Korrespondierender Autor, zweite Institution:</b>	
<b>Erstautor:</b>	Michael Becker, Dr.
<b>Erstautor, zweite Information:</b>	
<b>Reihenfolge der Autoren:</b>	Michael Becker, Dr. Nele McElvany, Prof. Dr. Oliver Lüdtke, Prof. Dr. Ulrich Trautwein, Prof. Dr.
<b>Reihenfolge der Autoren - zweite Information:</b>	
<b>Zusammenfassung:</b>	<p>Anhand der Berliner Leselängsschnitt-Studie (LESEN 3-6) wurde die schriftsprachliche Kompetenzentwicklung (Leseverständnis, Wortschatz, Dekodierfähigkeit) bei Grundschulern sowie Gymnasiasten längsschnittlich analysiert. Ein Fokus lag dabei auf der Frage, ob Kinder mit frühzeitigem Wechsel auf das Gymnasium (in Berlin nach Klasse 4) eine andere Lernentwicklung aufweisen als Kinder mit sechsjähriger Grundschulzeit. Insgesamt nahmen N=772 Schüler aus Berlin teil. Auf deskriptiver Ebene wurden keine differenziellen Effekte während der gemeinsamen Grundschulzeit zwischen den beiden Gruppen identifiziert. In der Zeit, in der die Kinder in unterschiedliche Schultypen aufgeteilt waren, zeigten sich für Wortschatz und Dekodierfähigkeit höhere Zugewinne für Gymnasiasten. Mit Wachstumskurvenanalysen wurde darüber hinaus überprüft, ob sich der differenzielle Zugewinn direkt auf die schulische Gliederung zurückführen ließ, wenn gleichzeitig die interindividuellen Unterschiede in intraindividuellen Veränderungen insgesamt berücksichtigt wurden. Unter Kontrolle der Gesamtentwicklung konnte in keiner der drei Domänen ein eindeutiger Effekt der schulischen Gliederung belegt werden. Inhaltliche und methodische Aspekte dieser Befunde werden diskutiert.</p> <p>Longitudinal analyses of data from the Berlin Longitudinal Reading Study examined the development of reading literacy (reading comprehension, vocabulary, decoding ability) in elementary and Gymnasium students. A focus of analysis was on whether students who transferred to Gymnasium after 4 years of elementary schooling showed</p>

differential learning trajectories from students who completed 6 years of elementary schooling. N=772 students participated in the study. On the descriptive level, no differential effects were found between the two groups while all students were still in elementary schooling. During the period in which they attended different school types, however, Gymnasium students showed higher learning gains in vocabulary and decoding ability. Growth curve analyses controlling for interindividual differences in intraindividual change were used to test whether these differential learning gains were directly attributable to ability tracking. When overall development was controlled, tracking was not found to have an effect in any of the three domains.

Running Head: LESEKOMPETENZEN UND FRÜHÜBERGANG IN GYMNASIEN

Lesekompetenzen und schulische Lernumwelten:  
Besondere Fördereffekte des Frühübergangs in Gymnasien?

Datum der Einreichung: 27. Januar 2013

Zusammenfassung

Anhand der Berliner Leselängsschnitt-Studie (LESEN 3-6) wurde die schriftsprachliche Kompetenzentwicklung (Leseverständnis, Wortschatz und Dekodierfähigkeit) bei Grundschülerinnen und Grundschülern sowie bei Gymnasiastinnen und Gymnasiasten längsschnittlich analysiert. Ein Fokus lag dabei auf der Frage, ob Schülerinnen und Schüler mit frühzeitigem Wechsel auf das Gymnasium – der in Berlin nach Klassenstufe 4 möglich ist – eine andere Lernentwicklung aufweisen als Kinder mit sechsjähriger Grundschulzeit. Insgesamt nahmen  $N = 772$  Schülerinnen und Schüler aus Berlin an der Studie teil. Auf einer deskriptiven Ebene wurden keine differenziellen Effekte während der gemeinsamen Grundschulzeit (Klassenstufe 3 und 4) zwischen den beiden Gruppen identifiziert. In der Zeit, in der die Kinder in unterschiedliche Schultypen aufgeteilt waren (5. und 6. Klasse), zeigten sich für den Wortschatz und – zumindest tendenziell – für die Dekodierfähigkeit höhere Zugewinne für die Gymnasiasten. Mit Wachstumskurvenanalysen wurde darüber hinaus überprüft, ob sich der differenzielle Zugewinn direkt auf die schulische Gliederung zurückführen ließ, wenn gleichzeitig die interindividuellen Unterschiede in intraindividuellen Veränderungen insgesamt berücksichtigt wurden. Unter Kontrolle dieser interindividuellen Unterschiede konnte in keiner der drei Domänen ein eindeutiger Effekt der schulischen Gliederung belegt werden. Inhaltliche und methodische Aspekte dieser Befunde werden diskutiert.

Schlüsselwörter:

Lesekompetenz, Wortschatz, Dekodierfähigkeit, Schereneffekt, Leistungsgruppierung, Längsschnittstudie, Wachstumskurvenmodelle

Abstract

Longitudinal analyses of data from the Berlin Longitudinal Reading Study (LESEN 3-6) examined the development of reading literacy (reading comprehension, vocabulary, and decoding ability) in elementary and Gymnasium students. A focus of analysis was on whether students who transferred to Gymnasium (the most academic track of the highly differentiated German secondary system) after 4 years of elementary schooling showed differential learning trajectories from students who completed 6 years of elementary schooling. A total of  $N = 772$  students from Berlin participated in the study. On the descriptive level, no differential effects were found between the two groups while all students were still in elementary schooling (grades 3 and 4). During the period in which they attended different school types (grades 5 and 6), however, the Gymnasium students showed higher learning gains in vocabulary and – by tendency – decoding ability. Growth curve analyses controlling for interindividual differences in intraindividual change were used to test whether these differential learning gains were directly attributable to ability tracking. When overall development was controlled, tracking was not found to have a clear effect in any of the three domains. Technical and practical aspects of the findings are discussed.

Key words: reading comprehension, vocabulary, word decoding ability, fan-spread effect, school tracking, longitudinal study, growth curve model.

## Lesekompetenzen und schulische Umwelten:

### Besondere Fördereffekte des Frühübergangs in Gymnasien?

Deutliche Unterschiede der Schulleistungen an den unterschiedlichen Schultypen des gegliederten deutschen Schulsystems sind durch jüngere Studien wie PISA gut dokumentiert (vgl. z.B. Klieme et al., 2010). Diese können theoretisch auf unterschiedliche Art entstanden sein: Zum einen unterscheiden sich Schülerinnen und Schüler unterschiedlicher Sekundarschulformen schon vor dem Übergang in die Sekundarstufe. Zum anderen ist es möglich, dass Leistungsdifferenzen zwischen den Kindern in der Sekundarschulzeit verstärkt werden (Baumert, Stanat & Watermann, 2006). Eine solche divergierende Entwicklung wird in der Literatur auch als Schereneffekt (Schneider & Stefanek, 2004) oder Matthäuseffekt („wer hat, dem wird gegeben“, Merton, 1968; vgl. für schulische Leistungsentwicklung u.a. Bast & Reitsma, 1997, 1998; Stanovich, 1986; Walberg & Tsai, 1983) bezeichnet.

Die Forschungslage ist uneinheitlich bezüglich der Frage, ob die Leistungsheterogenität am Ende der Sekundarschulzeit allein auf die unterschiedlichen Eingangsvoraussetzungen oder auch auf die Aufteilung in unterschiedliche Schultypen zurückzuführen ist (Baumert et al., 2006; Becker, 2009). Im vorliegenden Beitrag wird anhand der Berliner Leselängsschnitt-Studie (LESEN 3-6) untersucht, inwiefern sich divergierende Entwicklungen in schriftsprachlichen Kompetenzen in der Unterrichtssprache Deutsch im Primar- und frühen Sekundarschulbereich feststellen lassen. Die Studie LESEN 3-6 ist für die Beantwortung dieser Fragestellung geeignet, da sie ermöglicht, Schülerinnen und Schüler mit sechsjähriger Grundschulzeit mit Schülerinnen und Schülern zu vergleichen, die nach Klasse 4 frühzeitig auf das Gymnasium wechseln. Ein besonderes Augenmerk soll auf die konzeptuellen und methodischen Herausforderungen gerichtet werden, die bei der Untersuchung von Schulformeffekten auftreten.



### Lesekompetenz: Definition und Entwicklung

Der systematische Erwerb von Lesekompetenz wird in der Regel in der Grundschule eingeleitet, wobei vom orthographischen Lesenlernen, also dem Entziffern von Buchstaben und Worten, zum komplexen Textlesen, das oft auch als der Kern des Leseverständnisses verstanden wird, übergegangen wird (Duke & Pearson, 2002; Durkin, 1993; Groeben & Hurrelmann, 2002; zur Beziehung zwischen elementarem und Textverständnis vgl., Kintsch, 1998, S. 49-120). Der elementare Lernprozess des orthographischen Lesens wird in der deutschen Sprache bis spätestens Ende der dritten Klasse in der Regel abgeschlossen (Klicpera & Gasteiger-Klicpera, 1993, S. 7-77; Seymour, Aro & Erskine, 2003). In den nachfolgenden Jahren verändert sich der Fokus hin zu komplexeren Lese- und Verständnisprozessen (Snow, Scarborough & Burns, 1999). Insofern ist die Entwicklung von Lesekompetenz nicht mit der Grundschule abgeschlossen, sondern auch eine Aufgabe der Sekundarstufe (Artelt, Stanat, Schneider & Schiefele, 2001).

Wie sich der Prozess und Übergang vom Nicht-Leser zum kompetenten Leser vollzieht, ist hierbei eng an individuelle Voraussetzungen gebunden: Dies variiert in Abhängigkeit von psycholinguistischen und perzeptuellen Voraussetzungen, kognitiven und motivationalen Kompetenzen wie auch dem intellektuellen sowie sozialen (Lern-)Umfeld, was einerseits den schulischen Unterricht, andererseits auch das elterliche Vorbild und elterliche Unterstützung umfasst (Gee, 2001; Snow, 2002). Hiervon ausgehend wurde in der Literatur wiederholt vermutet, dass es zu einer divergierenden Leistungsentwicklung kommen sollte: Kinder unterscheiden sich schon vor der Einschulung in vieler Hinsicht in Hinblick auf die Voraussetzungen des Leseerwerbes (vgl. B. Becker & Biedinger, 2006; Lee & Burkam, 2002); Schülerinnen und Schüler mit günstigeren individuellen Eingangsvoraussetzungen sollten zu einem früheren Zeitpunkt die elementaren Erwerbsprozesse abgeschlossen haben und früher zur Einübung komplexerer Leseprozesse übergehen (vgl. Stanovich, 1986). Hierbei sollten unter anderem positivere Erfahrungen gesammelt werden, was eine häufigere Aufnahme von

lesebezogenen Aktivitäten und eine schnellere Akquise leserelevanten Wissens (insbesondere Wortschatz) begünstigt; dies müsste wiederum, in einer Art Rückkopplung, den Leseprozess und dessen Entwicklung positiv unterstützen (vgl. auch Helmke & Weinert, 1997; Huslander, Olson, Willcutt & Wadsworth, 2010; Morgan & Fuchs, 2007; Pfof, Dörfler & Artelt, 2010; Sternberg & Powell, 1983; siehe aber Baumert, Nagy & Lehmann, 2012).

Die Ergebnisse der Überprüfung dieser Annahmen sind jedoch uneinheitlich. Es wurde zwar mehrfach gezeigt, dass Kinder mit günstigeren Voraussetzungen insbesondere in den frühen Jahren des Erwerbsprozesses höhere Lernraten zeigten (Compton, 2000; McCoach, O'Connell, Reis & Levitt, 2006; Williamson, Appelbaum & Epanchin, 1991). Jedoch blieben divergierende Entwicklungen in anderen Studien aus (Catts, Hogan & Fey, 2003; Jordon, Kaplan & Hanich, 2002) beziehungsweise sprachen sogar für kompensatorische Prozesse (Aarnoutse & Leeuwe, 2000; Aarnoutse, Leeuwe, Voeten & Oud, 2001; Francis, Shaywitz, Stuebing, Shaywitz & Fletcher, 1996; Leppänen, Niemi, Aunola & Nurmi, 2004; Scarborough & Parker, 2003; Skibbe et al, 2008).

Bast und Reitsma (1997, 1998) fanden empirische Hinweise dafür, dass verschiedene Facetten der Lesekompetenz bei der Untersuchung von Entwicklungsverläufen berücksichtigt werden sollten. So berichteten sie in derselben Stichprobe Indizien, dass sich Unterschiede in der Dekodierfähigkeit vergrößerten, im Textverständnis hingegen keine divergierende Entwicklung beobachten ließ. Parrila, Aunola, Leskinen, Nurmi und Kirby (2005) fanden Hinweise darauf, dass diese (differenziellen) Entwicklungsmuster auch sprach- bzw. kulturspezifisch sein könnten: In einer vergleichenden Untersuchung zeigten sie für kanadische Kinder über die Domänen hinweg relativ konsistent konvergierende Leistungsentwicklungen, bei finnischen Kindern hingegen sowohl konvergierende (Leseverständnis) als auch divergierende Entwicklungen (Dekodiergeschwindigkeit).

### Lesekompetenzentwicklung im deutschen Sekundarschulwesen

In Hinblick auf empirische Arbeiten zur Entwicklung von Lesekompetenzen in Deutschland im Sekundarschulbereich zeichnet sich in ähnlicher Weise ab, dass die Befunde allein auf einer deskriptiven Ebene differenziert zu sehen sind. Für das Textverständnis liegen aus längsschnittlichen Studien kaum Indizien für Schereneffekte in der Sekundarstufe vor. Weder in der Bundesland übergreifenden Untersuchung Deutsch Englisch Schülerleistungen International (DESI; DESI-Konsortium, 2008), in der Hamburger Studie Aspekte der Lernausgangslage und Lernentwicklung (LAU; Caro & Lehmann, 2008; Lehmann, Peek, Gänsfuß & Husfeldt, 2001) noch in der Nachfolgeuntersuchung zu LAU Kompetenzen und Einstellungen von Schülerinnen und Schülern (KESS; Bos, Bensen, Gröhlich, Jelden & Rau, 2006; Bos & Pietsch, 2007) ließen sich differenzielle Entwicklungen zwischen leistungsstärkeren und –schwächeren Schülerinnen und Schülern bzw. zwischen den Schulformen nachweisen. In der schleswig-holsteinischen Untersuchung Lesen in der Sekundarstufe (LISA; Retelsdorf & Möller, 2008; Retelsdorf, Becker, Köller & Möller, 2012) deuteten sich Unterschiede zwischen Schülerinnen und Schülern unterschiedlicher Schulformen an, verfehlten jedoch statistische Signifikanz. Für Kinder zwischen der 4. und 6. Klasse an Berliner Schulen konnten in der Untersuchung Erhebung zum Lese- und Mathematikverständnis - Entwicklungen in den Jahrgangsstufen 4 bis 6 in Berlin (ELEMENT; Lehmann & Lenkeit, 2008) ebenfalls keine systematischen Unterschiede zwischen Schülerinnen und Schüler an Gymnasien und Grundschulen nachgewiesen werden (Baumert, Becker, Neumann & Nikolova, 2009, 2010). Eine Ausnahme hiervon stellt die Studie Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter (BiKS; Pfof, Karing, Lorenz & Artelt, 2010) dar, in der für Leseverständnis Unterschiede vor allem zwischen Hauptschülerinnen und Hauptschülern auf der einen und Schülerinnen und Schülern anderer Schulformen auf der anderen Seite identifiziert wurden.

Für andere Subfacetten der Lesekompetenz sprechen bislang vorliegende Ergebnisse etwas konsistenter für Schereneffekte: Zwar zeigten sich in LAU auch im Vokabel- bzw. grammatischen Wissen keine Schereneffekte (Lehmann, Peek et al., 2001). In DESI wiesen jedoch in einer ähnlichen Subfacette (Sprachbewusstheit) die Unterschiede zwischen den Schulformen das für einen Schereneffekt charakteristische Muster auf, nach dem für Schülerinnen und Schüler auf Gymnasien ein stärkerer Leistungszuwachs zu beobachten ist (vgl. Eichler, 2008). In der Münchener Longitudinalstudie zur Genese individueller Kompetenzen (LOGIK) analysierten Schneider und Stefanek (2004, 2007) im schriftsprachlichen Bereich die Veränderung verbaler Intelligenzleistungen und Rechtschreibkompetenzen während der Grund- und Sekundarschulzeit. Sie konnten hierbei Schereneffekte, die sich auch in die Sekundarstufe hinein fortsetzten, zwischen Schülerinnen und Schülern von Hauptschulen und Gymnasien belegen. Für schleswig-holsteinische Schülerinnen und Schüler zeigten sich in der LISA-Studie Schereneffekte für die Dekodierfertigkeit (Retelsdorf et al., 2012). Ähnliche Ergebnisse fanden sich in der BiKS-Studie für die Wortschatzentwicklung (Pfost et al., 2010).

Zusammenfassend betrachtet, ist die empirische Befundlage bezüglich differenzieller Entwicklungsverläufe von Lesekompetenzen während der Sekundarschulzeit in Deutschland nicht eindeutig. Im Hinblick auf das Leseverständnis blieben Schereneffekte zwischen den Schulformen weitgehend aus, in anderen Subfacetten der Lesekompetenz hingegen wurden wiederholt Unterschiede belegt.

#### Zur Trennung von Schulformeffekten und individuell bedingten Entwicklungsfaktoren

Verlässt man die deskriptive Ebene, also die reine Beschreibung von Veränderungen der Schulleistungen, stellt sich prinzipiell die Frage, inwiefern differenzielle Entwicklungsverläufe allein durch interindividuell unterschiedliche Voraussetzungen bedingt oder auch durch andere, kontextuelle Faktoren geprägt sind. Dies ist spätestens dann zentral, wenn untersucht werden soll, ob die schulische Gliederung die Lesekompetenzentwicklung beeinflusst. Neben den

interindividuellen Unterschieden, die schon vor Beginn der Sekundarschulzeit bestehen (vgl. Bos et al., 2004; Ditton, Krüsken & Schauenberg, 2005) und zumindest theoretisch unterschiedliche Lernkurven erwarten lassen, werden auch qualitativ-institutionelle Schul(form)unterschiede und Unterschiede in der Schülerkomposition als mögliche Erklärungsfaktoren für eine differentielle Leistungsentwicklung postuliert (vgl. vertiefend u.a. Becker, 2009; Becker, Lüdtke, Trautwein, Köller & Baumert, 2012).

Unter institutionellen Faktoren werden Aspekte wie etwa Unterschiede in den Curricula, in der Lehrerausbildung und -qualifikation sowie in Unterrichtskultur und -organisation zusammengefasst. In Deutschland wurden Schul(form)unterschiede dieser Art vielfach belegt: Gymnasiallehrerinnen und -lehrer unterrichten kognitiv aktivierender, während vor allem an Hauptschulen repetitive Unterrichtselemente größeres Gewicht haben (Baumert et al., 2004, 2010; Kunter & Baumert, 2006; siehe aber Klieme et al., 2008). Ebenso unterscheiden sich Lehrkräfte der verschiedenen Schulformen in ihrer Ausbildung und damit einhergehend in ihrem fachdidaktischen und fachlichen Wissen (Brunner et al., 2006).

Unterschiede in der Komposition, also der Zusammensetzung der Schülerschaft einer Klasse bzw. einer Schule, kommen ebenso in Frage unterschiedliche Leistungsentwicklungen zu bedingen. Hierbei lassen sich im Wesentlichen fünf Aspekte differenzieren: die Leistungskomposition (z.B. mittleres Vorwissen), die soziokulturelle Zusammensetzung (z.B. Bildungsstatus der Eltern), die Zusammensetzung hinsichtlich psychosozialer Risikofaktoren des Elternhauses (z.B. Erwerbslosigkeit von Eltern), die ethnisch-kulturelle Komposition (z.B. Anteil bilingualer Kinder) sowie der Anteil lernbiografischer Belastungsfaktoren einer Klasse (z.B. Klassenwiederholer, vgl. u.a. Baumert et al., 2006).

In der Forschung finden sich Hinweise sowohl für institutionelle als auch Effekte der Komposition (vgl. z.B. Köller & Baumert, 2001; Kunter, 2005; Neumann et al., 2007; Baumert et al., 2006; zur internationalen Debatte vgl. etwa Dronkers, 2010; Hanushek & Wößmann, 2006; Van Ewijk & Slegers, 2010). Aber national wie international sind Wirkmechanismen und

Effektgrößen institutioneller wie auch kompositionsbezogener Aspekte umstritten bzw. ist die Befundlage nur uneindeutig (vgl. Becker, 2009; Hattie, 2008; Thrupp, Lauder & Robinson, 2002). Die methodische Herausforderung bei der Untersuchung von differenziellen Leistungsentwicklungen besteht vor allem darin, den Beitrag individueller und institutioneller Faktoren voneinander zu trennen. Dies sei anhand eines stark vereinfachten Beispiels verdeutlicht, bei dem wir von vereinfachenden Modellannahmen (z.B. lineares Wachstum) ausgehen und Störhypothesen ausblenden (vgl. Raudenbush, 2001). Das Beispiel bezieht sich auf das Berliner Schulsystem, das auch den nachfolgenden Analysen zugrunde liegt (vgl. nachfolgender Abschnitt). In Berlin beträgt die reguläre Grundschulzeit sechs Jahre. Einem kleinen Teil der Schülerinnen und Schüler wird es nach der vierten Klasse ermöglicht, frühzeitig auf sogenannte grundständige Gymnasien überzugehen.

In Abbildung 1 werden für diese zwei Gruppen unterschiedliche, hypothetische Muster der Leistungsentwicklung in den Klassenstufen 2, 4 und 6 dargestellt. Die eine Gruppe wird von Schülerinnen und Schülern gebildet, die nach Klassenstufe 4 auf eine andere Schulform wechseln (hier: grundständige Gymnasien). Die zweite Gruppe verbleibt in der gleichen Schulform (hier: Grundschulen). In Abbildung 1a finden sich über drei Messzeitpunkte hinweg parallele Lernzuwächse in den zwei unterschiedlichen Populationen. Dies würde man als Hinweis darauf deuten, dass keine Schereneffekte auftreten.

In Abbildung 1b ist ein Schereneffekt dargestellt. Die Gruppe der Schülerinnen und Schüler, die die Schule wechselt, hat einen steileren Lernverlauf als die Gruppe derjenigen, die keinen Wechsel vollzieht. Allerdings sind die Unterschiede in den Lernkurven bereits zu einem Zeitpunkt vorhanden, zu dem alle Schüler noch die gemeinsame Schule, d.h. die Grundschule besuchten. Von daher liegt es nahe, insbesondere individuellen Faktoren (und nicht dem Besuch des Gymnasiums) eine divergenzsteigernde Funktion zuzusprechen.

Schließlich findet sich in Abbildung 1c ein Beispiel für einen akzelerierten Schereneffekt. Während der Leistungszuwachs zwischen den beiden Gruppen von der zweiten zur vierten

Grundschulklasse noch parallel ausfällt, geht die Leistungsentwicklung anschließend deutlich auseinander. Dies würde so interpretiert werden, dass Unterschiede in den individuellen Lernraten nach der vierten Grundschulklasse von einem institutionellen bzw. kompositionellen Effekt überlagert bzw. verstärkt werden.<sup>1</sup>

In der bisherigen Forschungslage zur Entwicklung von Lesekompetenz im Sekundarschulbereich liegen keine Daten vor, die abzubilden erlauben, inwiefern der Übergang von der Grundschule in die Schulen der Sekundarstufe verändernd auf den Entwicklungsverlauf wirkt. Dies gilt sowohl für die besondere Berliner Situation des Frühübergangs (d.h. Aufteilung der Schülerschaft in einen Teil, der in das Sekundarschulsystem übergeht, und einen anderen, der in der Primarstufe verbleibt) als auch für den „regulären“ Übergang in eine in mehrere Schulformen gegliederte Sekundarstufe.<sup>2</sup> Bisherige Analysen ermöglichen lediglich, den Verlauf während der Sekundarstufe abzuschätzen und Unterschiede zum Zeitpunkt des Übergangs statistisch zu kontrollieren, jedoch nicht diese in Relation mit den vorangehenden Entwicklungen zu stellen. Dies ist jedoch eine Voraussetzung, wenn auch differenziert werden soll, welcher Art die Effekte sind, ob also differenzielle Gewinne – oder sogar Verluste – im Vergleich zur vorangehenden Entwicklung entstehen und nicht nur geschätzt werden soll, ob sich über die individuellen Eingangsvoraussetzungen hinaus prinzipiell Unterschiede für die Leistungsentwicklung aufgrund eines Schul(form)wechsels ergeben (vgl. für eine umfassende Diskussion Becker, 2009).

---

<sup>1</sup> Theoretisch ließen sich noch weitere Verlaufsmuster unterscheiden, z.B. gegenläufige Gewinne und Verluste. Dies soll jedoch hier nicht weiter vertieft werden (siehe weiterführend Becker, 2009).

<sup>2</sup> Einzige Ausnahme hierzu bildet die LOGIK-Studie, die jedoch anderen methodischen Einschränkungen unterliegt, insbesondere da unterschiedliche Instrumente verwendet wurden und keine gemeinsame Metrik über die Messzeitpunkte etabliert wurde (vgl. für eine weiterführende methodische Diskussion auch Becker, Lüdtke, Trautwein & Baumert, 2006; Becker, 2009).

### Anlage der vorliegenden Studie und Fragestellung

Die Entwicklungsabschnitte vor und nach dem Übergang in unterschiedliche Schultypen sollen im Folgenden mithilfe des Berliner Leselängsschnitts (LESEN 3-6) verglichen werden. Die Studie umfasst die Entwicklung vom Ende der 3. Klasse bis zum Ende der 6. Klasse. Das Berliner Schulsystem weist die für die vorliegende Studie günstige Besonderheit auf, dass für die Klassenstufen 5 und 6 Gymnasiasten mit Grundschulern verglichen werden können (für eine differenziertere Darstellung vgl. Baumert et al., 2009, 2010, sowie Lehmann & Lenkeit, 2008). Die reguläre Grundschule in Berlin umfasst sechs Jahrgangsstufen. Den Schülerinnen und Schülern steht aber prinzipiell offen, nach der 4. Klasse in ein grundständiges Gymnasium überzugehen. Eine Minderheit der Schülerschaft (ca. 7 bis 8 Prozent eines Jahrgangs) nimmt dies wahr. Die anderen Schülerinnen und Schüler, die nach der 6. Klasse regulär in die Gymnasien übergehen (weitere ca. 35 Prozent), befinden sich also im betrachteten Zeitraum noch in den Grundschulen. Es wurden sowohl Schülerinnen und Schüler in Berlin berücksichtigt, die die gesamte Zeit eine Grundschule besuchten als auch Schülerinnen und Schüler, die bereits nach Klassenstufe 4 auf ein grundständiges Gymnasium wechselten.

Die Lesekompetenzentwicklung wurde zu drei Messzeitpunkten (Klasse 3, 4 und 6) in drei schriftsprachlichen Domänen (Leseverständnis, Wortschatz und Dekodierfähigkeit) erfasst, die über diese Messzeitpunkte auf einer einheitlichen Metrik interpretiert werden können. Somit kann geprüft werden, zu welchen absoluten Veränderungen es kommt, ob sich Schereneffekte feststellen lassen und inwiefern der frühzeitige Gymnasialbesuch eine direkt verstärkende Wirkung auf interindividuelle Unterschiede in der Leistungsentwicklung hat.

Neben der Deskription der Leistungsentwicklung besteht ein weiteres Ziel der vorliegenden Studie darin, die möglichen Ursachen für das Vorliegen von Schereneffekten genauer zu analysieren. Findet sich zwischen Schülerinnen und Schülern unterschiedlicher Schulformen eine divergierende Leistungsentwicklung, so stellt sich die Frage, ob diese Unterschiede auf früh angelegte Unterschiede in individuellen Lernkurven oder auf



Kompositions- bzw. institutionelle Effekte zurückgeführt werden können. Das in diesem Beitrag umgesetzte längsschnittliche Design ermöglicht es, die individuellen Lernraten aller Schülerinnen und Schüler vor dem Übertritt auf das Gymnasium zu bestimmen, da bereits zwei Messzeitpunkte vor dem Übertritt vorliegen. Mithilfe einer gemeinsamen Modellierung der Veränderung durch Wachstumskurvenanalysen soll daher überprüft werden, ob bereits Unterschiede in den Lernraten vorliegen („individueller Matthäuseffekt“), bevor für einen Teil der Kinder der Übertritt auf ein Gymnasium erfolgt, welches möglicherweise zusätzlich divergenzsteigernd ist („institutioneller/kompositioneller Matthäuseffekt“).

Wir untersuchen demnach im Folgenden drei Fragestellungen:

- a) In welchem Umfang nehmen die Leseleistungen der Schüler zwischen der dritten und sechsten Klasse zu?
- b) Ist hierbei auf einer deskriptiven Ebene ein Schereneffekt beobachtbar?
- c) Sind etwaige divergierende Entwicklungen auch dann noch nachweisbar, wenn die vorangehende Individualentwicklung berücksichtigt wird?

## Methode

### Stichprobe

Die LESEN 3-6-Studie wurde im Zeitraum von 2003 bis 2006 am Berliner Max-Planck-Institut für Bildungsforschung durchgeführt (vgl. McElvany & Artelt, 2007; McElvany, Becker & Kortenbruck, 2008). Ziel der Studie war, die Leseentwicklung von Kindern im Längsschnitt von der 3. bis zur 6. Jahrgangsstufe zu untersuchen und individuelle, soziale und institutionelle Einflussfaktoren zu ermitteln, wobei im sozialen Bereich vor allem familiäre Determinanten untersucht wurden. Die Messungen der vorliegenden Arbeit wurden zeitlich Ende der dritten Klasse im Juni 2003 (T1), Mitte der vierten Klasse im Januar 2004 (T2) und zum Ende der sechsten Klasse im Mai 2006 (T3) umgesetzt.

An der LESEN 3-6-Untersuchung nahmen insgesamt  $N = 883$  Schüler aus ursprünglich 54 Berliner Grundschulklassen an 22 öffentlichen und privaten Berliner Grundschulen teil. Von der Stichprobengrundgesamtheit wurde eine Reihe von Schülerinnen und Schülern aus den Analysen ausgeschlossen, da diese in eine andere Population übergingen (z.B. in eine Sonderschule) oder keine Informationen zur Leseleistung vorlagen. Insgesamt betraf dies  $N = 111$  Schülerinnen und Schüler, sodass  $N = 772$  Schülerinnen und Schüler in die Analysen einbezogen werden konnten. Von den 772 Kindern gingen  $N = 55$  (7.1 Prozent) vorzeitig in ein grundständiges Gymnasium über. Diese wurden im Folgenden für einen Vergleich mit den Grundschulern herangezogen. Ein kleiner Teil der Stichprobe,  $N = 116$ , nahm zusätzlich an einem Leseförderprogramm teil. Die Ergebnisse blieben im Wesentlichen gleich auch unter Berücksichtigung der Teilnahme an diesem Programm, weshalb auf eine Darstellung im Folgenden verzichtet wurde. 47.1 Prozent der Stichprobe waren Mädchen. Das Durchschnittsalter der teilnehmenden Schülerinnen und Schüler betrug zu Beginn der Studie in der dritten Klasse  $M = 9.1$  Jahre ( $SD = 0.5$ ). Der sozioökonomische Hintergrund, operationalisiert über den höchsten Internationalen Sozioökonomischen Index in der Familie (HISEI), lag bei  $M = 52.0$  ( $SD = 16.7$ ). Der Anteil an Kindern, die zu Hause eine andere Sprache als Deutsch sprechen, lag bei 28.4 Prozent. Die Stichprobe erschien in zentralen Merkmalen unter anderem zur ELEMENT-Untersuchung weitgehend vergleichbar.

Die Stichproben zwischen den Schülerinnen und Schülern, die den Frühübergang vollzogen, unterschieden sich systematisch von denjenigen, die in den Grundschulen verblieben. So ist die Leistung dieser gymnasialen Schülerschaft in allen drei hier betrachteten Leistungsdomänen höher (vgl. unten, Tab. 1), aber auch sozial sind die Gymnasien positiv selektiv. So fiel etwa der elterliche sozioökonomische Status (HISEI) an den Gymnasien höher aus (GS:  $M = 51.2$ ; GY:  $M = 63.2$ ). Die vorgefundenen Muster entsprechen den bei Baumert et al. (2009, 2010) berichteten Befunden. Die Bedeutsamkeit dieser Unterschiede wird im Diskussionsteil weiterführend erläutert.

Die Datenerhebung dauerte zu den ersten beiden Messzeitpunkten jeweils zwei Schulstunden. Zu Messzeitpunkt 3 wurde die Testzeit auf drei Schulstunden ausgeweitet. Die Schülerinnen und Schüler wurden in der Regel in ihrem Klassenverband untersucht. Kinder, die aus ihrem ursprünglichen Klassenverband ausschieden, darunter auch die Gymnasiasten, wurden in Gruppensitzungen am Max-Planck-Institut für Bildungsforschung untersucht.

### Instrumente

#### Leseverständnis

Zur Erfassung des Leseverständnisses wurden in den beiden ersten Wellen Itemsbstichproben in Form von Multiple-Choice-Verständnisaufgaben aus dem Hamburger Lesetest (HAMLET 3-4; Lehmann, Peek & Poerschke, 1997) eingesetzt, die jeweils fünf Texte mit je vier Aufgaben umfassten. In der letzten Erhebung, in der 6. Klasse, wurden zwei Texte mit jeweils neun Aufgaben aus dem Diagnostischen Test Deutsch (DTD; Nauck & Otte, 1980; Testversion A) verwendet. Die Texte wurden anhand der unterschiedlichen Itemschwierigkeiten der zugehörigen Aufgaben ausgewählt, sodass ein möglichst breites Fähigkeitsspektrum erfasst werden konnte. Die Komplexität der Verständnisfragen reichte von einfacher Informationsentnahme bis hin zu selbstständigen, schlussfolgernden Verständnisleistungen. Es wurden zu allen Testzeitpunkten zwei Pseudoparallelförmige Testhefte mit leicht variiertem Itemfolge verwendet.

Die Tests der drei Messzeitpunkte können als Schätzer einer gemeinsamen Leseverständnisdimension über die Zeit interpretiert werden. Die jeweiligen Testhefte überlappten sich nicht direkt, das heißt, zu jedem der Messzeitpunkte wurden jeweils unterschiedliche Leseverständnisaufgaben eingesetzt. Eine einheitliche Metrik wurde über zwei externe Kalibrierungsstichproben etabliert, anhand derer die Itemparameter der einzelnen Tests im Rahmen eines gemeinsamen eindimensionalen Raschmodells geschätzt wurden (Wellen 1 und 2 als equivalent group design mit  $N = 809$  Schülern; Wellen 2 und 3 als common item non-equivalent group design mit  $N = 828$  Schülern, vgl. Kolen & Brennan, 2004). Die über die

Kalibrierungsstichproben geschätzten Itemparameter wurden anschließend zur Skalierung der Leistungsdaten des Haupttestes herangezogen. Dies bietet den Vorteil, dass die Veränderung der Leistung ohne Erinnerungseffekte erfasst wird, da keine Items doppelt bearbeitet wurden. Als Parameterschätzungen wurden WLE (Warm, 1989) verwendet. Die Reliabilitäten der Tests zu den drei Messzeitpunkten waren gut bis befriedigend (Cronbachs  $\alpha_{T1} = .78$ ,  $\alpha_{T2} = .69$ ;  $\alpha_{T3} = .69$ ).

### Wortschatz

Der Wortschatz wurde mithilfe von Verbalanalogie-Tests erfasst. Diese Komponente kann sowohl als ein Wortschatztest als auch als Test verbaler Intelligenz interpretiert werden. In den ersten beiden Erhebungen wurde die Wortschatzkomponente des Grundintelligenztests (CFT; Weiß, 1998) verwendet; für die 3. Welle wurde die gleiche Komponente des kognitiven Fähigkeitstests (KFT 4-12+ R; Heller & Perleth, 2000) herangezogen. Die Aufgabenstellung beider Tests ist identisch: Zu einem Wortpaar, das in einem gewissen Verhältnis zueinander steht, muss ein zweites Wortpaar mit identischer Beziehung gebildet werden. Hierzu wird jeweils ein Wort des zu bildenden Paares vorgegeben und aus vier möglichen Antwortkategorien muss dasjenige herausgesucht werden, welches mit dem vorgegebenen dritten Wort die gleiche Beziehung wie die des ersten Wortpaares herstellt.

In insgesamt sechs Minuten mussten jeweils 15 Items bearbeitet werden. Die Tests wurden im Vergleich zu den originalen Testanweisungen leicht verkürzt (um fünf Aufgaben und eine Minute Testzeit), weshalb ein Vergleich mit den Populationsnormen der jeweiligen Tests nicht möglich ist. Auch der Wortschatztest wurde im Rahmen des Raschmodells auf eine gemeinsame Metrik über die drei Erhebungswellen gebracht, wofür die gleichen Kalibrierungsstichproben herangezogen wurden. Als Parameterschätzungen wurden WLE (Warm, 1989) verwendet. Die internen Konsistenzen sind gut bis befriedigend (Cronbachs  $\alpha_{T1} = .69$ ;  $\alpha_{T2} = .73$ ;  $\alpha_{T3} = .78$ ).

### Wortdekodiergeschwindigkeit

Die Wortdekodiergeschwindigkeit wurde mit der Würzburger Leise Leseprobe (WLLP; Küspert & Schneider, 1998) erfasst. Dieser Test misst die Dekodiergeschwindigkeit von Kindern, indem geschriebenen Worten jeweils vier Bildalternativen gegenübergestellt werden und aus diesen dasjenige Bild ausgewählt und angestrichen werden muss, welches das geschriebene Wort als Bild wiedergibt. Die Kinder werden in der Instruktion angehalten, möglichst zügig zu arbeiten.

Diese Aufgabe wurde ähnlich wie der Wortschatztest leicht verkürzt. Von den 140 vorgesehenen Aufgaben wurden in den ersten Erhebungen jeweils nur 70 Items dargeboten, wofür drei Minuten Zeit gewährt wurden. In der 6. Klasse wurde anschließend der volle Test mit 140 Aufgaben, aber in der ursprünglichen verkürzten Zeit vorgegeben, um Deckeneffekte auszuschließen. Die Aufgaben werden zur Auswertung zu einfachen Summenwerten aufaddiert. Die Testgüte wird im Manual des Tests als gut ausgewiesen. Für die 4. Klassen beträgt die Paralleltestreliabilität der Normierungsstichproben  $r = .82$  und die 14-Tage Reteststabilität  $r_{tt} = .78$ . In der vorliegenden Untersuchung betragen die Stabilitäten zwischen T1 und T2  $r_{tt} = .75$  und zwischen T2 und T3  $r_{tt} = .69$ .

### Statistische Auswertungen

#### Analysestrategie

Es wurde einerseits anhand der grundlegenden Veränderung der Mittelwerte und Varianzen zu allen Messzeitpunkten geprüft, inwieweit sich auf einer deskriptiven Ebene differenzielle Veränderungen in der 3. - 4. Klasse im Unterschied zur 4. - 6. Klasse zwischen den Grundschulern und Gymnasiasten zeigten. Inwiefern sich auch differenzielle Entwicklungen unter Kontrolle der vorangehenden individuellen Veränderung finden lassen, wurde in einer gemeinsamen Modellierung der Veränderung mit Wachstumskurvenmodellen überprüft (*growth curve analysis*, vgl. Bollen & Curran, 2006; Duncan, Duncan & Stryker, 2006). Die Modelle wurden univariat, mit unspezifiziertem Wachstumsfaktor bestimmt (vgl. Abb. 2), das heißt, zur

Identifikation des Modells wurde die erste Ladung  $\lambda_{T1}$  auf 0 und die Ladung zum dritten Messzeitpunkt  $\lambda_{T3}$  auf 1 fixiert;  $\lambda_{T2}$  für den Indikator zum zweiten Messzeitpunkt wurde frei geschätzt. Dies bietet den Vorteil, dass die Ladung zur zweiten Messung als proportionale Veränderung zu T3 interpretiert werden kann (Bollen & Curran, 2006). Diesen sogenannten Spline-Modellen wurde der Vorzug gegeben, da sie konsistent über alle drei Domänen hinweg einen guten Modellfit erbrachten im Unterschied zu Modellen mit linear spezifiziertem Wachstumsverlauf.

Um in den Wachstumskurven die differenzielle Veränderung der Gymnasiasten und Grundschüler abzubilden, wurde zusätzlich eine Dummyvariable für die Schulformzugehörigkeit in die Modelle aufgenommen (1 = Gymnasium). Da die Schulformzugehörigkeit die Variabilität erst zum letzten Messzeitpunkt beeinflussen konnte, wurde sie als unabhängige Variable zum dritten Messzeitpunkt eingeführt (vgl. Abb. 2). Die Modellierung erfolgte analog zur Spezifizierung zeitlich variierender Kovariaten, das heißt die Zugehörigkeit zum Gymnasium wurde als Prädiktor auf der Ebene des Indikators zu T3 in das Modell aufgenommen (vgl. Bollen & Curran, 2006; Duncan et al., 2006).

Die Notation der Wachstumskurvenmodelle entspricht derjenigen von Bollen und Curran (2006). Varianzen und Kovarianzen der Faktoren wurden so modelliert, dass sie als latente Varianzen und Kovarianzen interpretiert werden können. Es wurde in der methodischen Literatur kritisch angemerkt, dass Fitindizes von Wachstumskurvenmodellen im Unterschied zu klassischen SEM (z.B. konfirmatorische Faktorenanalyse) weniger aussagekräftig darüber sind, inwiefern die Modelle die individuellen Verläufe adäquat repräsentieren (vgl. Singer & Willett, 2003). Die Fitindizes ( $\chi^2$ -Werte, RMSEA und SRMR) sind in Tabelle 3 wiedergegeben, werden aber im Text nicht weitergehend interpretiert. Alle Modelle wurden in Mplus 6.11 geschätzt (Muthén & Muthén, 1998-2011).

## Umgang mit fehlenden Werten

Missing data stellt in Längsschnittstudien ein zentrales Problem dar. Die Anteile von Missings betragen in der vorliegenden Studie je nach Leistungstest zwischen 24.0 und 41.1 Prozent. In Dropoutanalysen wurde überprüft, inwiefern sich systematische Unterschiede zwischen den Schülern, die entweder vorzeitig aus der Stichprobe ausschieden, zu einem späteren Zeitpunkt hinzukamen bzw. vollständig an der Untersuchung teilnahmen, ergaben. Es fanden sich keine substantiellen Unterschiede zwischen den einzelnen Substichproben (für Details vgl. Kortenbruck, 2007).

Die fehlenden Daten wurden mit dem Programm Mplus 6.11 (Muthén & Muthén, 1998-2011) imputiert. Als Variablen in das Imputationsmodell gingen einerseits die hier betrachteten abhängigen Leistungsvariablen und die unabhängige Gymnasialvariable ein, andererseits wurden auch weitere verfügbare Variablen als Hilfsvariablen (v.a. Angaben zu Alter, Geschlecht, Noten, Leseverhalten, -motivation und -selbstkonzept sowie Variablen zum sozialen Hintergrund) hinzugenommen, um eine größere Präzision bei der Parameterschätzung zu erreichen (Collins, Schafer & Kam, 2001). Es wurden 25 vollständige Datensätze erzeugt, in denen die fehlenden Werte durch sogenannte plausible values ersetzt sind. Die Ergebnisse der Analysen wurden nach den Regeln von Rubin (1987) kombiniert, die sich auch auf Regressionskoeffizienten und Korrelationen übertragen lassen. Die entsprechenden Formeln sind in der Analyseoption Type = Imputation von Mplus (Muthén & Muthén, 1998-2010) implementiert.

## Ergebnisse

### Deskriptive Statistiken: Leistungszuwachs und Schereneffekt

Der Tabelle 1 sind die deskriptiven Statistiken für die Gesamtstichprobe sowie für die beiden Substichproben (Grundschüler und grundständige Gymnasiasten) zu entnehmen. Betrachtet man die Gesamtstichprobe, so kann man in allen drei Domänen einen statistisch signifikanten Zuwachs zwischen den Messzeitpunkten erkennen (jeweils  $p < .001$ ). Der Zuwachs auf der manifesten Ebene betrug über die drei Schuljahre im Leseverstehen  $d = 0.88$ , im Wortschatz  $d =$

1.32 und im Dekodieren  $\underline{d} = 1.82$  (jeweils relativiert an der Streuung zu T1). Dies entspricht umgerechnet auf ein Schuljahr einer Veränderung zwischen knapp einem und zwei Drittel einer Standardabweichung.

Die Gymnasiasten wiesen zu allen Messzeitpunkten statistisch signifikant höhere Mittelwerte in allen drei Domänen auf als die Grundschüler (vgl. Tabelle 1). Im Leseverständnis lag der Unterschied zwischen  $\underline{d} = 0.77$  und  $\underline{d} = 0.89$ , im Wortschatztest zwischen  $\underline{d} = 0.72$  und  $\underline{d} = 0.96$  sowie zwischen  $\underline{d} = 0.78$  und  $\underline{d} = 0.87$  in der Domäne der Dekodierfähigkeit (jeweils relativiert an der Standardabweichung der Gesamtstichprobe). Durch einen Vergleich der Zugewinne für Gymnasiasten und Grundschüler kann das Vorliegen eines Schereneffektes zwischen den beiden Gruppen überprüft werden; in Tabelle 2 sind die Zuwächse jeweils von T1 nach T2 sowie von T2 nach T3 wiedergegeben. In der Domäne des Leseverstehens fielen Leistungsveränderungen der beiden Stichproben nicht differenziell aus. Der Unterschied im Leistungszuwachs war weder in der gemeinsamen Primarschulzeit statistisch signifikant (Gymnasium:  $\underline{M}(\Delta) = 0.26$ ; Grundschule:  $\underline{M}(\Delta) = 0.28$ ;  $p = .919$ ), noch in der Zeit, in der Gymnasiasten und Grundschüler tatsächlich getrennt unterrichtet wurden (Gymnasium:  $\underline{M}(\Delta) = 0.87$ ; Grundschule:  $\underline{M}(\Delta) = 0.76$ ;  $p = .288$ ).

Auch im Wortschatz und der Dekodierfähigkeit waren die Leistungsveränderungen in der Primarschulzeit zwischen der 3. und 4. Klasse nicht statistisch signifikant unterschiedlich. Der Wortschatz der Gymnasiasten nahm im Mittel zwischen der 3. und 4. Klasse um  $\underline{M}(\Delta) = 0.45$  zu, der Wortschatz der Grundschüler um  $\underline{M}(\Delta) = 0.44$  ( $p = .955$ ). Ebenso wenig war der Unterschied in der Domäne der Dekodierfähigkeit in der gemeinsamen Grundschulzeit statistisch bedeutsam (Gymnasium:  $\underline{M}(\Delta) = 9.08$ ; Grundschule:  $\underline{M}(\Delta) = 9.43$ ,  $p = .847$ ). Im Unterschied zum Leseverstehen zeigten die Gymnasiasten allerdings zwischen der 4. und 6. Klasse sowohl im Wortschatz als auch im Dekodieren höhere Zugewinne. Die Veränderung der Gymnasiasten im Wortschatz war mit  $\underline{M}(\Delta) = 1.40$  höher als die der Grundschüler mit  $\underline{M}(\Delta) = 1.09$ . Die Unterschiede waren statistisch signifikant ( $p = .014$ ). In der Dekodierfähigkeit deuteten sich für



die Gymnasiasten höhere Zugewinne ( $M(\Delta) = 17.21$ ) als die Grundschüler ( $M(\Delta) = 13.84$ ) an. Der Unterschied war mit  $p = .062$  marginal statistisch bedeutsam.

Für die zweite Frage, ob sich Schereneffekte über die Schuljahre zeigen, lassen sich als ein weiteres Indiz Veränderungen in der Variabilität heranziehen (vgl. Tab. 1). Die Standardabweichungen unterschieden sich in den Domänen Leseverständnis und Wortschatz nicht statistisch signifikant zwischen dem ersten und letzten Messzeitpunkt (Leseverständnis:  $p = .257$ ; Wortschatz:  $p = .349$ ). Für die Dekodiergeschwindigkeit erhöhte sich die Standardabweichung für die Gesamtstichprobe zwischen dem ersten und dritten Messzeitpunkt statistisch signifikant (T1:  $SD = 12.98$ ; T3:  $SD = 15.71$ ;  $p < .001$ ). Insofern fanden sich zusätzliche Hinweise auf Schereneffekte in dieser Domäne.

#### Wachstumskurvenmodelle: Schereneffekt nach Kontrolle vorangehender Leistung

Im nächsten Schritt wurde geprüft, ob sich eine divergierende Leistungsentwicklung auch dann noch nachweisen lässt, wenn die vorangehende Individualentwicklung berücksichtigt wird. Es werden die Ergebnisse von Wachstumskurvenmodellen berichtet die einerseits die mittlere individuelle Veränderung der Schüler zwischen dem Ende des 3. und dem Ende des 6. Schuljahres abbilden. Andererseits kann mit Hilfe dieser Modelle geprüft werden, ob sich nach Kontrolle der Unterschiede in der individuellen Entwicklung eine zusätzliche Veränderung nachweisen lässt, die auf den Gymnasialbesuch nach der 4. Klasse zurückgeht.

Hinsichtlich der Fragestellungen 1 und 2 ergeben sich im Wachstumskurvenmodell ähnliche Ergebnisse wie auf manifester Ebene (vgl. Tab. 3). Es zeigte sich ein Zuwachs über die drei Jahre hinweg, der auf der latenten Ebene tendenziell etwas größer ausfiel als auf Indikatorebene (Leseverstehen:  $\underline{d} = 1.26$ ; Wortschatz:  $\underline{d} = 1.57$ ; Dekodieren:  $\underline{d} = 2.11$ ). Dies ist insofern plausibel, da die latente Varianz des Intercept im Wachstumskurvenmodell keine zeitspezifische Residualvarianz enthält und somit kleiner sein sollte. Die Modelle zeigen darüber hinaus, dass für alle drei Domänen eine statistisch bedeutsame Varianz in der Ausgangsmessung

vorlag (Leseverstehen:  $\psi_{\alpha\alpha} = 0.68$ ,  $p < .001$ ; Wortschatz:  $\psi_{\alpha\alpha} = 1.00$ ,  $p < .001$ , Dekodieren:  $\psi_{\alpha\alpha} = 125.21$ ,  $p < .001$ ), die Variabilität des Slopes jedoch nur für die Dekodierfähigkeit statistische Bedeutsamkeit ausfiel (Leseverstehen:  $\psi_{\beta\beta} = 0.06$ ,  $p = .726$ ; Wortschatz:  $\psi_{\beta\beta} = 0.05$ ,  $p = .730$ , Dekodieren:  $\psi_{\beta\beta} = 4.34$ ,  $p < .001$ ). Für keine der drei Domänen fand sich eine statistisch signifikante Kovarianz zwischen Intercept und Slope (Leseverstehen:  $\psi_{\alpha\beta} = -0.05$ ,  $p = .565$ ; Wortschatz:  $\psi_{\alpha\beta} = -0.12$ ,  $p = .208$ ; Dekodieren:  $\psi_{\alpha\beta} = 3.56$ ,  $p = .724$ ).

Zur Prüfung der Fragestellung 3 wurde der Koeffizient der Regression der Leistung zu T3 auf die Dummyvariable der gymnasialen Zugehörigkeit betrachtet. Der Effekt war für das Leseverständnis, wie aufgrund der Ergebnisse auf manifester Ebene zu erwarten war, zufallskritisch nicht abzusichern ( $b_{T3,GY} = 0.20$ ,  $p = .760$ ). Der Unterschied in der Wortschatzleistung erreichte jedoch ebenfalls keine statistische Signifikanz ( $b_{T3,GY} = 0.08$ ,  $p = .907$ ), und auch für die Dekodierfähigkeit konnte kein zusätzlicher institutioneller Effekt statistisch abgesichert werden ( $b_{T3,GY} = 3.32$ ,  $p = .279$ ). Obwohl die Regressionskoeffizienten allesamt einen positiven Effekt des Gymnasiums andeuteten, ist unter Berücksichtigung der vorausgehenden intraindividuellen Entwicklung kein statistisch signifikanter zusätzlicher Effekt der Schulform nachweisbar.

### Diskussion

Anhand von Berliner Grundschulern und Gymnasiasten wurde die Entwicklung von Lesekompetenzen untersucht. Auf deskriptiver Ebene zeigte sich, dass sich die Leistungen im Mittel deutlich verbesserten. Die Leistungen nahmen je nach Subfacette um mehr als ein bis zwei Drittel einer Standardabweichung pro Schuljahr zu. Im Vergleich zwischen Grundschulern und Gymnasiasten zeigten sich ausgeprägte Mittelwertsunterschiede in allen drei Domänen ( $\underline{d} \geq 0.72$ ). Im Unterschied zu fast allen anderen bislang verfügbaren Untersuchungen gewährt die vorliegende Untersuchung einen Einblick in die Lernraten unterschiedlicher Schülergruppen sowohl vor als auch nach dem Übergang auf einer zeitlich übergreifend vergleichbaren Metrik: Die Leistungszunahmen fielen teilweise differenziell zwischen Gymnasiasten und

Grundschulern aus. In der Zeit der gemeinsamen Grundschulzeit (3. und 4. Klasse) konnten in keiner der drei Domänen statistisch signifikante Unterschiede nachgewiesen werden. Während der Zeit getrennter Beschulung fiel die Leistungsentwicklung im Leseverstehen von Gymnasiasten und Grundschulern ebenfalls nicht unterschiedlich aus. Demgegenüber ließ sich ein statistisch signifikanter differenzieller Zuwachs sowohl hinsichtlich des Wortschatzes als auch der Dekodierfähigkeit zwischen der 4. und 6. Klasse nachweisen. Abschließend wurde mit Wachstumskurvenmodellen überprüft, ob sich unter Berücksichtigung der individuellen Unterschiede in den intraindividuellen Lernraten ein Effekt der unterschiedlichen Schultypen statistisch absichern lässt. Es konnten keine spezifischen zusätzlichen Effekte des grundständigen Gymnasiums im Vergleich zu den bereits in der Grundschule bestehenden Unterschieden belegt werden.

Die Befundmuster für Leseverstehen stehen hierbei in Einklang mit den Ergebnissen der von Baumert et al. (2009, 2010) berichteten Auswertungen für vergleichende Analysen zum gymnasialen Frühübergang in Berlin; für Wortschatz und Dekodieren liegen bislang keine vergleichbaren Daten für die Situation des Frühübergangs vor. Im weiteren Kontext der Forschungsfrage nach differenziellen Entwicklungen in unterschiedlichen Schulformen – unabhängig vom gymnasialen Frühübergang – stehen die hier vorgestellten Ergebnisse insofern in Einklang mit vorangehender Forschung, als dass bisherige Arbeiten eher keine Scheren- bzw. Schulformeffekte für Leseverständnis berichteten (z.B. Bos et al., 2006; Caro & Lehmann, 2008; DESI-Konsortium, 2008; Retelsdorf et al., 2012; siehe aber Pfof et al., 2010), entsprechende Effekte allerdings etwas konsistenter für Kompetenzen in den Bereichen Wortschatz und Dekodierfähigkeiten gefunden wurden (vgl. Pfof et al., 2010; Retelsdorf et al., 2012; Schneider & Stefanek, 2004, 2007; auch: Eichler, 2008; siehe aber Lehmann, Peek et al., 2001).

Worauf kann das Ausbleiben von deutlichen Hinweisen auf (institutionelle) Schereneffekte vor allem für das Leseverstehen im Allgemeinen (in der vorliegenden Studie sogar für alle drei Subdomänen lesebezogener Kompetenzen) zurückgeführt werden?

Hinsichtlich des gewählten Bereichs lesebezogener Kompetenz ist prinzipiell umstritten, welche Rolle der Schule und Schulqualität im Sekundarschulbereich hinsichtlich der Lesekompetenzen in der Muttersprache zukommt. Der Erwerb grundlegender Lese- und Schreibkompetenzen ist in der Regel in der Grundschule abgeschlossen (Klicpera & Gasteiger-Klicpera, 1993). Es wird diskutiert, ob nach dem Erwerb dieser grundlegenden Kompetenzen ausschließlich das individuelle (private) Leseverhalten entscheidend für die weitere Entwicklung der Kompetenzen ist (Guthrie, Wigfield, Metsala & Cox, 1999; Stanovich, 1986). Daher können sich durchaus, wie im vorliegenden Fall zumindest für Wortschatz und Dekodierfähigkeiten, Schereneffekte zeigen, diese sind aber weitgehend als individuelle Entwicklungsverläufe erklärbar. Allerdings sei vermerkt, dass dies nicht prinzipiell ausschließt, dass weiterhin institutionelle Effekte bestehen: Diese könnten indirekt verursacht werden, wenn z.B. Schule einen Einfluss auf Quantität und Qualität des Lesestoffs der außerschulischen Lektüre nimmt und damit die Kompetenzentwicklung mitbestimmt.

Aus methodischer-statistischer Sicht stellt sich allerdings auch die Frage, wie sensitiv unsere Schlussfolgerungen gegenüber der Wahl des Analysemodells sind. Insbesondere bei längsschnittlichen, nicht-experimentellen Daten werden in der methodischen Literatur eine Vielzahl von unterschiedlichen Analyseverfahren vorgeschlagen (z.B. Bollen & Curran, 2006; Hertzog & Nesselroade, 2003). Wir haben deshalb zusätzlich eine Reihe von alternativen Modellspezifikationen vorgenommen, die im Wesentlichen alle zu dem Ergebnis führten, dass es keine Hinweise auf einen Effekt der Schulform auf die Lesekompetenz gibt. Als alternative Modellierung im Rahmen der Wachstumskurvenmodelle wurde der interessierende Effekt des Gymnasiums nicht als Kovariate des Intercepts und des Slopes und Prädiktor der Messung zur dritten Erhebungswelle, sondern als direkter Prädiktor der Indikatoren zu allen drei Messzeitpunkten spezifiziert (vgl. Bollen & Curran, 2006, S. 192-195). Der Indikator für das Gymnasium wurde also verwendet, um schulformspezifische Variabilität direkt in den Indikatoren vorherzusagen. Entsprechend wurde in dieser Art der Modellierung geprüft,

inwiefern sich der (potenziell kausal zu interpretierende) Effekt des Gymnasiums zum letzten Messzeitpunkt von den (nicht kausal zu interpretierenden) Effekten auf zeitlich vor dem Gymnasialbesuch liegende Unterschiede unterscheidet. Diese Form der Modellierung erbrachte allerdings ähnliche Resultate wie die hier vorgestellten.

Andererseits wurden die Modelle erweitert, um Unterschiede zwischen den Stichproben aus Grundschulen und grundständigen Gymnasien zu modellieren, die zum Beispiel für nicht-lineare Effekte des Zuwachses relevant sein und in Abhängigkeit der Ausgangsleistung potenziell bestehen könnten (dies wurde etwa von Baumert et al., 2012, als Möglichkeit formuliert). In weiterführenden Analysen, hier nicht im Detail wiedergeben, wurden die LGM mit *Propensity Score Matching*-Analysen kombiniert, ein Vorgehen, das zusätzlich zur Modellierung der Entwicklung auch Unterschiede zwischen den Grundschulern und Gymnasiasten zu kontrollieren erlaubt (zentrale Hintergrundvariablen: elterlicher sozioökonomischer Status und elterliche Bildung, Migrationsstatus, Geschlecht; kognitive und motivationale Variablen: Leseverständnis, Wortschatz, Dekodierfähigkeit, Leseverhalten, Leseinteresse, jeweils längsschnittlich zu T1 und T2 in der Grundschule; zur Begründung des Vorgehens vgl. Ho, Imai, King & Stuart, 2007). Demnach wurden die Gymnasiastinnen und Gymnasiasten nur mit den ähnlich leistungsstarken und sozial positiv selektiven Grundschülerinnen und Grundschulern verglichen. Die Ergebnisse sind im Wesentlichen mit den vorgestellten identisch, insofern sich keine statistisch bedeutsamen Unterschiede zugunsten der Gymnasiastinnen und Gymnasiasten identifizieren ließen. Hierdurch wurde gleichzeitig auch überprüft, ob gegebenenfalls differenzielle Lernraten von Schülerinnen und Schülern unterschiedlicher Leistungsgruppen in Form nicht-linearer Entwicklungsverläufe das Ausbleiben von Schereneffekte begründen könnten (vgl. Baumert et al., 2012). Dies war jedoch nicht der Fall.<sup>3</sup>

---

<sup>3</sup> Aus grundlegend differenziellen Lernraten leistungsstärkerer und –schwächerer Schülerinnen und Schüler erklärt sich der (scheinbare) Widerspruch zwischen den Ergebnissen, dass sich deskriptiv zumindest für Wortschatz und

In einem letzten Kontrollschritt wurden die Effekte auch in konventionellen Regressionsverfahren überprüft. Hierbei ergaben sich keine Effekte des Gymnasiums bei direkter Kontrolle der im vorangehenden Abschnitt genannten Kovariaten in regressionsanalytischen Modellen. Um auszuschließen, dass dies durch die herabgesetzte Teststärke bei multiplen Prädiktoren bedingt wurde, wurde ein zweites Modell mit lediglich dem Propensity Score als zusammenfassende Kovariate gerechnet (mitunter auch als PS-ANCOVA bezeichnet, vgl., Steiner, Cook, Shadish & Clark, 2010). Ein Gymnasialeffekt blieb hierbei ebenso aus.

Darüber hinaus erscheint es uns auch wichtig darauf hinzuweisen, dass einige Aspekte unserer Arbeit die Chance, institutionelle Effekte nachzuweisen, eingeschränkt haben. Die wichtigste Beschränkung besteht sicherlich darin, dass die Stichprobe der Gymnasiasten mit nur  $N = 55$  Personen insgesamt relativ klein ist. Dieses Problem wird nochmals dadurch verschärft, dass die Teststärke von LGM in Hinblick auf die Identifikation differenzieller Entwicklungen vergleichsweise gering ist, obwohl sie Veränderung auf einer latenten, um den erhebungsspezifischen Messfehler bereinigten Ebene modellieren (Hertzog, Lindenberger, Ghisletta & von Oertzen, 2006). Insgesamt erhöht dies die Wahrscheinlichkeit von  $\beta$ -Fehlern. Zudem war der Zeitraum, in dem die Entwicklung in der Grundschule erfasst wurde, relativ kurz (sieben Monate). Dies könnte zu einer unzureichenden Erfassung der intraindividuellen Veränderung geführt haben. Die differenzielle Reliabilität der Maße, die für das Leseverständnis zu T2 und T3 und Wortschatz zu T1 mit Cronbachs  $\alpha = .7$  als nur hinreichend anzusehen ist, könnte darüber hinaus eine Erklärung der unterschiedlichen Veränderung in den drei Domänen

---

Dekodierfähigkeit größere Lernraten in der Zeit nach dem Übergang zugunsten der Gymnasien zeigten – im Unterschied zu den Entwicklungen vor dem Übergang –, in den LGM und den vorangehend erwähnten Prüfungen sich jedoch keine Hinweise auf Schulformeffekte ergaben. Dies ist darauf zurückzuführen, dass leistungsschwächere Lernende tendenziell größere Leistungsveränderungen (zumindest in der gemeinsamen Grundschulzeit) zeigten als leistungsstärkere. Letztere sind allerdings in den Gymnasien über- und den Grundschulen unterrepräsentiert. Insofern müssen aus dieser inhaltlichen Perspektive für die Schätzung differenzieller Effekte unterschiedlicher Bedingungen die vorangehenden Veränderungen (wie hier im LGM) bzw. zumindest die Ausgangsleistungen kontrolliert werden (wie etwa in PSM oder auch in einfachen Regressionsanalysen) (in der methodischen Literatur auch als sog. Simpson-Paradox, Simpson, 1951, diskutiert; vgl. auch Steyer 1992).

sein (vgl. auch Robitzsch, Dörfler, Pfof & Artelt, 2011) und einen Mangel an Teststärke, um differenzielle Effekte der Schulformen zu entdecken, bedingen.

Einige weitere Einschränkungen gelten für die vorliegende Arbeit, stellen aber auch generell Herausforderungen dar. So gilt es zu bedenken, dass statistische Modelle häufig nur eine Vereinfachung des tatsächlich bestehenden Beeinflussungsmusters darstellen. Im Theorieteil und der Ableitung der Fragestellung haben wir zwischen drei Erklärungsmustern (individuelle, kompositionelle und institutionelle Erklärungen) für Schereneffekte unterschieden und graphisch idealtypische Verläufe der Lesefähigkeit dargestellt, die helfen können, zwischen den wirksam werdenden Mechanismen zu unterscheiden. Es muss nochmals betont werden, dass dabei eine Reduktion der Komplexität vorgenommen wurde. So geht unsere Analysestrategie beispielsweise implizit von der Annahme aus, dass in den Grundschulen keine kompositionellen/institutionellen Effekte auftreten. Da aber Kinder in der Regel die geographisch nächst gelegene Grundschule besuchen, kann es gerade in Großstädten zu Grundschulklassen kommen, deren soziale und kognitive Zusammensetzung eher späteren Gymnasialklassen entspricht und hohe Übergangsraten in die Gymnasien aufweisen (Bos et al., 2004; Ditton & Krüsken, 2006; Weinert, 1997). Für solche Schüler ist die Möglichkeit, einen „zusätzlichen“ Gymnasialeffekt zu belegen, eingeschränkt, da Kompositionseffekte, die üblicherweise mit dem Gymnasialbesuch einhergehen, hier schon in der Grundschule vorhanden sind. Allerdings kann aber auch erwartet werden, dass der Gymnasialeffekt in denjenigen Fällen besonders stark ist, wenn Schülerinnen und Schüler aus eher leistungs- und sozial schwachen Grundschulen in Gymnasien wechseln.

Für den Nachweis eines Gymnasialeffektes könnte es auch ein Nachteil sein, wenn der Vergleich unter Einschluss der 6. Klasse geführt wird. Diese Klassenstufe ist in Berlin für die Grundschülerinnen und Grundschüler übergangsrelevant, sodass die Eltern die Leistungsentwicklung der Grundschülerinnen und Grundschüler in dieser Klassenstufe besonders umfassend fördern (Ditton & Krüsken, 2006; Moser & Rhy, 2000; Neumann et al.,

2007). Insofern könnte die Entwicklung der Grundschüler gerade im 6. Schuljahr optimiert und durch außerschulische Lerngelegenheiten, vor allem durch das Elternhaus, geprägt sein. Was als ausbleibender Gymnasialeffekt interpretiert würde, wäre dann (teilweise) ein kompensatorischer Elterneffekt in der Grundschule. Zudem könnte Ähnliches für Gymnasiastinnen und Gymnasiasten in der 4. Grundschulklasse gelten: Deren Leistungen dürften in dieser Jahrgangsstufe optimal verlaufen sein, was die Spezifikation „des“ Gymnasialeffektes erschweren dürfte. Die genannten Schwierigkeiten erfordern eine sorgfältige Auswahl von Untersuchungsdesign und Auswertungsverfahren. Will man vorübergehend differenziell beschleunigte Lernkurven untersuchen, um den Nachweis für nicht-lineare Entwicklungen vor dem Übergang zu erbringen, so wird man weitere Messzeitpunkte einplanen müssen.

Die Analyse differenzieller Lernverläufe im gegliederten Schulsystem stellt die Entwicklungspsychologie und pädagogische Psychologie vor spannende Herausforderungen, die nur unter Berücksichtigung der jeweiligen institutionellen Gegebenheiten und mit anspruchsvollen Forschungsdesigns gemeistert werden können. Trotz der Beschränkungen, die unser Datensatz aufwies, dürfte unsere Studie wichtige Hinweise für die weitere Forschung geben.



Literaturverzeichnis

- Aarnoutse, C. A. J. & Van Leeuwe, J. F. J. (2000). Development of poor and better readers during the primary school. Educational Research and Evaluation, 6, 251-278.  
doi:10.1076/1380-3611(200009)6:3;1-A;FT251
- Aarnoutse, C., Leeuwe, J. v., Voeten, M. & Oud, H. (2001). Development of decoding, reading comprehension, vocabulary and spelling during the elementary school years. Reading and Writing, 14, 61-89. doi:10.1023/A:1008128417862
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann, & M. Weiß (Hrsg.), PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich (S. 69-137). Opladen: Leske + Budrich.
- Bast, J. & Reitsma, P. (1997). Mathew effects in reading: A comparison of latent growth curve models and simplex models with structured means. Multivariate Behavioral Research, 32, 135 - 167. doi:10.1207/s15327906mbr3202\_3
- Bast, J. & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. Developmental Psychology, 34, 1373-1399. doi:10.1037//0012-1649.34.6.1373
- Baumert, J., Becker, M., Neumann, M. & Nikolova, R. (2009). Frühübergang in ein grundständiges Gymnasium: Übergang in ein privilegiertes Entwicklungsmilieu? Ein Vergleich von Regressionsanalyse und Propensity Score Matching. Zeitschrift für Erziehungswissenschaft, 12, 189-215. doi:10.1007/s11618-009-0072-4
- Baumert, J., Becker, M., Neumann, M. & Nikolova, R. (2010). Besondere Förderung von Kernkompetenzen an Spezialgymnasien? Der Frühübergang in grundständige Gymnasien in Berlin. Zeitschrift für Pädagogische Psychologie, 24, 5-22. doi:10.1024/1010-0652.a000001

- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133-180. doi:10.3102/0002831209345157
- Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W. & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (Hg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland: Ergebnisse des zweiten internationalen Vergleichs* (S. 314-354). Münster: Waxmann.
- Baumert, J., Nagy, G. & Lehmann, R. (2012). Cumulative advantages and the emergence of social and ethnic inequality: Matthew effects in reading and mathematics development within elementary schools? *Child Development*, 83, 1347-1367. doi:10.1111/j.1467-8624.2012.01779.x
- Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95-188). Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-90082-7\_4
- Becker, B. & Biedinger, N. (2006). Ethnische Bildungsungleichheit zu Schulbeginn. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 58, 660-684. doi:10.1007/s11577-006-0261-6
- Becker, M. (2009). *Kognitive Leistungsentwicklung in differenziellen Lernumwelten: Effekte des gegliederten Sekundarschulsystems in Deutschland*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Becker, M., Lüdtke, O., Trautwein, U. & Baumert, J. (2006). Leistungszuwachs in Mathematik.

- Evidenz für einen Schereneffekt im mehrgliedrigen Schulsystem. Zeitschrift für Pädagogische Psychologie, 20, 233-242. doi:10.1024/1010-0652.20.4.233
- Becker, M., Lüdtke, O., Trautwein, U., Köller, O. & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? Journal of Educational Psychology, 104, 682-699.  
doi:10.1037/a0027608
- Bollen, K. A. & Curran, P. J. (2006). Latent curve models: A structural equation approach. Hoboken, NJ: John Wiley & Sons.
- Bos, W. & Pietsch, M. (Hrsg.). (2007). KESS 4 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 4 in Hamburger Grundschulen. Münster: Waxmann.
- Bos, W., Bensen, M., Gröhlich, C., Jelden, D. & Rau, A. (2006). Erster Bericht zu den Ergebnissen der Studie „Kompetenzen und Einstellungen von Schülerinnen und Schülern – Jahrgangsstufe 7“ (KESS 7). Dortmund: Institut für Schulentwicklungsforschung (IFS).  
Internetressource: <http://www.hamburger-bildungsserver.de/schulentwicklung/qualitaet/kess/KESS7.pdf> (letzter Zugriff 06.02.2008).
- Bos, W., Lankes, E., Prenzel, M., Schwippert, K., Valtin, R. & Walter, G. (2004). IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich. Münster: Waxmann.
- Brunner, M., Kunter, M., Krauss, S., Baumert, J., Blum, W., Dubberke, T., Jordan, A., Klusman, U., Tsai, Y.-M. & Neubrand, M. (2006). Welche Zusammenhänge bestehen zwischen dem fachspezifischen Professionswissen von Mathematiklehrkräften und ihrer Ausbildung sowie beruflichen Fortbildung? Zeitschrift für Erziehungswissenschaft, 9, 521-544. doi:10.1007/s11618-006-0166-1
- Caro, D. H. & Lehmann, R. (2009). Achievement inequalities in Hamburg schools: how do they change as students get older? School Effectiveness and School Improvement, 25, 1–25.

doi:10.1080/09243450902920599

- Catts, H. W., Hogan, T. P. & Fey, M. E. (2003). Subgrouping poor readers on the basis of individual differences in reading-related abilities. Journal of Learning Disabilities, 36, 151-164. doi:10.1177/002221940303600208
- Collins, L. M., Schafer, J. L. & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychological Methods, 6, 330-351. doi:10.1037//1082-989X.6.4.330-351
- Compton, D. L. (2000). Modeling the growth of decoding skills in first-grade children. Scientific Studies of Reading, 4, 219-259. doi:10.1207/S1532799XSSR0403\_3
- DESI-Konsortium (Hrsg.) (2008). Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie. Weinheim: Beltz.
- Ditton, H. & Krüsken, J. (2006). Der Übergang von der Grundschule in die Sekundarstufe I. Zeitschrift für Erziehungswissenschaft, 9, 348-372. doi:10.1007/s11618-006-0055-7
- Ditton, H., Krüsken, J. & Schauenberg, M (2005). Bildungsungleichheit: Der Beitrag von Familie und Schule. Zeitschrift für Erziehungswissenschaft, 8, 285-303. doi:10.1007/s11618-005-0138-x
- Dronkers, J. (2010). Features of educational systems as factors in the creation of unequal educational outcomes. In J. Dronkers (Ed.), Quality and Inequality of Education (pp. 299-327). Dodrecht: Springer. doi:10.1007/978-90-481-3993-4\_12
- Duke, N. K. & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. E. Farstrup, & S. J. Samuels (Eds.), What research has to say about reading instruction (3rd ed., pp. 205–242). Newark, DE: International Reading Association. doi:10.1598/0872071774.10
- Duncan, T. E., Duncan, S. C. & Stryker, L. A. (2006). An introduction to latent variable growth curve modeling: Concepts, issues, and applications. Mahwah, NJ: Lawrence Erlbaum.
- Durkin, D. (1993). Teaching them to read (6th ed.). Boston, MA: Allyn & Bacon.

- Eichler, W. (2008). Sprachbewusstheit Deutsch. In DESI-Konsortium (Hrsg.), Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie (S. 112-118). Weinheim: Beltz.
- Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A. & Fletcher, J. M. (1996). Developmental Lag Versus Deficit Models of Reading Disability: A Longitudinal, Individual Growth Curves Analysis. Journal of Educational Psychology, 88, 3-17. doi:10.1037/0022-0663.88.1.3
- Gailberger, S. & Willenberg, H. (2008). Leseverstehen Deutsch. In DESI-Konsortium (Hrsg.), Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie (S. 60-71). Weinheim: Beltz.
- Gee, J. (2001). Reading as situated language: A sociocognitive perspective. Journal of Adolescent & Adult Literacy, 44, 714–725. doi:10.1598/JAAL.44.8.3
- Groeben, N. & Hurrelmann B. (Hrsg.) (2006). Lesekompetenz: Bedingungen, Dimensionen, Funktionen. Weinheim: Juventa.
- Guthrie, J.T., Wigfield, A., Metsala, J.L. & Cox, K.E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. Scientific Studies of Reading, 3, 231-256. doi:10.1207/s1532799xssr0303\_3
- Hanushek, E. A. & Wößmann, L. (2006). Does educational tracking affect performance and inequality? Differences- in-differences evidence across countries. Economic Journal, 116, C63–C76. doi:10.1111/j.1468-0297.2006.01076.x
- Hattie, J. A. C. (2002). Classroom composition and peer effects. International Journal of Educational Research, 37, 449-481. doi:10.1016/S0883-0355(03)00015-6
- Hattie, J. A. C. (2009). Visible learning: a synthesis of over 800 meta-analyses relating to achievement. Milton Park: Routledge.
- Heller, K.A. & Perleth, Ch. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision. Göttingen: Hogrefe.

- Helmke, A. & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistung. In F. E. Weinert (Hg.), Psychologie des Unterrichts und der Schule: Enzyklopädie der Psychologie. Serie Pädagogische Psychologie (Bd. 3, S. 71-176). Göttingen: Hogrefe.
- Hertzog, C., Lindenberger, U., Ghisletta, P. & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. Psychological Methods, *11*, 244-252. doi:10.1037/1082-989X.11.3.244
- Ho, D. E., Imai, K., King, G. & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. Political Analysis, *15*, 199-236. doi:10.1093/pan/mpi013
- Huslander, J., Olson, R. K., Willcutt, E. G. & Wadsworth, S. J. (2010). Longitudinal stability of reading-related skills and their prediction of reading development. Scientific Studies of Reading, *14*, 111-136. doi:10.1080/10888431003604058
- Jordon, N. C., Kaplan, D. & Hanich, L. B. (2002). Achievement growth in children with learning difficulties in mathematics: Findings of a two-year longitudinal study. Journal of Educational Psychology, *94*, 586-597. doi:10.1037//0022-0663.94.3.586
- Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge, UK: Cambridge University Press.
- Klicpera, C. & Gasteiger-Klicpera, B. (1993). Lesen und Schreiben - Entwicklung und Schwierigkeiten: Die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit. Bern: Huber.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., Schneider, W. & Stanat, P. (Hrsg.) (2010). PISA 2009: Bilanz nach einem Jahrzehnt. Münster: Waxmann.
- Klieme, E., Eichler, W., Helmke, A., Lehmann, R. H., Nold, G., Rolff H.-G. et al. (2006). Unterricht und Kompetenzerwerb in Deutsch und Englisch: Zentrale Befunde der Studie Deutsch-Englisch-Schülerleistungen-International (DESI). Im Internet verfügbar unter:

- [http://www.dipf.de/desi/DESI\\_Zentrale\\_Befunde.pdf](http://www.dipf.de/desi/DESI_Zentrale_Befunde.pdf) (letzter Zugriff: 28. 3. 2008).
- Klieme, E., Jude, N., Rauch, D., Ehlers, H., Helmke, A., Eichler, W., Thomé, G. & Willenberg, H. (2008). *Alltagspraxis, Qualität und Wirksamkeit des Deutschunterrichts*. In Desi-Konsortium (Hrsg.), Unterricht und Kompetenzerwerb in Deutsch und Englisch: Ergebnisse der DESI-Studie (S. 319-344). Weinheim: Beltz.
- Klieme, E., Schümer, G., & Knoll, S. (2001). *Mathematikunterricht in der Sekundarstufe I. „Aufgabenkultur“ und Unterrichtsgestaltung*. In E. Klieme & J. Baumert (Hrsg.), TIMSS – Impulse für Schule und Unterricht (S. 43–57). Bonn: Bundesministerium für Bildung und Forschung.
- Kolen, M. J., & Brennan, R. L. (2004). Test equating, scaling, and linking: Methods and practices. New York: Springer.
- Köller, O., & Baumert, J. (2001). *Leistungsgruppierungen in der Sekundarstufe I: Ihre Konsequenzen für die Mathematikleistung und das mathematische Selbstkonzept der Begabung*. Zeitschrift für Pädagogische Psychologie, 15, 99-110. doi:10.1024//1010-0652.15.2.99
- Köller, O., & Baumert, J. (2002). *Entwicklung schulischer Leistungen*. In R. Oerter & L. Montada (Hrsg.), Entwicklungspsychologie: Ein Lehrbuch (S. 756-786). Weinheim: Psychologie Verlags Union.
- Kortenbruck, M. (2007). Entwicklung der Lesekompetenz: Einfluss von Motivation und Geschlecht. Unveröffentlichte Diplomarbeit, Freie Universität, Berlin.
- Kunter, M. (2005). Multiple Ziele im Mathematikunterricht. Münster: Waxmann.
- Kunter, M. & Baumert, J. (2006). *Linking TIMSS to research on learning and instruction: a re-analysis of the German TIMSS and TIMSS Video data*. In S. J. Howie & T. Plomp (Hg.), Contexts of learning mathematics and science: lessons learned from TIMSS (S. 335-351). London: Routledge.
- Küspert, P. & Schneider, W. (1998). Würzburger Leise Leseprobe (WLLP). Göttingen: Hogrefe.

- Lee, V. E. & Burkam, D. T. (2002). Inequality at the starting gate: Social background differences in achievement as children begin school. Washington, DC: Economic Policy Institute.
- Lehmann, R. H., Peek, R. & Poerschke, J. (1997). Hamburger Lesetest für 3. bis 4. Klassen (HAMLET 3-4). Weinheim: Beltz Verlag.
- Lehmann, R. H., Peek, R., Gänsfuß, R. & Husfeldt, V. (2001). Aspekte der Lernausgangslage und der Lernentwicklung – Klassenstufe 9 – Ergebnisse einer längsschnittlichen Untersuchung in Hamburg. Hamburg: Behörde für Schule, Jugend und Berufsbildung, Amt für Schule.
- Leppänen, U., Niemi, P., Aunola, K. & Nurmi, J.-E. (2004). Development of reading skills among preschool and primary school pupils. Reading Research Quarterly, 39, 77-93. doi:10.1598/RRQ.39.1.5
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung: Probleme und Lösungen. Psychologische Rundschau, 58, 103-117. doi:10.1026/0033-3042.58.2.103
- McCoach, D. B., O'Connell, A. A., Reis, S. M. & Levitt, H. A. (2006). Growing readers: A hierarchical linear model of children's reading growth during the first 2 years of school. Journal of Educational Psychology, 98, 14-28. doi:10.1037/0022-0663.98.1.14
- McElvany, N. & Artelt, C. (2007). Das Berliner Eltern-Kind Leseprogramm: Konzeption und Effekte. Psychologie in Erziehung und Unterricht, 4, 314-332.
- McElvany, N., Kortenbruck, M. & Becker, M. (2008). Lesekompetenz und Lesemotivation: Entwicklung und Mediation des Zusammenhangs durch Leseverhalten. Zeitschrift für Pädagogische Psychologie, 22, 207-219. doi:10.1024/1010-0652.22.34.207
- Morgan, P. L. & Fuchs, D. (2007). Is there a bidirectional relationship between children's reading skills and reading motivation? Exceptional Children, 73, 165-183.
- Moser, U. & Rhyh, H. (2000). Lernerfolg in der Primarschule: Eine Evaluation der Leistungen am Ende der Primarschule. Aarau: Sauerländer.



- Muthén, B. O. & Muthén, L. K. (1998-2011). Mplus (Version 6.11) [Computer software]. Los Angeles, CA.
- Nauck, J. & Otte, R. (1980). Diagnostischer Test Deutsch (DTD). Braunschweig: Westermann.
- Neumann, M., Schnyder, I., Trautwein, U., Niggli, A., Lüdtke, O. & Cathomas, R. (2007). Schulformen als differenzielle Lernmilieus: Institutionelle und kompositionelle Effekte auf die Leistungsentwicklung im Fach Französisch. Zeitschrift für Erziehungswissenschaft, *10*, 399-420. doi:10.1007/s11618-007-0043-6
- Parrila, R., Aunola, K., Leskinen, E., Nurmi, J.-E. & Kirby, J. R. (2005). Development of individual differences in reading: Results from longitudinal studies in English and Finnish. Journal of Educational Psychology, *97*, 299-319. doi:10.1037/0022-0663.97.3.299
- Pfost, M., Dörfler, T. & Artelt, C. (2010). Der Zusammenhang zwischen außerschulischem Lesen und Lesekompetenz: Ergebnisse einer Längsschnittstudie am Übergang von der Grund- in die weiterführende Schule. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, *42*, 167–176. doi:10.1026/0049-8637/a000017
- Pfost, M., Karing, C., Lorenz, C. & Artelt, C. (2010). Schereneffekte im ein- und mehrgliedrigem Schulsystem: Differenzielle Entwicklung sprachlicher Kompetenzen am Übergang von der Grund- in die weiterführende Schule? Zeitschrift für Pädagogische Psychologie, *24*, 259-272. doi:10.1024/1010-0652/a000025
- Raudenbush, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. Annual Review of Psychology, *52*, 501-525. doi:10.1146/annurev.psych.52.1.501
- Retelsdorf, J. & Möller, J. (2008). Entwicklungen von Lesekompetenz und Lesemotivation: Schereneffekte in der Sekundarstufe? Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, *40*, 179-188. doi:10.1026/0049-8637.40.4.179
- Retelsdorf, J., Becker, M., Köller, O. & Möller, J. (2012). Reading Development in a Tracked

- School System: A Longitudinal Study over 3 Years Using Propensity Score Matching. British Journal of Educational Psychology, 82, 647-671. doi:10.1111/j.2044-8279.2011.02051.x
- Robitzsch, A., Dörfler, T., Pfost, M. & Artelt, C. (2011). Die Bedeutung der Itemauswahl und der Modellwahl für die längsschnittliche Erfassung von Kompetenzen. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 43, 213-227. doi:10.1026/0049-8637/a000052
- Scarborough, H. S. & Parker, J. D. (2003). Matthew effects in children with learning disabilities: Development of reading, IQ, and psychosocial problems from Grade 2 to Grade 8. Annals of Dyslexia, 53, 47–71. doi:10.1007/s11881-003-0004-6
- Seymour, P. H., Aro, M. & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. British Journal of Psychology, 94, 143–174. doi:10.1348/000712603321661859
- Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society. Series B*, 13, 238–241.
- Skibbe, L. E., Grimm, K., Stanton-Chapman, T., Justice, L., Pence, K. & Bowles, R. (2008). Reading trajectories of children with language difficulties from preschool through fifth grade. Language, Speech, and Hearing Services in Schools, 39, 475-486. doi:10.1044/0161-1461(2008/07-0016)
- Snow, C. E. (2002). Reading for understanding: Toward a research and development program in reading comprehension. Santa Monica: RAND.
- Snow, C.E., Scarborough, H.S. & Burns, M.S. (1999). What speech-language pathologists need to know about early reading. Topics in Language Disorders, 20, 48–58. doi:10.1097/00011363-199911000-00006
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21, 360-407.

doi:10.1598/RRQ.21.4.1

Steiner, P. M., Cook, T. D., Shadish, W. R. & Clark, M.H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. Psychological

Methods, 15, 250-267. doi:10.1037/a0018719

Sternberg, R. J. & Powell, J. S. (1983). Comprehending verbal comprehension. American

Psychologist, 38, 878-893. doi:10.1037//0003-066X.38.8.878

Steyer, R. (1992). Theorie kausaler Regressionsmodelle. Stuttgart: Fischer.

Thrupp, M., Lauder, H. & Robinson, T., (2002). School composition and peer effects.

International Journal of Educational Research, 37, 483-504. doi:10.1016/S0883-0355(03)00016-8

Van Ewijk, R. & Slegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. Educational Research Review, 5, 134-150.

doi:10.1016/j.edurev.2010.02.001

Walberg, H.J. & Tsai, S.-L. (1983). 'Matthew' effects in education. American Educational

Research Journal, 20, 359-373. doi:10.3102/00028312020003359

Weinert, F. E. (Hg.) (1997). Entwicklung im Grundschulalter. Weinheim: PVU.

Weiß, R. H. (1998). Grundintelligenztest Skala 2 (CFT 20) mit Wortschatztest (WS) und

Zahlenfolgentest (ZF). 4. überarb. Aufl. Göttingen: Hogrefe.

Williamson, G. L., Appelbaum, M. & Epanchin, A. (1991). Longitudinal analyses of academic achievement. Journal of Educational Measurement, 28, 61-76.

doi:10.1111/j.1745-3984.1991.tb00344.x

## Tabellen

Tabelle 1:

Mittelwert und Standardabweichungen der drei Leistungsdomänen: Gesamtstichprobe, Grundschüler und Gymnasiasten (imputierte Stichprobe).

		Gesamtstichprobe (N = 772)				Grundschüler (N = 700)				Gymnasiasten (N = 72)				d <sup>1</sup>	p
		<u>M</u>	<u>SE</u> <u>(M)</u>	<u>SD</u>	<u>SE</u> <u>(SD)</u>	<u>M</u>	<u>SE (M)</u>	<u>SD</u>	<u>SE</u> <u>(SD)</u>	<u>M</u>	<u>SE (M)</u>	<u>SD</u>	<u>SE</u> <u>(SD)</u>		
Leseverständnis	T1	-0.35	0.05	1.17	0.04	-0.44	0.05	1.15	0.04	0.46	0.17	1.01	0.14	0.77	<.001
	T2	-0.09	0.04	1.03	0.03	-0.18	0.04	1.00	0.04	0.70	0.14	0.96	0.10	0.85	<.001
	T3	0.68	0.04	1.11	0.04	0.58	0.05	1.08	0.04	1.57	0.20	1.02	0.12	0.89	<.001
Wortschatz	T1	-0.64	0.05	1.19	0.04	-0.73	0.05	1.18	0.04	0.18	0.15	0.98	0.17	0.76	<.001
	T2	-0.21	0.05	1.26	0.04	-0.28	0.06	1.25	0.04	0.63	0.18	1.09	0.14	0.72	<.001
	T3	0.93	0.05	1.14	0.04	0.83	0.05	1.10	0.04	1.92	0.23	1.06	0.16	0.96	<.001
Dekodierfähigkeit	T1	38.89	0.55	12.98	0.40	37.90	0.56	12.70	0.42	48.51	1.73	11.57	1.20	0.82	<.001
	T2	48.29	0.53	13.19	0.40	47.33	0.56	13.03	0.43	57.58	1.61	10.69	1.25	0.78	<.001
	T3	62.45	0.69	15.71	0.52	61.17	0.68	15.36	0.51	74.79	2.68	13.40	1.41	0.87	<.001

Anmerkungen:<sup>1</sup> Cohens d: Mittelwertsabstand zwischen Gymnasiasten und Grundschülern relativiert an der Streuung der Gesamtstichprobe zum jeweiligen Messzeitpunkt.

Tabelle 2:  
 Durchschnittliche Leistungszunahme (Differenzwert  $\Delta$ ) bei Grundschulern und Gymnasiasten  
 in den drei Domänen (imputierte Stichprobe).

		Grundschüler		Gymnasiasten		diff( $\Delta_{GY}-\Delta_{Gs}$ )	z	p
		M( $\Delta_{GS}$ )	SE	M( $\Delta_{GY}$ )	SE			
Leseverständnis	$\Delta_{T2-T1}$	0.26	0.05	0.24	0.19	-0.023	-0.10	.919 <sup>a</sup>
	$\Delta_{T3-T2}$	0.76	0.05	0.87	0.19	0.113	0.56	.288 <sup>b</sup>
Wortschatz	$\Delta_{T2-T1}$	0.44	0.05	0.45	0.17	0.018	0.06	.955 <sup>a</sup>
	$\Delta_{T3-T2}$	1.09	0.05	1.40	0.13	0.169	2.23	.014 <sup>b</sup>
Dekodierfähigkeit	$\Delta_{T2-T1}$	9.43	0.44	9.08	1.76	-0.355	-0.19	.847 <sup>a</sup>
	$\Delta_{T3-T2}$	13.84	0.61	17.21	2.09	3.367	1.55	.062 <sup>b</sup>

<sup>a</sup> p-Werte unter zweiseitigem Signifikanztest.

<sup>b</sup> p-Werte unter einseitigem Signifikanztest.

Tabelle 3:  
 Univariate Wachstumskurvenmodelle für Leseverständnis, Wortschatz und  
 Wortdekodiergeschwindigkeit (unspezifiziertes Modell: Faktorladung zu T2 frei geschätzt;  
 imputierte Daten; Notation nach Bollen & Curran, 2006).

	Leseverständnis			Wortschatz			Dekodierfähigkeit		
	Koef. <sup>1</sup>	SE	p	Koef. <sup>1</sup>	SE	p	Koef. <sup>1</sup>	SE	p
Faktorladungen									
$\lambda_1$	0	-		0	-		0	-	
$\lambda_2$	0.27	0.05	<.001	0.27	0.03	<.001	0.40	0.02	<.001
$\lambda_3$	1	-		1	-		1	-	
Varianzen									
$\Psi_{\alpha\alpha}$	0.68	0.07	<.001	1.00	0.10	<.001	125.21	10.57	<.001
$\Psi_{\beta\beta}$	0.06	0.16	.726	0.05	0.14	.730	58.55	4.34	<.001
Kovarianz									
$\Psi_{\alpha\beta}$	-0.05	0.09	.565	-0.12	0.10	.208	3.56	10.09	.724
Mittelwerte									
$\mu_\alpha$	-0.36	0.05	<.001	-0.63	0.05	<.001	38.97	0.49	<.001
$\mu_\beta$	1.01	0.08	<.001	1.53	0.06	<.001	23.22	0.747	<.001
Kovariante Gymnasium									
$b_{T3,GY}^2$	0.20	0.65	.760	0.08	0.65	.907	3.32	3.06	.279
$\underline{\epsilon}_{\alpha,GY}$	0.08	0.02	<.001	0.08	0.02	<.001	0.89	0.17	<.001
$\underline{\epsilon}_{\beta,GY}$	-0.01	0.06	.870	0.01	0.06	.866	-0.02	0.36	.953
Residualvarianzen									
$\theta_{T1=T2=T3}$	0.57	0.05	<.001	0.52	0.06	<.001	43.17	3.41	<.001
Modellfit-Indizes									
$\chi^2$	7.59			3.71			4.97		
Df	2			2			2		
RMSEA	.060			.033			.044		
SRMR	.048			.033			.030		

Anmerkungen:

$\alpha$  = Intercept;  $\beta$  = Slope; RMSEA: Root Mean Square Error of Approximation; SRMR: Standardized Root Mean Square Residual.

<sup>1</sup> Unstandardisierte Lösung.

<sup>2</sup> p-Werte unter einseitigem Signifikanztest..

## Abbildungen

### Abbildung 1:

Schaubilder möglicher Entwicklungsverläufe von Schülerinnen und Schülern an Gymnasien (GY) und anderen Schulformen (GS) (Hypothetischer Übergang nach Klasse 4).

#### Abbildung 1a:

Parallele Leistungsentwicklung zwischen 2. und 6. Klasse.

#### Abbildung 1b:

Konstant divergierende Leistungsentwicklung zwischen 2. und 6. Klasse.

#### Abbildung 1c:

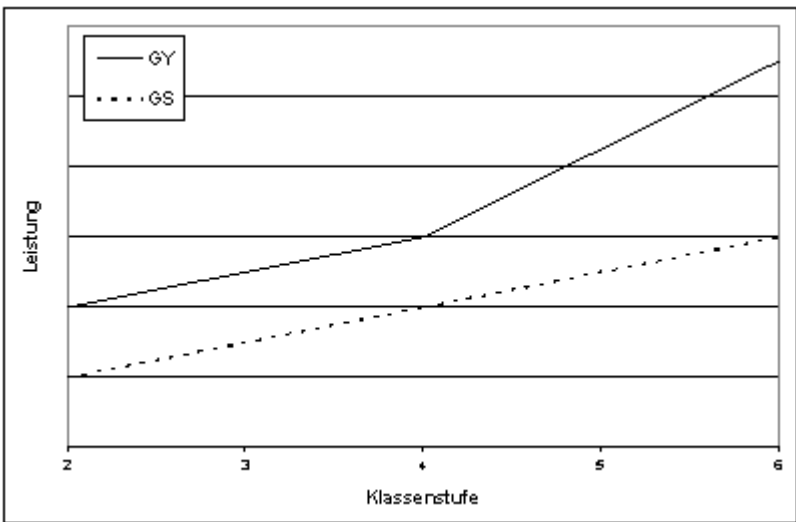
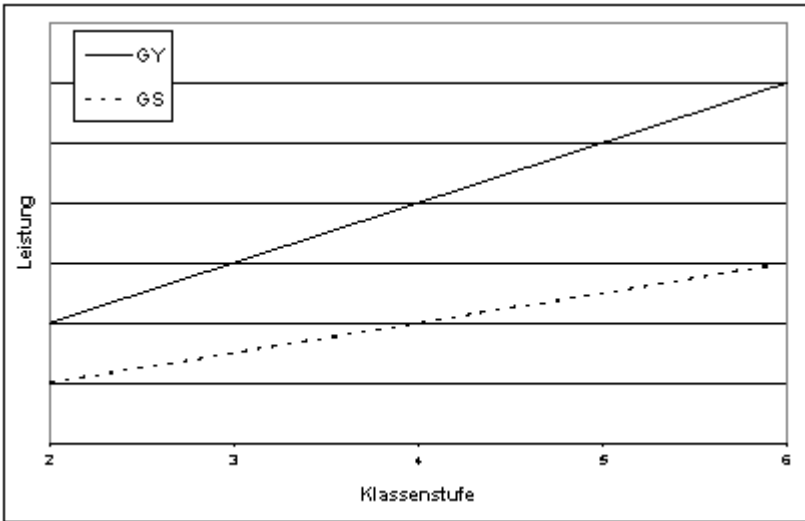
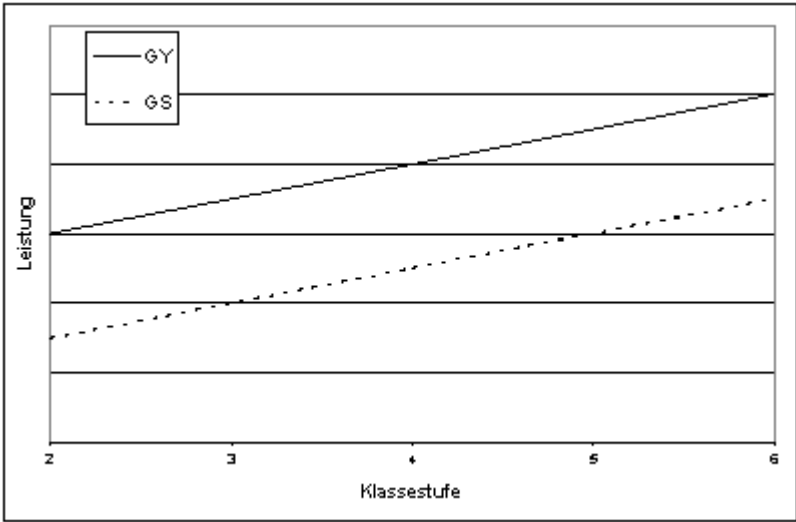
Akzeleriert divergierende Leistungsentwicklung zwischen 4. und 6. Klasse.

### Abbildung 2:

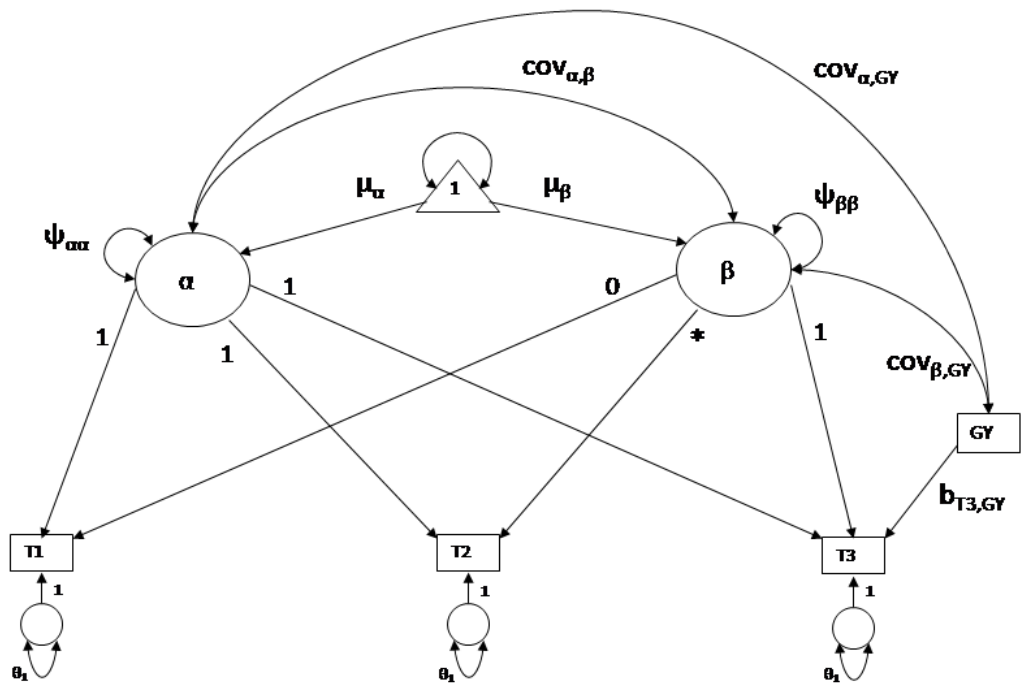
Univariates Wachstumskurvenmodell mit latentem Level- und Slope-Faktor mit zusätzlicher Regression der Testleistung zu T3 auf das Gymnasium (Notation nach Bollen & Curran, 2006).

#### Anmerkungen:

$\alpha / \beta$  = latenter Intercept bzw. Slope;  $\mu_{\alpha} / \mu_{\beta}$  = Mittelwerte Intercept bzw. Slope;  $\psi_{\alpha\alpha} / \psi_{\beta\beta}$  = Varianzen für Intercept bzw. Slope; latenter Faktoren Intercept bzw. Slope  $\theta$  = Residualvarianzen.







Autorenhinweis:

Die Autorin und die Autoren danken Herrn Jürgen Baumert für die Möglichkeit, die Berliner Leselängsschnittstudie zu realisieren, sowie Frau Cordula Artelt für ihre Beratung bei der Studiendurchführung. Herzlicher Dank gilt auch Frau Susannah Goss, Frau Kathrin Jonkmann, Frau Michaela Kropf und Frau Monika Oppong für die Unterstützung bei der Datenerhebung und Manuskripterstellung.